Eye-tracking Analysis of Reading in People with Aphasia

by

Mark Paulo Mendoza

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science

in

Speech-Language Pathology

Department of Communication Sciences and Disorders

University of Alberta

Abstract

Reading impairments (*acquired alexia*) commonly co-occur in people with aphasia (PWA). The ultimate goal of reading treatment is to read connected text accurately and quickly. Traditional treatments of acquired alexia have typically targeted single word reading, which has seen little generalization to text reading. Text treatments, though relatively rare, have demonstrated an increase in oral reading rate and accuracy. Investigators have proposed that text is facilitative for PWA as it allows them to use "top-down" processing to integrate syntactic and contextual information (*context effect*). However, it is unclear exactly *what* linguistic information PWA are utilizing and *how* it is being used. A potentially useful approach is to utilize eye-tracking methodology to investigate reading behaviour. In this thesis, the impact of a central language impairment on eye-movements during reading is explored. This investigation occurred in two parts. First, a systematic review of reading studies utilizing eye-tracking methodology was performed to describe the eye-movements of PWA relative to healthy readers. Six studies, including observational and experimental studies, with a total of 60 PWA were found. Compared to healthy readers, PWA had longer processing times, were less likely to skip words, and were more likely to reread previous parts of text. Furthermore, these studies revealed that similar to healthy readers, the reading of PWA can be mediated by bottom-up factors (i.e., lexical variables including length and frequency) and top-down factors (i.e., syntactic and contextual variables including sentence complexity and predictability). Second, a retroanalysis of Kim & Bolger (2012) was performed using linear mixed effects models to analyze the effects of context (a top-down factor) and word frequency (a bottom-up factor) on text reading. The eye-movements of participants were tracked as they read target words of varying frequency (High, Low) embedded in sentences with varying predictability (High, Low). Models were built for the response variables of total fixation duration, total number of fixations, and total number of regressions. Group effects were found for all response variables reflecting that PWA had longer processing times than healthy readers. The main effect of context was found in total fixation duration indicating that PWA, similar to healthy readers, utilize context to predict upcoming words. The results of both the retroanalysis and systematic review are consistent in that they show that PWA are utilizing linguistic information within text to facilitate their reading, and that this facilitation is evident at the level of eye-movements.

Acknowledgements

I would like to express my gratitude to my supervisor, Dr. Esther Kim. Your guidance and mentorship throughout this project has helped shaped my understanding of what it means to be a clinical researcher. Thank you to the members of my supervisory and examining committee, Dr. Melissa Skoczylas, Dr. Monique Charest, and Dr. Trelani Chapman, who have helped shape this project with their feedback. Finally, many thanks to the members of the Aphasia Lab who supported me during my defense preparation. This thesis was supported by the Frederick Banting and Charles Best Canada Graduate Scholarship-Master's (CIHR CGS-M) and Walter H. Johns Graduate Fellowship.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

Aphasia is an acquired language impairment resulting from damage to the language-dominant hemisphere, and may affect receptive and expressive language modalities. The underlying impairment can impact more than one modality, including oral speech, auditory comprehension, reading, and writing. It should be of no surprise that most individuals with aphasia may present with some degree of reading impairment (*acquired alexia*) (Brookshire, Wilson, Nadeau, Gonzalez Rothi, & Kendall, 2014). Acquired alexia can have a substantial effect on quality of life, because literacy skills are used to interact and navigate through society: we stay in touch with the world by reading books, newspapers, and signs; we connect with people around us with social media, phone text message, and emails; we navigate our world by reading bus schedules, signs, and television/computer screens. It is evident that the world is increasingly becoming more dependent on electronic devices to stay connected with those around us and to provide services and information. Those with acquired alexia are at a disadvantage as using these devices generally requires intact reading skills.

Reading is complex task involving the translation of print into comprehensible linguistic information (Hoover & Gough, 1990). The Simple View of Reading model describes reading comprehension as the result of *decoding* (i.e., semantic retrieval at word level from graphemic information) and *linguistic comprehension* processes (i.e., interpretation of semantic information into syntactic and discourse information) (Hoover & Gough, 1990). A number of treatments for acquired alexia have focused on decoding, as these target remediation of single word reading (Cherney, 2004). These treatments have generally resulted in greater accuracy and performance in reading single-words, but generalization to reading at the text level is usually absent. Text-based treatments are relatively uncommon. These are advantageous as they target both decoding and linguistic comprehension through sentence reading. Two that have been investigated are Multiple Oral Re-reading (MOR; Moyer, 1979) and Oral Reading for Language in Aphasia (ORLA; Cherney, Merbitz, & Grip, 1986). The application of both of these treatment approaches has resulted in increased oral reading fluency (i.e., faster reading speeds and decreased errors; e.g., Beeson & Insalaco, 1998; Cherney, 2004; Cherney, 2010). One hypothesis that has been proposed for these reading gains is that the semantic and syntactic context in text provides top-down facilitation for lexical retrieval (Beeson & Insalaco, 1998; Tuomainen & Laine, 1991). Facilitated reading caused by semantic and contextual constraints has also been referred to as a *context effect* (Tuomainen & Laine, 1991). Only a few studies have systematically investigated the context effect and have had little agreement on the underlying mechanisms responsible for it.

