Eye movement control during single-word reading in dyslexics

Manfred MacKeben
The Smith-Kettlewell Eye Research Institute, San Francisco, California, USA

Susanne Trauzettel-Klosinski
Department of Pathophysiology of Vision and Neuro-ophthalmology, University Eye Hospital, Tübingen, Germany

Jens Reinhard
Department of Pathophysiology of Vision and Neuro-ophthalmology, University Eye Hospital, Tübingen, Germany

Ute Dürrwächter
Department of Child and Adolescent Psychiatry and Psychotherapy, University of Tübingen, Germany

Martin Adler
Department of Pathophysiology of Vision and Neuro-ophthalmology, University Eye Hospital, Tübingen, Germany

Gunther Klosinski
Department of Child and Adolescent Psychiatry and Psychotherapy, University of Tübingen, Germany

We investigated whether dyslexics make instantaneous automatic adjustments of reading saccades depending on word length. We used a single-word reading paradigm on 10 dyslexic and 12 normally reading children aged 11-15 years. Eye movements were recorded by scanning laser ophthalmoscope (SLO) while subjects read single words of different length aloud. All subjects passed standardized prescreening tests, which included a reading test, to exclude those with discernible deficits of eyesight, oculomotor skill, or intellectual development. We measured number, direction, frequency, and amplitude of saccades, as well as the durations of inter-saccadic intervals, as functions of word length. The results show that word length influences the number and amplitude of reading saccades in both subject groups, but this relationship showed quantitatively significant group-specific differences: Both groups showed a gradual increase of the saccade amplitudes in either direction dependent on word length, but the gain of this function was significantly lower in the dyslexics. The durations of holding phases between saccades were significantly longer in the dyslexics, and accordingly, we found a lower rate of occurrence of saccades per unit time in the dyslexics. Forward saccade amplitudes showed no correlation with the duration of the preceding or following holding phases in either group. The data show that the mechanisms enabling dyslexics to make instantaneous adjustments of reading saccades depending on word length are present but quantitatively impaired. This supports the view that these adjustments may help dyslexics to increase reading speed, but that they cannot utilize them to the same extent as normal readers.

Keywords: dyslexia, perceptual span, saccades, inter-saccadic, language processing

Introduction

Reading has been investigated in numerous studies and fundamental insights regarding the contributing brain mechanisms have been gained (Pugh et al., 2000; Salmelin, Helenius, & Service, 2000; Simos et al., 2000). Much of the previous work concentrated on reading eye movements as the most observable manifestation of this complex process (e.g., Bouma, 1973; Bouma & de Voogd, 1974; Legge, Pelli, Rubin, & Schleske, 1985, Legge, Ahn, Klitz, & Luebker, 1997; McConkie & Rayner, 1976; O’Regan, 1980; Rayner & McConkie, 1976; Rayner & Pollatsek, 1981).

We have used eye movement recordings as a tool to investigate fundamental aspects of oculomotor control during reading and their differences between normal readers and dyslexics. The most common measure of reading performance is speed, and accomplished readers can reach levels of 200 words per minute (WPM) and more while reading silently, whereas dyslexics hardly exceed 100 WPM, even after years of training. During the time-consuming process of learning to read, some children have to fight overwhelming odds before they get diagnosed as “dyslexic.” Because of the great psychological, societal, and economic impact of a reading impairment, the problem has been researched extensively, and the results support different explanations for...
dyslexia. These include deficits affecting phonemic awareness (Bradley & Bryant, 1983; Rudel, Denckla, & Broman, 1978; Snowling, 1980; Swan & Goswami, 1997; for a review, see Nijokiikijten, 1994), the grapheme to phoneme conversion (Golden & Zemhauer, 1983), processing of visual signals (Di Lollo, Hanson & McIntyre, 1983; Jacobs, 1986; Lennerstrand & Ygge, 1992), selective deficits in the magnocellular stream of visual input (Best & Demb, 1999; Borsting, Ridder III, Dudeck, Kelly, Matsui, & Motoyama, 1996; Demb, Boynton, Best, & Heeger, 1998a; Demb, Boynton, & Heeger, 1998b; Stein & Walsh, 1997), asymmetric crowding (Geiger & Lettvin, 1987), deficits of attention (Stein & Walsh, 1997; Steinman, Steinman & Garzia, 1998), as well as the much debated role of oculomotor deficits (Rubino & Minden, 1973; pro: Biscaldi, Fischer, & Aiple, 1994; Biscaldi, Fischer, & Hartnegg, 2000; Eden, Stein, Wood, & Wood, 1994; Pavlidis, 1991; contra: Brown, Haegerstrom-Portnoy, Adams, Yingling, Galin, Herron, & Marcus, 1983; Black, Collins, De Roach, & Zubrick, 1984; De Luca, Di Pace, Judica, Spinelli, & Zoccolotti, 1999; Olson, Conners, & Rack, 1991; Stanley, Smith, & Howell, 1983).

Consequently, many subtypes of dyslexia have been described (Boder, 1973; Nijokiikijten, 1994; Stark, Giveen, & Terdiman, 1991). A recent study of our own also has provided experimental evidence that a phonemic deficit may underlie the problem in all dyslexics, which can be made worse in some who show an additional visual/eidetic deficit that impairs pictogram-naming performance (Trauzettel-Klosinski, MacKeben, Reinhard, Feucht, Dürrwächter, & Klosinski, 2002). For the current study, however, we did not distinguish between subtypes.

