universität
wien

MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master‘s Thesis

„Eye-tracking as Diagnostic Method in Dyslexia“

verfasst von / submitted by

Bernadette Maul BA

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Master of Arts

Wien, 2018 / Vienna 2018

Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet:	A 066 867
Studienrichtung lt. Studienblatt / degree programme as it appears on the student record sheet:	Allgemeine Linguistik: Grammatiktheorie und kognitive Sprachwissenschaft
Betreut von / Supervisor:	Ao.Univ. Prof. Dr. Chris Schaner-Wolles

Table of contents

List of abbreviations

AD	Attentional Deficit
ADHD	Attentional Deficit Hyperactivity Disorder
AOI	Area of Interest
CD	Cluster Density
CG	Control Group
DYS	Dyslexia
EEG	Electroencephalogram
ERP	Event-Related Potential
EVS	Eye-Voice Span
FLP	First Landing Position
Hz	Hertz
ICD10	International Classification of Diseases 10
IQ	Intelligence Quotient
L	Parafoveally Loaded
L1	First Language
L2	Second Language
LPC	Late Positive Component
ms	Milliseconds
nL	Parafoveally Not Loaded
nP	Parafoveally Not Previewed
NRD	No Reading Difficulties
OVP	Optimal Viewing Position
P	Parafoveally Previewed
PVL	Preferred Viewing Location

RAN	Rapid Automatized Naming
RD	Reading Difficulty
RSE	Rapid Serial Encoding
SLI	Specific Language Impairment
SLO	Scanning Laser Ophthalmoscope

List of figures

List of tables

1. Introduction

The interest in how people move their eyes in various tasks is not only one of nowadays researchers but actually goes back to ancient times. Over the last 200 years, remarkable advances in eye movement research were achieved. One of them, eye-tracking, proved an adequate and efficient methodological tool to examine cognitive processes that are active during various tasks. Special focus was set on the investigation of eye movements during reading. As soon as there was research on regular eye movements in reading, an interest in irregular eye movements in reading emerged. While there is remarkable work done on eye movements in attentional deficits and Specific Language Impairment, also research on ocular movements in dyslexia yielded important results (Pavlidis, 1985).

While the body of research in eye movements in dyslexia is growing remarkably, so is the need for comparing and contrasting their results. In this master thesis, 18 eye movement studies on dyslexia published in peer-reviewed journals in the last 15 years are compared and contrasted with regard to three main parameters which I explain in more detail below. This extensive comparison should shed light on what general assumptions can be made regarding eye movements in dyslexia, in how far results differ and what possible reasons for varying results there are. Furthermore, the question of how eye-tracking as methodology could be used in dyslexia, especially in diagnosis, are addressed and discussed based on one particular study by Benfatto et al. (2016). The problem of efficient, objective screening for dyslexia becomes more and more important and eye-tracking might prove to be a step into the right direction.

As methodology, the systematic comparison of eye-tracking studies was chosen. The author first read and summarized eye movement studies with regard to their focus and hypotheses, their experiment settings (including the samples of participants, stimuli, tasks and procedures, the eye trackers used and the eye movement parameters investigated), and their key results and interpretations. In a next step, all 18 eye-tracking studies were compared with regard to one parameter at a time. At this point, first similarities are found; however, also the extent of research in this field becomes apparent. In this part of the comparison, the relevance of various aspects, such as sample size or task, is discussed. Further, the eye-tracking studies were grouped according to their foci in order to examine whether or not studies with similar foci and/or similar methodology yield comparable results or not and if not, why so. These synchronic and diachronic analyses should provide substantial insight into these studies, their results, their similarities and dissimilarities.

The outline of the paper is as follows. First, a short introduction to dyslexia as condition is given. The symptoms and different theories on its origin are discussed, as well as screening for dyslexia and current research. Further, an overview of eye-tracking as research method is given. The history of eye-tracking is shortly addressed before eye movements are discussed and research on eye-tracking in reading is summarized in order to give insight into how eye movements are used to investigate cognitive processes. The next part forms the main part of this thesis. Eighteen eye-tracking studies in dyslexia are shortly introduced and then extensively compared and contrasted. In a first step the foci, hypotheses and expectations of each study are summarized. Then, the experiment settings are compared as follows: first, the samples of participants, their number, age, and inclusion/ exclusion criteria are compared. In a second step, the possible relevance of the number and age of the participants, of different inclusion/ exclusion criteria and of different languages and orthographies on the results are analyzed and backed up with research publications. Furthermore, the stimuli, tasks and procedures are summarized and compared before the possible relevance of task on the results is discussed. The final aspects of experiment setting to be compared are the eye-tracking devices which are used in the experiments and the eye movement parameters measured. This part is again followed by a short analysis of the possible relevance of eye-trackers and their sampling frequencies. The final and maybe most crucial step of this synchronic analysis is the comparison of the results and their interpretation. What follows is a rather diachronic analysis in that papers with similar foci or methodology were grouped together and compared in order to find similarities and dissimilarities. If dissimilarities are found, possible reasons for the varying results are discussed. A final yet significant step is the analysis and discussion of eye-tracking in the diagnosis of dyslexia. For this purpose, one particular study by Benfatto et al. (2016) is summarized and discussed before a future prospect is given. The thesis is rounded up by a conclusion in which all main findings and aspects are summarized.

2. Dyslexia

a. The attempt of a definition

Defining dyslexia is not an easy task and neither is giving a short yet sufficient overview of the condition. However, in order to analyze eye movement data with respect to the heterogeneity of readers, an introduction in dyslexia is inevitable.

According to the Statistical Classification of Diseases and Related Health Problems published by the World Health Organization (ICD10), and a number of researchers, dyslexia is defined as mental developmental reading disorder (Benfatto et al., 2016: 1; Goldberg, Shiffman, and Bender, 1983: 1f.). Generally, it can be said that children with dyslexia show great difficulties in learning to read, write and spell which is unexpected with regard to their cognitive abilities and formal instruction (Fawcett &Nicolson, 2008: 77). Lyon and colleagues introduced a working definition in 2003:

> “Dyslexia is a specific learning disability that is neurobiological in origin. It is characterized by difficulties with accurate and/or fluent word recognition and by poor spelling and decoding abilities. These difficulties typically result from a deficit in the phonological component of language that is often unexpected in relation to other cognitive abilities and the provision of effective classroom instruction. Secondary consequences may include problems in reading comprehension and reduced reading experience that can impede growth of vocabulary and background knowledge“ (Lyon, Shaywitz, & Shaywitz, 2003: 2).

As Reid and Fawcett stress, there is need for a not yet reached consensus regarding the definition of dyslexia (2004: 3). While there are individual differences in how dyslexia can be manifest, there are still some general conclusions that seem to apply: It is a primarily language-based condition, which is neurological and partly neurodevelopmental in origin, and it is partly genetic (Gilger, 2006: 32). According to Stein, there are some risk factors for dyslexia, including teaching, IQ, heredity and developmental speech and/or motor problems (2008: 54-56). In their long term study from 2008, Lyytinen and colleagues found that the parents’ dyslexia notably increases the likelihood of dyslexia in their children. About 50% of the children with dyslexia in the family were faced with reading problems. However, they also observed a positive effect of a supportive familial environment on the language development of dyslexic and non-dyslexic children. The researchers further stress the importance of early identification and note that delayed development of expressive language was the earliest difference between children with and without reading difficulties later among the group of children who were at familial risk of dyslexia (Lyytinen et al. 2008: 138).

Word decoding can be seen as a very complex process that does not come naturally but has to be learned (Stein, 2008: 53). Reading ability and disability occur as poles along a continuum without clear limits. Thus estimates of 5-15% of the population being affected by dyslexia can only be seen with regard to the difficulty and variation in defining and diagnosing this condition (Benfatto et al., 2016; Breznitz, 2008; Shaywitz & Shaywitz, 2005; Rosen, Wang, Fiondella, LoTurco, 2009). Despite the problem of clear definitions, seeing reading ability as continuum helps to understand the heterogeneity of readers, as every reader is located at a different position in this continuum. Regardless of their definition and diagnosis, reading difficulties can have a big influence on school performances and success in academia. This can further lead to psychological stress, which can be a huge burden for affected persons (Benfatto et al., 2016: 2).

b. Comorbidity

What further blurs the definition of dyslexia is its high comorbidity rate and association with other conditions, such as ADHD (Attentional Deficit Hyperactivity Disorder) or SLI (Specific Language Impairment). “These co-occurrences […] suggest strongly that in some proportion of cases the distinctions between disorders […] may, at times, be more artificial than real, and it may also muddle a complete understanding of the individual presenting a complicated symptom profile” (Gilger, 2008: 34). As Deponia (2004) claims, a label can only be applied as soon as a specific difficulty is identified. This means that there has to be agreement on the indicators of a specific condition, however, the indicators for dyslexia, apraxia, SLI and ADHD show commonalities, which makes clear identification difficult. It is important not to match specific behavior to certain indicators for one particular condition but to examine the child as whole in order to provide reliable assessment. The relation of dyslexia with ADHD seems to be the most prominent one with the greatest body of evidence of co-occurrence, followed by SLI (Deponia, 2004: 324-327). As Deponia puts it:

> “The variation in the percentage overlap suggested by the above researchers is not important. What is important is the fact that research consistently demonstrates that a percentage of pupils displaying discrepancies in their learning and performance do not meet the criteria for only one specific learning difficulty. Examination of indicators of dyslexia common to other specific difficulties demonstrates how difficult it can be to reach an appropriate ‘diagnosis’ [original emphasis]” (2004:328).

To sum up, it is important to consider comorbidity and to try differentiating dyslexic children and children with more general conditions (Reid & Fawcett, 2004:4). One further factor disturbing a clear picture of dyslexia is the question of orthographic depth, as dyslexia seems to be manifest differently according to the orthographic depth of a language (Beaton, 2004: 90-93).[1]

c. Theories on dyslexia

The reasons for dyslexia are widely debated and not completely clear yet. However, there seems to be a genetic component with children having dyslexic parents or siblings being more at risk (Shaywitz & Shaywitz, 2005). The debate on possible reasons for dyslexia has yielded a number of theories, of which the most important ones were summarized and discussed in a study by Ramus et al. (2003). Only a shortened summary of Ramus et al.’s discussion is given here (841-844).

The phonological theory

First, the theory that is most widely accepted, the phonological theory, is summarized. In this theory, the difficulties of dyslexics are attributed to “a specific impairment in the representation, storage and/or retrieval of speech sounds” (Ramus et al., 2003: 842). The foundation of learning to read in alphabetic systems is learning

---

[1] Orthographic depth describes the regularity with which language symbols (usually letters) are matched to speech sounds (phonemes. For further information and clarification see “relevance of language and orthography”.

the correspondence between symbols/letters and sounds. However, if there are deficits in the representation, storage and/or retrieval of these sounds, the development of the grapheme-phoneme correspondences will be affected as well. According to Reid and Fawcett, “phonological awareness is a meta-linguistic skill involving knowledge about the sounds that make up words, at both the syllable and the phoneme level” (2004:5). There seems to be consensus on the crucial role of phonology in dyslexia, even though there are different perspectives on the causes for these phonological problems. Neurological research indicates that the origin of the disorder seems to be a dysfunction in the left-hemisphere brain areas attributed to phonological representation or connecting phonological to orthographic representations. Dyslexic individuals’ poor performance on phonological awareness tasks provides support for this theory. Anatomical and functional brain imaging studies provide evidence for a left hemisphere dysfunction as basis for the phonological deficit. As Rosen and colleagues put it, there are three main symptoms for a phonological deficit: poor phonological awareness, poor verbal short-term memory, and slow lexical retrieval (2009: 21). However, a weakness of the phonological theory seems to be its inability to account for sensory and motor disorders at least in some dyslexic individuals. Followers of the theory seem to explain these disorders as not being causal but only being markers of dyslexia (Ramus et al., 2003: 842).

The rapid auditory processing theory

The rapid auditory processing theory seems more basal than the phonological theory. It locates the deficit in the perception of short or rapidly varying sounds. Indeed, there seems to be evidence that dyslexics perform poorly on a number of auditory tasks and that they may have less efficient categorical perception of certain contrasts. In this sense, the auditory deficit causes the phonological deficits and leads to difficulties in learning to read (Ramus et al., 2003: 842).

The visual theory

According to the visual theory, dyslexia is primarily a visual deficit giving rise to difficulties in processing elements of written text. It can be manifested in unstable fixations, difficulties in vergence control or visual crowding. While the visual theory does not exclude a phonological deficit, it puts stress on the importance of visual skills in reading problems, at least in some individuals. The biological basis for the visual theory stems from the division of the visual system into the magnocellular and parvocellular pathways of which the first is believed to be disrupted in at least some dyslexic individuals. This might lead to deficiencies in visual processing and abnormal binocular control. Indeed, anatomical and psychophysical studies provide evidence for this theory (Ramus et al., 2003: 842).

The cerebellar theory

The claim of the cerebellar or automaticity theory is that the cerebellum is slightly dysfunctional in dyslexics. Since the cerebellum is important in motor control and articulation, a dysfunction of the same would cause deficiencies in phonological representations. Furthermore, the cerebellum is also involved in the automatization of tasks, such as driving, and thus weak capacities could affect letter-to-sound correspondences, for example. Dyslexic individuals indeed often show poor performance in a number of motor tasks. Furthermore, anatomical, metabolic and activation differences in the cerebellum of dyslexics could be illustrated by brain imaging studies. Similar to the phonological theory, the cerebellar theory seems to fail to explain sensory disorders in dyslexics. Furthermore, the causal link between articulation and phonology, on which this theory relies, is itself based on the outdated and overcome motor theory of speech. Finally, it is still not completely clear how many dyslexics indeed show motor deficits (Ramus et al. 2003: 843).

The magnocellular theory

Finally, the last theory to be discussed here is the magnocellular theory which tries to include aspects of all of the above mentioned theories. It can be seen as more general than the visual theory in that it postulates the magnocellular dysfunction not to be restricted to the visual pathways. The cerebellum is also predicted to be influenced by this defect as it is exposed to extensive input from various magnocellular systems in the brain. It is the only theory accounting for all manifestations of dyslexia. Evidence comes from magnocellular abnormalities in the medial and lateral geniculate nucleus of dyslexics’ brains as well as the poor performance of some dyslexics in visual, auditory and tactile tasks. In their criticism, Ramus and colleagues have taken the visual and auditory theory introduced above together with the magnocellular theory. Criticism on this theory is mainly based on failures to replicate findings of auditory disorders and visual deficits in dyslexics. Thus, the magnocellular theory seems to fail to account for missing auditory and visual deficits in dyslexia. Furthermore, it has been argued that auditory deficits do not necessarily predict phonological deficits (Ramus et al. 2003: 843).

There is no complete consensus on which theory is correct. However, according to Shaywitz and Shaywitz (2005), the phonological theory is widely supported. This theory seems to account for the fact that speech is natural while reading has to be learned. In order to successfully learn reading, letters have to be understood as representations of spoken sounds and that spoken words can be divided into elements of speech (Shaywitz & Shaywitz, 2005: 1301f).

According to Ramus et al. (2003) it could be possible that the different theories account for different dyslexic individuals, or that there is indeed one theory accounting for all realizations of dyslexia with some manifestations being causal and others correlational. Further multi-modal research is needed.

d. Subtypes of dyslexia

The heterogeneity of the group of dyslexics has led to the assumption of subtypes of developmental dyslexia. The idea of dyslexic subtypes goes back to the 1960s and developed through the last 60 years (Beaton, 2004: 80-86). Today, there are at least two types that are generally described, phonological and surface dyslexia. Readers with phonological dyslexia seem to be more impaired in phonological decoding and processes related to it, such as phonological awareness. Adults with phonological dyslexia cannot pronounce unfamiliar regular words, as their sub-lexical route is impaired (see dual-route model on page 43). Research indicates that the phonological dyslexia profile is “more likely a developmentally deviant pattern with strong biological influences” compared to the surface dyslexia profile which seems to be a developmentally delayed pattern with mostly unspecified origins (Manis &Bailey, 2008: 153). Readers with surface dyslexia show difficulties in reading and pronouncing exception words, as their lexical route seems impaired. However, research indicates that there is only a relatively small number of “pure” cases for each type. Many dyslexics show symptoms fitting both profiles, which again reflects the need for individual profiles of affected persons (Manis & Bailey, 2008: 149-s153).

e. Recent research

In the next part, more recent research on dyslexia is introduced. In their study from 2012, Deacon and colleagues examined two recruitment strategies and their impact on the sample of high-functioning dyslexics, dyslexics who have compensated for most of their reading difficulty. They compared the performance of students who reported reading acquisition difficulties on a self-report questionnaire with the performance of students who were recently diagnosed on standardized measures of word and non-word reading and fluency, passage comprehension and reading rate, and phonological awareness. Furthermore, both groups were compared to a control group without reading acquisition difficulties. Both groups with reading difficulties showed similar performance in timed reading comprehension, word-level reading, and phonological awareness. These similarities suggest that reading disabilities might be under-identified in the university setting. Furthermore, these findings indicate that the two recruitment methods probably sample from the same underlying population. However, there were also differences between the two groups: The diagnosed group performed better on untimed reading comprehension while the self-report group performed better on reading rate. According to the authors, these results suggest that further research on possible differences in adaptive strategies between the two groups is needed. With regard to the performance of the control group, both groups with a history of reading difficulties were outperformed on almost all measures indicating that full compensation for early reading difficulties seems rare. Both groups of high-functioning dyslexics showed remaining deficits in phonological awareness which is surprising considering their level of untimed reading comprehension. This finding suggests that effective reading comprehension might not fully rely on phonological awareness (Deacon, Cook, & Parrila, 2012).

The above introduced study shows that notwithstanding the phonological deficits in developmental dyslexia persisting into adulthood, there are dyslexic adults studying at university level. Since there is little known about how they manage to do so, Cavalli, Duncan, Elbro, El Ahmadi and Colé (2017) investigated whether the development of morphological knowledge can be preserved and dissociated from the development of phonological knowledge and might thus serve as compensatory mechanism. The authors tested reading, phonological and morphological abilities in 20 dyslexic and 20 non-dyslexic university students. Dyslexic participants showed persisting deficits in phonological but not in morphological abilities, which indicates a dissociation of the two skills. Second, there was a correlation observed between the magnitude of the dissociation and the reading level. These findings support the hypothesis that university students with reading disabilities may use morphological abilities as compensatory mechanism in reading (Cavalli, Duncan, Elbro, El Ahmadi, & Col??, 2017).

In their study from 2013, Hasko and colleagues investigated which steps in processing might be degraded in dyslexic children in order to better understand reading speed deficits in those individuals. The authors used the electroencephalogram method (EEG) to examine three particular reading related event-related potentials (ERPs), namely the N170, N400 and LPC (late positive component), in 52 children with developmental dyslexia and twenty-nine children without reading difficulties. Each participant had to conduct a phonological lexical decision task in which they had to decide whether the presented stimulus sounded like an existing German word or not. Among the stimuli were words, pseudo-homophones, pseudowords and false fonts. In all the investigated ERPs, dyslexic children exhibited deficits compared to the control group. At first, in the time window of the N170, a smaller area under the curve for the word material-false font contrasts was found, which indicates a reduced degree of print sensitivity in dyslexics. Secondly, the authors observed declined N400 amplitudes, which are suggested to reflect the access to the orthographic lexicon and grapheme-phoneme conversion. Finally, the results for the LPC indicated that phonological access was impaired as well in dyslexic children. Also in the LPC, processing differences dependent on the linguistic material in children without dyslexia were observed, which suggests that regardless of the orthographic familiarity, similar reading processes were adopted. According to the authors, these results indicate that effective treatment should include orthographic as well as phonological training (Hasko, Groth, Bruder, Bartling, & Schulte-Körne, 2013).

In their study from 2015, Helland and colleagues aimed at identifying neurocognitive precursors of literacy development in the first language L1(in this case Norwegian) and a second language L2(English) in children before reading instruction, during the emergent literacy stage and literacy stage. They compared a group of dyslexic children to a typical group of children without reading difficulties. Dyslexic children could only be identified at the age of 11 years and data of the children at the beginning of the project, when they were 5 years old, were analyzed in retrospect. For first language literacy, there were two early precursors, namely visuospatial recall and Rapid Automatized Naming (RAN). For the second language literacy, phonological awareness was observed as early precursor. Verbal long term memory seemed to be important in both L1 and L2 skills in the final literacy stage. Interestingly, group differences in the literacy scores increased by literacy stage while group differences in neurocognitive scores decreased by literacy stage. This result possibly mirrors the inconsistencies found in dyslexia research. The authors argue for early identification and training to be essential in order to avoid academic failure. They claim visuo-spatial memory and RAN to be potential suitable early markers at least in transparent orthographies. On the other hand, phonological awareness was observed as early precursor only in L2 English (Helland & Morken, 2016).

f. Diagnosis

Finally, there is the important question of diagnosis in dyslexia. Due to varying definitions and different thresholds for dyslexia, there is no standardized uniform diagnosis (Beaton, 2004: 7). As has been mentioned above, one definition of dyslexia is an unexpected poor performance in reading despite normal intelligence and instruction. However, it is much less clear what an unexpected level of poor performance really is. Traditionally, it is defined by normative data comparing individuals to the performance of their age-matches (Wagner, 2008: 174). Thus, currently, the identification of dyslexic individuals is mostly based on standardized reading and/or writing tests. Benfatto and colleagues (2016) criticize that these methods often rely on oral or written tests based on overt output of the participant and its evaluation, often conducted under time pressure, which cannot be completely objective. Further, the researchers criticize that the values achieved might show the performance in a specific reading-related task but that they are not predictive for actual reading processes (Benfatto et al. 2016). Further, Nicolson and Fawcett criticize the informative value of a “snapshot” of the abilities of individuals or groups at one point in time (2008: 196). Another point of criticism is the lack of control for comorbidity (Deponia, 2004).[2]

Despite being difficult, an early diagnosis is of great importance in order to provide professional and effective support for affected children (Lyytinen et al, 2008: 122). In Sweden, the average age at diagnosis is 13 years, which is seven years after the initiation of formal instruction (Benfatto et al., 2016). According to Lindsay, useful screenings for dyslexia should provide high sensitivity and specificity in order to really identify only those at risk but, on the same hand, not to “oversee” affected children (2004: 279). One method, which could prove promising and helpful in an objective, efficient and adequate diagnosis, is eye-tracking (Benfatto et al., 2016). Its potential and importance are discussed below.

3. Eye-tracking

Before eye-tracking studies in dyslexia are reviewed, a short introduction in the field of eye-tracking is given here. Eye-tracking generally describes the measurement and recording of eye movements with appropriate devices. In the following, a summary on the history of eye-tracking and a brief introduction into general eye movements is given before eye movements and eye movement measures in reading, and eye-tracking as research method are discussed. This chapter should serve as preparation for the comparison in order to understand certain eye movement related concepts.

---

[2] Götzinger-Hiebner (2014) offers a summary of the most often used standardized reading tests for German.

a. Historic background

To catch a glimpse on the history of eye-tracking, the way goes far back to Aristotle, who can be seen as a pioneer in the study of eye movements during reading. However, it was not until the late 19th century that groundbreaking progress had been made especially due to Hermann Helmholtz as well as Ewald Hering, to name only two of many. In 1867, Helmholtz introduced the bite bar in order to control for head position and to enable more precise observation. Twelve years later, in 1879, Hering could demonstrate, contrary to common belief, that eye movements do not run as smooth as thought but rather run discontinuously with forward and backward movements. These advances triggered more studies and further development of eye-tracking devices, including photographic eye-trackers as well as light reflection devices (Wade& Tatler, 2011). Today, eye trackers work with sampling frequencies of up to 1250H and very high spatial resolution. The more precise and qualitative the data, the less is needed for the analysis.[3]

b. Eye movements

Generally, human visual perception consists of three parts. The central area of the retina comprises the region of the highest visual resolution. It is called fovea and it is the locus where most information can be extracted. Surrounding the fovea is the parafovea extracting less precise yet important information for example for saccadic computation that is eye movements from one fixation to another. The periphery describes the region outside of the parafovea where only blurred vision is possible. The area of effective vision is called perceptual span and varies from person to person. In order to move stimuli that capture the attention into the fovea, people move their eyes. There are mainly two basic types of eye movements to be distinguished: fixations and saccades. Fixations describe resting positions of the eyes and last from 100-500 milliseconds (ms). It is during fixations that information can be extracted. Eye movements between fixations are saccades, which last for approximately 30-50 ms. It is not possible to acquire new information during saccades (Lai et al., 2013; Hyönä, 2011: 819). Word recognition needs high visual

---

[3] See “relevance of eye trackers and sampling frequency”.

acuity. In order to successfully identify words in a text, readers have to locate the fovea in a way that allows for light reflected from the fixated word to fall directly on it.

c. Eye-tracking as research method

In psychology, eye-tracking as method is already widely used in order to observe information processing and basic cognitive processes during reading in particular. The reason for its intense use is the idea that this method is capable of recording online cognitive processes (Lai et al., 2013). Just and Carpenter developed the eye-mind assumption based on the theory of eye movements providing a trace where attention has been shifted to (Just & Carpenter, 1980). It is based on the observation of the difficulty of attending to a stimulus without shifting the gaze towards it. Thus, attention shift presumably plays an important role in the preparation of eye movement computation. Before a saccadic movement can lead to a next fixation, attention has to be shifted in order to calculate the saccadic landing position (Kristjánsson, 2011). Saccade computation is definitely subject to research. As Ludwig puts it, it is the product of a decision making process of if, where, and when to move the eyes (2011: 425). According to Lai and colleagues, it is meanwhile widely agreed that eye movements and attention are connected in complex information processing, such as reading (2013). This might be due to the same neural resources that eye movements and attention share to a notable extent (Kristjánsson, 2011).

i. Eye movement measures

As has been mentioned above, there are two general eye movements: fixations during which information is extracted and saccades which move the eyes. As for eye movement measures, there are three scales of measurement: temporal, spatial and count. In the following table, the most important measures are defined, following Lai et al. (2013).