Investigators studying reading in aphasia have typically used oral reading accuracy as a measure of reading performance. This measure is informative as it allows for characterization of reading profiles in people with aphasia (PWA). However, oral speech may or may not be spared in PWA. A more direct approach to investigating reading behaviour may be in the eyes themselves through the use of eye-tracking methodology. When we read, our eyes capture visual (i.e., letter and word) information which is then processed by language networks in the brain. Eye-tracking methodology tracks the movement of the eyes during specific tasks such as reading. It has been demonstrated in healthy readers that patterns of eye-movements reflect cognitive processes, including attention and comprehension, during reading (Rayner, 1998). Therefore, examining eye-movements in individuals with reading impairments may be a useful tool for discerning the nature of the impairment and could potentially offer insights for remediation of reading disorders.

Eye-tracking methodology is a powerful tool to investigate how a language impairment affects reading behaviour in people whose oral language may be impaired. It has only been recently applied in studies with PWA, and has been used to study comprehension of auditory information (Dickey, Choy, & Thompson, 2007; Meyer, Mack, & Thompson, 2012) as well as cognitive processing (Heuer & Hallowell, 2015; Ivanova & Hallowell, 2012). Studies involving reading behaviour are relatively few (e.g., Schattka, Radach, & Huber, 2010; Huck, Thompson, Cruice, & Marshall, 2017a). Eye-tracking methodology can provide a potentially beneficial means of studying readers with aphasia, because moment-to-moment processes during reading can be discerned, without the need for an overt response.

The purpose of this thesis was to investigate reading behaviour in PWA by examining their eye movements. Reading is a complex behaviour. At the word level, it involves the decoding of visual information, accessing phonological, lexical and semantic information. When reading sentences, this becomes more complicated. Readers integrate syntactic information across the sentence, to derive a holistic meaning. Properties of the word ("bottom up factors" including word length, word frequency, and word class) can affect how we process information (Rayner, 1998). Likewise, properties of the sentence ("top-down factors" including grammatical complexity, word order, and context) can similarly affect reading behaviour (Rayner, 1998). Generally, when text gets more difficult, reading is slower and more effortful. In contrast, reading can be facilitated by making the text simpler and more predictable.

This thesis explored the impact of language impairment on reading behaviour through answering two related research questions:

1. How do eye-movements of PWA reflect their reading processes? In a systematic review of reading studies involving eye-tracking, I summarize how a central language impairment affects the reading of single words and sentences relative to healthy readers.
2. What can eye-tracking measures tell us about the reading behaviour in PWA when reading is facilitated by word frequency (a bottom-up factor) and sentence predictability (a top-down factor)? This question is explored through a retroanalysis of a conference paper by Kim and Bolger (2012). These data were re-analyzed using linear mixed effects modeling to investigate frequency and contextual effects on word reading between PWA and demographically matched controls.

In Chapter 2, I provide a general background to acquired alexia, text treatments, and eye-movements during reading. In Chapter 3, I report the results of a systematic review of eye-tracking reading studies involving PWA. In Chapter 4, I investigate how bottom-up and top-down factors affect reading in PWA through a retroanalysis of a conference paper by Kim and Bolger (2012). Finally, a general discussion of eye-movements in PWA is provided in Chapter 5.

Chapter 2: Background

In an early paper, Webb and Love (1983) described reading difficulty to be a common occurrence in PWA. More recently, Brookshire et al. (2014) reported in a convenience sample of PWA (n=99), 68% had impaired reading abilities. A similar result was found by Leff and Starrfelt (2013) who found in sample of 212 stroke patients with aphasia that approximately 67% had reading difficulties. These studies outline that aphasia and reading difficulties overlap considerably. The reasons for this may be explained by how the underlying language impairment potentially affects multiple modalities. The central linguistic processes of phonology, semantics, and orthography contribute to language processing across all language modalities, including oral language, auditory comprehension, reading, and writing (Beeson, Rising, & Rapscak, 2011; Harm & Seidenberg, 1999). When one of these components becomes impaired, it may affect multiple domains of language. It should be of no surprise then that people with an oral language impairment may have some degree of difficulty in other language modalities, including reading.