Readers scan the text by a pattern of eye movements that is characterized by horizontal saccades and holding phases between them. (We elect not to call the latter “fixations,” because they lack the characteristic fixational eye movements (Cunitz & Steinman, 1969). Neither the saccades nor the holding phases are always quantitatively the same, so that the question of what controls their variations has been the focus of much discussion. It is widely accepted that linguistic and phonemic factors exert a dominating influence, but visual characteristics are important also (Bouma & de Voogd, 1974; Jacobs, 1986; Rayner, Sereno, & Raney, 1996; Snowling, 1980; Whittaker & Lovie-Kitchin, 1993; Wolf & Bowers, 1999).

A central factor in trying to understand a reading impairment that is common in countries using alphabetic languages is the fact that words consist of strings of letters. Groups of letters have to be identified as syllables, which together form words, whose phonemic equivalent must be retrieved from memory. Although the use of letters seems to be at the heart of the problem, the ability to quickly recognize single letters will not make a person a fluent reader. This is supported by the fact that reading letter-by-letter slows down reading, which can be seen in normal beginning readers or in patients with extremely restricted visual fields (Trauzettel-Klosinski & Reinhard, 1998). Rather, it is the ability to let the eye jump from one group of letters to the next that can make the process fast. Reading speed and fluency are thus achieved by sequentially placing strings of letters (parts of words, complete words, or even groups of words) on a central, asymmetric horizontal strip of the central visual field, a dynamic behavioral measure called the “perceptual span” (McConkie, & Rayner, 1975; McConkie & Rayner, 1976; Taylor, 1957; Tinker, 1958).

In addition, readers need the ability to let their gaze rest in one place just long enough to take in the information needed for recognition and for gauging the next eye movement. This process requires simultaneous acquisition of visual information from more than one letter and, thus, possibly parallel processing (Rayner, 1983, p.102; Schroyens, Vitu, Brysbaert, & d’Ydewalle, 1999) of which even some dyslexics seem to be capable (Van Strien, Bouma, & Bakker, 1993), or parallel saccade programming (Henderson & Ferreira, 1993; Morrison, 1984). The duration of the oculomotor holding phases between saccades have been found to be longer in dyslexics than in controls performing search tasks (Olson et al., 1991), in a subgroup of dyslexics in an oculomotor task (Biscaldi et al., 1994) and during reading (Adler-Grinberg & Stark, 1978). In addition, Ciufreda, Bahill, Kenyon, and Stark (1976) reported that one "slow reader" (case 2, page 391) had increased "fixation durations." On the other hand, Rubino and Minden (1973) found no such differences. Clearly, the question whether inter-saccadic holding phases are longer in dyslexics warrants further research.

Thus, where and when to place gaze on the text during reading might be a function of the width of the perceptual span and of gauging each saccade amplitude accordingly (Jacobs, 1986). It has been shown that linguistic and visual aspects of the reading material contribute to the control of this process (O’Regan, 1979, 1980; Rayner & Pollatsek, 1981). A particularly interesting fact is that the brain can adjust saccadic amplitudes according to word length: O’Regan (1980) reported that normal readers tend to use longer saccades within the word when reading a long word and shorter ones to read a short word. The mechanism controlling this process has not been identified yet, nor has it been shown whether it is present in dyslexics also. As an instantaneous and automatic adjustment of saccade amplitude has the potential to make reading faster and more fluent, one could speculate that a lack of such a mechanism might be one of the reasons underlying diminished reading speed in dyslexics. Thus, our investigation was undertaken to find answers to the following questions:

1. Is there an instantaneous and automatic adjustment of reading saccade amplitudes to word length in dyslexics?

2. If yes, is the mechanism quantitatively operating at the same level as in normally reading children of comparable chronological age?
3. Is the increased absolute number of reading saccades reported in dyslexics, at least among others, caused by an increase of the number of saccades per unit time?

4. Are the holding phases of the dyslexics’ eye movements during reading longer than in controls?

We base the rationale for our study design on the facts outlined above. Hence, we hypothesize that we will be able to manipulate the saccade gauging mechanism directly by varying word length to create a base for comparison with previous work (O’Regan, 1980). Restricting the reading material to single words will minimize linguistic control, because it minimizes guessing or lexical inference based on knowledge of the dictionary (Mansfield, & Legge, 1999) or on knowledge of the context. The fact that the words will be read aloud will monitor in real time whether the word was read correctly.

We concentrated on subjects fulfilling the criteria defined by ICD10 (World Health Organization, 1996) and those for "specific dyslexia" (see Stark et al., 1991) to limit the heterogeneity of the experimental group. We did not include those whose poor reading performance could be accounted for by any other explanation (e.g., deficits of the visual or oculomotor system, a primary speech impediment, or an impeded development of intelligence or general cognitive functions). Second, we used text reading speed as the distinguishing criterion between subject groups.

Partial and preliminary results of this study have been previously communicated (Trauzettel-Klosinski, Klosinski, Sadowski, & Tornow, 1999; Trauzettel-Klosinski, Reinhard, Klosinski, & MacKeben, 1999).

The experimental design included four steps:

**Step 1** tested reading speed for continuous text with subsequent assignment to the dyslexic or the control group.

**Step 2** consisted of a psychological and psychiatric evaluation of all subjects assigned to the dyslexic group.

**Step 3** investigated the visual capabilities by a complete ophthalmological and neuro-ophthalmological examination of the members of the both groups.