Table 1: Eye movement measures following Lai et al. 2013

Temporal	
Total fixation duration	Total time spent on fixations
Gaze duration	Total fixation duration within a word or an AOI (Area of Interest)
Average fixation duration	Mean of fixation duration on each AOI. (i.e., Gaze duration mean)
First fixation duration	Time spent on the first fixation
Saccade duration	Sum of saccadic time spent within an AOI
Total reading time	Total time spent for a reading task or spent within an AOI
First pass time	Time spent for the first entering of an AOI until leaving
Re-reading time	Sum of revisited time spent within an AOI
Spatial	
Fixation position	Location of a fixation
Saccade length	Distance between two consecutive fixations
Count	
Total fixation count	Total number of fixations counted in an AOI or in a task
Average fixation count	Average fixation count on each AOI
Revisited fixation count	Sum of revisited fixation count within an AOI
Saccade count	Total number of saccades counted within an AOI
Regression count	Number of regressions
Percentage of regressions	The percentage of regressions

The terminology used in the papers discussed below might differ from these, however the measures are still clear.

ii. Eye movements in reading

Eye-tracking proved to be specifically helpful in the research on reading. In the following part, an introduction to the research on eye-tracking in reading is given based on three articles published in the Oxford Handbook of Eye Movements. First, children’s eye movements during reading are discussed based on Blythe and Joseph (2011), followed by a summary of linguistic and cognitive influences on eye movements in reading based on Rayner and Liversedge (2011) before foveal and parafoveal processing during reading is analyzed based on Hyönä (2011).

1. Eye movements in the reading child

The development of eye movement behavior was investigated thoroughly in several studies. Basically, children’s eye movements in reading change with chronological age and are characterized by a decrease in fixation duration, refixation probability, number of fixations and regressions and overall sentence reading times, whereas saccade amplitudes and word skipping probability increase (Blythe & Joseph, 2011: 647). Furthermore, lexical identification is shown to be slower in children. Reading is not only determined by oculomotor skills but also extensively by linguistic skills. Evidence for this assumption is provided by the observation that a child’s linguistic skills at the age of 8 years can predict sentence reading times of the same child at the age of 10 years, however not its oculomotor skills (Blythe & Jospeh, 2011: 651). Around the age of 11, children’s eye movements do not extensively differ from adult level anymore. One of the first skills that reach adult level is saccade targeting. Very soon, children target their saccades at word centers. If the initial fixation fails the center, not enough visual information can be extracted and refixations are more likely. Children differ from adults in that their refixation saccades seem less efficient. Furthermore, children have smaller perceptual spans and, thus, less parafoveal information, which is mirrored in word identification difficulties (Blythe & Joseph, 2011: 652). However, improvements in reading skills broaden the perceptual span. The left-right asymmetry in perceptual spans observed in adults reading in a left-to-right writing system is already established at an age of 7 years. This asymmetry enables readers to locate their attention to the characters to the right and, thus, enable more efficient parafoveal processing (Blythe & Joseph, 2011: 654). Finally, word length and word frequency effects are more pronounced in young readers (Blythe & Joseph, 2011: 657).

2. Linguistic and cognitive influences on eye movements during reading

The question of what controls eye movements in reading has yielded a lot of research and debate. It led to the question of how linguistic and cognitive variables influence eye movements. Generally, there are two perspectives on this question: the oculomotor view, according to which fixation durations are assumed to be independent of moment-to-moment cognitive processing; and the linguistic/cognitive view, which sees contextual properties as main reason for fixation durations. The shorter the time spent on processing a word, the easier its meaning is accessed. Even though there seems to be no consensus on the question above, it seems clear that fixation duration and saccade length are both determined by aspects of the word currently processed, such as linguistic properties. Rayner and Liversedge argue that the when question is driven by linguistic/ cognitive processing, while the where question is determined by low-level visual processes (2011). “Very consistent with this distinction is the finding that linguistic variables, such as word frequency, word predictability, and age-of-acquisition have major influences on when readers move their eyes, while word length information has a major influence on where readers move their eyes” (Rayner & Liversedge 2011: 753f.). The authors stress that the relationship between linguistic processes and eye movement measures is not always entirely transparent and straight forward. The question of the different kinds of linguistic processing showed that lexical processing is followed by syntactic and semantic processing, for which successful word identification is necessary. In the investigation of lexical influences on fixation times, the well-known frequency effects and predictability effects were observed. Further influences can be the age of acquisition of a word, word familiarity and the number of meanings a word has. The disappearing text paradigm proved an interesting and helpful research design in order to investigate word identification. The most interesting findings are that there is little or no effect on overall reading when the fixated word disappears after 50ms. This does not mean that 50-60ms are enough to fully process a word but that this time is sufficient to retrieve enough visual information in order to further process the word. Interestingly, even though the word disappears, the eyes of the readers still remain in place until the word is fully processed. Additionally, even the word frequency effect is still observable. These results are often taken as evidence for eye movements being driven by linguistic/cognitive processes (Rayner & Liversedge, 2011: 756).

3. Foveal and parafoveal processing during reading

Finally, foveal and parafoveal processing during reading are discussed following Hyönä (2011). In reading, orthographic and phonological coding lead to lexical and meaning activation. Visual perception is divided into parafoveal and foveal processing. In parafoveal processing, orthographic and phonological information is already retrieved. While it is well-known that reading without foveal vision is almost impossible and that important information is also extracted in parafoveal vision, one key question of research is still how much textual information around fixations readers can really extract (Hyönä, 2011: 819f.). In order to discuss this question, foveal and parafoveal processing are discussed in turn.

Among the factors influencing foveal processing are orthographic, oculomotor, phonological, lexical and semantic factors (Hyönä, 2011: 820). The first observation to be discussed is the location of first fixations. Readers tend to position their first fixation close to the word’s center, which is the preferred viewing location. This allows for word identification. If the first fixation is further away from the optimal viewing location, refixations are more likely. Furthermore, word length seems to have a significant influence on foveal processing. The longer the word, the longer are gaze durations. This is probably due to visual crowding. When it comes to orthographic factors influencing word identification, there are two particularly interesting observations. First, the frequency of letter clusters influences foveal processing in that more frequent clusters require less fixation time (Hyönä, 2011: 821f.). The second observation was documented by White, Johnson, & Rayner (2008). They could demonstrate that foveal processing times increased when letters in words were switched. Words with transposed letters evoked longer gaze duration than words with correct spelling. Furthermore, the transposed letter effect was stronger when external letters were transposed instead of internal letters (e.g. rpoblem vs. probelm). Reading speed was slowed by 11% which indicates that word recognition seems to be rather flexible with regard to letter position. The influence of phonological coding was observed in a regularity effect in first fixation duration. Irregular words required longer first fixation durations than regular words. The regularity effect was also observed in gaze duration for low frequency words (Hyönä, 2011: 822). As has already been mentioned above, word frequency and the age of acquisition also influence foveal processing. Longer gaze durations were observed for low frequency words and for late acquired words. The effect of the age of acquisition is likely to have a semantic origin. Word meaning influences foveal processing especially in ambiguous words, where meaning dominance seems to determine the order in which the meanings are accessed. The influence of morphological structure is discussed as well. While long compounds consisting of two free morphemes are decoded serially by their constituents, it was also demonstrated that short compounds can be processed holistically, which again refers to the length of words and foveal crowding (Hyönä, 2011: 824f.). Finally, contextual predictability seems to influence foveal processing in that contextually predictable words require shorter fixation times. If syntactic prediction is violated, processing will be more difficult (Hyönä, 2011:825.)

In the following part, influences on parafoveal processing are discussed. First, the closer the reader's fixation is to the parafoveal word, the more parafoveal processing is done. Furthermore, research on influences of word length on parafoveal processing showed that readers retrieve word-length information up to 15 character positions to the right of the fixation which particularly affects saccade computation. With regard to orthographic coding, in the parafoveal word readers were observed to get information on the shape of letters and also partly letter identity information. Eye movement studies suggest an early involvement of phonological codes in word processing. This is supported by evidence of first fixation duration being shorter when the target word was preceded by a homophone preview instead of a visually matched non-homophone preview. With regard to lexical-semantic and morphological influences, the results are inconclusive, mixed and/or meagre (Hyönä, 2011: 826-830).

There is one particularly interesting effect on parafoveal processing. The relative difficulty of processing the foveal word seems to influence parafoveal processing in that less parafoveal information is retrieved when foveal processing is difficult. This finding suggests that foveal processing difficulty restrains the parafoveal attentional span in that less attentional resources are available (Hyönä, 2011: 831).

To sum up, Hyönä summarized the results in a table (Hyönä, 2011: 832):

Table 2: factors influencing foveal and parafoveal processing following Hyönä 2011

	Foveal processing	**Parafoveal processing**
Location of initial fixation	Yes	N/a
Word length	Yes	Yes
Orthographic coding	Yes	Yes
Phonological coding	Yes	Yes
Word frequency	Yes	Mixed
Age of acquisition	Yes	?
Word meaning	Yes	Mixed
Morphological structure	Yes	Mixed (language-dependent?)
Contextual predictability	Yes	Mixed

In conclusion, only a very short and comprised introduction into eye movement research in reading could be given here. Generally, it can be said that effective letter identification requires not only high visual acuity but also efficient word processing on many levels. As could be seen, there are several factors influencing word processing.[4]

---

[4] For further information on eye movements and eye tracking as methodology see Duchowski and Andrew (2007), Holmqvist and colleagues (2011), or Munoz and colleagues (2007).

4. Eye-tracking in studies on dyslexia

In the previous sections of this thesis, aspects of dyslexia have been discussed as well as the questions of what eye-tracking is and what it can be used for. Combining those two fields can shed light on yet another important issue: eye movements of dyslexia. The following part of this thesis focuses on current research on dyslexia using the eye-tracking method. As Al Dahhan et al. (2014) point out in their study, it is near to impossible to really compare the findings of different eye-tracking studies. This is partly due to the fact that different studies use varying research paradigms, partly because they show differences in the characteristics of the participants, and finally because the eye movement parameters and measures used are different (Al Dahhan et al., 2014:139f.). On the other hand, it might also be interesting to see if similar results are achieved despite the different methods and research settings. Recent papers and former research are summarized and analyzed in order to point out similarities as well as differences in the findings, to look for possible reasons for the differences, and to stress the necessity of this kind of research.

Before the comparison of recent studies is presented, a summary of one specific paper might serve as a useful introduction. In 2013, Bellocchi, Muneaux, Bastien-Toniazzo, and Ducrot published a paper with the title “I can read it in your eyes: What eye movements tell us about visuo-attentional processes in developmental dyslexia”. What they did was rather similar to this work: they compared various studies to find parallels and differences in an effort to contribute to the bigger picture.

In the past, studies on reading abilities have led to further methodological approaches, of which eye movement recording was a very important one. Research on eye movements notably increased the knowledge about visuo-attentional processes involved in reading. Furthermore, these studies could also (at least partly) reveal to which extent impairments are specific to one disorder. One particularly important finding with regard to reading was that readers do not only process the fixated word. A perceptual span was observed in proficient readers which approximately includes 3-4 spaces to the left and 14-15 spaces to the right (in orthographies written from left to right). However, when the text is hard to read, the perceptual span is smaller. It is divided into the fovea, which is the area of clear vision, close to the fixation point in which readers identify words, and the parafovea, which is beyond that region and which gives readers grosser information on words. One major question in eye movement research is that of what cognitive processes are involved in eye movement control. According to the authors, one of the strongest mechanisms of saccade generation and computation seems to be visuo-spatial attention, so whatever captures a person's attention is foveated. Further results of previous research are that there are two important viewing locations: the preferred viewing location (PVL) and the optimal viewing position (OVP), which can of course differ in a person.

In order to successfully decode written words, children need to develop good visual skills. It has been suggested that there is a mechanism of graphemic selection based on the automatization of visuo-attentional processes that enables children to segment new words in order to apply grapheme-phoneme conversion rules (sub-lexical reading strategy). Rapid serial encoding (RSE) treats letters as parts of single objects and enables to automatically achieve a sub-lexical representation. For such focusing operations, attentional processes very likely play an important role. This leads to the hypotheses that reading will not be optimal or even possible if attentional or parafoveal processes are deficient. When it comes to dyslexia, there is still no consensus on the main cause. One widely accepted hypothesis suggests a core deficit at the phonological level of processing (phonological theory). However, there is also evidence indicating the presence of visual and oculomotor deficits in dyslexic readers.[5] When it comes to eye movements of dyslexics, findings may vary from one another. It is not entirely clear whether the eye movement patterns of dyslexics in reading, which usually show more and longer fixations, more regressions, shorter saccades, etc., do really differ from normal readers, since they often resemble those of inexperienced readers. However, they have still been interpreted as an indication for a failure of orthographic whole-word recognition and as inability to use the lexical route efficiently which might lead to a reliance on the sub-lexical route. Eye movement recordings might prove very helpful in clarifying these questions. Furthermore, with regard to saccadic computation, research showed that dyslexic children did not choose an optimal initial fixation position in words which might lead to more re-fixations(Bellocchi et al., 2013). To sum up, all of these results have shown,

---

[5] For further information on theories of dyslexia, see the introduction to dyslexia above.

that there are several factors, such as orienting, focusing, saccadic computation, deficits in visuo-attentional span, or attention shifting, which are probably linked to developmental dyslexia. However, the question arises which of these are specific for developmental dyslexia? Especially with regard to comorbidity, this question might prove very important for future research.

a. Selection of papers

Papers for the comparison were selected according to the following criteria. First and foremost, only studies focusing on eye-tracking as their primary method to investigate eye movements were considered. All papers were published in peer-reviewed journals. A threshold for the timeliness of the papers was set at 15 years of age in order to ensure informative value and to enable comparison. Thus, the earliest paper discussed here was published in 2002(De Luca et al., 2002), the latest in 2017 (Kim & Wiseheart 2017). Furthermore only studies focusing on developmental dyslexia instead of acquired dyslexia were included.[6]

List of papers included:

1. Al Dahhan, N.; Georgiu, G. K.; Hung, R.; Munoz, D.; Parrila, R.; Kirby, J. R. 2014. “Eye movements of university students with and without reading difficulties during naming speed tasks”. *Annals of Dyslexia* (2014) 64, 137-150.
2. Bucci, M. P.; Brémond-Gignac, D.; Kapoula, Z. 2008. “Latency of saccades and vergence eye movements in dyslexic children”. *Experimental Brain Research* 188(1), 1-12.
3. De Luca, M.; Borrelli, M.; Judica, A.; Spinelli, D.; Zoccolotti, P. 2002. “Rapid communication. Reading words and pseudowords: An eye movement study of developmental dyslexia”. *Brain & Language* 80, 617-626.
4. Dürrwächter, U.; Sokolov, A. N.; Reinhard, J.; Klosinski, G.; Trauzettel-Klosinski, S. 2010. „Word length and word frequency affect eye movements in dyslexic children reading in a regular (German) orthography”. *Annals of Dyslexia* (2010) 60, 86-101.

---

[6] For recent research on eye movements in acquired dyslexia, see Ablinger et al. 2014 and Schattka, Radach & Huber 2010.

5. Hawelka, S.; Gagl, B.; Wimmer, H. 2010. „A dual-route perspective on eye movements of dyslexic readers”. *Cognition* 115, 367-379.
6. Hutzler, F.; Wimmer, H. 2004. “Eye movements of dyslexic children when reading in a regular orthography”. *Brain and Language* 89(1), 235-242.
7. Hutzler, F.; Kronbichler, M.; Jacobs, A. M.; Wimmer, H. 2006. „Perhaps correlational but not causal: No effect of dyslexic readers‘ magnocellular system on their eye movements during reading”. *Neuropsychologica* 44, 637-648.
8. Jainta, S.; Kapoula, Z. 2011. “Dyslexic children are confronted with unstable binocular fixation while reading”. *PLOS One* 6, 1-10.
9. Kim, S.; Lombardino, L. J.; Cowles, W.; Altmann, L. J. 2014. “Investigating graph comprehension in students with dyslexia: An eye-tracking study”. *Research in Developmental Disabilities* 35, 1609-1622.
10. Kim, S.; Wiseheart, R. 2017. “Exploring Text and Icon Graph Interpretation in Students with Dyslexia: An Eye-tracking Study”. *Dyslexia* 23, 24-41.
11. Pan, J.; Yan, M.; Laubrock, J.; Shu, H.; Kliegl, R. 2013. „Eye-voice span during rapid automatized naming of digits and dice in Chinese normal and dyslexic children”. *Developmental Science* 16(6), 967-979.
12. Pan, J.; Yan, M.; Laubrock, J.; Shu, H.; Kliegl, R. 2014. „Saccade-target selection of dyslexic children when reading Chinese”. *Vision Research* 97, 24-30.
13. Prado, C.; Dubois, M.; Valdois, S. 2007. “The eye movements of dyslexic children during reading and visual search: Impact of the visual attention span”. *Vision Research* 47, 2521-2530.
14. Silva, S.; Faísca, L.; Araújo, S.; Casaca, L.; Carvalho, L.; Petersson, K. M.; Reis, A. 2016. “Too little or too much? Parafoveal preview benefits and parafoveal load costs in dyslexic adults”. *Annals of Dyslexia* 66, 187-201.
15. Thaler, V.; Urton, K.; Heine, A.; Hawelka, S.; Engl, V.; Jacobs, A. M. 2009. „Different behavioral and eye movement patterns of dyslexic readers with and without attentional deficits during single word reading”. *Neuropsychologia* 47(12), 2436-2445.
16. Trauzettel-Klosinski, S.; Koitzsch, A. M.; Dürrwächter, U.; Sokolov, A. N.; Reinhard, J.; Klosinski, G. 2010. „Eye movements in German-speaking

children with and without dyslexia when reading aloud”. *Acta Ophthalmologica* 88, 681-691.

17. Vagge, A.; Cavanna, M.; Traverso, C. E.; Iester, M. 2015. “Evaluation of ocular movements in patients with dyslexia”. *Annals of Dyslexia* 65, 24-32.
18. Yan, M.; Pan, J.; Laubrock, J.; Kliegl, R., Shu, H. 2013. „Parafoveal processing efficiency in rapid automatized naming: A comparison between Chinese normal and dyslexic children”. *Journal of Experimental Child Psychology* 115, 579-589.

b. Comparison

In the comparison of the papers, the following differences in the studies and their possible effects on the results are discussed. The first aspect, in which the studies differ from each other, is their focus, their hypotheses, and expectations. Second, as has already been mentioned above, these studies differ considerably in their experiment settings. First, there is variation in the inclusion-exclusion criteria applied for the choice of participants. Usually these include scores on standardized reading tests, IQ tests with differing thresholds and varying definitions of dyslexia and reading-difficulties. Second, there are differences in the languages spoken by the participants. Since German, for example, is considered to have a regular orthography in contrast to English, which serves as example for a deep orthography, the language in which the study is conducted can certainly influence the results. Furthermore, the number and age of the participants can vary considerably. Age might be an especially important factor, as reading proficiency is considered to increase with the number of years of reading instruction and practice. Apart from the sample of participants, the tasks and stimuli used in an experiment certainly make for additional factors of variation to be considered. They can include letter reading, single-word reading as well as text reading, reading of real words as well as reading of pseudowords, and reading silently as well as reading aloud. Furthermore, there are differences in the eye-tracking device used in the experiment and the eye movement parameters investigated. There is not only a variety of devices, such as goggles as well as fixed devices with chin rests, but they also differ in their technical specs, such as their sampling frequency.

Taken together, all these aspects might generate different key results and interpretations. In the following part, the eighteen studies introduced above are compared and analyzed according to the prior mentioned parameters.

i. Synchronic analysis

Before the studies are grouped and compared according to their experiment settings and results, a synchronic analysis might shed light on the extent to which these papers differ. First, their different foci, hypotheses and expectations are compared before the samples of participants and the experiment setting of each is shortly summarized. Finally, the findings and interpretations conclude the synchronic analysis.

1. *Focus, Hypotheses, Expectations*

The reason why the focus of a study has influence on the research itself and the results might seem trivial. Since hypotheses lead the way of studies, in this part only the different foci are discussed, with no additional analysis as to why this might influence the findings.

One specific aspect, on which some studies focus, is saccade and vergence control and computation. Jainta & Kapoula (2011) investigated saccade and vergence control during real text reading of dyslexic and normal readers in order to evaluate the maintenance of the vergence angle appropriate for both groups. They expected if oculomotor deficits in the maintenance of fixations actually exist in dyslexic readers, this could disturb the fusional process and it might, thus, be necessary to distinguish which particular aspect of binocular coordination is deficient in reading.

More generally, Bucci, Brémond-Gignac, & Kapoula (2008) focused on the latency of saccades and vergence in eye movements in dyslexic children in a non-reading task. They examined the latency of saccades at far and near distance, of convergence and divergence, and of combined saccade-vergence movements in dyslexics and age-matched controls.

In a rather different field is the work of Pan et al. (2014) who examined saccade-target selection of dyslexic children when reading Chinese. Their aim was to provide evidence that FLP (First Landing Position) acts as indicator of parafoveal word segmentation, which is necessary in Chinese, as there are no inter-word spaces. The authors hypothesized that if dyslexics are more affected by the absence of overt word boundaries, there should be more pronounced differences in the saccade-targeting of first-fixation landing positions in single-fixation cases than multi-fixation cases due to a smaller perceptual span of dyslexic children.

The importance of parafoveal vision is another focus some studies set. Staying with Chinese, Yan et al. (2013) investigated parafoveal processing efficiency in rapid automatized naming (RAN). Their main goal was to determine whether dyslexics operate with smaller perceptual spans when performing RAN based on the question of whether normal readers obtain larger amounts of parafoveal information in a task with oculomotor and saccade programming demands similar to normal reading. The authors expected dyslexics to spend more attentional resources on the task of translating symbols to phonology if automaticity played a key role. Dyslexics were expected to show restricted perceptual spans, as fewer resources are available for parafoveal processing, presumably due to local processing difficulties. Thus, dyslexics should suffer less from removing parafoveal preview.

More generally, Silva et al. (2016) tested two different hypotheses on parafoveal dysfunction in dyslexics with RAN. The reduced parafoveal preview benefits hypothesis indicates that dyslexics have “too little” parafovea available. If this was the case, the authors expected dyslexics to ignore parafoveal input in serial RAN. On the other hand, the increased parafoveal load cost indicates that there is “too much” parafoveal vision in dyslexics. Thus, they would expect dyslexic participants to be confused by the presence of parafoveal input in serial RAN. Furthermore, Silva et al. aimed at determining whether phonology is involved in each of the two hypothetical parafoveal dysfunctions in a silent letter-finding task.

A further RAN study was conducted by Pan et al. (2013) who investigated the eye-voice spans (EVS) during RAN of digits and dice in Chinese normal and dyslexic children. In this experiment, digits served as alphanumeric and dice as symbolic stimuli which evoke the same phonological response. Thus they have identical output demands with different degrees of automaticity of mapping symbols to phonological output. The authors expected larger group differences in digit RAN, larger EVS for digit RAN, and larger EVS for normal readers.

In order to find out how naming speed is related to meaning, Al Dahhan et al. (2014) conducted an eye-tracking study and investigated eye movements of university students with and without dyslexia during naming speed tasks. The expectations were shorter fixations and longer saccades for non-impaired readers and that fixation duration and saccade length predicted individual differences in reading.

Another aspect some studies focused on are the effects of word frequency, word length and word type. For example, De Luca et al. (2002) compared eye movement patterns of dyslexic and normal readers in short and long word and pseudoword reading in Italian. The authors hypothesized that dyslexics might show similar eye movement patterns regardless of the lexical value of the letter strings and that they would experience a length effect in both words and pseudowords. In contrast, controls were expected to discriminate between words and pseudowords and that they would only experience a length effect for pseudowords.