Alexia Subtypes

Impairments to orthographical, semantic, and/or phonological processes can result in central alexias: surface alexia, phonological alexia, and deep alexia (Cherney, 2004). In cognitive models of reading, these alexia subtypes are described based on the types of errors produced. The dual route cascaded model of reading (Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001) is one such model in which reading profiles of acquired alexia subtypes are described. A modified version of this model is presented in Figure 2.1. According to this model, oral reading of single words occurs via one of two routes: a lexical route and a non-lexical route. In the lexical route, the visual word forms activate representations in the orthographic input lexicon, which then activates semantic and phonological representations, and finally the word is assembled phonologically where it can be spoken. In the non-lexical route, visual forms activate grapheme-phoneme representations, which then are converted to their corresponding phonological representations, and finally assembled where it can spoken. Damage to either of these routes results in an over-reliance on the intact reading route, arising in specific errors that characterize alexia subtypes.

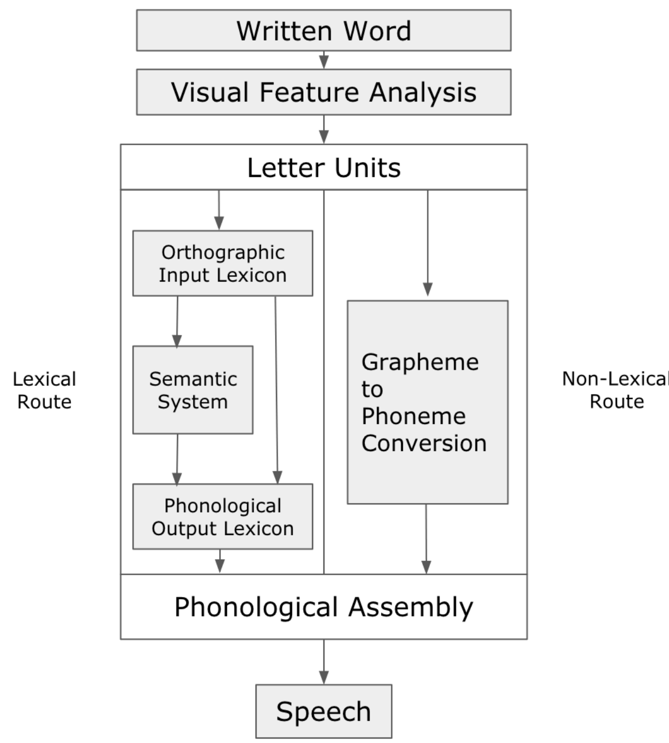


Figure 2.1: A dual route cascaded model of reading (Coltheart et al., 2001) modified by Riley and Kendall (2017).

Individuals with *surface alexia* have an impairment in accessing orthographic representations, and tend to produce regularization errors due to an over-reliance on the non-lexical route. For example, a reader may pronounce an irregular word such as ‘choir’ as ‘chore’ (/tʃɔr/), because he/she attempts to read the word through regular grapheme-phoneme conversion rules.

Individuals with *phonological alexia* have an impairment in accessing phonological representations, and tend to produce errors when reading nonwords. For example, a nonword like “blaf” may be read as a real word like “black”. This is explained through an over-reliance on the lexical route, where the reader attempts to read the nonword utilizing their orthographic lexicon.

Deep alexia represents broad impairments to semantic, phonological and orthographic processing resulting in visual errors (e.g. saying “park” when the word was “dark”), semantic errors (e.g. saying “dog” when the word was “cat”), impaired reading of nonwords, impaired reading of words with low imageability (e.g. “justice”), and difficulty in reading functor words (e.g., “for”).