**Step 4** consisted of the main experiment: reading single words by dyslexic and control subjects.

## General methods

### Group assignment and pretesting

The diagnosis of dyslexia was based on the criteria defined by ICD10 (20), so that the results of reading and writing tests were at least 2 SDs below the level to be expected based on the age and IQ of the child.

The candidates for the dyslexic group were further pretested to eliminate additional disorders as explanations for their poor reading performance. The assessment was based on standardized reading tests, such as the Zürich reading test (Linder & Grissemann, 1980), intelligence tests (the Wechsler Intelligence Scale for Children [WISC-R, Tewes, 1983] or the Colored Progressive Matrices [Raven, Court, & Raven Jr., 1980], spelling tests (the Diagnostic Orthography Test, 1980) for the appropriate age groups, and attention tests (e.g., “Test d2”, Brickenkamp, 1981).

### Subjects

As a result of the pretests for the different deficits, three children were excluded from the list of subjects due to the presence of a brain tumor that could affect reading, congenital nystagmus, and unstable fixation. Two groups of subjects were selected for further experiments, 10 dyslexics and 12 normally reading children, aged 11 to 15.5 years.

### Instrumentation

All data on reading text and single words were obtained during monocular viewing using a scanning laser ophthalmoscope (SLO) (model 101, Rodenstock). The stimuli were scanned directly onto the retina by a laser beam that was modulated by an acousto-optic modulator. Thus, the SLO records the absolute position of the fovea and that of the stimulus with a spatial resolution of < 5 arcmin. This is based on the relative accuracy by comparison with a reference frame, which can be made with +/- 1 pixel accuracy, corresponding to +/- 2 arcmin. The video frequency of 50 fields/s allowed field-by-field analysis with a temporal resolution of 20 ms. The main advantage of this method is that it allows reliable recordings that are not dependent on a calibration, because fovea and stimulus are visible simultaneously and their distance can easily be converted to units of visual angle off-line. In addition, the method provides an automatically synchronized sound recording.

### Reading material

For the preliminary reading tests, the SLO scanned paragraphs of text into the eye that came from a simple German story of approximately fourth-grade equivalent. The subjects saw the text as black on a bright red background of 3.6 × 10^4 trolands at 97% Michelson contrast and read the text aloud. Each paragraph contained 6–7 lines of between 18.1 deg- and 19.7-deg line length. The font was Times New Roman and capital letters subtended 0.64 deg (38.4 arcmin), which corresponds to 1.4 times the size of newspaper print at 25-cm distance.

### Statistics

All statistical tests were performed by StatView (v 4.51) software on a personal computer. We used means and SDs only for the analysis of the preliminary oculomotor experiments, because they are sensitive to outliers, which was desirable here, and because their use for this purpose is widely accepted.
For the statistical analysis of the text reading test and the main experiment, we made no assumptions regarding a normal distribution of the data and, thus, used non-parametric methods. This precaution was confirmed by visual comparison with model Gaussian curves, which showed that some of the data deviated substantially from normal distributions. Hence, all results of the main experiment are reported as medians with the inter-quartile range (IQR) as measure of variability. Group comparisons were performed by the Mann-Whitney U-test (MWU), and Spearman’s rho was used as a measure of correlation.

In cases where printed scattergrams indicated a strong linear relationship, we used linear regression analysis to calculate the coefficient of determination $r^2$.

**Procedure**

The recordings were performed on the dominant eye, as determined by a peephole test. In the preliminary reading tests, the children read continuous text aloud. All experiments were performed at the Department of Pathophysiology of Vision and Neuro-Ophthalmology, University Eye Hospital, Tübingen, Germany.

### Preliminary experiments

Two preliminary experiments were designed (1) to separate the subject groups by their defining variable under controlled conditions and (2) to quickly assess whether any of the dyslexic group showed an obvious deficiency of oculomotor control, which has been reported as a possible origin of reading problems (see Biscaldi et al., 1994; Pavlidis, 1991). In Experiment 1, each subject read text aloud that was scanned onto the retina by the SLO, while the movements of the retina were recorded. In Experiment 2, subjects fixated a central mark for 10 s to test fixation stability, made 2–4 goal-directed saccades to each of three marks at 5° right and 5° left and back to the center to test saccade accuracy for 10–20 s, and maintained fixation on each mark for at least 1 s.

#### 1. Text reading speed

Both groups of children read aloud a minimum of three 6–8 line paragraphs of German text. Members of the control group read the texts at a speed of 114–177 WPM with a median of 143.0 (IQR = 33.0). The speeds were, as expected, much slower in the dyslexic children, ranging from 24 to 89 WPM with a median of 45.0 WPM (IQR = 27.25). Due to our selection criteria, the data from the two groups showed no overlap, and the difference was statistically highly significant (MWU, $p = .0001$).

While the normal readers showed a moderate correlation of reading speed with age ($\rho = 0.593$), which could be expected according to McConkie, Zola, Grimes, Kerr, Bryant, & Wolff (1991), the dyslexic showed only a very weak one using the same texts ($\rho = 0.192$).