Dürrwächter et al. (2010) investigated how the difficulty of reading material affects the eye movement patterns of young German dyslexic readers and their controls in order to uncover word length, word frequency and length-by-frequency effects. Furthermore, they compared the results to other regular languages and English which leads yet to another focus for research in this field: the focus on orthography and the comparison of different languages and orthographies.

An example would be Hutzler & Wimmer (2004) who examined eye movements of dyslexic children when reading in a regular orthography. They expected dyslexics to reveal abnormal eye movement patterns in everyday reading situations but differences between dyslexics and non-dyslexics to be reduced in pseudoword reading. Further, the authors compared results for German, English and Italian.

Trauzettel-Klosinski et al. (2010) investigated the influence of different levels of phonological difficulty of reading material in German on reading strategies of dyslexic and normal readers. They expected dyslexic readers to perform better on easier texts. The authors compared their findings to English results in order to find out if different reading strategies are used in different languages.

Pursuing the question of different reading strategies, Hawelka, Gagl, & Wimmer (2010) examined a dual-route perspective on eye movements of dyslexic readers using two different reading models (dual-route cascaded model and E-Z reader model). They used these models to predict and evaluate dyslexic readers' eye movements. The authors expected systematic "overshoots" of saccades for short words and "undershoots" for long words in dyslexic readers. Furthermore Hawelka et al. found it likely that orthographic recognition failures may have resulted in a general tendency to target the beginnings of words.

In order to investigate the source of dyslexia and to explore possible visual and oculomotor deficits in dyslexic children, Hutzler et al. (2006) studied the effects of the magnocellular system on dyslexic readers' eye movements. The authors used a task of immediate relevance, which means that its oculomotor and perceptual demands are identical or functionally equivalent to those required for reading. For this task, they formed consonant strings by replacing vowels in pseudowords by consonants. The perceptual task was to search for two adjacent identical letters, the other task was to read the original pseudowords. The research team hypothesized that if dyslexic readers do have poor oculomotor control and visual perception, they should perform worse than normal readers in both tasks; however if dyslexic readers perform as well as unimpaired readers during string processing but worse during reading, the source of the problem is not likely to be found at the level of oculomotor control or visual perception.

Kim et al. (2014) and Kim & Wiseheart (2017) have a similar focus with different perspectives. The former examined and compared various aspects of graph comprehension in college students with and without dyslexia. In particular, they investigated the role of graphic properties, question types and the viewer's characteristics on graph comprehension. In their study, the authors aimed to explore the hypothesis that reading skill is a core component of graph tasks. Dyslexic students were predicted to be slower in answering questions on the graphs and to show longer fixations on verbal areas.

The latter compared text and icon graph interpretation in students with dyslexia. Their goal is to explore whether orthographic processing difficulties that underlie inefficient text reading in dyslexia can also account for inefficient graph interpretation. The authors hypothesized that students with dyslexia might use graphs as effectively as their peers as long as graphic displays were free of orthographic information.

In their paper, Prado, Dubois, & Valdois (2007) investigated the impact of the visual attention span on dyslexic children’s eye movements during reading and visual search. The purpose of this study is to provide evidence for a visual attention span dysfunction as a potential source of eye movement disorders in dyslexia. The authors hypothesized that if reading familiar words involves a larger visual attention span than visual search or unfamiliar letter strings, visual attention span reduction should more severely impact performance in reading than in visual search. Dyslexic participants were expected to display a higher number of rightward fixations in reading. The difference between dyslexics and controls, however, was expected to be smaller in visual search.

A final potential aspect to focus on in eye-tracking studies on dyslexia is the prospect of finding dyslexia-specific patterns and, thus, to propose future use of eye-tracking in the diagnosis. One particular study, which is discussed in more detail below, is the paper by Benfatto et al. (2016). Another study to be discussed here is Thaler et al. (2009) who analyzed different eye movement patterns of dyslexic readers with and without attentional deficits, of readers with attentional deficits only and of control readers during single word reading. They hypothesized that the different groups of participants would differentially be affected by word length and cluster density.

Similarly, Vagge et al. (2015) evaluated ocular movements of children with dyslexia to analyze the relationship between dyslexia and eye movements and to assess whether an analysis of eye movements can be useful in the identification of dyslexia in children.

It might be evident now that there are many different directions of research in the field of eye-tracking in dyslexia. It may again seem trivial to mention that different hypotheses require different experiment settings.

2. Experiment setting

As has been mentioned above, there are several aspects of experiment setting in which research can differ from one another. In the following, the sample of participants, the tasks and stimuli, the eye-tracking devices and the eye movement parameters investigated are discussed in more detail.

Participants

The eighteen studies to be analyzed vary notably in their samples of participants as well as their inclusion and exclusion criteria. In particular, they have different numbers of participants, different ratios of males and females, different mean ages, different thresholds for the diagnosis of dyslexia, different tests to measure reading skills and intelligence, different exclusion criteria, and different languages. However, all studies controlled for normal intelligence or cognitive abilities and almost all studies controlled for visual acuity.

Seven of the studies investigated 20-30 participants. Bucci et al. (2008) had 30 participants of which 16 were diagnosed with dyslexia and 14 controls. There is no indication about the number of males and females. A total number of 22 participated in De Luca et al. (2002), of which 12 were in the dyslexia group and 10 controls. All participants were male. Also a total of 22 participants were surveyed in Hutzler & Wimmer (2004). The dyslexia and control groups both counted 11 participants, and all participants were male. In Vagge et al. (2015), the number of children participating was 22 as well. 11 were dyslexic and the other half regular readers. In the dyslexic group, there were 7 males and 4 females, in the control group 6 males and 5 females. Hutzler et al. (2006) conducted two experiments with 22 participants in the first and 26 participants in the second experiment. Exactly half of the participants were dyslexic, the other half regular readers, one half female and one half male. Jainta & Kapoula (2011) had the smallest sample size of 20 participants in their study, of which 13 were dyslexic and 7 control readers. 10 of the 13 dyslexic readers and 3 of the 7 controls were male. Finally, Prado et al. (2007) conducted their experiment with 28 participants, half dyslexic and half normal readers. There were 11 males and 3 females in the dyslexic participants, and 9 males and 5 females in the control group.

Five of the former mentioned studies conducted their experiments with 30-40 participants. For example, Dürrwächter et al. (2010) counted 16 participants in the dyslexia group and 16 in the control group, which makes a total of 32 participants. 10 of the dyslexic participants and 12 of the controls were male. In Hawelka et al. (2010), the number of 36 participants is also evenly spread across the two groups, dyslexic and control. All participants were male. Kim et al. (2014) counted a total of 35 participants, of which 15 were dyslexic and 20 regular readers. There were 5 males and 10 females in the dyslexic, and only 2 males and 18 females in the control group. Silva et al. (2016) surveyed 34 participants, again equally divided into 17 dyslexic and 17 control participants. In each group, there were 7 males and 10 females. Finally, Trauzettel-Klosinski et al. (2010) had 32 participants, 16 dyslexic and 16 normal readers. In the dyslexia group there were 12 males and 4 females, in the control group there were 10 males and 6 females.

Six of the studies had larger samples. Al Dahhan et al. (2014) had a sample of 47 participants, of which 20 were diagnosed with dyslexia and 27 served as control group. 8 of the participants in the dyslexia group were male and 12 female, while there were 9 male and 18 female participants in the control group. In Pan et al. (2013), there were a total of 56 participants, of which 30 were dyslexic and 26 normal readers. The dyslexic group was equally divided into male and female; the control group counted 11 males and 15 females. In their (2014) study, Pan et al. counted 62 participants, of which 33 were diagnosed with dyslexia and 29 were controls. The dyslexia group consisted of 18 males and 15 females, while the control group contained 13 males and 16 females. A similar number of participants were surveyed in Yan et al. (2013), in particular 63. The dyslexic group, 20 males and 15 females, counted a total of 35, while in the control group there were 13 males and 15 females, 28 taken together.

Finally, there are two studies with even bigger samples of participants. In their recent study, Kim & Wiseheart (2017) counted a total of 77 participants, of which 29 were dyslexic and 48 normal readers. In both groups there were 10 male participants, the rest were female. Due to their specific focus, Thaler et al. (2009) needed 74 participants for their experiment. They were divided into four groups: 20 children with dyslexia only, of which 13 were boys and 7 girls; 20 children with dyslexia and an attentional deficit, of which 16 were boys and 4 girls; 14 children with attentional deficits only, of which 10 were boys and 4 girls; and 20 children in the control group, of which 12 were boys and 8 girls.

Taken together, a total of 740 participants were surveyed in these studies. Even though not all studies controlled for exact same group sizes (dyslexic and controls), in total numbers the exact half of the participants, namely 370, had been diagnosed with dyslexia. With regard to the ratio of male and female participants, one study had to be excluded as no information on the sex of the children was given (Bucci et al., 2008). Of a total of 710 participants, 397 were male and 313 female. This makes a ratio of 56:44.

Another important aspect of the sample constellation is the mean age of the participants. The majority of the papers, in particular 13, concentrated on children under the age of 14 years. Four studies concentrated on participants in primary school age from 8.9 to 9.6 years (Dürrwächter et al., 2010: DYS: 9.5, CON: 9.6; Thaler et al., 2009: DYS: 9.5, CON: 9.5; Trauzettel-Klosinski et al., 2010: DYS+CON: 9.5; Vagge et al., 2015: DYS+CON: 9.4). Further nine studies conducted their research with children in early high school age from 11 to 14 years (Bucci et al., 2008: DYS: 11.12 CON: 12.08 ; De Luca et al., 2002: DYS: 13.1 CON: 12.4; Hutzler et al., 2006: DYS: 13.6 CON: 13.3; Hutzler & Wimmer, 2004: DYS: 13.6 CON: 13.3; Jainta & Kapoula, 2011: DYS: 11.7 CON: 12.7; Pan et al., 2013, 2014: DYS: 10..7, CON: 10.6; Prado et al., 2007: DYS: 11.1, CON: 10.8; Yan et al., 2013: DYS: 10.75, CON: 10.65). Since Hutzler et al. (2006) conducted two studies with two groups of participants, the mean ages of the second group of participants should be given here. For the dyslexia group the mean age was 15.9 years and 15.3 years for the control group. Finally, there are five studies which were conducted with young adult participants aging 17.8-26.35 years (Al Dahhan et al., 2014: DYS: 24.59, CON: 21:52; Hawelka et al., 2010: DYS: 17.8, CON: 17.6; Kim et al., 2014: DYS: 20.89, CON: 19.71; Kim & Wiseheart, 2017: DYS: 21, CON: 21.48; Silva et al., 2016: DYS: 26.35, CON: 26.12).

Taken together, the mean age of all dyslexic participants is 14.25 years while the mean age of the control participants is 13.51.

One specifically crucial factor in the choice of participants for these studies, are the inclusion criteria for the dyslexia groups. In all of the studies, some kind of reading test was conducted first with the potential participants in order to group them into dyslexic and normal readers, although it is not completely clear if all used standardized reading tests. Among the standardized reading tests used were the L2MA battery for French (Bucci et al., 2008; Jainta & Kapoula, 2011), the Zürcher Lese Test (Dürrwächter et al., 2010; Trauzettel-Klosinski et al., 2010), the Würzburg Silent Reading Test (Dürrwächter et al., 2010), The Salzburger Lese Rechtschreibtest (Dürrwächter et al., 2010; Trauzettel-Klosinski et al., 2010), the reading test of the Wechsler Adult Intelligence Scale (Hawelka et al., 2010), the Test of Word Reading Efficiency (Kim & Wiseheart, 2017), the “Alouette Reading Test” (Prado et al., 2007), the ODEDYS test (Prado et al., 2007), the MT reading test (De Luca et al., 2002), and the Battery for Assessment of Developmental Dyslexia and Dysorthographia-2 (DDE-2) (Vagge et al., 2015). The thresholds set for these reading tests differed as well. While in some studies dyslexic participants were defined by scores at least 2 standard deviations below the mean (Bucci et al., 2008; Vagge et al., 2015), others used less strict thresholds of 1.5 SD (Dürrwächter et al., 2010; Pan et al., 2013, 2014; Silva et al., 2016; Trauzettel-Klosinski et al., 2010) or even 1 SD below the mean (Al Dahhan et al., 2014; Kim et al., 2014; Kim & Wiseheart, 2017). Others put the thresholds for results indicating dyslexia at or below the 16th percentile (Trauzettel-Klosinski et al., 2010), the 15th percentile (Hutzler et al., 2006; Hutzler & Wimmer, 2004) or the 10th percentile (Hawelka et al., 2010; Hutzler et al., 2006; Hutzler & Wimmer, 2004). Yan et al. (2013) tested their participants with a standardized test based on a character recognition task but did not inform on the specific test nor on the results. Thaler et al. 2009 set the threshold at a reading quotient below 85 on a standardized German reading test by Mayringer & Wimmer (2003). De Luca et al. (2002) included children with a marked delay on the MT reading test. Four of the studies also required external diagnosis of dyslexia for their participants. In Bucci et al. (2008), children received a complete evaluation of their dyslexia state by a pediatric hospital. Jainta & Kapoula (2011) only included children with diagnoses from specialized schools, medical centers or children’s hospital services. Also Prado et al. (2007) required a clinical examination and Silva et al. (2016) a formal diagnosis of dyslexia. Only three studies required (self-) reported histories of reading or spelling difficulties (Al Dahhan et al., 2014; Kim et al., 2014; Silva et al., 2016).

Most studies used adaptions of the Wechsler Scale of Intelligence to control for normal intelligence in the participants (WISC III, HAWIK-II, HAWIK-III, C-WISC, WAIS III; WAIS-R) with normal Intelligence Quotient (IQ) being defined as over 85 (Al Dahhan et al., 2014; Bucci et al., 2008; Dürrwächter et al., 2010; Pan et al., 2013, 2014; Silva et al., 2016; Trauzettel-Klosinski et al., 2010; Yan et al., 2013) and once over 90 (Hawelka et al., 2010). Four studies used tests of cognitive abilities, namely the Primary Test of Cognitive Skills (Hutzler et al., 2006; Hutzler & Wimmer, 2004) and the Woodcock-Johnson III Test of Cognitive Abilities (Kim et al., 2014; Kim & Wiseheart, 2017). In three of the studies colored progressive matrices were used, in particular Raven's Coloured Progressive Matrices (De Luca et al., 2002; Prado et al., 2007) and Coloured Progressive Matrices, which were not further described (Thaler et al., 2009). Unfortunately, there is no indication how Jainta & Kapoula (2011) evaluated the IQ of their participants.

Finally, the last important factor in which these study differ, is the language in which an experiment is conducted. Most of the studies discussed here were conducted in German (Dürrwächter et al., 2010; Hawelka et al., 2010; Hutzler et al., 2006; Hutzler & Wimmer, 2004; Thaler et al., 2009; Trauzettel-Klosinski et al., 2010), followed by English (Al Dahhan et al., 2014; Kim et al., 2014; Kim & Wiseheart, 2017), Chinese (Pan et al., 2013, 2014; Yan et al., 2013), French (Bucci et al., 2008; Jainta & Kapoula, 2011; Prado et al., 2007), Italian (De Luca et al., 2002; Vagge et al., 2015), and finally Portuguese (Silva et al., 2016).

The following table offers an overview of the various factors discussed above. In the following part, the relevance of these factors are evaluated and discussed.

Table 3: sample of participants- differing factors

PAPER	TOTAL	DYS: CON.	MALE: FEMALE	MEAN AGE	DYSLEXIA	IQ	LANG.
Al Dahhan et al. 2014	47	20:27	DYS: 8 M, 12 F CON: 9 M, 18 F	DYS: 24.59 CON:21.52	Self-reported history; reading fluency score at least 1 SD below mean	Wechsler Abbreviated Scale of Intelligence	English
Bucci et al. 2008	30	16:14	DYS: CON:	DYS: 11.12 CON: 12.08	Pediatric hospital-complete evaluation of dyslexia state; scores beyond 2 SD on L2MA battery (standard test by Applied Psychology Centre of Paris	WISC III	French
De Luca et al. 2002	22	12:10	DYS: 12 M CON:10 M	DYS: 13.1 CON: 12.4	Marked reading delay on standard reading test	Raven's Coloured Progressive Matrices	Italian
Dürrwächter et al. 2010	32	16:16	DYS: 10 M, 6 F CON: 12 M, 4 F	DYS: 9.5 CON: 9.6	Zürcher reading test, Würzburg silent reading test, Salzburger Lese Rechtschreibtest at least 1.5 SD below expected	IQ>85 HAWIK-III	German
Hawelka et al. 2010	36	18:18	DYS: 18 M CON: 18 M	DYS: 17.8 CON: 17.6	Reading test Wechsler Adult Intelligence Scale below percentile 10	IQ>90 WAIS-R	German
Hutzler &Wimmer 2004	22	11:11	DYS: 11 M CON: 11 M	DYS: 163 M CON:160 M	Reading rate lower than percentile 15 in individually administered reading test in Grade 3 and present (Grade 7)reading score lower than percentile 10	IQ>85 Primary Test of Cognitive Skills	German
Hutzler et al. 2006	22/26	11:11 13:13	DYS: 11 / 13 M CON: 11/13 M	DYS: 163/ 191M CON:160/ 183M	Reading rate lower than percentile 15 in individually administered reading test in Grade 3 and present (Grade 7)reading score lower than percentile 10	IQ>85 Primary Test of Cognitive Skills	Presum. German
Jainta& Kapoula 2011	20	13:7	DYS: 10 M, 3 F CON: 3 M, 4 F	DYS: 11.7 CON: 12.7	Classified by specialized schools, medical centres or children's hospital services; L2MA battery	Normal IQ	French
Kim et al. 2014	35	15:20	DYS: 5 M, 10 F CON: 2 M, 18 F	DYS: 20.89 CON: 19.71	Scored at or below 1 SD of mean on Test of Word Reading Efficiency and on at least one phonological processing subtest of Comprehensive Text of Phonological Processing; Reported history of spelling difficulties	Woodcock Johnson III Test of Cognitive Abilities	English
Kim&Wiseheart 2017	77	29:48	DYS: 10 M, 19 F CON: 10 M, 38 F	DYS: 21 CON:21.48	Score below 1 SD of mean on word reading from Test of Word Reading Efficiency	Woodcock-Johnson III Test of Cognitive Abilities	English
Pan et al. 2013	56	30:26	DYS: 15 M, 15 F CON: 11 M, 15 F	DYS: 10.7 CON: 10.6	Standard character recognition test at least 1.5 SD below respective age means	IQ matched controls (Picture completion in the Wechsler Intelligence Scale for Chinese Children); dyslexic score>85 in C-WISC with two exceptions	Chinese
Pan et al. 2014	62	33:29	DYS: 18 M, 15 F CON: 13 M, 16 F	DYS: 10.7 CON: 10.6	Standard character recognition test at least 1.5 SD below respective age means	Normal IQ (WISC)	Chinese
Prado et al. 2007	28	14:14	DYS: 11 M, 3 F CON: 9 M, 5 F	DYS: 11.1 CON: 10.8	Clinical examination, 45 months between chronological and reading age	Normal IQ Raven Matrices	French
Silva et al. 2016	34	17:17	DYS: 7M, 10 F CON: 7 M, 10 F	DYS: 26.35 CON: 26.12	Self-reported history, formal diagnosis as child, 1.5 SD below mean on test for adult reading problems	Normal IQ (WAIS III)	Portuguese
Thaler et al. 2009	74	DYS: 20 DYS+ AD: 20 AD:14 CG:20	DYS: 13 M, 7 F DYS+AD: 16 M 4 F AD: 10 M, 4 F CG: 12 M, 8 F	DYS: 113.4 M DYS+AD: 113.5 AD: 106.9 M CG: 113.4 M	Reading quotient below 85 on standardized German reading test	Coloured Progressive Matrices, IQ>85	German
Trauzettel-K. et	32	16:16	DYS: 12 M/4 F	9.5	Zürcher Lese Test, spelling subtest of Salzburger Lese Rechtschreib Test,	Normal intelligence (HAWIK II)	German

al. 2010			CON: 10 M/ 6F		score below 16th percentile and at least 1.5 SD below performance expected		
Vagge et al. 2015	22	11:11	DYS: 7 M, 4 F CON: 6 M/ 5 F	DYS: 9.4 CON: 9.2	Diagnosed based on DSM-IV, reading speed at least 2 SD below expected, Test of reading and writing skills from battery for assessment of Developmental Dyslexia and Dysorthographia-2 (DDE-2)	Normal IQ>85 Wechsler	Italian
Yan et al. 2013	63	35:28	DYS: 20 M, 15 F CON: 13 M, 15 F	DYS: 10.75 CON: 10.65	Standard test based in character recognition task	Matched IQ (C-WISC)	Chinese

Relevance of number and age of participants

In the former chapter the participant samples of 18 eye-tracking studies were compared with regard to the number of participants, their age and further factors, such as the inclusion criteria and orthography, the relevance of which is discussed below. As for the number and age of participants, some papers differed notably. While there were studies with a total of 20 participants (e.g. Jainta & Kapoula, 2011), others tested over 60 participants (e.g. Thaler et al., 2009); some included younger children (e.g. Dürrwächter et al., 2010), others tested college students (e.g. Kim & Wiseheart, 2017). In the following part, possible effects of number and age of the participants on an experiment and its results are discussed.

As with regard to sample size, depending on the research field, there can be certain conventions, however in the end it is the research team that has to decide on the number of participants. As Ryan points out, there is still a remarkable amount of research conducted with inaccurate sample sizes, which certainly affects the results of an experiment (2013). On the other side, as larger sample sizes often imply higher costs, not all research teams have the opportunity of testing large samples (Ryan, 2013: 17f.). In 2017, Elston wrote a letter from the editor on sample size for the Journal of the American Academy of Dermatology, which he started with the words:

> “The appropriate sample size depends on the magnitude of difference that is clinically relevant, the degree of variability in the attributes being measured, and the level of precision of the measurements. Greater variability (a more heterogeneous population) requires a larger sample size”.

The relevance of the heterogeneity of dyslexic readers becomes more and more evident when investigating different aspects of dyslexic reading. Furthermore, Elston noted that even large effects possibly do not reach significance if the sample size is too small or variability too high (2017). To conclude, defining accurate sample sizes is difficult as researchers have to specify unknown parameters and this has to be done on assumptions. This does not mean that the eye-tracking studies discussed here included inaccurate sample sizes. However, for analysis and comparison, it can be helpful to bear the importance of sample size in mind (Elston, 2017).

Certainly, when investigating reading skills, the age of the participants can be crucial. Generally, reading skills develop and increase with age, instruction and practice. During the early stages of formal reading instruction, the basic visuo-attentional processes involved in reading are well established. Improvements of reading skills cause an increase of the amount of information that can be extracted and processed during a single fixation, which inevitably changes a person's eye movements during reading. Dyslexic readers often show a similar reading behavior as less experienced or younger readers or adults reading a challenging passage. Thus, defining clear dyslexia-specific characteristics in reading is not an easy task, yet an important one in order to guarantee objective identification of dyslexic readers (Bellocchi et al., 2013, Blythe & Joseph, 2011).

Adult readers with dyslexia are expected to have acquired certain mechanisms to compensate for their reading difficulties. However, children at the beginning of their formal reading instruction are less likely to have developed such compensational strategies yet (Trauzettel-Klosinski et al., 2010). Depending on which aspect a research team is focused on, an age span for the participants can be set. As it bears chances for defining research, it also bears dangers. Investigating dyslexia in young children can be very difficult due to the above mentioned aspect, as age only gives a rough estimation of a child's cognitive status and not all children are at the same reading level albeit being of equal age. Sometimes participants are not even at the same age, yet their reading skills are still compared. While most studies chose age-matched participants for their control groups, in some cases it might be interesting to compare dyslexic readers to reading-level matched controls.

Relevance of inclusion/exclusion criteria

Without having a clear and commonly accepted definition of dyslexia, it is difficult to set thresholds for diagnostics. Depending on different reading tests, participants scoring in the last 10th or 15th percentile are often identified as dyslexic readers. However, just because a reader's score is not yet in the official range of dyslexia, this does not mean that this person does not show reading difficulties. It is important to bear in mind that these are standardized tests and can rather be seen as snapshots of a person's reading skills, which can be influenced by a number of factors, such as test setting or personal condition (Benfatto et al., 2016). Furthermore, since at least some of these standardized tests are easily accessible in the Web, it cannot always be guaranteed for them to be administered by professional or otherwise qualified personnel.