Beeson and colleagues (Beeson, Rising, & Rapcsak, 2011) have described how aphasia and alexia subtypes may overlap because of underlying linguistic deficits caused by the lesion site. There is no one-to-one relationship between aphasia and alexia type; however, a few generalizations can be made. Phonological information is represented by neural regions in the perisylvian zone, while semantic and orthographic regions are contained within the extrasylvian zones. Aphasia syndromes associated with damage to perisylvian zone, such as Broca’s aphasia, may co-occur with phonological / deep alexia. Likewise, aphasia syndromes associated with the extrasylvian zones, such as anomic or transcortical sensory aphasia, may co-occur with surface alexia. The aphasia and alexia profiles are affected by the lesion site and the extent of the damage. This is rarely isolated to a single region. What is most likely to occur in PWA is a *mixed alexia* profile (Leff & Starrfelt, 2013), reflecting multiple levels of impairment within language processing. Indeed, in an analysis of 64 PWA with acquired alexia, Leff & Starrfelt (2013) argued that 78% of cases were mixed profiles, and only about 22% were observed to have "pure" profiles. This raises the importance of understanding how a central linguistic impairment overall affects reading behaviour, because clients will likely present with symptoms of mixed alexia.

Text Reading in Acquired Alexia

The ultimate goal of reading therapy is for the individual to be able to read functionally during their activities of daily living. Often, the long term goal is to regain the ability to read connected text quickly and accurately. However, traditional treatment approaches have focused on remediating single-word oral reading (Cherney, 2004). These treatments often target the level of the breakdown, which can be at the phonological, semantic, and/or orthographic level. Single-word reading is a foundational skill to remediate, but it may not be informative of a patient's reading performance when words are placed into context.

Two text-reading treatment protocols that have been developed are Multiple Oral Re-reading (MOR; Moyer, 1979) and Oral Reading of Language for Aphasia (ORLA; Cherney, Merbitz, & Grip, 1986). Both protocols involve the repeated oral reading of text. These protocols have been applied to a range of acquired alexia profiles. Indeed, efficacy for MOR has been demonstrated through an increase in text reading rates for individuals with pure alexia (Moyer, 1979; Tuomainen & Laine, 1991), phonological alexia (Beeson & Insalaco, 1998; Cherney, 2004), deep alexia (Kim & Russo, 2010), and mixed alexia (Kim, Rising, Rapcsak, & Beeson, 2015). Similarly, ORLA has also been used as a treatment approach for deep alexia (Cherney, 2004), mixed-alexias along a range of aphasia severities (Cherney, Merbitz, & Grip, 1986; Cherney, 2010), and apraxia of speech (Cherney, 1995). Individuals treated using MOR and ORLA have demonstrated generalization to reading novel material, including untrained reading passages (Beeson & Insalaco, 1998) and even text on the TV and newspapers (Tuomainen & Laine, 1991). Still, the underlying mechanisms that caused the rehabilitative effect remain unclear. It is hypothesized that top-down processes facilitate reading by using knowledge of the surrounding semantic and syntactic context to constrain lexical selection (Beeson & Insalaco, 1998; Tuomainen & Laine 1991). This improved reading performance explained by contextual and semantic constraints will be referred to as a *context effect* (Tuomainen & Laine 1991).

The hypothesis that PWA integrate syntactic and contextual information during text reading is a reasonable one. Constraint based approaches argue that language occurs within specific contexts, meaning that word meanings and their surrounding syntax are related to each other (Seidenberg & MacDonald, 1999). Early studies have demonstrated that healthy readers utilize lexical information from verb meaning to create biases of upcoming syntactic structures (i.e., lexical bias) (Trueswell & Tanenhaus, 1994; Trueswell, Tanenhaus, & Kello, 1993). Later studies by Gahl and colleagues have suggested that PWA, similar to healthy readers, are able utilize lexical bias during sentence processing (Gahl, 2002; Gahl et al., 2003).

Experimental studies investigating the mechanisms of the context effect in PWA are relatively few and each have hypothesized a different underlying mechanism. Silverberg and colleagues (Silverberg, Vigliocco, Insalaco, & Garrett, 1998) proposed a production based hypothesis using Garrett's (1982) model of sentence production. In this model, comprehension and production processes interact during sentence reading. The word retrieval of closed class words (i.e., function words) are facilitated in sentence production through syntactic constraints. In their study, the oral reading of single words and sentences was compared in 3 people with deep dyslexia. Each participant demonstrated relatively intact sentence production as observed in language samples. Sentence comprehension was confirmed through standardized testing (i.e., Psycholinguistic Assessments of Language Processing in Aphasia, Philadelphia Comprehension Battery for Aphasia). Their results revealed that all three participants read closed class words significantly more accurately in sentences than in single words. Benefits were also seen in the reading of open class words (i.e., content words). One participant had a significant increase in accuracy in reading open class words in sentences, while the other two participants saw numerically greater improvements in accuracy. This evidence supported their hypothesis that context effects may be the result of sentence production skills facilitating word retrieval, especially that of closed class words.