#### 2. Stability of fixation and accuracy of ± 5-deg saccades

Fixation stability was measured as the SD of the mean eye position calculated as $(x^2 + y^2)^{1/2}$ for 10-s fixation intervals (i.e., in a continuous sequence of 500 video fields). The ranges of SDs overlapped widely (dyslexics: 0.075 – 0.213 deg; controls: 0.050 – 0.166 deg), group averages were close (dyslexics: ± 0.135 deg; controls: ± 0.097 deg), and the difference was not statistically significant (MWU-test, $p = .065$).

Saccadic accuracy was measured as the mean eye position after 5-deg saccades to the right and left and middle. Group averages were for the dyslexics ($Left = -5.23$ deg; $Middle = -0.10$ deg; $Right = 4.91$ deg) and for the controls ($Left = -5.09$ deg; $Middle = -0.07$ deg; $Right = 4.99$ deg). The difference for none of the positions was statistically significant ($p > .3$).

Gaze stability after 5-deg saccades was measured as mean SDs of the mean horizontal components at the same three positions: 5° left, middle, and 5° right: Group averages of the SDs showed small but not statistically significant differences (MWU, $p > .4$): dyslexics: $Left = 0.29$; $Middle = 0.24$; $Right = 0.18$; controls: $Left = 0.20$; $Middle = 0.23$; $Right = 0.21$.

These results allow the tentative conclusion that the small and not statistically significant differences of oculomotor accuracy found here are not likely to be the cause of these children’s dyslexia and cannot account for the large differences in reading performance.

### The main experiment

#### Additional methods

**Targets**

In the main experiment, only single German words were used. Each subject read three words of each of the following lengths: 2, 3, 4, 5, 6, 7, 8, 10, 12, and 14 letters per word. Thus, each subject read 30 words aloud while they were scanned onto the retina by the SLO. Before each trial, the subjects fixated a central fixation mark, which allowed us to vary the alignment of the target words. Thus, one of the three words at each of the 10 word lengths was right aligned, one was left aligned, and one was centered on the fovea. Figure 1 shows the SLO picture of a subject’s retina and a target word as seen by the examiner.

**Instructions**

The subjects were instructed to read each word aloud. No special instructions regarding speed or fluency were given. Each trial started with the appearance of the word and ended when the subject had articulated the word correctly.
Time-to-word recognition

The differences in reading speed were obvious not only while reading text but also during reading single words. This was measured as time to recognition (target appearance to beginning of articulation). The overall time that subjects needed to read all 30 words varied greatly in the dyslexics (50 to 122 s, median = 71.5, IQR = 35.0), but little in the controls (40 to 50 s, median = 45.0, IQR = 9.0). The two distributions touched, but did not overlap, and the group difference was statistically highly significant (MWU, \( p = .0001 \)).

When the short words that are very frequent in German were excluded from the analysis, the coefficient of determination turned out to be only 0.068, indicating a negligible influence of word frequency.

The control subjects read each word in 1-2 s, regardless of length. The values for the dyslexics, on the other hand, varied substantially. To reduce the noise in the data, we compared recognition times for short words (2-4 letters/word), for medium length words (6-8 letters/word), and for long words (10-14 letters/word). The group differences were highly significant in all three word-length classes, with \( p < .0014 \) to \( p < .0001 \). Comparing the medians of the time taken to read the three shortest versus the three longest words [in s/word], the value for the control group increased from 1.3 to 1.8 s/word (IQR = 0.14 and 0.28, respectively). The dyslexics showed a steeper increase from 1.5 to 3.6 s/word (IQR = 0.26 and 1.91, respectively).

Data analysis

Continuous images of the retina and stimuli were recorded on videotape together with a vertical interval time code (VITC). This allows giving every frame of video its own unique identifying number.

Analyses of the data were performed off-line in two stages: The primary analysis used a semiautomatic computer program based on retinal landmarks (vessel branchings). It calculated the horizontal and vertical coordinates of the foveola relative to the target image. The video frequency of 50 fields/s allowed field-by-field analysis with a temporal resolution of 20 ms. Between 2,000 and 6,100 video fields were analyzed for each subject, depending on the time spent to read all 30 words.

In the second stage, time functions of the horizontal eye position component were printed and analyzed graphically by ruler in millimeters (estimated accuracy of judgment, 0.2 mm), which were then translated into minutes of arc. Thus, the accuracy of measurement was equivalent to 2 arcmin, which is below SLO resolution. In addition, the width of a lower case "n" (here called n-space = 28.8 arcmin) was universally used as an average letter space. A regressive (backward = leftward) saccade was analyzed only if it could be considered part of the reading strategy (i.e., if it was either followed by at least one more forward saccade and if it was not a "reset" saccade at the end of an episode that returned the eye to the position where it had been at target appearance). Figure 2 illustrates the data analysis schematically.

Results of the main experiment

We report a detailed analysis of the characteristics of eye movements made during reading of 30 words of different lengths by 22 subjects (i.e., 660 words all together). Responses of the different variables will be reported one at a time below.
The absolute maximum was 18 s for a 14-letter word, during which the subject’s eye “stepped” through the entire word four times (three of which are shown in Figure 3). This strategy of going back to start over at or near the word beginning was seen in 7 of the 10 dyslexics, whereas the other 3 tended to spend more time making many small forward saccades.