For inclusion criteria in the studies compared here, the authors often defined their dyslexic participants by reading test scores 1-2 standard deviations below the mean. When choosing two standard deviations as threshold, the participant sample is likely to show more severe cases of dyslexia. On the contrary, if one standard deviation is set as threshold, the participant sample might be more heterogeneous including more and less severe cases of dyslexia. Reid and Fawcett stress the importance of similar samples in order to enable results to be generalized (2004: 4). For future research it might be interesting to account for the heterogeneity of the partaking dyslexics and to divide the participant sample into further subgroups.

Relevance of language and orthography

The languages in which the former mentioned studies were conducted range from English and German to Italian, French, Swedish, Portuguese as well as Chinese. The reason why differences in language are important to be considered is that each language follows an underlying orthography. Orthographies vary in how they represent the phonology of a language and may thus be more or less transparent (Everatt & Elbeheri, 2008: 427f.). Different kinds of orthography require different reading strategies. While German is considered to conform to a regular or narrow orthography, due to its relatively clear grapheme-phoneme correspondence, English , on the contrary, adheres to a deep orthography (Hutzler & Wimmer, 2004). Thus, pronunciations in German are relatively easy to produce by grapheme-phoneme, that is spelling-to-sound, conversion rules while same spellings in English can have different phonological realizations, as in *read* (Present simple tense of ‘to read’: /riːd/) and read (Past simple tense of ‘to read’: /rɛd/).

In 2001, Coltheart and colleagues introduced the dual-route cascaded model of visual word recognition and reading aloud (figure 1). Both routes start the same way by identifying visual feature units and letter units. Then, however, the routes are split. Following the arrows to the left describes the direct lexical route by which, after identifying letter units, the orthographic input lexicon is activated and linked to the semantic system as well as the phonological output lexicon which then gives information to the phoneme system leading to speech. Due to its deep orthography, this is the route presumably used by English speaking readers. By applying the lexical route, readers identify letter units and then recognize the whole word in order to get information on its pronunciation. This is where English dyslexics show difficulties. English dyslexic reading is characterized by a slow reading pace, a high number of regressions as well as relatively high error rates probably due to an inability to identify whole words caused by unstable eye movements (Hutzler & Wimmer, 2004; Everatt& Elbeheri, 2008: 431). The second route presented in the model is the sub-lexical route, also starting off with visual feature unit and letter unit identification, however then leading to the grapheme-phoneme rule system and further to the phoneme system, from which connections are also made to the phonological output lexicon and the semantic system. This is presumably the route applied by speakers of languages with regular orthographies, such as German or Italian. This is in accordance with German dyslexic readers displaying slow reading, with an extraordinarily high number of fixations and shorter saccades, but with relatively few errors (De Luca et al., 2002; Everatt& Elbeheri 2008: 431).

To conclude, the manifestation of dyslexia may vary from language to language due to different orthographies and is, thus, not language-independent. Possible different manifestations of dyslexia should also be considered in assessment measures used for identification of dyslexia. In two different languages, other measures might be most informative. However, as Everatt and Elbeheri (2008) argue, measures of phonological processing are still the best tool for predicting literacy difficulties regardless of the orthography.

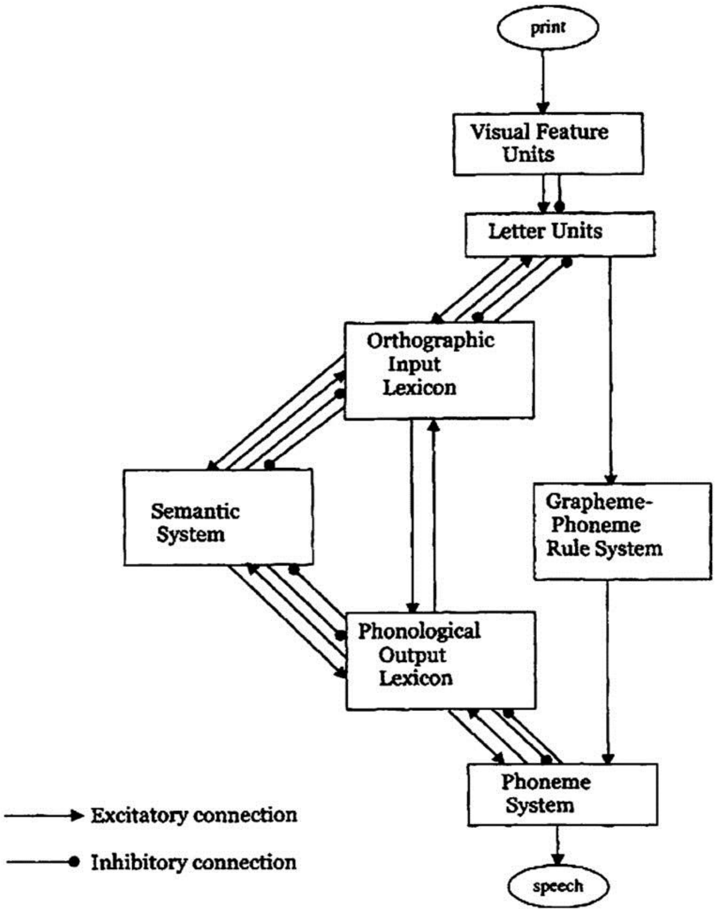

Figure 1: The dual-route cascaded model of visual word recognition and reading aloud, (Coltheart et al., 2001: 214)

As has been mentioned above, different orthographies require different reading strategies which are reflected in varying patterns of eye movements. However, the regularity of the orthography is not the only crucial aspect that can influence results. Even though German, English and Italian show different levels of regularity in their orthographies, they are similar in the sense of how written words are constructed: letter strings are divided into bigger units by blank spaces. Chinese, on the contrary, includes one-character and multi-character words which, in written form, do not show overt word boundaries. Thus, readers have to set the word boundaries in preprocessing, in order to successfully read the text, as Pan et al. (2014) point out.

Stimuli, tasks, procedures

Adjusted according to the varying hypotheses, the discussed eye-tracking studies certainly show differences in their stimuli, tasks and procedures. However, most of the studies included a reading task, in the sense of matching symbols to sounds.

Some of those studies included Rapid Automatized Naming tasks (RAN) and two studies in particular used letter naming as experimental task. Al Dahhan et al. (2014) included three letter naming speed tasks: one original task (a, d, o, p, s), one manipulated with regard to visual similarity (o replaced by q), and one manipulated with regard to phonological similarity (o replaced by v). For each naming speed task there were ten repetitions of five letters and the participants were instructed to name all letters as fast as possible.

Also one of the tasks used in Silva et al. (2016) is a modified RAN task. Their stimulus set included 28 RAN matrices containing the letters a-o-s-p-d pseudo-randomly ordered. Half of the letter items stood in the right parafoveal span of the preceding item and could be parafoveally previewed (P). The other half stood beyond the parafoveal span of N-1 and was not previewed (nP). Half of the items were followed by parafoveal inputs at the right and thus had parafoveal load (L) and the other half were not, so they had no parafoveal load (nL). Therefore, the task included four cases: PL (o **s** a), PnL (p **a**), nPL (**o** s), nPnL (**d**) with 98 letters each. The participants were instructed to name the letters in rows as fast as possible. The second task was a search task conducted on the same stimulus material. The participants were requested to look silently at the items, look for the sequence s-p, and respond with “yes” or “no” after each row.

Two further studies included digit RAN tasks. In their experiment, Yan et al. (2013) differentiated between continuous RAN, where the digits were presented on the screen simultaneously, and discrete RAN, where only the currently fixated item was displayed and no preview was available, in order to study the automaticity of the perceptual span in the RAN task. The participants were instructed to read all items on the screen as quickly and accurately as possible.

In contrast, Pan et al. (2013) compared eye-voice spans (EVS) in digit and dice RAN in Chinese dyslexics. The digits included 1,2,3,4,5 while the dice surfaces included the same values with 1-5 dots. One matrix included 30 items of one stimulus type and the participants were instructed to name these 30 items as rapidly as possible. There were 10 screens for each stimulus type, which makes a total of 300 digits and 300 dice surfaces.

The RAN tasks discussed above are not yet completely equivalent to reading. However, the following are. Dürrwächter et al. (2010), for example, included single word reading in their study. Participants had to read 40 single words of basic vocabulary, of which half were short words (4-5 letters), half long (7-12 letters), half high frequency and half low frequency (high freq., short; low-freq., short; high-freq., long; low-freq., long) as quickly and accurately as possible.

Thaler et al. (2009) chose a relatively similar approach. Their experiment also included single word reading of 30 items which were manipulated with regard to their length and cluster density (CD). Thus, they had four groups of stimuli as well: short-low CD, short-high CD, long-low CD, long-high CD. The instructions were to name the word aloud if and only if the participant knew the pronunciation of the whole word.

In contrast, De Luca et al. (2002) chose silent single word and pseudoword reading to approach their hypotheses. The stimuli contained 64 high frequency words and 64 pseudowords, of which half were short (4-5 letters) and half long (8-10 letters). The word lists were fully displayed and there was no time limit given. The experimenter read four words aloud after each trial and the participant had to decide if these were part of the list.

Pseudoword reading was also part of the experiment setting of Hutzler et al. (2006), besides a search task. Both experiments conducted by this research group are discussed here. In their first experiment, they started with a string processing task, in which the participants had to read through 60 consonant strings of four characters each and should find two adjacent identical letters, as in VPLL. If a target was found, they were instructed to respond with “yes”. The second part of this experiment was silent pseudoword reading of 60 stimuli, of which only the 20 short items (monosyllabic with four letters) were included in the analysis. The second experiment was very similar in design. The string processing task was the same, only with longer stimuli of 5-6 letters and 3 adjacent identical letters to be found, in order to elicit a higher number of fixations. This task included 120 items. The pseudowords to be read were also longer (tri-syllabic with 6 letters) and had to be read aloud.

In Hutzler & Wimmer (2004) there were two different stimuli as well. Participants were instructed to silently read through 60 pseudowords and two text passages with a total of 60 simple words.Hawelka et al. (2010) also chose silent sentence reading of 144 sentences (5-11 words) for their experiment.

In their 2014 study, Pan et al. asked their participants to read aloud 60 sentences of 15-23 Chinese characters with word lengths varying from 1-4 characters, of which 40 sentences were age-appropriate and 20 taken from the Bejing Sentence Corpus. A randomly selected third of the sentences was followed by yes-no questions.

Jainta & Kapoula (2011) included text reading as well, however silent. Participants had to silently read the text “L’alouette” at two viewing distances, 40cm and 100 cm with text sizes rescaled to the distances. Another study including silent text reading is Vagge et al. (2015). They instructed participants to silently read through an age appropriate text of 67 words.

In contrast to silent text reading, Trauzettel-Klosinski et al. (2010) instructed their participants to read two texts aloud. Both texts were taken from the Salzburger Lese- und Rechtschreibtest and differed in difficulty. The longer text with longer words and more composite nouns was presented first to not induce a training effect.

The last study also including text reading is Prado et al. (2007). They presented their participants with a 39 words long, age appropriate paragraph to be read aloud without time limit. However, similar to Hutzler et al. (2006) and Silva et al. (2016), the research team also included a visual search task. For this, Prado et al. (2007) replaced the vowels in the text by consonants and asked the participants to count the number of occurring “r”s without having a time limit.

Two of the papers have a very concrete focus, namely graph comprehension. Kim et al. (2014) conducted a graph comprehension task on three types of graphs (line, vertical bar, and horizontal bar graphs), two types of graph patterns (single or double pattern), and two types of questions (point location or comparison question). Each participant was confronted with a total of 144 graphs, which each included a question with two possible answers. The participants were instructed to answer the questions as rapidly and accurately as possible by pressing one of two buttons.

The experiment design used in Kim & Wiseheart (2017) is rather similar. This time, there were 3 information formats: a mixed modality condition (word-graph), an orthography free condition (icon graph), and an orthography-only condition (text). Again, the stimuli varied in information complexity (single-double graphs/ single-double sentence text), and question complexity (point location or comparison). The participants were instructed to view the 72 stimuli and to answer the questions as rapidly and accurately as possible, again without having time limit.

Finally, there is one study, which investigated pure eye movements without any textual stimuli. Bucci et al. (2008) placed LEDs in two isovergence circles at 20cm and 150cm distance, with 3 LEDs in the closer circle and 5 LEDs in the more distant circle. At the beginning of each trial, a fixation LED lit up at the center of one circle. If the target was in the same circle, it evoked a pure saccade. If the target was at the center of the other circle, it evoked pure vergence and if the target was lateral and on the other circle, it evoked combined saccade and vergence eye movement. The experiment was conducted in two paradigms: gap, in which there was a short pause between the disappearance of the fixation LED and the appearance of the target LED, and simultaneous, in which the disappearance of the fixation LED and the appearance of the target LED happened at the same time. There were 144 trials for each participant.

The following table offers an overview of the various tasks used in the eye-tracking studies discussed here.

Table 4: Overview of tasks and instructions

	Silent	Aloud|fast& accurately	Aloud
Letters		Al Dahhan et al. 2014 Silva et al. 2016	
Digits		Yan et al. 2013 Pan et al. 2013	
Dice		Pan et al. 2013	
Words	De Luca et al. 2002	Dürrwächter et al. 2010 Thaler et al. 2009	
Pseudowords	De Luca et al. 2002 Hutzler&Wimmer 2004 Hutzler et al. 2006		Hutzler et al. 2006
Sentences	Hawelka et al. 2010		Pan et al. 2014
Text	Hutzler&Wimmer 2004 Jainta&Kapoula 2011 Vagge et al. 2015		Prado et al. 2007 Trauzettel-K. et al. 2010
Search task	Hutzler et al. 2006 Prado et al. 2007 Silva et al. 2016		
Pure EM not reading	Bucci et al. 2008		
Graph comprehension→ questions fast& Accurately	Kim et al. 2014 Kim&Wiseheart 2017		

Relevance of task

As Reid and Fawcett put it, in order to compare results of various studies, research teams should have shared methodologies since even simple changes in the experiment design or evaluation can notably affect the results (2004:4). However, with regard to their hypotheses, each research team has to decide on an accurate task which yields the necessary results. In the eye-tracking studies discussed here almost all studies chose some sort of reading or naming task. Only one decided to include a non-reading task, which was discussed above in more detail (Bucci et al., 2008). With regard to the reading tasks, the reading mode required in a task can have an influence on eye movements. In their paper Krieber et al. (2017) evaluated intra-individual eye movements in oral and silent reading in a regular orthography. They tested 22 German-speaking children at the age of 13 without reading difficulties and analyzed their eye movements in silent and oral text reading. Their results show a significant influence of the reading mode on spatial and temporal eye movement parameters. The analysis of the data suggests silent reading to be faster and less demanding. Furthermore, in silent reading, increasing reading proficiency leads to less time needed to read a text, than in oral reading (Krieber et al., 2017). Although this study was conducted among unimpaired children, influence of the reading mode on eye movements can also be assumed in dyslexic children. Furthermore, these results are in accord with Ashby et al. who found that perceptual spans in university students are smaller in oral reading than in silent reading (Ashby, Yang, Evans, & Rayner, 2012). However, in their paper published in 2004, Hutzler and Wimmer claim that the reading rate deficit of dyslexic readers in silent reading corresponds to the rate deficit they show in reading aloud (Hutzler & Wimmer, 2004: 237). It seems not to be clear yet if the reading mode, silent or oral, influences dyslexic readers as much as normal readers.

Eye-trackers and eye movement parameters

The final aspects of the experiment settings discussed here are the eye trackers used for the experiments and the eye movement parameters investigated. Even though most eye trackers work similarly, there are still differences in their technical specifications. The vast majority of eye trackers is video-based and works with infrared light. In about half of the studies using video-based trackers, eye-tracking systems by Eyelink were used. Eyelink I, for example, was used in Prado et al. (2007) and is a head mounted device which works at a sampling frequency of 250Hz. The following device, the Eyelink II, is also head mounted but can be used with a sampling frequency of 250Hz (Hutzler et al., 2006) and 500Hz (Al Dahhan et al., 2014; Kim et al., 2014). The most used device is the Eyelink 1000/2K which works with a sampling frequency of 1000Hz and is available as tower mount or desktop device. Among the studies that used this device are Hawelka et al. (2010); Pan et al. (2013), (2014); Thaler et al. (2009); and Yan et al. (2013). Further video based infrared eye trackers are the ET4 AMTech device used in De Luca et al. (2002), with a sampling frequency of 500 Hz, and the head mounted device by Chronos Vision used in Jainta & Kapoula (2011), with a sampling frequency of 200Hz. Vagge et al. (2015) used the Ober 2 System goggles, which are also video based infrared eye trackers, with a rather modest sampling frequency of 100 Hz. The eye tracker with the highest sampling frequency is the SMI Hi-speed device with a rate of 1250Hz. It is a video based infrared tower mount device used in Silva et al. (2016). In contrast, the headfree desktop eye follower by LC Technologies used in Kim & Wiseheart

(2017) works at a sampling frequency of 120Hz. Finally, two research teams used an ISCAN desktop product with a 50Hz sampling rate (Hutzler et al., 2006; Hutzler & Wimmer, 2004). There are two eye-tracking systems that work differently to the above mentioned. First, there is the oculometer Dr. Bouis which works with a high number of photodetectors which are presented with an infrared image of the eye in order to compute the exact location of the pupil center. Its sampling frequency is 500Hz and it is used in Bucci et al. (2008). Second, two studies used a Scanning Laser Ophthalmoscope (SLO) which images the retina of the participant and the stimuli simultaneously (Dürrwächter et al., 2010; Trauzettel-Klosinski et al., 2010).

Before the relevance of the eye tracker and the sampling frequency is discussed, the various eye movement parameters investigated are analyzed. The following table offers an overview of the eye-tracking devices used and some of their technical specifications.

Table 5: Eye-tracking devices

Eye tracker	Sampl. freq.	Video based	Infrared	Device	Study
Eyelink I	250Hz	Yes	Yes	Head mounted	Prado et al. 2017
Eyelink II	250/ 500Hz	Yes	Yes	Head mounted	Al Dahhan et al. 2014 Hutzler et al. 2006 Kim et al. 2014
Eyelink 1000/2K	1000	Yes	Yes	Tower mount/ desktop system	Hawelka et al. 2010 Pan et al. 2013 Pan et al. 2014 Thaler et al. 2009 Yan et al. 2013
AMTech ET4	500	Yes	Yes		De Luca et al. 2002
Chronos Vision	200	Yes	Yes	Head mounted	Jainta & Kapoula 2011
Ober 2 System	100	Yes	Yes	Goggles	Vagge et al. 2015
SMI hi-speed	1250	Yes	Yes	Tower mount	Silva et al. 2016
LC Technologies eye follower	120	Yes		Headfree/ Desktop	Kim& Wiseheart 2017
ISCAN	50	Yes		Desktop	Hutzler et al. 2006 Hutzler&Wimmer 2004
SLO: Scanning laser ophthalmoscope		Yes		Images retina and the stimuli simultaneously	Dürrwächter et al. 2010 Trauzettel- Klosinski et al. 2010
Oculometer Dr. Bouis	500	This device presents an infrared image of the eye to a dense array of photodetectors which computes the location of the pupil centre			Bucci et al. 2008

With regard to the eye movement parameters investigated in these studies, three major types can be distinguished: temporal measures, spatial measures and total numbers and percentages.[7]

First, the temporal measures are analyzed. The most common eye-tracking measure in this type is some sort of fixation duration, which includes mean fixation duration (Al Dahhan et al., 2014; De Luca et al., 2002; Dürrwächter et al., 2010; Hawelka et al., 2010; Hutzler & Wimmer, 2004; Prado et al., 2007; Silva et al., 2016; Trauzettel-Klosinski et al., 2010), single fixation duration (Hawelka et al., 2010; Pan et al., 2013; Thaler et al., 2009; Yan et al., 2013), first fixation duration (Hutzler et al., 2006; Hutzler & Wimmer, 2004; Pan et al., 2013; Yan et al., 2013),first of multiple fixation duration (Hawelka et al., 2010), successive fixation duration (Hawelka et al., 2010), gaze duration (Hawelka et al., 2010; Hutzler et al., 2006; Kim & Wiseheart, 2017; Pan et al., 2013, 2014; Yan et al., 2013), first viewing time (Kim et al., 2014), and total viewing time (Kim et al., 2014). Other temporal measures, which are not necessarily eye movement measures, are the reading speed or reading time (De Luca et al., 2002; Thaler et al., 2009; Trauzettel-Klosinski et al., 2010), and the overall reaction time in one region (Kim et al., 2014; Kim & Wiseheart, 2017).

Second, there is a number of spatial measures analyzed in these studies. Probably most importantly, saccade size or amplitude is often investigated (Al Dahhan et al., 2014; De Luca et al., 2002; Hawelka et al., 2010; Pan et al., 2013, 2014; Trauzettel-Klosinski et al., 2010; Yan et al., 2013). Furthermore, saccade landing positions (Pan et al., 2013; Yan et al., 2013) and saccade launch sites (Pan et al., 2014) are analyzed. Other spatial measures are pure saccades (Bucci et al., 2008), pure vergence (Bucci et al., 2008), combined saccades and vergence (Bucci et al., 2008), and eye-voice-span (EVS) (Pan et al., 2013).

Third, there are simple count measures analyzed in the papers, usually referring to fixations or saccades and regressions. The ones referring to fixations include number of fixations (Al Dahhan et al., 2014; Hutzler et al., 2006; Prado et al., 2007; Thaler et al., 2009; Trauzettel-Klosinski et al., 2010) or number of fixations per word (Hawelka et al., 2010; Hutzler & Wimmer, 2004; Pan et al., 2014), number of rightward fixations

---

[7] For an overview and explanation of common eye movement parameters, see “Eye movement measures”

(Prado et al., 2007), number of forward re-fixations (Hawelka et al., 2010), number of skipped words (Hawelka et al., 2010), number of singly fixated words (Hawelka et al., 2010), number of multiply fixated words (Hawelka et al., 2010), and number of letters per rightward fixation (Prado et al., 2007). The ones referring to saccades or regressions include number of saccades (Al Dahhan et al., 2014; Dürrwächter et al., 2010), number of rightward saccades (De Luca et al., 2002; Trauzettel-Klosinski et al., 2010), number of regressions (Dürrwächter et al., 2010; Prado et al., 2007; Trauzettel-Klosinski et al., 2010), percentage of regressions (Dürrwächter et al., 2010; Hutzler & Wimmer, 2004; Trauzettel-Klosinski et al., 2010), number of regressive movements (De Luca et al., 2002), and number of between words regressions (Hawelka et al., 2010). Two further measures were the number of backward eye movements to the beginning of the next line (return sweeps) (Trauzettel-Klosinski et al., 2010), and general number of all eye movements (Kim & Wiseheart, 2017).

The following table offers an overview of the eye movement parameters analyzed in the studies.

Table 6: Eye movement measures in analysis

Measures	**Studies**
Temporal	
Fixation duration	Al Dahhan et al. 2014 De Luca et al. 2002 Dürrwächter et al. 2010 Hawelka et al. 2010 Hutzler&Wimmer 2004 Prado et al. 2007 Silva et al. 2016 Trauzettel-Klosinski et al. 2010
Gaze duration	Hawelka et al. 2010 Hutzler et al. 2006 Kim& Wiseheart 2017 Pan et al. 2013 Pan et al. 2014 Yan et al. 2013
First viewing time	Kim et al. 2014
Total viewing time	Kim et al. 2014
Single fixation duration	Hawelka et al. 2010 Pan et al. 2013 Thaler et al. 2009 Yan et al. 2013
First fixation duration	Hutzler&Wimmer 2004 Hutzler et al. 2006 Pan et al. 2013 Yan et al. 2013
First of multiple fixation duration	Hawelka et al. 2010

Successive fixation duration	Hawelka et al. 2010
Reading speed/time	De Luca et al. 2002 Thaler et al. 2009 Trauzettel-Klosinski et al. 2010
Overall reaction time in one region	Kim et al. 2014 Kim& Wiseheart 2017
Spatial	
Saccade size/amplitude	Al Dahhan et al. 2014 De Luca et al. 2002 Hawelka et al. 2010 Pan et al. 2013 Pan et al. 2014 Trauzettel-Klosinski et al. 2010 Yan et al. 2013
Saccade landing position	Pan et al. 2013 Yan et al. 2013
Launch Site	Pan et al. 2014
Pure saccades	Bucci et al. 2008
Pure vergence	Bucci et al. 2008
Saccades and vergence combined	Bucci et al. 2008
Eye-Voice Span	Pan et al. 2013
Number	
No of saccades	Al Dahhan et al. 2014 Dürrwächter et al. 2010
No of rightward saccades	De Luca et al. 2002 Trauzettel-Klosinski et al. 2010
No of regressions	Dürrwächter et al. 2010 Prado et al. 2007 Trauzettel-Klosinski et al. 2010
Percentage of regressions	Dürrwächter et al. 2010 Hutzler&Wimmer 2004 Trauzettel-Klosinski et al. 2010
No of between words regressions	Hawelka et al. 2010
No of regressive movements	De Luca et al. 2002
No of fixations	Al Dahhan et al. 2014 Hutzler et al. 2006 Prado et al. 2007 Thaler et al. 2009 Trauzettel-Klosinski et al. 2010
No of fixations per word	Hawelka et al. 2010 Hutzler&Wimmer 2004 Pan et al. 2014
No of rightward fixations	Prado et al. 2007
No of forward refixations	Hawelka et al. 2010
No of skipped words	Hawelka et al. 2010
No of singly fixated words	Hawelka et al. 2010
No of multiply fixated words	Hawelka et al. 2010
No of eye movements	Kim& Wiseheart 2017
No of letters per rightward fixation	Prado et al. 2007
No of backward eye movement to beginning of next line→ return sweep	Trauzettel-Klosinski et al. 2010
Reading	
Accuracy	De Luca et al. 2002 Thaler et al. 2009

Two studies, however, did not completely follow this pattern and adopted less conventional measures. Jainta & Kapoula (2011) calculated conjugate and disconjugate eye movements. For each saccade, they further calculated the amplitude and extracted the change in vergence between the saccade onset and offset. Furthermore, for each fixation period the research team calculated the absolute minimum amount of binocular fixation error (the moment in time at which the vergence error in respect to the actual viewing distance was smallest), the disconjugate drift in vergence (the change in vergence between the beginning of a fixation period and the minimum fixation disparity), the standard deviations of fixation disparity across the whole fixation period, and the fixation duration. Vagge et al. (2015) conducted a stability analysis while the participants were fixating a still target. Furthermore, they conducted an analysis of tracking saccades and of fixation pauses, speed reading, saccades and regressions.

Relevance of eye trackers and sampling frequency

As discussed above, there are several different eye-tracking devices used in these studies, even though a vast majority of them were infrared video based eye trackers. One exception is the scanning laser ophthalmoscope (SLO), which scans the retina and stimulus simultaneously. Generally, there are goggles, head mounted, tower-mounted, and desktop devices set up at different viewing distances, usually spanning from 45cm to 120cm. Often, there are head and chin rests in order to stabilize the participants head and to ensure better spatial resolution. However, the temporal resolution is also crucial, depending on the measures. Indicator for temporal resolution is the sampling frequency at which a device operates, which ranges from 50Hz (Hutzler & Wimmer, 2004) to 1250Hz (Silva et al., 2016). The usage of different eye-tracking devices working at different sampling frequencies makes direct comparison difficult: “Such comparisons are fraught with uncertainties such as noise in the temporal resolution of the eye-tracking systems, which only can be avoided in direct comparative studies” (Hutzler & Wimmer, 2004: 241). An interesting finding with regard to sampling frequency was that doubling the sampling frequency leads to a remarkable decline of data requirements. In particular, if the same average temporal sampling error should be maintained, data requirements would be lowered to one fourth. This indicates that a higher temporal resolution yields more exact results. Furthermore, one-point measures require only half the amount of data compared to two-point measures (e.g. gaze duration), again if the sampling error should be maintained (Andersson, Nyström, & Holmqvist, 2010). When comparing eye-tracking studies, it might be helpful to bear in mind that different eye trackers might yield different results. Direct comparison using the same devices might be interesting in future research.

3. *Key results and interpretations*

The final aspect of comparison, and probably the most important one, is the key results and their interpretations of the studies. In the following part, the aim of each paper is quickly revised and its findings are discussed.

Al Dahhan and colleagues (2014) investigated the eye movements of university students with and without dyslexia during three naming speed tasks. They expected shorter fixations and longer saccades for the non-dyslexic group, however longer fixations and shorter saccades for both groups in visually similar naming speed tasks. Furthermore, the research team expected fixation duration and saccade length to predict individual differences in reading. The results show that the NRD (no reading difficulties) group performed significantly better on each reading fluency measure, on phoneme elision and phonological choice. With regard to naming speed performance, the findings indicate that the NRD group was indeed faster in naming time but neither the naming speed task nor its interaction with group was significant. Both groups were, not significantly, slower on the visually similar task, which suggests that the visual similarity of the letters in the task have greater influence than the phonological similarity. When the strength of the correlation between naming speed task and reading were compared, there were no differences between the three tasks, which suggests that the mechanisms underlying the naming speed-reading relationship are not dependent on the item composition in the naming speed tasks. With regard to eye movement parameters, the research team could show that the groups significantly differed on each parameter. The NRD group displayed shorter fixation durations, fewer fixations, longer saccades and fewer saccades than the RD (reading difficulty) group. The lack of any significant group by task interaction implies that the task manipulations did not have a contrasting impact on the processing times of the dyslexic group. The RD group seemed to require more time to get the same amount of information, indicated by their longer fixation durations, and seemed to be either less efficient in their parafoveal processing or during the fixations, indicated by shorter and more frequent saccades. With regard to the correlations between naming speed times, eye movement parameters and reading outcomes, the authors found that total times on all three tasks significantly correlated with text reading speed and sight word efficiency for both groups. Fixation duration significantly correlated with text reading speed and sight word efficiency in both groups and is the only measure that predicted every reading outcome. Thus, the authors suspect this measure to be the best predictor for reading fluency in individuals. They assume access to phonological representations not to be the critical component in the naming speed-reading relationship, due to fixation duration being independent of phonological processing. However, the work of Al Dahhan and colleagues (2014) also comes short in some respects. The three naming speed tasks only differed in one letter which might be the reason why no significant differences among the tasks were found. Furthermore they conducted their research with a relatively small sample of young adults and not children, who would be more likely not to have learned any compensatory mechanisms yet.

Bucci et al. (2008) investigated the latency of saccades at near and far distance, the latency of divergence and convergence, and of combined saccade-vergence movements in dyslexic children and age-matched controls in a non-reading task. The first of their key results was that, for dyslexic children, the mean latency of all eye movements tended to be longer. However, the only significant difference is for saccades at far distance for both paradigms (gap and simultaneous), which means that distance seems to have an influence. The authors find it likely that triggering a saccade at far might require more involvement of the parietal-frontal pathways. Those are believed to be used for voluntary saccade initiating. If so, this might suggest dyslexics to possibly have problems with initiating voluntary eye movements. Relying on other research, the authors suspect visual attention to play an important role for dyslexics in switching between reflexive and voluntary eye movement initiation. The second key finding was that the gap paradigm significantly reduced the mean latency of all eye movement types for both groups. The third result showed that the occurrences of anticipatory and express latencies were significantly more frequent in the gap paradigm. The gap paradigm has been known to have a double effect, namely the emergence of express movements and the decrease of mean latencies. The decrease of mean latency due to the gap paradigm, in this study, occurs for all eye movement types and is similar in both groups, the dyslexic and control group. In accord with previous research, the express latencies for convergence are rare in both groups. However, the underlying mechanisms of express saccades are still widely debated. Interestingly, findings on the cortical structures involved in vergence control showed that in the initiation of saccades and vergence eye movements, the same cortical oculomotor areas are activated. The fourth and novel finding of this study was that express latencies occur for a larger variety of eye movements for dyslexics (especially in divergence), and mostly for saccades at near for controls. In conclusion, dyslexic participants showed longer latencies for saccades at far and have more frequent express divergence latencies. Bucci and colleagues (2008) attribute these results to difficulties in reflexive and voluntary control of attention to targets in a three-dimensional space.

In their study from (2010), Dürrwächter and colleagues investigated the word length and word frequency effect on eye movements of dyslexic children reading in a regular orthography. Furthermore, they compared their results to findings in other languages. First, dyslexic participants displayed a significantly higher number of saccades, in particular three times as high as the results for normal readers. Especially for long and low-frequency words, dyslexic readers made more saccades. In both groups, main effects of word length and word frequency are attributed to an increase of saccades. As for the number of regressions, the findings show a similar pattern. Dyslexic participants tripled the results of controls. For both measures, the word frequency effect was separately significant for both levels of word lengths and the word length effect was separately significant for both levels of frequency. With regard to the percentage of regressions, only the main effect of word frequency reached significance, which indicated that decreasing word frequency lead to an increased percentage of regressions, while word length did not. Finally, dyslexic participants showed 100 ms longer average fixation durations than controls. In either group, low frequency words evoked 50 ms longer mean fixation durations. Increasing word length only had an effect on dyslexic fixation durations, which prolonged on average by 50ms. So taken together, the number of saccades and regressions was larger in dyslexics, they showed longer fixation durations, but there were word length and word frequency effects in both groups with an additional increase for long low-frequency words, which was more pronounced in dyslexics. The authors assume, that the reason for the increase of eye movements in dyslexic participants are the smaller units that can be analyzed at once and not the inability of eye movement control. In this respect, the differences between German and Italian readers could be due to the use of different reading strategies depending on the orthography of a language. This might also explain why the percentage of saccades and regressions are higher in dyslexic children but, in regular orthographies, comparable to those of controls, as in contrast to irregular orthographies. Dürrwächter and colleagues (2010) noticed word length and word frequency effects in both groups. In particular, the percentage of regressions increased for low frequency words, however more pronounced in dyslexics. A possible explanation would be that low frequency words increase the need to return to areas previous to the current fixation. According to the authors, word frequency effects are likely to reflect information processing. In this respect, prolonged fixation durations of dyslexics, especially for long words, indicate information processing difficulties evoked by differences in word length and word frequency. To sum up, dyslexic and normal readers were both affected by word frequency, word length and a combination of both. Results can vary from language to language, due to differences in the regularity of grapheme-phoneme correspondence and differing reading strategies. According to Dürrwächter et al. (2010), their findings are in accord with other research indicating that reading difficulties in dyslexics reflect problems at higher word processing. One constraint to this study is the stimuli presentation in form of a word list that might have enabled a preview on the upcoming words.

Hawelka, Gagl, & Wimmer (2010) used two reading models (dual-route cascaded model and E-Z reader model) to predict and evaluate dyslexic reader’s eye movements. They expected “overshoots” of landing positions for short and “undershoots” for long words, as well as orthographic recognition failures resulting in a tendency to target word beginnings. Their findings show that dyslexic readers exposed twice the number of fixations in comparison to controls. Furthermore, dyslexics skipped a much smaller number words and the frequency of more than one fixation per word was increased. They showed frequent progressive fixations, which is due to frequent progressive re-fixations. As was partly expected by the authors, dyslexic readers tended to massively “undershoot” the centers of longer words. This means that their first fixation locations were less affected by word length than controls. Generally word length and word frequency affected the number of fixations and gaze durations of dyslexics more than of controls. The reading performance of dyslexic participants on short words resembled the performance of unimpaired readers on the longest words. A similar effect can be seen with the frequency effect. Furthermore, the dyslexic group exhibited an abnormally strong length by frequency effect for the number of fixations and gaze duration. While the length effect on high frequency words was close to normal, the length effects on low and medium frequency words were not. However, there was no word length effect noticed on the gaze duration for singly fixated words. Both groups showed the shortest fixation durations for words of 6-7 letters. Taken together, only for a reduced number of words, dyslexic readers showed efficient word processing. The authors interpreted this finding as a sign for a deficient orthographic lexicon. The prolonged single fixations were assumed to reflect an impaired access to the whole-word-phonology, and the prolonged multiple fixations slow access to the sub-lexical phonology. Further, Hawelka et al. (2010) suggest that the higher number of fixations may be an indicator for the reliance on smaller sub-lexical units.

The aim of the study of Hutzler et al. (2006) was to investigate dyslexic reader’s eye movements and the effects the magnocellular system might have on them. They expected dyslexics to perform worse than controls on search tasks and reading tasks if their oculomotor control and visual perception was indeed poor. Otherwise, if dyslexics would perform similar to controls on string processing but worse on reading, the problem might not be at the level of visual perception or oculomotor control but beyond. To test their hypotheses, the research team conducted two experiments. The results of the first experiment did not show obvious differences between the two groups in the string processing task with regard to the number of fixations, first fixation duration and gaze duration. Both groups showed a reliable length effect with more fixations for longer items. In contrast, the group differences in pseudoword reading were pronounced. Dyslexic readers did show a higher number of fixations, longer first fixation durations (about 290ms longer) and longer gaze durations (about 610ms longer). A correlation analysis showed that the correlation of reading speed with the three eye movement measures is reliable. However, even though there were no group differences in the string processing task, it might still be possible that inter-individual differences could be related to the functioning of each participant's magnocellular visual system, which is why Hutzler et al. (2006) conducted the second experiment with a similar design. The results for the second string processing task showed no main effect of group but a reliable main effect of length, similar to the first experiment. For first fixation duration, there are two interesting findings: there was a significant main effect of length in the subject based analysis and an effect of borderline reliability. Normal readers displayed longer first fixation durations on long items. With regard to gaze duration, participants tended to look at long items about 60ms longer than on short items. When comparing these results to pseudoword reading, a different picture can be seen: dyslexic participants exhibited 2.6 more fixations per item than the control group. Furthermore, for both groups first fixation duration was approximately 180ms longer for pseudowords than for the consonant strings. Third, during pseudoword reading dyslexics exhibited 1200ms longer gaze durations than controls. There was a high correlation between number of fixations and gaze duration with reading speed. According to the authors, these results suggest that the fixation durations of children approximate those of adults faster and might not be as strongly linked to reading skill as the number of fixations. The dyslexic group did not perform significantly different on the coherent motion task which indicates that the differences in eye movements between dyslexic and normal readers cannot be explained in group differences in the magnocellular system and its functionality. Furthermore, dyslexic readers might not have difficulties in the accurate perception of letters or oculomotor control. The eye movement patterns dyslexic children showed on the reading task rather reflect their difficulties in the reading process. All in all, these results indicate that there is no correlation between the visual perception and oculomotor control, and the functioning of the magnocellular system. For further research, the authors suggest this research to be replicated with other samples of participants and different string processing tasks.

In their 2004 study, Hutzler & Wimmer investigated the eye movements of dyslexic and normal children in text and pseudoword reading in a regular orthography.

Generally, their findings show that dyslexic readers needed more fixations of longer duration and more regressions. Especially the number of fixations of normal readers compared to dyslexic readers showed that dyslexics skipped words much less often. While controls skipped 36% of the words, dyslexics only skipped 18%. A close analysis of short versus long words showed that the number of fixations increases from short to long words more in dyslexics than in normal readers. For dyslexic readers, the duration of first fixations is almost twice as long as that of controls. There is a similar picture for gaze duration: poor readers exhibit a greater increase from short to long words. With regard to pseudoword reading, the results are similar. There was also a larger increase in the number of fixations from short to long words in dyslexics, longer first fixations and larger gaze durations. These German findings partly resembled Italian findings, which the authors explained with their similarity in relatively simple grapheme-phoneme correspondence. Unlike English-based results, the number and percentage of regressions was small. Further, the results suggest that German dyslexics, similar to Italian, differed from normal readers especially due to shorter forward eye movements. However, unlike the Italian findings, the German participants showed notably prolonged fixation durations. Hutzler & Wimmer (2004) interpreted their findings as being due to the different nature of German and Italian syllables. Even though comparing their results to other research, the authors also notice that comparisons of these kinds can never be completely accurate due to too many differences in the experiment settings.

Jainta & Kapoula (2011) investigated saccade and vergence control in dyslexic and normal readers during real text reading. They aimed at evaluating the maintenance of the appropriate vergence angle. If visual control deficits exist in this area in dyslexics, they could disturb the fusional process, and it might be helpful to differentiate which part of binocular coordination might be deficient. In summary, their findings of dyslexic eye movements showed an increased number of fixations as well as regressions, prolonged fixations, and a tendency for larger saccade amplitudes. However, the new finding of this study is the increased saccade disconjugacy in dyslexics, an increased disconjugate drift after saccades, and an uncorrelated saccade and post-saccadic drift conjugacy in text reading. Generally, these results indicate that dyslexics exhibit poor binocular coordination during and after saccades. Usually, disconjugate drifts occur after saccade disconjugacy in order to correct it.

However, dyslexic participants did not show this typical pattern for vergence, thus the disconjugacy due to saccades could not be restored by the subsequent disconjugate drift. According to the authors, this might be the reason for the poor binocular control of dyslexics in saccades. In order to establish a single percept in a fusional process, sensory-driven vergence adjustment in fixations should provide and maintain the necessary requirements. A slight tendency for larger fixation disparities in dyslexics could be observed, which means that dyslexics have to cope with slightly larger residual disparities when fusing the information sent from both eyes. This might strain fusional capacities and cause fatigue. However, the results also showed slightly increased fixation disparity for all participants when reading distance was reduced. The most important finding, according to the authors, is that dyslexic participants exhibited increased standard deviations of their fixation disparity during fixations, which means changing disparities for the same item. The effect was bigger for the close reading distance. This might reflect the remarkable demand on the fusional processes in order to obtain single clear vision and might disturb word identification. Even in freely exploring a painting, the dyslexic participants exhibited larger standard deviations for fixations. Taken together, these results support the hypothesis of visual motor deficits possibly disturbing the fusional process in dyslexics.

Kim et al. (2014) and Kim & Wiseheart (2017) follow similar research designs in their studies. Both investigated graph comprehension in students with dyslexia. The first study (Kim et al., 2014) was designed to compare dyslexics and normal readers reaction times and eye movements on tasks with varying graph types, graphic patterns and question types. They expected dyslexics to spend more time fixating verbal areas and to show slower reaction times. An analysis of variance (ANOVA) yielded the following results: reaction times were significantly longer for dyslexics, longer for line graphs than for vertical bar graphs, and they were significantly longer for comparison questions and for double graphic patterns. Differences in the reaction times for question types and for graphic patterns were significantly larger for dyslexic participants. Furthermore, differences in the reaction times for the two question types were larger for the double graphic pattern than for the single graphic pattern. Finally there was one significant three-way-interaction between graph type, graphic pattern and question type, indicating that reaction times were significantly shorter for the vertical bar graph in the single graphic pattern with a point location question. The eye movement analysis of the first viewing time and total viewing time of the different regions exhibits that dyslexic participants spent significantly more time on the question and answer region. Taken together, the level of complexity of the stimulus had a stronger effect on the reaction times of dyslexics, which means the more difficult the stimulus, the more pronounced was the difference in reaction times. These results suggest that dyslexics might be more challenged by the cognitive resources necessary for graph comprehension, especially for more difficult stimuli. In accord with the expectations, dyslexic participants exhibited longer fixation times on verbal regions than controls. One particularly interesting result was that dyslexics exhibited longer fixation times in non-linguistic regions only in the total viewing time measure, whereas the first viewing time was similar to controls. This indicates that dyslexic participants returned attention to already examined areas. Possible explanations would be that they might have had difficulties in processing the specific area itself, or, after re-examining the verbal areas, had to refresh their memories of the graphic data in order to answer the question. If so, the reaction times may be a result of inefficient text processing rather than difficulties in processing graphic information. The authors also included a working memory framework in their analysis. Every person's working memory has limited capacities. An increase in information load is likely to result in longer fixation times and, thus, in slower reaction times. Difficulties in holding the question in the working memory during examining the graphic information might possibly the reason for the longer fixation times and slower reaction times of dyslexics. Similar performances for graph types were observed in both groups. Both reacted faster to questions on vertical bar graphs, which the authors associate with the cognitive naturalness principle and the information processing principle.

In Kim & Wiseheart (2017), the focus was on the question whether orthographic processing difficulties could also account for inefficient graph interpretation. They hypothesized dyslexic students to show similar patterns as controls if graphic displays would lack orthographic information. First and foremost, the results showed no difference in accuracy scores between typical and dyslexic readers. Generally, text graphs evoked longer reaction time than icon graphs, which in turn evoked longer reaction times than word graphs. An ANOVA showed that in all three presentation formats, reaction times of dyslexics were significantly longer.

Furthermore, with regard to information complexity, dyslexic readers were slower in the single and the double condition, however the effect was more pronounced in the double form. Dyslexic participants needed more time in both question types, however the difference was bigger in comparison questions. The ANOVA results for information types and group show that dyslexics were significantly slower in all information types. The authors further investigated whether the longer reaction times in dyslexics resulted from the question area, the stimulus area, or both. Their findings showed that dyslexics spent more time on the stimulus area only in the text condition and significantly more time on the question area in all three presentation formats. To sum up, dyslexics exhibited longer reaction times which mainly resulted from text-related areas. In the stimulus area, there were no group differences (besides the text condition). Furthermore, similar comprehension accuracy indicates that dyslexic readers are capable of accurate but slow text processing. Eye movement data on further areas shows that the dyslexic participants exhibited longer gaze durations on the x-axis of word graphs, while there could be no differences observed between the two groups of participants for the icon graph. For the legend area, only a marginal difference between the different information formats could be observed for the typical readers but not for dyslexic readers. When the stimulus was presented in form of a word graph, compared to an icon graph, dyslexics exhibited more eye movements between the question and the stimulus area. No significant interaction between stimulus area, group and information format could be observed. For both graph types, the number of eye movements of dyslexics was increases within the question area. In conclusion, the results indicate that group differences in eye movement patterns mainly resulted from processing orthographic information. With regard to eye gaze durations, there were no group differences observed in the orthography-free condition, whereas dyslexics exhibited longer gaze durations in the mixed-modality condition. Furthermore, dyslexic participants exhibited more eye movements in written questions than controls, and more eye movements between areas in word-graph conditions. The results support the hypothesis that graph interpretation problems in dyslexia are related to a deficit in orthographic processing rather than to difficulties in processing graphic information. The research team attributes the fact that students generally processed word graphs faster than icon graphs to familiarity effects, as icon graphs might not be a typical representation. For instructors, the authors suggest to be aware of the fact that there are multiple reasons why some students may underperform on graphic tasks, for example that the complexity of the graphics used might affect text processing in dyslexic students. As shortcomings of their research, Kim & Wiseheart (2017) used easy graphs and easy questions, whereas future studies might want to use more age appropriate stimuli. Also, they used a closed-answer-question system, while open-ended questions might yield different results. As for the inclusion criteria for the dyslexic group, the authors chose a less conservative threshold which could be a reason why there were no clear group differences for certain measures. Finally, the authors measured the reading abilities of the dyslexic participants but did not evaluate their visuospatial skills.

In their study from 2002, de Luca et al. compared the eye movements of dyslexic and control participants when reading short and long words and pseudowords. They expected dyslexics to exhibit similar eye movements regardless of the lexical value of the letter strings, while controls would show different patterns for words and pseudowords. They expected length effects to show in dyslexics for both types of stimuli, while controls were expected to show length effects only in pseudowords. With regard to reading performance, dyslexics showed pronounced deficits in speed and accuracy for both types of stimuli, while comprehension is barely affected. This pattern points to an extensional use of grapheme-phoneme decoding. The eye movement analysis also showed significant results. The analysis of the number of rightward saccades showed that, for controls, there is not a big difference between short words and pseudowords. However, they showed more saccades for long pseudowords than for words. The dyslexic participants exposed a contrasting pattern: while there was no difference between long words and pseudowords, they used more saccades for short pseudowords than words. Both groups needed more saccades for long pseudowords than for short. Dyslexics, but not controls, also displayed more saccades for long words. Generally it can be said, that there were more saccades observed in dyslexics than in controls, in reading pseudowords than words and for long rather than for short stimuli. With regard to saccade amplitude, the results for controls show no differences in amplitude for short words and pseudowords, but larger saccades for long words than pseudowords. They showed a notable difference between short and long words and a smaller yet significant difference between short and long pseudowords. Dyslexics exposed no difference between words and pseudowords for neither length and thus no length effect for either stimulus type. Taken together, saccade amplitudes were smaller in dyslexics than in controls, in reading pseudowords than words and for short rather than long stimuli. Whereas regressions were basically rare, there were still different patterns observable between the two groups of participants. While controls showed no differences for short stimuli, they exposed more regressions for long pseudowords than for words. Dyslexic participants displayed an opposite pattern, as they exposed more regressions for short pseudowords than for words, but no difference for long stimuli. Generally more regressions could be observed for dyslexics, for pseudowords and for long stimuli. Finally, the analysis of fixation duration showed longer fixations for pseudowords than for words and for dyslexics than for controls. Proficient readers in this study seem to adjust their saccades to the length of the word to be read. For long words, that would mean they used larger saccades while the number of saccades remained the same. For pseudowords, there was a length effect observable in the increasing number of saccades dependent on stimulus length. In contrast, dyslexic participants exposed a marked length effect regardless of the lexical value of the stimuli. The number of saccades increased with stimulus length while the amplitude stayed small. The data presented indicates that dyslexics mainly rely on sub-lexical grapheme-phoneme conversion. The eye movement of dyslexics shows that, for short stimuli, they differentiate between words and pseudowords which implies a certain degree of sighting vocabulary. Controls scan short stimuli in a similar way which implies that reading these stimuli is easy for them. When processing long stimuli, dyslexic participants exhibited a more parceled pattern with a higher number of saccades of smaller amplitudes, compared to controls. Taken together, these results indicate that dyslexic readers rely on a slow sub-lexical mode of processing.