Number of forward saccades per word

We investigated the number of forward (progressive) saccades (“F-saccades”) necessary to correctly read all 30 words. The dyslexics made more F-saccades to read them (21 to 232, median = 90, IQR = 127) than the control group (26 to 111, median = 54, IQR = 42). The number of F-saccades was a function of word length. The dependence of the number of F-saccades per word is well described by a linear relationship ($r^2 = 0.98$ for the dyslexics and $r^2 = 0.96$ for the controls), with a steeper slope for the dyslexics than for the controls. Group differences were tested for each word length separately, and they were found to be statistically highly significant for all words of five or more letters (MWU, $p < .007$).

This could, at least in part, be caused by the cumulative effect of the many times that some dyslexics made a return sweep of 80% or more of the word length and started over trying to read the word. Therefore, we separately counted the F-saccades made only during the first pass through a word for the dyslexics. If a backward saccade that did not qualify as a return sweep occurred, subsequent F-saccades were counted only if they led the fovea further rightward than it had been before. The dependence of the number of F-saccades on word length was still linear. To simplify comparison, we formed three groups of word lengths: short (2–4 letters), medium (5–8), and long (10–14). The median numbers of F-saccades were found to be almost identical in both subject groups for short words (0.29 [controls] vs. 0.30 [dyslexics]). For medium and long words, however, the medians in the dyslexics (0.41 and 0.43) were approximately double of those for the controls (both 0.23).

Rate of occurrence of forward saccades

To investigate whether the higher number of F-saccades could be due to a higher absolute rate of occurrence per unit time, we divided the total number of F-saccades made by each subject by the sum of seconds spent reading all 30 words. This quotient showed an increase from the shortest to the longest words by a factor of 4 in the dyslexics (0.45 to 1.89 saccades/s) and a factor of 2.3 in the controls (0.84 to 1.94 saccades/s). The relationship was again approximately linear ($r^2 > 0.87$ for both groups), but the difference between the groups was not statistically significant (MWU, $p = .82$).

For further conclusions regarding the rate of occurrence per unit time, see also Duration of holding phases.

Forward saccade amplitudes

The amplitudes of all F-saccades were measured for both groups (dyslexics: $N = 1076$; controls: $N = 619$). The medians at the different word lengths varied between 17 and 68 arcmin in the dyslexics (equivalent to 0.6 and 2.2 n-spaces), and between 20 and 106 arcmin in the controls (0.7 to 3.5 n-spaces). Their dependence on word length (see Figure 4) was positive in all subjects of either group and...
could be described as linear in the controls ($r^2 = 0.98$), and best for the dyslexics as a second-order polynomial ($r^2 = 0.96$). The slopes of these functions varied more widely in the dyslexics (1.69 to 8.3, median = 3.3, IQR = 2.89) than in the controls (6.35 to 9.99, median = 7.03, IQR = 1.45). They showed no appreciable correlation with either text reading speed or subject age. The group differences at word lengths of 10 letters per word and above were statistically highly significant ($p < .003$).

**Eye movement stops near the word end**

None of the subjects moved the fovea to the last letter of left-aligned and centered words, but rather stopped some distance before. We measured this distance-to-word-end (DWE) for these words from four-letter length on upward. These data contain only 16 such episodes per subject (eight word lengths, two kinds of alignment). A schematic example of this behavior is shown in Figure 2, whereas Figure 3, top left, shows data from a control subject (see first three words).

The median DWE across subjects for each word length ranged from 34 to 51 arcmin (1.2 to 1.8 n-spaces) in the dyslexics, and from 41 to 125 arcmin (1.4 to 4.4 n-spaces) in the control group. The group difference was statistically significant ($p < .0025$). For the longest words of $\geq 10$ letters length, DWE in the controls was always at least double as large as in the dyslexics. DWE correlated with word length only weakly in the dyslexics (rho $= 0.373$), much less than F-saccade amplitudes, which was not so in the control group (rho $= 0.82$), where the relationship could be reasonably fit by a linear function ($r^2 = 0.76$).

**Number of backward saccades**

As can be expected, the dyslexics made many more backward (regressive) saccades ("B-saccades") during the reading episodes of all 30 words (369 vs. 170), with a median of 35 (IQR = 37) per subject (range, 9 to 72). The corresponding values for the control group were 15 (IQR = 7) per subject (range, 1 to 27). While the number of B-saccades increased with word length in the dyslexics, it did not in the control group. Group comparison showed statistically significant differences for all word lengths from seven letters upward ($p < .05$).

**Backward saccade amplitudes**

As in the F-saccades, the median amplitudes of the B-saccades were larger in the control group and ranged from 10 to 92 arcmin, whereas those for the dyslexics ranged
from 20 to 61 arcmin. In both groups the amplitude increased with word length, a tendency that ended at a word length of 10 letters in the dyslexics (see Figure 5).

### Duration of holding phases

The dyslexic subjects showed longer median durations of the holding phases between forward saccades (124 to 877 ms, median = 297, IQR = 158, N = 629) than in the controls (74 to 581 ms, median = 250, IQR = 110, N = 372), although the distributions overlapped (see Figure 6). The difference was statistically highly significant (MWU, \( p < .0001 \)). This shows that the rate of occurrence of F-saccades was, in fact, lower in our dyslexics than in the control group.

Investigating the relationship of holding durations to word length like O'Regan (1980), we found a moderate dependence only in 10/12 members of our control group, with rho ranging between 0.02 and −0.66 (median = −0.36, IQR = 0.485). In the dyslexics, on the other hand, we found that the dependence of duration of holding phases on word length was entirely inconsistent: Five subjects had a mild to moderate negative, four had a similar positive correlation, and one had none (rho < 0.06). This correlation varied between 0.595 to −0.717 (median = 0.111, IQR = 0.485), which was larger than in the control group. The medians for the subgroups of dyslexics with positive/negative rho were 0.349 and −0.367, respectively. Furthermore, word length dependence correlated moderately with text reading speed (rho = −0.538), but not with subject age.