The purpose of the study of Pan et al. (2013) was to explore the eye-voice spans of dyslexic and normal Chinese readers in rapid automatized naming of digits and dice. The two stimuli differ in their automaticity of mapping symbols to sounds, but have identical phonological output demands. The authors expected larger group differences in the alphanumeric digit-RAN, larger Eye-Voice Spans (EVS) for digit-RAN and larger EVS for normal readers. The results show, that the two groups of participants differed significantly in reading fluency, digit- RAN and dice-RAN scores, as well as in their eye movement patterns. In digit-RAN, the groups significantly differed on all eye movement measures, whereas in dice-RAN they only differed in gaze duration and EVS. Generally, smaller group differences for dice than for digit-RAN could be observed. With regard to the question, whether differential RAN scores can be predicted with gaze duration and EVS, Pan et al. (2013) found that, for controls, EVS are predictive of psychometric digit-RAN, while gaze duration was significantly predictive in both conditions and reader groups. Taken together, these findings suggest that EVS should rather be seen as correlated indicator of the RAN-reading relationship instead of the cause. The results were further discussed with regard to two possible dyslexia-related deficits: a phonological deficit and an automaticity deficit. According to the authors, the data seems more in accord with an automaticity deficit. A large degree of the relationship between RAN and reading might be due to the level of automaticity in print-to-sound conversion in different stimuli. This is based on the assumption that symbol-to-sound conversion is an initially demanding process, which is automatized with practice. Dice-to-phonology translation is probably less well practiced than digit-to-phonology conversion, which might be reflected in longer RAN times during dice-RAN for both groups. Digit-RAN is expected to be a more reliable predictor of normal reading, if reading speed is indeed linked to the automaticity of naming. Further, the authors analyzed EVS and perceptual spans. EVS increase with reading competence which indicates that normal readers buffer more information when stimulus-to-sound translation is easy. Indeed, the results showed EVS for digit-RAN, for normal children and the group difference was also larger in digit-RAN. According to Pan et al. (2013), dyslexic eye movement patterns indicate a processing difficulty in the perceptual span. A wide perceptual span seems possible when naming is automatized and more visuo-attentional resources are available. In conclusion, digit-RAN seems to better predict a dyslexic status than dice-RAN. Furthermore, apart from the phonological representation per se, print-to-sound conversion might also be important to understand dyslexia and reading development.

Pan et al. (2014) investigated saccade target selection in dyslexic children when reading Chinese. They wanted to provide evidence that the first landing position might act as indicator of parafoveal word segmentation. Due to their reduced perceptual spas, they expected dyslexic readers to show larger differences in saccade targeting, given they are more affected by the absence of word boundaries.

As expected, the dyslexic participants generally exhibited fewer single fixations and skipped words less often. They needed more and prolonged fixations per word, and single fixations landed significantly closer to the beginning of the word. However, no significant difference in the landing position could be observed for multi-fixation cases. Basically, the control group exhibited longer incoming saccade amplitudes and their launch sites were further away from the word to be fixated in multi-fixation cases. Words which received a single fixation and words with multiple fixations showed significant differences in word length and word frequency. The analysis of first-fixation landing positions showed that saccades landed further into single-fixated words and further into long than into short words. While dyslexics tended to undershoot word centers of multi-character words, controls typically landed on the middle of the word. In order to process words, dyslexics exhibited an increased number of prolonged fixations. Furthermore, word length effect on the number of fixations and gaze duration was more pronounced in dyslexic participants. Generally, words that are difficult to recognize receive more fixations of longer durations. With regard to incoming saccade amplitude and launch site, the results indicate that single fixations are associated with shorter launch sites. This is probably because information about the boundaries of a word might be easier accessible if a fixation is closer to the word to be fixated. Thus, readers might be able to compute larger saccades in order to target the center of words. The findings of this study support the hypothesis that higher level linguistic processing can affect saccade-target selection. In scripts without explicit word boundaries, readers use parafoveal word segmentation in order to process word length. While this might not be difficult for skilled readers with larger perceptual spans, it can be very challenging for developing readers as, for them, it requires more resources. One specifically critical finding of this study is a three way interaction between fixation type, word length and subject group. This indicates that the two groups actually program their saccades differently. Controls parafoveally segmented easy words and targeted a position almost in the center of the word and processed the word in a single fixation. If preprocessing in the parafovea was not possible, they targeted the beginning of the word, regardless of its length. Dyslexics did not expose such efficient word segmentation and often “undershot” the center in single-fixation cases. The authors suggest that this might be due to their uncertainty about the word boundaries because of their reduced perceptual spans. Thus, they might be too careful not to “overshoot”. The findings suggest that the dyslexic participants were able to parafoveally process the word to some extent and targeted as far into the word as they could. However, these are the results for Chinese dyslexics, who might be more careful in their saccade-target selection due to the lack of word boundaries.

In their paper from 2007, Prado et al. investigated eye movements of dyslexic children during reading and visual search. They aimed at providing evidence for visual attention span deficits as potential reason for dyslexic eye movement patterns. Dyslexic participants were expected to show different results in reading but less so in visual search. The performance analysis yielded the following results: for the visual search task, indeed no significant difference between dyslexics and controls could be observed, neither for the number of target letters identified, nor for the time needed. All participants needed more time for the visual search task. However, dyslexic readers needed more time for the reading task. Also with regard to eye movement measures, results were similar in visual search but differed a lot in reading, which indicates that eye movement patterns of dyslexics cannot be explained with general deficits in oculomotor control or visual perception. For rightward fixations, dyslexic participants exhibited more rightward fixations in reading, while results for the visual search task were similar. Generally, all participants needed more fixations in the visual search task, which indicates that dyslexics need a lot of fixations regardless of the task. Dyslexics processed fewer letters during reading than controls, and fewer letters were processed in the visual search task than in reading. This finding indicates that normal readers adapt the number of letters to be processed to the task while dyslexics seem unable to increase the number of letters being processed simultaneously in reading. The analysis of rightward fixation duration showed that it was higher in dyslexics than in controls in reading, whereas not differing in visual task. Generally rightward fixation durations were higher in the visual search task than in reading. Again, dyslexics exhibited rightward fixations of the same duration regardless of the task, in contrast to normal readers. For leftward fixations, the number was higher in the visual search task. Generally, dyslexics exhibited more leftward fixations than controls, and controls showed a larger percentage of regressions in visual search, while the proportion for dyslexics was the same regardless of the task. The results indicate that leftward fixations are not triggered by linguistic deficits but are rather due to oculomotor landing errors. The duration of leftward fixations was also longer in the visual search task and dyslexics exposed longer leftward fixations in reading but not in visual search. Finally, a correlation analysis showed that the visual attention span is significantly and negatively correlated with the number of rightward fixations in reading but not in the visual search task. The more reduced the visual attention span of a reader is, the more fixations in reading are necessary, the more likely is the text to be read analytically. This finding might indicate that a smaller visual attention span prevents dyslexics to process as many letters simultaneously as controls, however only in reading. Further it was correlated with the number of leftward fixations in reading, however again not in visual search. Even though the number of rightward fixations was correlated with the visual attention span, as expected, to claim this relationship to be causal, it requires proof that a visual attention span reduction is not merely due to the poor reading level of dyslexics. Thus, further research is needed. The findings of this study do not rule out the possibility of other factors being involved in eye movement control. However, they strongly indicate the heterogeneity of dyslexic people to be a relevant dimension. Furthermore, the results suggest that even similar tasks can pose different demands on visual perception. Taken together, the results indicate that dyslexic eye movement patterns do not seem to be due to an oculomotor disorder. According to the findings, it might be possible that visual span deficits in dyslexics contribute to their atypical eye movement patterns, however only in reading.

As there are two different forms of parafoveal dysfunction that have been hypothesized to be the core deficit of dyslexics, namely increased parafoveal load costs and reduced parafoveal preview benefits, Silva et al. (2016) wanted to test each hypothesis against the other in an eye-tracking RAN study. Further, they aimed at investigating whether phonology is involved in each of the two hypotheses. For this purpose, they included a second eye-tracking task, where naming was suppressed and replaced by silent letter finding. Group differences were expected to be smaller in the silent letter finding task than in the naming task. First, the effects of parafoveal preview potential and parafoveal load cost on fixation time were analyzed. In the RAN task, dyslexics exposed longer fixation times, whereas no group differences could be observed in the second task. When preview was given (P), for both groups, the fixation times on loaded items (L) increased, whereas an opposite pattern could be observed for non-loaded items (nL), for which fixation times decreased. Therefore, preview benefits could only be observed for non-loaded items. Generally, the parafoveal preview benefit of dyslexics was only marginally significant in the naming task, whereas that of controls was significant. Neither of the two groups showed a significant effect in the letter finding task. For both groups, parafoveal load cost was not significant in the letter-finding task, however it proved significant in the naming task. Only evidence for reduced preview benefits could be observed in the comparison of the two groups of participants, which supports the relation of reading performance and parafoveal dysfunction in dyslexics. This effect was only observable in the naming task but not in the letter-finding task. Neither dyslexics nor controls exhibited significant parafoveal load cost effects in the letter finding task, which indicates that parafoveal load might only be costly when phonological processes are involved. Indeed, dyslexics as well as controls exhibited significant load cost effects in the naming task. One possible explanation is that dyslexics are influenced by parafoveal input in the sense of crowding where it may cause interference but they cannot use it as preprocessing target. So the main finding is that dyslexic adults exhibited reduced parafoveal preview benefits while no increased parafoveal load costs could be observed. This may partly explain why dyslexic readers are impaired in serial RAN tasks. For further research, the authors suggest to investigate parafoveal preview benefits of typically developing and dyslexic children as well as to compare the dyslexic participants of their study with reading level matched controls, instead of age matched controls.

In their 2009 study, Thaler et al. investigated and compared eye movement patterns of dyslexic readers with and without comorbid attentional deficits during single word reading. They expected children with attentional deficits only (AD) to be affected by word length but not cluster density in reading accuracy, children with dyslexia only (DYS) to be affected by both in reading time but not reading accuracy, and children with dyslexia and an attentional deficit (DYS+AD) to be affected by both and to exhibit longer reading times and impaired reading accuracy. Results show that, except for the DYS+AD group, reading accuracy was generally high. The interaction between group and cluster density showed that DYS and DYS+AD were not influenced by density in terms of error rates, while, interestingly, the error rates of the control group (CG) and AD were influenced by density. These findings support the hypothesis that dyslexic children in regular orthographies show little problems with reading accuracy. If other studies yield other results, it might be due to the fact that DYS+AD children did indeed exhibit significant errors. In a post-hoc analysis, the authors investigated the orthographic characteristics of the consonant clusters. They compared (1) the effect of di(tri)graph clusters to phonologically based clusters, (2) clusters within a syllable versus clusters across syllables, and (3) clusters within one morpheme to clusters across morphemes. The results show that the participants needed the same number of fixations for both cluster classes (1). With regard to the second point (2), CG seemed to use a reading strategy on syllable structure, while DYS and DYS+AD showed greater difficulty in reading clusters spanning across two syllables. Finally, clusters spanning across morphemes generally needed more fixations in affected participants. These results indicate that unimpaired readers seem to profit from splitting larger parts into syllables and/or morphemes. However, the authors refuse to make firm conclusion as this is only their post hoc analysis.

Contrary to the authors' expectations, children with attentional deficits only did not show problems with reading accuracy, which the authors explain with single-word-reading possibly not requiring a lot of attention. The results for reading time shows that both dyslexia groups exhibited significantly longer reading times, and that word length and cluster density had effects on all groups. With regard to the number of fixations, dyslexics exhibited more fixations than all other groups. Children with dyslexia and an attentional deficit showed significantly more fixations in short word reading than children with an attentional deficit only. Again for all groups, words with high cluster density led to an increased number of fixations. The results for the last eye-tracking measure shows that all dyslexic participants had longer mean fixation durations than the control group, while children with a comorbid attentional deficit exposed also longer fixation durations than children with an attentional deficit only. All groups showed longer fixation durations for short rather than for long words, however fixation duration was prolonged for long words with high cluster density. These findings indicate that dyslexics showed a massively reduced reading speed. Furthermore, the eye movement patterns confirm that dyslexics with and without an attentional deficit show different reading profiles. While children with dyslexia only primarily use serial decoding and fixate almost every letter, dyslexics with a comorbid attentional deficit analyze larger chunks but show anomalies in their fixation duration. According to the authors, this indicates that their reading strategy per se differs less from normal readers than that of children with dyslexia only in that they try to analyze larger bits of information. However this strategy seems not to be adequate for their reading skills, as they show a relatively high error rate and prolonged fixation durations. All groups showed a word length effect. While children in the DYS+AD group were more affected than normal readers, children with dyslexia only were most affected. These findings indicate that dyslexics might have difficulties in building up an orthographic lexicon. With regard to short words, children with dyslexia and an attentional deficit differed even more from controls. The authors suggest that in these children comorbid deficits have caused additional problems in building up an orthographic lexicon even for short words. The results of an analysis of the orthographic characteristics of consonant clusters seem to support the idea that children with reading disabilities mainly rely on serial decoding whereas controls seem to use a syllable-based reading strategy, although not yet morpheme-based. Finally, all groups showed a large unit effect, which means that the reading times of all participants were affected by cluster density. This finding might indicate that in the context of the German orthography, children use sub-lexical units above the grapheme-phoneme level.

In 2010, Trauzettel-Klosinski and colleagues investigated the eye movements of German children with and without dyslexia when reading aloud. They wanted to find out how different levels of phonological difficulty of the reading material might influence performance. Results were compared to English findings in order to find out whether different reading strategies were used or not. The findings show that dyslexic participants read less than half the words per minute than their controls, exhibited more saccades, regressions and fixations, and their saccade amplitudes were smaller than in controls. The oral reading speed results of English speaking students correspond with those results. Generally, reading speed can be influenced by age, skill, silent or oral reading, and the difficulty of the stimulus material. Both groups of participants were influenced by text difficulty since both groups were significantly faster in reading the easier text and needed fewer saccades per word. The saccade amplitude of dyslexic participants was shorter in both texts, thus the increasing text difficulty only showed an effect on controls in this respect. Basically, controls read faster and exposed fewer saccades per word and of longer amplitude. Text difficulty seemed to influence the number of saccades per word in dyslexics, whereas it seemed to influence saccade amplitude in controls. The increased number of saccades in dyslexics mirrors the possible preference of German developing readers to use grapheme-phoneme-conversion as reading strategy. The slightly larger saccade amplitude showed by controls when reading the harder text might indicate that they adapt their saccade amplitude to word length, while dyslexics are less able to do so, which is reflected in their relatively steady saccade amplitudes in both texts. Dyslexics further exhibited shorter saccade amplitudes than controls which might be an indicator for a reduced perceptual span. The percentage of regressions was slightly increased in dyslexics, which stands in contrast with English findings, in which dyslexics exposed a higher percentage of regressions. These results can be explained with the relatively regular grapheme-phoneme-correspondence in German and Italian which allows for serial decoding. Yet, at some points it seems necessary for dyslexic readers to be reassured of the meaning of a word in text reading, which causes regressions. With regard to mean fixation duration, the German results differed notably from the English results. Fixation duration was prolonged in dyslexics but did not seem to be dependent on text difficulty. They tended to use more fixations instead of prolonging them. However, the English results rather show an opposite pattern which indicates that English readers prefer a direct decoding strategy for which longer fixations are more adequate than a higher number of eye movements. An increased number of regressions in dyslexics necessarily leads to an increased number of fixations. Also controls exhibited a higher number of fixations per word compared to English speaking children which reflects the grapheme-phoneme conversion strategy of German children. The increased number of additional regressions during return sweeps observed in dyslexics seemed independent of text difficulty but reflects uncertainties in text comprehension. As expected, the dyslexic participants performed better on the easier task which was reflected in the reading speed and number of saccades and regressions. The results indicate that in increasing level of text difficulty primarily has an effect on the number of eye movements in dyslexics, and on saccade amplitude in controls. Taken together, the findings of this study suggest that phonological difficulty indeed influences reading speed and the number of eye movements. Shared findings with English studies are slowed reading speed and an increased number of saccades and regressions in dyslexics. However, this study differed from the English findings in that the percentage of regressions was only slightly increased in dyslexics but the number of fixations higher than in English participants. Further, an increase in text difficulty evoked a higher number of eye movements in German dyslexics, while it led to longer fixations in English dyslexics. Finally, saccade amplitudes were smaller in German dyslexic children. Taken together, these results indicate that according to the orthography, German children rely on a sub-lexical analysis, whereas English children favor a direct lexical analysis.

In 2015, Vagge et al. evaluated dyslexic readers eye movements in order to assess whether their analysis can be useful in the identification of dyslexic children. They found that the strategy that dyslexics adopted in text reading differed notably from that of controls. Dyslexic participants showed a significant increase in the number of fixations, which was even more pronounced for long and less common words. This increase is mainly due to progressive but also regressive saccades. Generally, in both groups improvements of fixation stability with age could be observed. However, dyslexic participants showed developmental deficits, which have been argued to be a possible consequence of a deficit in the magnocellular pathway. The findings of this study are in accord with previous research. Dyslexic participants seemed to adapt rather a sub-lexical strategy, which is reflected in the increased number of shorter saccades. They analyzed letter by letter or syllable by syllable rather than a word-level unit. Further, their impaired reading fluency can be explained in the increased number of regressions they made. According to the authors, those findings support the hypothesis that dyslexic eye movements cannot be attributed to dysfunctions in oculomotor control but rather reflect deficient visual processing of linguistic material. Furthermore, they support the idea that the assessment protocol for identifying dyslexia might profit profoundly by thorough examinations of visual functions.

Yan et al. 2013 compared Chinese dyslexic and unimpaired children with regard to their parafoveal processing efficiency in rapid automatized naming (RAN). Their main goal was to find out whether or not dyslexics operate with smaller perceptual spans in RAN. If they do, they should be less affected from removing parafoveal preview than normal readers. To test their hypothesis, the authors used discrete and continuous RAN. The results showed that both groups performed slower in the discrete RAN with a larger effect on the control group. This finding suggests, that controls show more efficient parafoveal information processing and are, thus, more affected by the lack of preview. Overall, dyslexic participants needed more time for the task indicated by longer single fixation durations and longer gaze durations. There was a significant main effect of condition observable for all three measures, single fixation duration, gaze duration and first fixation duration, which indicates that the processing difficulty in discrete RAN was higher due to the absence of parafoveal preview. Furthermore, there was a significant interaction between condition and group for all three measures. In discrete RAN the group difference was reduced for single fixation duration and gaze duration. The analysis of landing positions showed that in continuous RAN normal readers landed further into areas of interest than dyslexics. However, in discrete RAN no group differences were observed. According to the authors, differences in landing positions were due to differing inter-item saccade amplitudes. In conclusion, there were three main findings. First, a general slowdown was observable in discrete RAN, which might be due to an increase of processing difficulty caused by the absence of parafoveal preview. This was indicated by longer fixation durations and shorter saccades. Finally, dyslexics seem to extract less parafoveal information. One particular finding was that dyslexic and normal readers show perceptual spans of different sizes. While normal readers acquire more information in parafoveal vision and show stronger preview benefits, dyslexics seem to need most of their attentional resources for print-to-sound conversion. However, there still was a preview benefit observable in dyslexics, indicating that they could actually use parafoveal information but with reduced efficiency. Generally, dyslexic participants exhibited prolonged fixation durations. The group difference in this respect was larger in continuous RAN what the authors attribute to better parafoveal processing of controls. They believe larger perceptual spans to be consequences of more automatic foveal processing of unimpaired readers. The more automatized the processes involved in reading, the more attentional resources are available for parafoveal processing. Also the results of differing saccade amplitudes and landing positions in continuous RAN point to differences in the perceptual span. In conclusion, dyslexic readers extract less parafoveal information than normal readers, what the authors believe to be rather a symptom of the disorder. Due to their less automatized translation skills from letters to phonological representations, there are less attentional resources available for parafoveal processing.

ii. Diachronic analysis

The synchronic analysis and comparison of the 18 eye-tracking studies on dyslexia gives insight into the diversity of the research done in this field. This diversity certainly comes with advantages, such as the vast array of different foci, and disadvantages, such as the difficulty of comparing and contrasting these studies. So far, the focus of the comparison in this thesis was on contrasting the papers and highlighting how they differ. The following part aims at grouping and comparing similar papers and at finding similarities. As has been pointed out so far, this is a very difficult and sensitive work as it yields the danger of neglecting the various factors influencing research. However, bearing this in mind, it can also contribute to the bigger picture of eye-tracking studies in dyslexia. In order to find typical eye movements of dyslexic persons, see if they are causal or correlational, and in order to test eye-tracking as possible diagnostic method, a comparison of this kind seems even necessary. The studies are grouped according to their foci in the following manner:

First, two particular studies on parafoveal processing in RAN tasks of dyslexics are analyzed (Silva et al., 2016; Yan et al., 2013) and further compared to a third study, which similarly concentrates on attentional spans in dyslexics (Prado et al., 2007). Second, there are three studies on saccade and vergence computation and control in dyslexics, two of them in reading (Jainta & Kapoula, 2011; Pan et al., 2014) and one in an LED-stimulus-task (Bucci et al., 2008). Third, two papers focusing on the comparison of different orthographies are analyzed (Hutzler & Wimmer, 2004; Trauzettel-Klosinski et al., 2010), followed by two papers concentrating on dyslexia specific eye movement patterns (Thaler et al., 2009; Vagge et al., 2015). Further, there are two studies investigating graph comprehension of dyslexics (Kim et al., 2014; Kim & Wiseheart, 2017), and two studies focusing on the effects of word length, frequency, and type (De Luca et al., 2002; Dürrwächter et al., 2010). Finally, the results of visual search tasks are shortly summarized.

1. *Parafovea and attentional span*

Research suggests that attention is an important factor in eye movement control in reading and that attentional resources enable wider perceptual spans (Bellocchi et al., 2013; Bucci et al., 2008; Pan et al., 2013; Yan et al., 2013). While Yan et al. (2013) aim at investigating whether or not dyslexic readers show reduced parafoveal processing efficiency and smaller perceptual spans in RAN, Silva et al. (2016) are one step further and test two hypotheses of deficient parafoveal processing in dyslexics against each other with RAN. The first hypothesis is that of reduced parafoveal preview benefits, and the other increased cost of parafoveal load. The two studies differ particularly in their samples of participants. While Yan et al. (2013) examined 35 native Chinese dyslexic children with a mean age of 10.75 years, Silva et al. (2016) examined 17 native Portuguese dyslexic adults with a mean age of 26.35 years. However, both studies used eye trackers with relatively high temporal resolution (1000Hz and 1250Hz). While both research groups measured fixation times, Yan et al. (2013) additionally measured fixation landing positions and saccade amplitudes. As an adequate task, the authors chose continuous RAN (with parafoveal preview) and discrete RAN (without parafoveal preview) with 5 digits (1-5) randomly ordered as stimuli. Instead of digits, Silva et al. (2016) used letters (a,o,s,p,d). Half of the letters could be parafoveally previewed (P), which was not possible for the other half (nP). One half of the items was parafoveally loaded (L), while the other half was not (nL). Thus, there were four types of stimuli: PL (o **s** a), PnL (p **a**), nPL (**o** s), and nPnL (**d**). Even though the two studies have very different samples of participants and slightly different stimuli, they yield similar results. First, both show that dyslexic participants exhibited longer fixation durations. Second, both observed a parafoveal preview benefit, which was stronger for controls. These findings suggest that dyslexics were able to use parafoveal information, however not as efficiently as normal readers. Yan et al. (2013) and Silva et al. (2016) suggest that dyslexics show difficulties in translating visual symbols to phonological representations and, thus, have less attentional resources available for parafoveal pre-processing. Since, according to the authors, the foveal processes run more automatically in controls, they are likely to have more attentional resources available to devote to parafoveal vision, which leads to wider perceptual spans.

In the following part the results discussed above are compared to yet another study with a different experiment setting but a similar focus. In 2007, Prado et al. examined the impact of the visual attention span on eye movements of French dyslexic children during reading. Their aim was to provide evidence for visual attention span dysfunction being a potential cause for dyslexic eye movement patterns. For this purpose, they examined 14 native French dyslexic children, with a mean age of 11.1 years, in reading a text aloud without time limit. With an Eyelink I 250Hz device they measured the total number of fixations, number of rightwards fixations, percentage of regressive fixations, mean fixation duration, and the total number of letters divided by the number of rightward fixations as an approximate estimation of the visual attention span. The results showed that dyslexics performed much slower in text reading than their controls. They exhibited more rightward fixations of longer duration, and more of leftward fixations of longer duration. These findings indicate that the participants with reading difficulties processed fewer letters per fixation than unimpaired readers, who seem able to increase the number of letters processed according to the task (reading versus visual search). This blends in well in the hypothesis of reduced visual attention spans in dyslexics. The more reduced the visual attention span, the more fixations are needed and the more likely is the text read analytically. Thus, a dysfunction in the visual attention span prevents impaired readers to simultaneously process as many letters as controls, however only in reading. A significant correlation analysis of the number of fixations and visual attention span provides further support. The authors conclude that dyslexic eye movement patterns cannot be explained in terms of an oculomotor disorder but rather by reduced visual attention span abilities.

It seems that all three studies refer to the same phenomenon with slightly different perspectives. While Prado et al. (2007) examine the whole visual attention span, Yan et al. (2013) and Silva et al. (2016) already divide it into fovea and parafovea. However, what they have in common is the idea that there are not enough attentional resources available in dyslexic readers to expand their perceptual spans in order to process more information at a time.

2. *Saccade and vergence computation and control*

The next three studies to be compared here are not completely similar in their focus. While all of them primarily investigate saccades in dyslexics, they do so in different ways. Jainta & Kapoula (2011) concentrated on saccade and vergence control in dyslexic readers during silent text reading, while Pan et al. (2014) focused on saccade-target selection of Chinese dyslexic children. In contrast to the two studies on eye movements in reading, Bucci et al. (2008) investigated the latency of saccades and vergence eye movements in dyslexic children in a non-reading task. First, the two reading task based studies are compared and further contrasted to the paper by Bucci et al.

With regard to the sample of participants, Jainta & Kapoula (2011) examined 13 French dyslexic children with a mean age of 11.7 years, while Pan et al. (2014) worked with a larger sample of 33 Mandarin dyslexic children at a slightly younger mean age of 10.7 years. While both studies measured the number of fixations, gaze duration, saccade amplitude and saccade launch site, Jainta & Kapoula (2011) additionally measured conjugate and disconjugate eye movements, change in vergence between saccade onset and offset, and for each fixation they measured the minimum fixation disparity, the smallest vergence error, disconjugate drift in vergence and the standard deviation of fixation disparity across the whole fixation period. Both studies further differ in the eye trackers being used, a Chronos Vision Berlin 200Hz device in Jainta & Kapoula (2011) and a 1000Hz Eyelink 2K device in Pan et al. (2014), and the tasks, including silent text reading at two distances in Jainta & Kapoula (2011) and oral sentence reading in Pan et al. (2014). In accord with other research, the results of both studies showed that dyslexics exhibited more and longer fixations. Due to their different foci, the two studies also yield different results, such as an increased saccade disconjugacy in dyslexics or that the passive disconjugate drift at the beginning of a fixation after a saccade was not sufficient to counterbalance this saccade disconjugacy in dyslexic children (Jainta & Kapoula, 2011). In contrast, Pan et al. (2014) found that dyslexic readers undershot the word center of multi-character words, while controls tended to land in the middle of a word. The authors suggest that dyslexics do not seem to segment words as efficiently as controls and thus undershoot word centers, which the authors attribute to reduced perceptual spans in dyslexics. However, both studies somehow suggest that dyslexics show difficulties in saccade control.

The study of Bucci et al. (2008) certainly differs in some aspects, first and foremost the kind of task involved. In order to test pure saccades, pure vergence, and pure saccade-vergence movements in dyslexics, they set up LEDs in two circles at 20cm and 150cm distance and always lit up one fixation LED and then a target LED in gap and simultaneous paradigm. Nevertheless, there are also similarities to the other two studies, for example the sample of participants which contained 16 dyslexic children at a mean age of 11.12 years, or the eye movements measured. In order to track the eye movements, the authors used an Oculometer by Dr. Bouis with a sampling frequency of 500Hz. In accord with the other two studies, Bucci et al. (2008) found longer mean latencies especially for saccades starting at a far point in dyslexic participants. The authors suggest that visual attention plays a crucial role in switching from reflexive to voluntary eye movement initiation and that problems in voluntary and reflexive control of attention to targets in a three-dimensional space might cause these prolonged latencies.

3. *Comparison of orthographies and dyslexia specific patterns*

The reason for grouping papers addressing the comparison of different orthographies and papers investigating dyslexia specific eye movement patterns together is that the two papers of the former group also focus on dyslexia specific symptoms. Thus, it would be more helpful to group them together. First, Hutzler & Wimmer (2004) and Trauzettel-Klosinski et al. (2010) are compared, then Thaler et al. (2009) is compared to Vagge et al. (2015). Dürrwächter et al. (2010) compared their results to other languages as well. However, this paper is discussed below.

Once again, the two papers being compared here have slightly different foci. Hutzler & Wimmer's (2004) aim is to extend the empirical basis on eye movement patterns of dyslexic children when reading in a regular orthography compared to irregular orthographies. Trauzettel-Klosinski et al. (2010) focused on how orthographic regularity and phonological difficulty of a language influence reading strategies, also by comparing German to English. Nevertheless, both studies investigate the influence of regularity in orthographies and compared a regular orthography (German) to a deep orthography (English). Their sample of participants mainly differed in age: Trauzettel-Klosinski et al. (2010) chose slightly younger participants (16 dyslexic children, mean age 9.5 years) than Hutzler & Wimmer (2004) (11 dyslexic boys, 13.58 years). Among the eye movement parameters measured in both studies were the percentage of regressions, fixation duration and number of fixations. Especially Trauzettel-Klosinski et al. (2010) had additional measures, such as saccade amplitudes or return sweeps. While Hutzler & Wimmer (2004) used a video-based eye-tracking system with a sampling frequency of 50Hz (ISCAN), Trauzettel-Klosinski et al. (2004) chose a less common device, a scanning laser ophthalmoscope (SLO) to image retina and the stimuli simultaneously. Both studies included text reading, silent (Hutzler & Wimmer 2004) and oral (Trauzettel-Klosinski et al. 2010). Hutzler & Wimmer (2004) additionally used pseudoword reading. The results of the two studies show that dyslexic readers exhibited more fixations of longer duration in text reading, which is in accord with other research, and also an increased number of regressions. Hutzler & Wimmer (2004) found no group differences in pseudoword reading. Summarizing the results of Trauzettel-Klosinski et al. (2010), phonological difficulty influences reading speed and the number of eye movements but not fixation duration, which reflects a grapheme-phoneme strategy typical for reading in regular orthographies. The most interesting aspect of their analyses however is, in how far their results differed from Italian and English results. Hutzler & Wimmer (2004) found a small number and proportion of regressive eye movements which was similar to Italian findings and different to English-based results. Trauzettel-Klosinski et al. (2010) found the percentage of regressions only slightly increased in dyslexics as well, but interpreted these findings as being similar to English language studies, where reading speed of dyslexics had been found to be slowed by an increased number of saccades and regressions. Both studies found an increased number of fixations in German dyslexics, which was higher than in English speaking dyslexics, which indicates that German dyslexic children, similar to Italian, exhibited shorter forward movements of the eyes. Although showing a number of similarities, Hutzler & Wimmer (2004) also found differences to Italian findings, namely massively prolonged fixation durations in German dyslexic participants, which the authors attribute to the different natures of German and Italian syllables. This might indicate that, despite both having regular orthographies, Italian and German require slightly different reading strategies. One major difference Trauzettel-Klosinski et al. (2010) found between English and German dyslexics was that especially with increasing text difficulty, German participants exhibited an increased number of eye movements and smaller saccade amplitudes, while in English participants' fixation durations were prolonged. The findings suggest that, due to the relatively simple grapheme-phoneme-relations, German children favor the indirect sub-lexical route, while English speaking children prefer the direct lexical route. Unfortunately, there is no information or analysis on the influence of reading instruction on reading strategies in these studies.[8]

With regard to dyslexia specific eye movement patterns, there are two specific studies to be discussed here. While Vagge et al. (2015) take a rather general approach and aim at analyzing the relationship between dyslexia and eye movements and whether their analysis can be useful in the identification of dyslexics, Thaler et al. (2009) systematically investigate the reading behavior of children with dyslexia only, with an attentional deficit only, with dyslexia and a comorbid attentional deficit and of control children. While Thaler et al. (2009) investigated a total of 74 German speaking children (in all four groups) aged 7-11.5, the sample of Vagge et al. (2015) was slightly smaller with 22 Italian speaking children, of which 11 were dyslexic, aged 8-13 years. Thaler et al. (2009) used an eye tracker with a relatively high sampling frequency of 1000Hz (Eyelink-1000), while Vagge et al. (2015) used an OBER 2 System with a sampling frequency of only 100Hz. The studies also differed in their tasks. In Vagge et al. (2015) participants were instructed to silently read through two texts, while Thaler et al. (2009) asked their participants to read the word on the screen aloud as soon as they know how it is pronounced (naming task). A more detailed analysis of the results and their interpretation of both studies can be found above. Here, only the most important aspects and especially what both studies could confirm are discussed. Maybe most importantly, Thaler et al. (2009) could show that children with dyslexia only and with a comorbid attentional deficit use different reading strategies, at least at the young age of the participants. Further, in accord with Vagge et al. (2015), they could show that the reading speed of dyslexic readers is massively reduced due to the increased number of eye movements they make. With regard to word length effects, both studies showed that dyslexic readers

---

[8] For a recent analysis of the influence of reading instruction on reading strategies in university students in an L2-context, see Genç & Ünal (2017).

were even more affected by word length than controls. Taken together, these results indicate that children with reading difficulties seem to rely on serial decoding. These current findings support the idea that a thorough examination of visual functions in children can add crucial information to the assessment protocol for dyslexia. Furthermore, according to Vagge et al. (2015), it is not essential to know whether atypical eye movements are causal or an effect of dyslexia in order to use their evaluation for the diagnosis.

4. Graph comprehension

The next two papers to be compared are not only very similar in their focus but also share one author. Kim et al. (2014) and Kim & Wiseheart (2017) both investigated graph comprehension in college students with and without dyslexia. In particular, they both investigated the influence of graphic and orthographic properties. Kim et al. (2014) only had 15 dyslexic students in their participant sample, while Kim & Wiseheart (2017) had a larger sample of 29. Their mean ages were 20.89 and 21 years respectively, and all of them were native English speakers. Generally, in both studies reaction times and gaze durations were measured. Despite all the common aspects, there were two different eye-tracking devices used. Kim et al. (2014) used an Eyelink II 250Hz eyetracker, while Kim & Wiseheart (2017) decided for an LC Technologies head-free Eye Follower operating at 120Hz. In both studies, stimuli graphs were manipulated in three aspects. In both studies, there were two question types (point location and comparison), and two graph patterns (single or double). In Kim et al. (2014), there were three types of graphs (line graphs, vertical bar graphs and horizontal bar graphs), whereas in Kim & Wiseheart (2017) there were three different information formats (a mixed modality condition with orthographic and graphic information, an orthography free condition with an icon graph, and an orthography-only condition with texts only). The findings of both studies are mainly coherent. In both experiments, students with dyslexia exhibited significantly longer reaction times. According to the authors, these pronounced differences between dyslexic students and controls in reaction times can be attributed to the time spent on orthographic information, as for the stimulus area, if graphic, no differences in first viewing time could be found. Further, both research groups argued that slow but accurate text processing is consistent with the profile of dyslexia. These findings suggest that difficulties in graph comprehension in dyslexia are mainly related to orthographic processing difficulties rather than to difficulties in processing graphic information.

5. *Effects of word length, word frequency, and word type*

There were two studies in particular which investigated the effect of word length, word frequency, and word type. De Luca et al. (2002) focused on eye movement patterns of 12 Italian dyslexic children (mean age: 13.1 years) in short and long word and pseudoword reading, while Dürrwächter et al. (2010) investigated how word length and word frequency affected 16 German speaking children with dyslexia (mean age 9.5 years). Thus, both studies include reading in a regular orthography. Furthermore, Dürrwächter et al. (2010) compared their results to Italian and English findings. The eye movement parameters both studies measured were number of saccades, number of regressions and fixation duration. Like Trauzettel-Klosinski et al. (2010), Dürrwächter et al. (2010) also used a scanning laser ophthalmoscope, whereas De Luca et al. (2002) decided for an AMTech Eye-tracking system with a sampling frequency of 500Hz. According to their focus, De Luca et al. (2002) manipulated their stimuli in length, short with 4-5 letters versus long with 8-10 letters, and in their lexical value, so words and pseudowords. They instructed their readers to read silently and all words were high frequency. Due to their slightly different focus, Dürrwächter et al. (2010) manipulated their stimuli in word length as well as word frequency. Contrary to De Luca et al. (2002), participants were instructed to read the presented word aloud as quickly and accurately as possible. Before going into further detail about the results, both studies showed that dyslexic children seem to be affected by word length. In accord with former research, an increased number of saccades and regressions was found in dyslexic readers. The results of De Luca et al. (2002) suggest that proficient readers adjust their saccades to word length and only showed an increased number of saccades for pseudowords, whereas dyslexic participants exhibited a marked length effect regardless of the lexical value of the stimulus. For both groups, fixation duration was shorter when the target was meaningful. Generally, the length effects were more pronounced in dyslexics, indicating that dyslexic readers process words and pseudowords more parceled than controls do. These findings are in accord with the hypothesis of dyslexic readers' reliance on sub-lexical word processing in regular orthographies. Dürrwächter et al. (2010) noticed word length effects and word frequency effects as well, which were both more pronounced in dyslexic readers. The word length and word frequency effects were indicated by an increase of eye movements and prolonged fixation durations. As the authors claim, this increase in the number of eye movements is rather due to smaller units of analysis than an inability to control eye movements. The prolonged fixaion durations reflect information processing difficulties. Thus, the authors conclude that reading difficulties and inefficient eye movements in dyslexics rather reflect problems at higher-level word processing.

Generally, the comparison of orthographies in Dürrwächter et al. (2010) is in accord with the two studies discussed above in that there are small differences between German and Italian readers, which could be due to the slightly different processing modes of grapheme-phoneme conversion and the processing of larger units. However, English results differed more, which reflects whole-word decoding via the lexical route. This difference in orthographies is especially mirrored in the percentage of regressions, which , in regular orhtographies, is comparable to controls but which is significantly higher in English findings. These results can be attributed to the much higher regularity in grapheme-phoneme conversion of German and Italian.

6. Visual Search Tasks

In this very short abstract, the results of visual search tasks are summarized. There are three papers in particular, which are not discussed in detail here. Hutzler et al. (2006) compared eye movements of dyslexic and control participants in a visual search task and found no differences in the number of fixations or in gaze duration. They concluded that dyslexic readers do not seem to have difficulties in the accurate perception of letters or general problems in eye movement control. Similarly, Prado et al. (2007) found no significant difference between dyslexics and controls in visual search and see this as support for the hypothesis that the atypical eye movement patterns of dyslexics in reading cannot be explained in terms of oculomotor or visual perception problems. Finally, the findings of Silva et al. (2016) are in accord with the other results mentioned here, in that no differences in fixation times could be found. As has already been noted, these results indicate that the reason for inefficient eye movement patterns of dyslexics in reading are not due to an inability to control eye movements generally, but are more likely to derive from difficulties at higher level processing.

c. Eye-tracking in the diagnosis of dyslexia

The comparison and analysis given above shows that, notwithstanding the various different experiment settings in eye-tracking studies on dyslexia, the results of those studies often show similar patterns. Most studies could confirm that dyslexic participants generally exposed a higher number of eye movements (De Luca et al., 2002; Dürrwächter et al., 2010; Hutzler & Wimmer, 2004; Jainta & Kapoula, 2011; Pan et al., 2014; Prado et al., 2007; Silva et al., 2016; Trauzettel-Klosinski et al., 2010; Yan et al., 2013). Furthermore, there seems to be consent that dyslexics exhibit eye movements of longer duration (Bucci et al., 2008; De Luca et al., 2002; Dürrwächter et al., 2010; Hutzler & Wimmer, 2004; Jainta & Kapoula, 2011; Kim et al., 2014; Kim & Wiseheart, 2017; Pan et al., 2014; Prado et al., 2007; Silva et al., 2016; Trauzettel-Klosinski et al., 2010; Yan et al., 2013). The questions, which have not yet been answered, are in how far these eye movement patterns are dyslexia specific and in how far they can be used for diagnosing dyslexic readers. The first question is also approached in Thaler et al. (2009), but further research is certainly needed in order to give an educated answer. The second question initiated a study conducted by Benfatto and colleagues in (2016). In the following part, their study and results are summarized and its implications for future research on diagnostics are discussed.

i. Benfatto et al. 2016: Screening for dyslexia using eye-tracking during reading

In 2016, Benfatto and colleagues published a paper in which they argued for eye-tracking being a promising technique in the objective identification of dyslexic readers. In their paper, they criticize that arbitrary and subjective cutoffs are necessary in the diagnosis of dyslexia which is due to the lack of a specific definition of the reading deficit. This, however, does not mean that dyslexia is not real. Yet it makes an early identification very difficult, which, however, would be very important in order to support children with reading disabilities from early on. Benfatto and colleagues (2016) stress the need for a fast and objective means of identification, as current methods always require some explicit response by the subjects and often only measure individual cognitive skills but not the processes and functions actually active in reading. For this purpose, the authors introduced eye-tracking as technique to objectively measure real-time reading processes without the need of overt responses by the subjects. They combined eye-tracking with machine learning and predictive modeling to yield individual-level predictions with high sensitivity and specificity. By applying statistical cross-validation techniques, Benfatto et al. (2016) achieved a classification accuracy of 96%. In a second step, the authors aimed at identifying crucial features that differentiate high risk subjects from low risk subjects.

For their study, Benfatto and colleagues (2016) used the eye-tracking data of 185 subjects participating in a longitudinal research project from 1989 until 2010. 103 of the participants were high risk subjects aged 8 to 9 years, for which the inclusion criteria were Swedish being their first language, a performance lower than the 5th percentile of the full cohort on two word decoding tests, and an independent assessment of reading problems by their teacher. Control participants were matched pairwise to the high risk subjects. Further studies with the same sample of participants showed that reading problems often persisted and notably influenced school performance, academic achievement but also other domains of life.

To record eye movements, the authors used an infrared goggle-based Ober-2 system with a sampling frequency of 100Hz. The participants were instructed to silently read a text on a single page of white paper and to answer three questions on the content afterwards.

After collecting eye-tracking data, the raw recording signals of eye positions were analyzed and eye movement features extracted to be used as input for training a classification model. The following features were extracted: progressive and regressive saccades, progressive and regressive fixations, duration of the event, distance spanning the event, average eye position during the event, standard deviation of average position, maximum range between any two positions, and accumulated distance over all subsequent positions. These features capture information on eye movements in reading, such as their amplitude, stability, duration, direction and symmetry.

With the eye-tracking data, Benfatto et al. (2016) trained a learning algorithm. Its ability to identify high and low risk participants was assessed. The authors removed features with little or no predictive information which reduced noise and facilitated the identification of eye movement features that give best predictive performance.

Best classification accuracy was observed with 95,6 ±4,5% which shows that an automatic feature analysis can actually be used to differentiate more informative eye movements from less predictive eye movements. A comparison to conventional screening tests is more complicated as their accuracy is often not well known. The accuracy of screening instruments used for children before reading instruction is estimated at 70-80%. However, in these instruments levels sensitivity and specificity are highly imbalanced. If specificity is high but sensitivity low, this can mean that a test easily excludes children without dyslexia but may not really identify the ones with dyslexia, which makes those tests not very effective. As soon as children receive reading instruction, the accuracy of screening tests increases up to 80-90%. However, there is still the problem of imbalanced levels of sensitivity and specificity which, at the moment, can only be avoided by administering multiple tests. Multiple tests require more resources which might be a barrier that prevents schools from implementing routine dyslexia screenings.

In order to select those features that contributed more useful information to accuracy, the authors analyzed the frequency with which features were selected in 1000 internal training folds. The higher the frequency of a feature, the more likely it is more predictive. 24% of the selected features were progressive fixations, 26% progressive saccades, 21% regressive fixations, and 29% regressive saccades. According to the authors, progressive eye movements seem to be more informative. Some of their findings were in accord with previous findings on eye movements of dyslexic and normal readers, such as prolonged fixation durations and shorter saccades.

To sum up, early and objective identification is important in order to support children with dyslexia as soon as possible. However, conventional screening methods are not objective and often lack scientific support. Therefore, Benfatto et al. (2016) introduced and analyzed eye-tracking as potential screening method which yields individual predictions with high specificity and sensitivity in one minute of tracking time. The advantages would be that no overt response by the subject is required, as eye movement signals are the only response. Eye movements cannot be right or wrong which makes this method objective. Furthermore, a relatively natural reading situation could reduce the stress level of participants in testing. However, children with reading difficulties can follow various neuropsychological profiles. There is, for example, extensive symptom overlap and a high comorbidity rate between dyslexia, attentional deficits and specific language impairment. Furthermore, the authors suggest that a differentiation between surface and phonological dyslexia might be important. While eye movement analysis could be an efficient and helpful means of early identification, follow up screenings in order to get more information on an individual’s cognitive profile still seem necessary.

ii. Future prospect

The study of Benfatto et al. (2016) gives interesting insight into how eye-tracking could be used in the diagnosis of dyslexia. Yet, there are is a crucial question still to be answered: What exactly is needed to include eye-tracking data in the identification of dyslexics? In the following part, this question is elaborated.

First and foremost, in order to successfully train learning algorithms to detect dyslexia, most likely a lot of eye-tracking data is needed. From Benfatto et al. (2016) can only be assumed that their amount of eye-tracking data was sufficient, however it remains unclear what the lowest threshold could be. Eye movement recordings would have to be language specific, since children with different native languages expose slightly different eye movements in reading (Hutzler & Wimmer, 2004;

Trauzettel-Klosinski et al., 2010). Furthermore, eye-tracking data, and further each classifier, would have to be age specific. Eye movements in reading change and develop over time and with formal instruction and practice (Bellocchi et al., 2013).Thus, a 12 year old dyslexic reader might exhibit different eye movement patterns than an 8 year old dyslexic reader. Furthermore, dyslexic readers might develop compensatory mechanisms over time. In order to obtain sufficient eye movement data as quickly as possible, modern eye-tracking technology with a higher sampling frequency of 1000Hz or more might be advisable.

In addition, in order to make a model as efficient as possible, further research on the predictive value of specific eye movements might be needed. Research in this field could also be language specific, as different orthographies evoke different eye movement patterns and, thus, different specific eye movements might be more or less predictive in different orthographies.

Furthermore, there is a rather basic and widely discussed question that has to be dealt with. In dyslexia, a relatively high comorbidity rate and symptom overlap with attentional deficits could be observed (Thaler et al., 2009). This makes a clarification of dyslexia specific eye movement necessary. Without having clear eye movement patterns of dyslexic readers and readers with a reading disorder and an attentional deficit, for example, a clear diagnosis based on eye movements might not be possible.

Benfatto et al. (2016) show how promising eye-tracking as diagnostic method is. As soon as the background work is done, eye-tracking is a relatively simple, fast, and objective means of identification. However, until this point is reached, a lot of research is still needed. Furthermore, the study of Benfatto et al. (2016) was only conducted in Swedish. It is still to be shown if eye-tracking data is as useful and informative in other languages as well.

5. Conclusion

In conclusion, this master's thesis aimed at giving insight into one particular branch of research in dyslexia. The comparison of eye-tracking studies in dyslexia was supposed to shed light on the variety of research in this field as well as to filter differences and similarities between those studies. In doing so, implications for future research and maybe for diagnosis could be elaborated.

Dyslexia is a highly controversial condition. Despite the extensive body of research on dyslexia, no consensus with regard to the definition could be achieved in the past few decades. The question of dyslexia-specific symptoms is subject to debate as well. While some argue indeed for a dyslexia-specific reading behavior, others claim this behavior to mirror reading-level-matched controls. Factors further blurring a clear image of dyslexia are high comorbidity rates especially with attentional deficits and SLI, as well as the question of subtypes of dyslexia. All these factors influenced the theories elaborated to explain the origin and symptoms of dyslexia. While all theories seem to account at least for a part of the symptoms of dyslexia, there is one approach more widely accepted. The phonological theory locates the source of dyslexia in difficulties of representing, storing and/or retrieving speech sounds. One final yet crucial factor in dyslexia research is the question of diagnosis. Screening methods in dyslexia are often criticized for various reasons. Diagnosis certainly depends on definition; however there is no uniform definition of dyslexia. It mostly relies on overt output of a participant which can be influenced by a number of factors. Yet, early identification of dyslexic children seems important in order to provide early support. One particular approach is to investigate children's eye movements in reading in order to use this method for identification of dyslexic readers.