We hypothesized that a long F-saccade might be fostered by a long preceding holding duration, which allowed processing of more letters that lay ahead. In this case, any F-saccade amplitude should show a correlation with the duration of the preceding holding phase. We tested this relationship in 1,146 pairs of values and found that the results varied strongly between individuals, regardless of subject group. Only weak correlation coefficients were found in both, the dyslexics (rho = −0.357 to 0.262) and the controls (rho = −0.481 to 0.355).

This hypothetical relationship was then reversed, and the same calculation was performed for each forward saccade and the duration of its following holding phase. The rationale was that a longer F-saccade brings more letters into the perceptual span for processing during the next holding phase than a short one, and that this might necessitate a longer holding phase to allow adequate time for processing.

Hence, we tested the correlation of all F-saccade amplitudes with the durations of the following holding phases. We printed the individual data sets out as scattergrams and visually checked for obvious nonlinear relations, but none was found. Neither group showed an appreciable correlation, with rho between −0.264 and 0.280 for the dyslexics and between −0.334 and 0.219 for the controls.

![Figure 6](https://jov.arvojournals.org/pdfaccess.ashx?url=/data/journals/jov/932832/) Figure 6. Group comparison of median holding durations (intersaccadic intervals) while reading single words (yellow bars = dyslexics, blue bars = controls). Note the overlap between the groups. The difference between the group medians is statistically highly significant (MWU, \( p < .0001 \)). Regarding word length in relation to frequency and familiarity, see “Discussion” and the legend of Figure 4.
The influence of phonemic/linguistic factors

We varied the influence of these factors here only by manipulating word length, while tables assessing the vocabulary of young German teenagers were not available to us at the time. The frequency of occurrence of each word in the language is likely to be of lesser importance, because Inhoff and Rayner (1986) used a comparable paradigm in their “oneword-window condition” and found no influence of word frequency on saccade size or first fixation duration (see their p. 435). However, as their paradigm was not strictly identical to ours, we cannot completely rule out the possibility that differences in word frequency may have influenced the measured eye-movement parameters. This will be investigated in a subsequent study.

Familiarity with a word is not likely to have played as important a role here as it would for children reading English, due to the straight-forward rules of pronunciation in German. This was shown by experiments on children reading German (Wimmer, 1996; Wimmer & Goswami, 1994) and similarly unambiguous languages such as Italian (De Luca et al., 1999) and Spanish (Goswami, Gambert, & de Barrera, 1998). However, a slight influence on the eye movement parameters investigated here cannot be ruled out.

Time to recognition

It is not surprising that the sum of the durations of single-word reading episodes in our dyslexic group was significantly longer than that in the control group. It reflects the “struggle” to read (see Figure 3), which often involved starting over at the beginning of the word. Although the median durations that were needed to read words were significantly longer in the dyslexics (p < .002), they were close for words up to four letters (1.48 s [dyslexics] vs. 1.28 s [controls]), and then parted increasingly dependent on word length. As this was achieved by an increasing number of saccades, this may mean that having to combine information from two or more holding phases constituted an additional level of difficulty for the dyslexics.

Saccade strategy and frequency

The almost identical F-saccade/word ratios for short words in both groups correlated well with the fact that the recognition times for these words were quite similar for both groups. On the other hand, a dependence on word length could be expected (Tinker, 1958), and the much higher ratios in the dyslexics for medium and long words agree with the fact that the recognition times for longer words were significantly longer in the dyslexics.

The high coefficient of determination (r²) of the number of F-saccades versus word-length functions for 8 out of 10 dyslexic subjects confirms that the oculomotor strategy chosen for long words has a strong influence on reading text. Any opinion about why the two children with the smallest F-saccade/word ratios had so strongly divergent reading performance can, at this time, only be speculative. It is of interest, though, to note that the subject with the lowest ratio showed, clearer than any other subject, the tendency to perform reverse reading eye movements, as described by Zangwil and Blakemore (1972).

The fact that there was a significant difference between the groups in regard to the rate of occurrence of saccades was a consequence of the significantly longer holding durations. The fact that the absolute number of saccades performed by the dyslexics to solve the same task was much higher was caused by the longer time they spent to read each word with many shorter saccades. However, the relative proportion of their median number of forward saccades to the median number of backward saccades was only insignificantly higher than that for the control group (68.9% vs. 64.4%, respectively). In addition, the constituent individual values showed a wide range of inter-individual variation, both of which agree with Olson, Conners, and Rack (1991), who found no substantial difference in this regard between groups, albeit matched by reading-level.

Forward saccade amplitudes

The adaptation of F-saccade amplitude to word length in dyslexics is a new finding that constitutes the most important result of this study. The function turned out to be approximately linear for the control group, which confirms and adds much detail to the original finding by O’Regan (1980) in normal adults. The same function in the dyslexics saturates at an amplitude of approximately 1 deg (2 n-spaces, see Figure 4), which, nonetheless, shows that a mechanism for adapting the gain of reactive, stimulus-triggered saccades (Deubel, 1995) is active during reading in dyslexics also, but that it is quantitatively less developed. It might have an important practical consequence, because it opens the possibility to develop perceptual or attentional training methods to increase the gain of this function goal directly, from which dyslexic children could benefit (Beard, Levi, & Reich, 1995; Chung, Legge, & Cheung, 2004; Fahle, & Henke-Fahle, 1996; Fahle & Luberichs, 1995; Geiger, Lettvin, & Fahle, 1994).