Eye-tracking, in its most primitive sense, goes back to ancient times, which shows that the interest in how people move their eyes is not only a recent one. Particularly, advances in 19th century pathed the way to nowadays state of the art. Today, eye-tracking is a widely used method, especially in cognitive psychology, used in order to investigate cognitive processes active during various tasks. Visual perception can be divided into two main parts: the area of clear vision where most information can be retrieved- the fovea- and the parafovea, the area surrounding the fovea where less particular information for example on word length can be retrieved. Eye-tracking research is mostly built on the eye-mind assumption according to which our eyes move to where our attention is. There are mainly two kinds of eye movements, fixations and saccades. During fixations, eyes remain relatively still and clear vision enables the retrieval of information. Saccades describe the eye movements between fixations, during which humans are “blind” to new information. There is a large number of eye movement parameters which can be investigated and discussed, spanning three scales of measurement: temporal, spatial and count. Eye-tracking research in unimpaired reading soon triggered research in impaired reading, such as dyslexia.

Eye-tracking research in dyslexia yielded a high number of studies in the last few decades. However, these studies vary in many respects and direct comparison is difficult. Due to the large body of research and the different research paradigms, the question of generalization arose. In order to find similarities and difficulties in the paradigms and the results, 18 peer-reviewed eye-tracking studies published between 2002 and 2017 were systematically compared and contrasted. In the synchronic analysis, all studies were compared with regard to their foci and hypotheses, their experiment settings and their results and interpretation.

In a second step, the studies were grouped according to similar foci and/or experiment settings. Studies on parafovea and attentional span yielded similar results in that they found parafoveal vision in dyslexics to be not as efficient as in unimpaired readers. The results of three eye-tracking studies focusing on saccade and vergence computation and control in dyslexics revealed that dyslexic participants indeed show difficulties in saccade control and longer saccade latencies. In accord with the other studies, the studies focusing on orthography and language found that German dyslexic readers exhibited more fixations of longer duration in text reading as well as an increased number of regressions. In comparison to other languages, the three studies found an increased number of fixations in German dyslexics, which was higher than in English speaking dyslexics. This indicates that German dyslexic children, similar to Italian, exhibited shorter forward movements of the eyes, which could reflect a grapheme-phoneme strategy typical for reading in regular orthographies. While German dyslexics show similarities to Italian dyslexics, one major difference are the prolonged fixation durations in German participants, which is attributed to the differing syllable structures of German and Italian. While English speaking dyslexics exhibit prolonged fixation durations when reading a difficult passage, German participants exhibited an increased number of eye movements and smaller saccade amplitudes. The findings suggest that, due to the relatively simple grapheme-phoneme-relations, German children favor the indirect sub-lexical route, while English speaking children prefer whole word decoding via the direct lexical route. Another focus was set on dyslexia specific reading patters. Two studies indeed found differences in reading strategies in children with dyslexia only, with dyslexia and an attentional deficit, with an attentional deficit only and a control group. Further, in accord with the studies on different orthographies, reduction in reading speed was attributed to a serial decoding strategy. Two particular studies focused on graph comprehension with similar results. In both experiments, students with dyslexia exhibited significantly longer reaction times which the authors attributed to the time spent on orthographic information. Their findings suggest that difficulties in graph comprehension in dyslexia are mainly related to orthographic processing difficulties rather than to difficulties in processing graphic information. The studies on word length and word frequency showed more pronounced word length and word frequency effects in dyslexic readers. Finally, there were visual search tasks conducted in three studies. The results show no differences in the number of fixations or gaze durations between dyslexic and unimpaired readers. These findings indicate that atypical eye movement patterns in dyslexics in reading cannot be explained in terms of visual perception problems.

Generally, most studies could provide evidence for an increased number of eye movements, fixations and saccades, of longer duration in dyslexic readers. Further, most studies attribute these atypical eye movement patterns not to visual perception problems but to deficits in higher level processing. How these eye movement patterns of dyslexic readers could be used for identification was addressed in one particular study by Benfatto and colleagues (2016).

In 2016, Benfatto and colleagues published a paper in which they argued for eye-tracking being a promising technique in the objective identification of dyslexic readers. For this purpose, the authors introduced eye-tracking as technique to objectively measure real-time reading processes without the need of overt responses by the subjects. They combined eye-tracking with machine learning and predictive modeling to yield individual-level predictions with high sensitivity and specificity. For their study, Benfatto and colleagues (2016) used the eye-tracking data of 185 subjects participating in a longitudinal research project from 1989 until 2010. By applying statistical cross-validation techniques, the authors achieved a classification accuracy of 96% with only one minute of eye-tracking data. A comparison to conventional screening tests is more complicated as their accuracy is often not well known. In a second step, the authors aimed at identifying crucial features that differentiate high risk subjects from low risk subjects. According to them, progressive eye movements seem to be more informative. Some of their findings were in accord with previous findings on eye movements of dyslexic and normal readers, such as prolonged fixation durations and shorter saccades. Early and objective identification is important in order to support children with dyslexia as soon as possible. The advantages of this method would be that no overt response by the subject is required, as eye movement signals are the only response. Eye movements cannot be right or wrong which makes this method objective. Furthermore, a relatively natural reading situation could reduce the stress level of participants in testing. While eye movement analysis could be an efficient and helpful means of early identification, follow up screenings in order to get more information on an individual's cognitive profile still seem necessary to control for comorbidity and types of dyslexia.

In order to successfully use eye-tracking data in the identification of dyslexics, there are still a few steps needed. First, a large amount of eye-tracking data is required for training learning algorithms in order to detect dyslexia. These eye movement recordings would have to be language specific, since children with different native languages show different eye movement behavior in reading. Furthermore, the recordings would also have to be age specific as eye movements in reading develop over time with formal instruction and practice. In order to obtain sufficient eye movement data as quickly as possible, modern eye-tracking technology with a high sampling frequency of 1000Hz or more would be advisable. With regard to the efficiency of a model, further research on the predictive value of specific eye movements might be needed. Research in this field could also be language specific, as different orthographies evoke different eye movement patterns and, thus, different specific eye movements might be more or less predictive in varying orthographies. One particular challenge will be finding dyslexia specific eye movements. Due to high comorbidity rates and symptom overlap with other conditions, clear eye movement patterns are needed for efficient diagnosis. The study of Benfatto and colleagues (2016) shows how promising eye-tracking as diagnostic method can be. Even though preparations and ground work seem to require huge efforts and resources, as soon as they are done, eye-tracking might proof a relatively simple, fast, and objective means of identification. However, until this point is reached, a lot of research is still needed.

6. References

Ablinger, I.; von Heyden, K.; Vorstius, C.; Halm, K.; Huber, W.; Radach, R. 2014. „An eye movement based reading intervention in lexical and segmental readers with acquired dyslexia". *Neuropsychological Rehabilitation* 24(6), 833-867.

Al Dahhan, N.; Georgiu, G. K.; Hung, R.; Munoz, D.; Parrila, R.; Kirby, J. R. 2014). "Eye movements of university students with and without reading difficulties during naming speed tasks". *Annals of Dyslexia* 64, 137-150.

Andersson, R.; Nyström, M.; Holmqvist, K. 2010. "Sampling frequency and eye-tracking measures: how speed affects durations, latencies, and more"*. Journal of Eye Movement research* 3(3), 1-12.

Ashby, J.; Yang, J.; Evans, K. H. C.; Rayner K. 2012. "Eye movements and the perceptual span in silent and oral reading". *Attention, Perception, and Psychophysics* 74(4), 634-640.

Beaton, A. A. 2004. *Dyslexia, Reading and the Brain: A sourcebook of Psychological and Biological Research*. Hove: Psychology Press.

Bellocchi, S.; Muneaux, M.; Bastien Toniazzo, M.; Ducrot, S. 2013. "I can read it in your eyes: What eye movements tell us about visuo-attentional processes in developmental dyslexia". *Research in Developmental Disabilities* 34(1), 452-460.

Benfatto, M. N.; Seimyr, G. Ö.; Ygge, J.; Pansell, T.; Rydberg, A.; Jacobson, C. 2016. "Screening for dyslexia using eye-tracking during reading". *PLOS One* 11(12), 1-17.

Blythe, H. I.; Joseph, H. S. S. L. 2011. "Children's eye movements during reading". In Liversedge, S. P., Gilchrist, I., Everling, S. (eds.). *The Oxford handbook of eye movements.* Oxford: Oxford University Press, 643-662.

Breznitz, Z. 2008. "The Origin of dyslexia: The Asynchrony Phenomenon". In Reid G.; Fawcett, A.; Manis, F.; Siegel, L. (eds.). *The SAGE handbook of dyslexia. 1st edition*. Los Angeles: SAGE, 11-29.

Bucci, M. P.; Brémond-Gignac, D.; Kapoula, Z. 2008. "Latency of saccades and vergence eye movements in dyslexic children". *Experimental Brain Research* 188(1), 1-12.

Cavalli, E.; Duncan, L. G.; Elbro, C.; El Ahmadi, A.; Colé, P. 2016. "Phonemic-morphemic dissociation in university students with dyslexia: an index of reading compensation?". *Annals of Dyslexia* 67, 63-84.

Coltheart, M.; Rastle, K.; Perry, C.; Langdon, R.; Ziegler, J. 2001. "DRC: A dual-route cascaded model of visual word recognition and reading aloud". *Psychological Review* 108(1): 204-256.

De Luca, M.; Borrelli, M.; Judica, A.; Spinelli, D.; Zoccolotti, P. 2002. “Rapid communication. Reading words and pseudowords: An eye movement study of developmental dyslexia”. *Brain & Language* 80, 617-626.

Deacon, S. H.; Cook, K.; Parrila, R. 2012. “Identifying high-functioning dyslexics: is self-report of early reading problems enough?”. *Annals of Dyslexia* 62, 120-134.

Deponia, P. 2004. “The co-occurrence of specific learning difficulties: implications for identification and assessment”. In Reid, G.; Fawcett, A. J. (eds.). *Dyslexia in Context: Research, Policy and Practice*. London: Whurr Publishers, 323-335.

Duchowski, A. 2007. *Eye-tracking methodology: theory and practice*. London: Springer-Verlag.

Dürrwächter, U.; Sokolov, A. N.; Reinhard, J.; Klosinski, G.; Trauzettel-Klosinski, S. 2010. „Word length and word frequency affect eye movements in dyslexic children reading in a regular (German) orthography”. *Annals of Dyslexia* 60, 86-101.

Elston, D. M. 2017. “Sample size”. *Journal of the American Academy of Dermatology*

Everatt, J.; Elbeheri, G. 2008. “Dyslexia in different orthographies: variability in transparency”. In Reid, G.; Fawcett, A.; Manis, F.; Siegel, L. (eds.). *The SAGE handbook of dyslexia. 1st edition*. Los Angeles: SAGE, 427-438.

Fawcett, A.J.; Nicolson, R. I. 2008. “Dyslexia and the cerebellum”. In Reid, G.; Fawcett, A.; Manis, F.; Siegel, L. (eds.). *The SAGE handbook of dyslexia. 1st edition*. Los Angeles: SAGE, 77-98.

Genç, N.; Ünal, D. (2017). “Die Vermittlung von Lesestrategien und ihr Einfluss auf den Strategiegebrauch“. *Moderna Sprak* 111(1), 35-58.

Gilger, J. W. 2008. “Some special issues concerning the genetics of dyslexia: revisiting multivariate profiles, comorbidities, and genetic correlations”. In Reid, G.; Fawcett, A.; Manis, F.; Siegel, L. (eds.). *The SAGE handbook of dyslexia. 1st edition*. Los Angeles: SAGE, 30-52.

Goldberg, H. K.; Shiffman, G. B.; Bender, M. 1983. *Dyslexia: Interdisciplinary approaches to reading disabilities*. New York, NY: Grune& Stratton.

Götzinger Hiebner, M. 2014. *Die Bedeutung phonologischer Prozesse für den Leseerwerb.* Wien: Universität Wien.

Hasko, S.; Groth, K.; Bruder, J.; Bartling, J.; Schulte-Körne, G. 2013. „The time course of reading processes in children with and without dyslexia: an ERP study”. *Frontiers in Human Neuroscience* 7(570), 1-19.

Hawelka, S.; Gagl, B.; Wimmer, H. 2010. „A dual-route perspective on eye movements of dyslecic readers”. *Cognition* 115, 367-379.

Helland, T.; Morken, F. 2015. “Neurocognitive development and predictors of L1 and L2 literacy skills in dyslexia: A longitudinal study of children 5-11 years old”. *Dyslexia* 22, 3-26.

Holmqvist, K.; Nyström, N.; Andersson, R.; Dewhurst, R.; Jarodzka, H.; Van de Weijer, J. (eds.). 2011. *Eye-tracking. A comprehensive guide to methods and measures*. Oxford: Oxford University Press.

Hutzler, F.; Wimmer, H. 2004. “Eye movements of dyslexic children when reading in a regular orthography”. *Brain and Language* 89(1), 235-242.

Hutzler, F.; Kronbichler, M.; Jacobs, A. M.; Wimmer, H. (2006). „Perhaps correlational but not causal: No effect of dyslexic readers‘ magnocellular system on their eye movements during reading”. *Neuropsychologica* 44, 637-648.

Hyönä, J. 2011. “Foveal and parafoveal processing during reading”. In Liversedge, S. P.; Gilchrist, I.; Everling, S. (eds.). *The Oxford handbook of eye movements.* Oxford: Oxford University Press, 819-838.

Internationale statistische Klassifikation der Krankheiten und verwandter Gesundheitsprobleme, 10. Revision-BMG-Version 2014

Jainta, S.; Kapoula, Z. 2011. “Dyslexic children are confronted with unstable binocular fixation while reading”. *PLOS One* 6, 1-10.

Just, M. A.; Carpenter, P. A. 1980. “A theory of reading: From eye fixations to comprehension”. *Psychological Review* 87, 329-355.

Kim, S.; Lombardino, L. J.; Cowles, W.; Altmann, L. J. 2014. “Investigating graph comprehension in students with dyslexia: An eye-tracking study”. *Research in Developmental Disabilities* 35, 1609-1622.

Kim, S.; Wiseheart, R. 2017. “Exploring text and icon graph interpretation in students with dyslexia: An eye-tracking study”. *Dyslexia* 23, 24-41.

Krieber, M.; Bartl-Pokorny, K. D., Pokorny, F. B.; Zhang, D.; Landerl, K.; Körner, C.; Pernkopf, F.; Pock, T.; Einspieler, C.; Marschik, P. B. 2017. „Eye Movements during silent and oral reading in a regular orthography: Basic characteristics and correlations with childhood cognitive abilities and adolescent reading skills”. PLOS One 12(2), 1-16.

Kristjánsson, A. 2011. “The intriguing interactive relationship between visual attention and saccadic eye movements”. In Liversedge, S. P.; Gilchrist, I.; Everling, S. (eds.). *The Oxford handbook of eye movements.* Oxford: Oxford University Press, 455-470.

Lai, M. L.; Tsai, M. J.; Yang, F. Y.; Hsu, C. Y.; Liu, T. C.; Lee, S. W. Y.; Lee, M. H.; Chiou, G. L.; Liang, J. C.; Tsai, C. C. 2013. “A review of using eye-tracking technology in exploring from 2000 to 2012”. *Educational Research Review* 10, 90-115.

Lindsay, G. 2004. “Baseline assessment and early identification of dyslexia”. In Reid, G.; Fawcett, A. J. (eds.). *Dyslexia in Context: Research, Policy and Practice*. London: Whurr Publishres, 278-287.

Liversedge, S. P.; Gilchrist, I.; Everling, S. (eds.). 2011. *The Oxford handbook of eye movements.* Oxford: Oxford University Press.

Ludwig, C. J.H. 2011. “Saccadic decision-making”. In Liversedge, S. P.; Gilchrist, I.; Everling, S. (eds.). *The Oxford handbook of eye movements.* Oxford: Oxford University Press, 425-437.

Lyon, G. R.; Shaywitz, S. E.; Shaywitz, B. A. 2003. “A definition of Dyslexia”. *Annals of Dyslexia* (53), 1-14*.*

Lyytinen, H.; Erskine, J.; Ahonen, T.; Aro, M.; Eklund, K.; Guttorm, T.; Hintikka, S.; Hamalainen, J.; Ketonen, R.; Laakso, M. L.; Leppanen, P. H. T.; Lyytinen, P.; Poikkeus, A. M.; Puolakanaho, A.; Richardson, U.; Salmi, P.; Tolvanen, A.; Torppa, M.; Viholainen, H. 2008. “Early identification and prevention of dyslexia: Results from a prospective follow-up study of children at familial risk for dyslexia”. In Reid, G.; Fawcett, A.; Manis, F.; Siegel, L. (eds.). *The SAGE handbook of dyslexia. 1st edition*. Los Angeles: SAGE, 121-146.

Manis, F.; Bailey, C. E. 2008. “Exploring heterogeneity in developmental dyslexia: A longitudinal investigation”. In Reid, G.; Fawcett, A.; Manis, F.; Siegel, L. (eds.). *The SAGE handbook of dyslexia. 1st edition*. Los Angeles: SAGE, 149-173.

Munoz, D. P.; Armstrong, I.; Coe, B. 2007. “Using eye movements to probe development and dysfunction”. In: Van Gompel, Roger (eds). *Eye movements – a window on mind and brain*. Amsterdam: Elsevier GmbH, S. 99-125.

Nicolson, R. I.; Fawcett, A. J. 2008. “Learning, cognition and dyslexia”. In Reid, G.; Fawcett, A.; Manis, F.; Siegel, L. (eds.). *The SAGE handbook of dyslexia. 1st edition*. Los Angeles: SAGE, 192-211.

Pan, J.; Yan, M.; Laubrock, J.; Shu, H.; Kliegl, R. 2013. „Eye-voice span during rapid automatized naming of digits and dice in Chinese normal and dyslexic children”. *Developmental Science* 16(6), 967-979.

Pan, J.; Yan, M.; Laubrock, J.; Shu, H.; Kliegl, R. 2014. „Saccade-target selection of dyslexic children when reading Chinese”. *Vision Research* 97, 24-30.

Pavlidis, G. Th. 1985. “Eye movements in dyslexia: Their diagnostic significance”. *Journal of Learning Disabilities* (18)1, 42-50.

Prado, C.; Dubois, M.; Valdois, S. 2007. “The eye movements of dyslexic children during reading and visual search: Impact of the visual attention span”. *Vision Research* 47, 2521-2530.

Ramus, F.; Rosen, S.; Dakin, S. C.; Day, B. L.; Castellote, J. M.; White, S.; Frith, U. 2003. “Theories of developmental dyslexia: insights from a multiple case study of dyslexic adults”. *Brain* 126, 841-865.

Ramus, F., & Ahissar, M. 2012. “Developmental dyslexia: The difficulties of interpreting poor performance, and the importance of normal performance”. *Cognitive Neuropsychology* 29 (1-2), 104-122.

Rayner, K.; Liversedge, S. P. 2011. “Linguistic and cognitive influences on eye movements during reading”. In Liversedge, S. P.; Gilchrist, I.; Everling, S. (eds.). *The Oxford handbook of eye movements.* Oxford: Oxford University Press, 751-766.

Reid, G.; Fawcett A.J.2004. “An overview of developments in dyslexia”. In Reid, G.; Fawcett, A. J. (eds.). *Dyslexia in context: Research, Policy and Practice*. London: Whurr Publishers, 3- 19.

Reid, G.; Fawcett, A. J. (eds.). 2004. *Dyslexia in context: Research, Policy and Practice*. London: Whurr Publishers.

Reid, Gavin (ed.). 2008. The *SAGE handbook of dyslexia. 1st edition.* Los Angeles: SAGE.

Rosen, G. D.; Wang, Y.; Fiondella, C. G.; LoTurco, J. J. 2009. „The Brain and developmental dyslexia: Genes, anatomy, and behavior“. In Pugh, K.; McCardle, P. (eds.). *How children learn to read*. New York, NY: Psychology Press, 21-42.

Ryan, T. P. 2013. “Methods of determining sample sizes”. In Ryan, T. P. *Sample size determination and power*, New Jersey: John Wiley & Sons Inc., 17-55.

Schattka, K. I.; Radach, R.; Huber, W. 2010. „Eye movement correlates of acquired central dyslexia”. *Neuropsychologia* 48(10), 2959- 2973.

Shaywitz, S. E.; Shaywitz, B. A. 2005.” Dyslexia (specific reading disability)”. *Biological Psychiatry* 57(11), 1301-1309.

Silva, S.; Faísca, L.; Araújo, S.; Casaca, L.; Carvalho, L.; Petersson, K. M.; Reis, A. 2016. “Too little or too much? Parafoveal preview benefits and parafoveal load costs in dyslexic adults”. *Annals of Dyslexia* 66, 187-201

Stein, J. 2008. “The neurobiological basis of dyslexia”. In Reid, G.; Fawcett, A.; Manis, F.; Siegel, L. (eds.). *The SAGE handbook of dyslexia. 1st edition*. Los Angeles: SAGE, 53-76.

Thaler, V.; Urton, K.; Heine, A.; Hawelka, S.; Engl, V.; Jacobs, A. M. 2009. „Different behavioral and eye movement patterns of dyslexic readers with and without attentional deficits during single word reading”. *Neuropsychologia* 47(12), 2436-2445.

Trauzettel-Klosinski, S.; Koitzsch, A. M.; Dürrwächter, U.; Sokolov, A. N.; Reinhard, Klosinski, G. 2010. „Eye movements in German-speaking children with and without dyslexia when reading aloud”*. Acta Ophthalmologica* 88, 681-691.

Vagge, A.; Cavanna, M.; Traverso, C. E.; Iester, M. 2015. “Evaluation of ocular movements in patients with dyslexia”. *Annals of Dyslexia* 65, 24-32.

Wade, N. J.; Tatler, B. W. 2011. “Origins and applications of eye movement research”. In Liversedge, S. P.; Gilchrist, I.; Everling, S. (eds.). *The Oxford handbook of eye movements.* Oxford: Oxford University Press, 17-44.

Wagner, R. K. 2008. “Rediscovering dyslexia: New approaches for identification and classification”. In Reid, G.; Fawcett, A.; Manis, F.; Siegel, L. (eds.). *The SAGE handbook of dyslexia. 1st edition*. Los Angeles: SAGE, 174-191.

White, S. J.; Johnson, R. L.; Rayner, K. 2008. “Eye movements when reading transposed text: The importance of word-beginning letters”. *Journal of Experimental Psychology, Human Perception and Performance* 34(5), 1261-1276.

Yan, M.; Pan, J.; Laubrock, J.; Kliegl, R., Shu, H.2013. „Parafoveal processing efficiency in rapid automatized naming: A comparison between Chinese normal and dyslexic children”. *Journal of Experimental Child Psychology* 115, 579-589.

7. Appendix

Kurzzusammenfassung

Während die Dyslexie-Forschung sich seit vielen Jahren mit den Ursachen und Symptomen auseinander setzt, wurde zunehmend auch die Diagnostik in den Mittelpunkt der Forschung gerückt. Die Frage, inwiefern sich Kinder mit Dyslexie in ihrem Leseverhalten von Kindern ohne Lese-Rechtschreib-Problemen unterscheiden wurde unter anderem in Eye-Tracking Studien untersucht, die tatsächlich ein unterschiedliches Blickverhalten in den beiden Gruppen aufweisen konnten. Da dieses Forschungsfeld relativ neu ist, wurden Studien oft in unterschiedlichen Sprachen, mit unterschiedlichen Teilnehmergruppen und Zahlen und divergierenden Experiment-Settings durchgeführt. Dadurch entstand die Notwendigkeit, diese Studien mit einander zu vergleichen, um Gemeinsamkeiten und Unterschiede speziell in den Ergebnissen herauszuarbeiten. In dieser Masterarbeit wurden 18 Eye-Tracking Studien zu Dyslexie auf drei Hauptparameter hin verglichen und kontrastiert. Der Fokus lag dabei mehr auf den Gemeinsamkeiten und den Implikationen, die sich durch den Vergleich nicht nur für zukünftige Forschung, sondern speziell auch für Eye-Tracking als mögliches Diagnostikverfahren in Dyslexie ergaben. Während Eye-Tracking allein vermutlich keine hinreichende Methode darstellt, um Dyslexie in Kindern zu identifizieren, so ist es doch ein vielversprechender Ansatz, dessen mehr Forschung bedarf.

Abstract

While research in dyslexia has been focusing on the causes and symptoms of dyslexia for many years now, special focus was recently also put on diagnostics. Especially eye-tracking studies started elaborating on the question in how far children with dyslexia differ from unimpaired children in their reading behavior, and indeed many found diverging eye movement patterns in both groups. Since this field of research is relatively new, studies have been conducted in different languages, with different samples of participants and different experiment settings. Hence, there emerged the need for comparison in order to find similarities as well as differences especially in the results. In this Master’s thesis, 18 eye-tracking studies on dyslexia

were compared according to three main parameters. Focus was rather set on the similarities and the implications for future research and for eye-tracking in the diagnosis of dyslexia. While eye-tracking alone is presumably no sufficient means of identification, it proved a promising approach which definitely deserves more research.