Tentatively, the adjustment of F-saccade amplitude to word length may reflect a reading strategy based on the density of information in short words, which can be much higher than in a long word: Note the differences between words like “then,” “them,” “thin,” and “than,” analogous to “über,” “üben,” “oben,” and “eiben” in German. This makes reading short words in small saccadic steps more advisable for those who stress accuracy over speed, such as the children in our test situation, especially the dyslexics. In contrast, longer words permit higher reliance on recognition of entire suffixes or prefixes such as “-ous,” “-ward,” “con,” and “sub,” analogous to “-keit,” “-lich,” “ent-,” and “ver-” in Ger-
man. As our use of single words ruled out grammatical and contextual inferences, our results highlight the importance of word length for gauging reading eye movements particularly.

According to McConkie and Rayner (1976), the perceptual span is a behavioral measure based on the number of letters in a gaze-contingent display and the resulting reading speed. In normal, skilled readers, it can extend up to 15 letters to the right of the fovea. Two factors are likely to play a role in limiting F-saccade amplitudes in our subjects: (1) reading skills due to lower age, and (2) the fact that our subjects did not read continuous text. It is noteworthy, though, that the largest F-saccade amplitudes in the dyslexics were approximately half as large as the ones in the control group (see Figure 4). We tentatively interpret the shorter F-saccades as an indication of a more timid gaze strategy, which in the paradigm used by McConkie and Rayner would have shown as a narrower perceptual span. Consequently, it is likely that the number of letters processed in any one holding phase in the dyslexics is smaller. As our dyslexics had times to recognition for short words (2–4 letters) quite comparable to those in the controls, it is not likely that this might be caused by an increased "foveal processing difficulty" (Henderson & Ferreira, 1990). Furthermore, the relative statistical distribution of saccade amplitudes was found to be approximately the same as that reported for skilled readers (Morris & Rayner, 1991).

The lack of a relationship of the F-saccade amplitude/word length functions with text reading speed must be considered on the basis of the fact that they were derived from different paradigms: text reading versus single-word reading. Continuous reading of text presumably requires more skills dealing with the relationships between words, an element that is missing in single-word reading. We hypothesize that the mechanisms of between-the-words control might be more important in determining text reading speed than those of within-one-word control.

The fact that these functions also did not show any dependence on age correlated well with the fact that none was seen in this group for text reading speed either. It is conceivable that this could have been caused by the limited cohort investigated here.

DWE and DWB

The fact that we found the final landing points of saccades not at the end of words but rather some distance before can again be interpreted as a manifestation of the rightward extent of the perceptual span that made further progression simply not necessary due to parafoveal processing (McConkie & Rayner, 1976; Rayner, Well, Pollatsek, & Bertera, 1982).

The same can be said about the distances to the word beginning in the right-aligned words, where the leftward extent of the perceptual span could be used. Together with the findings regarding the F-saccade amplitudes (see above), these data further support the hypothesis that the dyslexics investigated here had a significantly narrower perceptual span than the children in the control group. Note that a similar result was also reported by Rayner (1986) in children who are learning to read, from which he concluded that this "...indicated that the size of the perceptual span is variable and can be influenced by the difficulty of the text."

**Durations of holding phases**

It has been shown that there are considerable variations in the durations of intersaccadic intervals during reading as well as other visual tasks in normally reading adults (Hooge & Erkelens, 1996; Jacobs, 1986; Osaka, 1992; Rayner & McConkie, 1976) as well as in normally reading children (McConkie et al., 1991). The results from our control group are in good agreement with these findings. O'Regan (1980) reported a negative dependence of duration of holding phases on word length in normal readers. That we could confirm O'Regan's finding only in a majority of our control group (10/12) does not diminish the fact that the group median still showed a moderate negative dependence.

The significantly longer median durations of the holding phases in dyslexics confirms earlier findings by Adler-Grinberg and Stark (1978) and conflicts with Rubino and Minden (1973). It is noteworthy that individual medians within the groups showed almost the same ranges (53 ms [dyslexics] vs. 55 ms [controls]), but an overlap of values in only five subjects (see Figure 6). Although the longer group median in the dyslexics contributes to their reading impairment (Olson et al., 1991; Rayner et al., 1996), it does not allow the conclusion that it causes the impairment. This is supported by our finding that the individual holding duration medians showed no significant correlation with the subject's reading speed in the dyslexics. The longer holding phases more likely reflect prolonged attentional dwell times demonstrated in dyslexics by Hari et al. (1999).

The higher variability of individual holding durations found in the dyslexics is based on the fact that the shortest values showed only small differences between the groups, whereas the longest durations were 200-ms longer in the dyslexics. In addition, only 3 values from the control group exceeded 500 ms, but 46 did in the dyslexics (i.e., 0.08% vs. 7%).

**Forward saccade amplitude correlations with durations of holding phases**

Models of reading impairments stress differences between a phonemic/linguistic and a visual/oculomotor component of the reading process (see Stark et al., 1991). With respect to the sensory input, it could be argued based on work by Salthouse and Ellis (1980) that the holding duration may depend on the amplitude of the preceding saccade. In the case of reading, a longer F-saccade amplitude might cause a longer duration of the following holding phase, because more letters take a longer time to process. The fact that we found only loose and inconsistent correla-
tions between amplitudes and subsequent holding durations allows rejecting this hypothesis for reading. This also confirms the findings of Rayner and McConkie in normal readers (1976), although they found a smaller range of variation. We conclude that in the dyslexics, the phonemic/linguistic factors must have the stronger influence on either variable.

Our findings agree with those reported by Rayner et al. (1996), who used the durations of the holding phases in normal readers as an indicator for which component tells the eyes to move on. Their findings support the view that this decision is primarily influenced by "on-line language processing" rather than by visual or oculomotor factors.

Further support of this finding in dyslexics comes from a previous study investigating the relationship of text reading with pictogram-naming performance in dyslexics. We found that both, a phonemic/linguistic and a visual/eidetic deficit, can contribute to the phenomenology of an individual's dyslexia. While the deficit in the phonemic/linguistic component may be common to all dyslexics, which may be the main reason for their poor text reading, the visual/eidetic component was compromised only in some dyslexic subjects (Trauzettel-Klosinski et al., 2002). Thus, the latter turned out to be poor pictogram namers, whereas the others performed as well in this task as the normally reading children.

Regarding the correlation of holding duration with word length, O'Regan (1980) found large differences between adult normal readers, which was in conflict with findings by Rayner and McConkie (1976). We found such differences also between our young normal readers, but particularly between the dyslexics. This makes it likely that a negative correlation - the longer the word, the shorter the holding duration - may be an acquired skill that promotes reading speed and fluency (see also the discussion of forward saccade amplitudes, above). We found it in most of our control group (10/12) and in 6/10 of the dyslexics. The others may still need more practice to recognize frequently occurring syllables in long words and, thus, show a positive correlation. The fact that the dependence of this correlation on age was weak could be due to the small group size. Nonetheless, the dyslexic child with the highest positive correlation coefficient ($\rho = 0.595$), whose eye rested the longest on long words, was the second oldest of the group, the slowest text reader and had the longest times to word recognition. On the other hand, the two dyslexics with the strongest negative correlation ($\rho = -0.72$ and $\rho = -0.48$) had the shortest holding durations, and were numbers 1 and 3 in text reading speed.

Our results also fit the larger conceptual framework of saccades performed in other visual tasks, where past experience as well as prediction can play a role. For instance, Hooge and Erkelens (1996) found in a search task that "control of fixation duration appears to be indirect" (i.e., not exclusively dependent on the currently fixated target). In a later study, they reported that fixation durations are not commensurate with the current task at hand, but are controlled by a mechanism that uses estimations of the foveal analysis time of previous fixated stimulus elements (Hooge & Erkelens, 1998; we added the italics). Corroborating evidence for endogenous eye-movement control was also presented by Andrews and Coppola (1999), who found "idiosyncratic characteristics" in saccadic eye movements that could not be explained otherwise. Moreover, the current results can be linked to reports of low-level deficits in the magnocellular stream of visual neurons in dyslexics (Borsting et al., 1996; Demb et al., 1998a, 1998b; Stein & Walsh, 1997; Steinman et al., 1998), which can affect the control of attention (Fischer, 1987; Heinze & Mangun, 1995; Kowler et al., 1995; MacKeben & Nakayama, 1993) via a frontoparietal network (Corbetta, 1998). This might then affect preprogramming of reading saccades (Morrison, 1984) as well as control of the sequential flow of attention during reading, which has been shown to be an important component in modeling reading eye movements (Henderson & Ferreira, 1993; Vidyasagar, 1999; Radach, Inhoff, & Heller, 2002).

Conclusions

Our results show that the mechanism allowing normal readers to adjust forward saccade amplitudes to word length is present in dyslexics, but that it is amplitude-limited during reading of words of five-letter length and longer. Thus, the dyslexic cannot fully take advantage of this adjustment that promotes reading speed. A tentative interpretation of this finding could be that the rightward extent of the perceptual span used for reading longer words is limited in dyslexics. These data allow no conclusion whether this reflects a saccade control problem possibly caused by a magnocellular deficit, or whether longer words simply pose a higher degree of decoding difficulty. We also found that the increased number of saccades in either direction made by the dyslexics during reading reflect only the longer time they have to spend deciphering words, but not a higher absolute rate of occurrence of saccades per unit time. Furthermore, in some dyslexics, single-word reading speed is reduced by longer durations of the holding phases between saccades, which could reflect the need for longer processing time of individual syllables or a speed deficit in finding their phonological equivalents. Three parameters measured here reflected the difficulties the dyslexics encountered while reading single words: The longer durations of holding phases, the slower progression through shorter forward saccades, and the tendency in some to start the reading process over again after failing to read a word correctly. We hypothesize that these deficits during single-word reading likely also contribute to the reduced text reading speed in dyslexics.
Acknowledgments

The authors wish to thank Björn Krämer for his help with the analysis of the SLO videotapes and two anonymous reviewers for many helpful comments.

Commercial relationships: none.
Corresponding author: Manfred MacKeben.
Email: mm@ski.org.
Address: The Smith-Kettlewell Eye Research Institute, San Francisco, California, USA.

References