Mitchum, Haendiges, and Berndt (2005) tested Silverberg's production based hypothesis in 5 people with chronic, fluent aphasia. Sentence production ability was determined through the telling of a common fairy tale (i.e., Cinderella story) and was scored using the Quantitative Production Analysis. This assessment revealed that 4 out of 5 PWA had relatively intact sentence production (<2 SD from norms), while 1 PWA was relatively more impaired (>2 SD from norms). PWA read target words (nouns, verbs, adjectives, and function words) as single words and in sentences. In their results, only 2 of the 4 PWA with relatively intact production read target words in sentences more accurately than single words. This did not support the production based hypothesis because relatively intact sentence production did not seem to contribute to the context effect.

To identify what mediated the context effect, Mitchum and colleagues analyzed whether target word accuracy was affected by imageability, semantic build-up, local context, and grammatical class. Imageability did not impact the accuracy of the sentence reading for all participants. The effects of semantic build-up and local context are based on assumptions that context provides constraint on the target word. Semantic build-up is the accumulation of informational cues over the course of the entire sentence, while local context refers to the informational cues that come from the word immediately preceding the target word. Neither semantic build-up nor local context were found to mediate target word accuracy in sentence reading. What differentiated the 2 participants with the observed context effect from the rest of the group was their response pattern to grammatical class. Within their reading errors, their responses were congruent with the grammatical class of the target word (e.g. if the word was the noun "car", an erroneous response that would have the correct grammatical category would be "train"). Mitchum and colleagues posit that PWA utilize informational cues in sentences to determine grammatical class to assist word retrieval. Unlike the theory proposed by Silverberg et al. (1998), this hypothesis did not rely on intact sentence production. These studies illustrate that there is little consensus about the underlying mechanisms of the context effect. Therefore it is necessary to further explore and identify the possible contributing factors.

In summary, PWA have been observed to improve their reading abilities following text based treatments, and these results have even generalized to new texts and other environments. Studies investigating the contextual effect are sparse with results that are unclear on the underlying mechanisms. It is suspected that elements within text facilitate the reading of words, however, there is no agreement on specifically what these elements are.

Eye-movement Characteristics in Reading

Eye-tracking methodology has been used to describe reading in healthy people (Rayner, 1998). This method uses eye-movements to infer language processes without the need for an overt response. Expressive language is not always spared in PWA, making eye-tracking methodology a potentially useful approach. This section will now go over the theory linking eye-movements to language processing and how these eye-movements are measured.

Video-based eye-tracking. Modern eye-tracking devices track eye-movements through video-based methods. A common method is to send infrared light into the eyes of the participant while an infrared camera captures reflections back from the center of the pupil and the cornea. Patterns of these reflections are used to track numerous eye movements, including: acceleration, distance traveled, location of the landing site, and the time spent on the movement.

Eye-movements and language processing. Eye-tracking technology has been of interest to researchers because of the link that has been proposed between eye-movements and language processing (Just & Carpenter, 1980; Rayner, 1998). Just and Carpenter (1980) postulated two assumptions that described the relationship between eye-gaze and reading comprehension: the ***immediacy assumption*** and the ***eye-mind assumption***. The immediacy assumption suggests that language processing begins immediately upon encountering a word and occurs until information is accessed at the semantic level. Just and Carpenter referred to language processing as the mechanisms involved in decoding a word, assigning meaning, and interpreting it with respect to its context. The second assumption posits that the "eye" and "mind" are coupled, meaning that both of their functions are inextricably linked. That is, what is being fixated upon is what is being processed in the mind. This implicates that measures of eye gaze provide moment-to-moment information of the language processes. Current reading models describe the relationship of eye-movements and language processing to be more complicated (e.g., Reichle, Pollatsek, & Rayner, 2006). Words within the parafoveal view can be preprocessed (Schotter, Angele, & Rayner, 2012) and the processing of words can spillover into the next fixation (Rayner, 2009). Regardless, Just and Carpenter's basic assumptions remain foundational to using eye-tracking methodology in reading studies.

Foveal and parafoveal views. Eye movements specific to reading have largely been characterized by studies conducted over the past several decades by Rayner and colleagues. The following description of eye movements is summarized from Rayner's (1998; 2009) review of reading research. The eye is made up of foveal, parafoveal, and peripheral views. The fovea is the center of fixation. It has a visual angle of 2 degrees and has the most visual acuity. Surrounding it to either side is the parafoveal view and has less visual acuity. This extends to about 5 degrees of the visual angle from the point of fixation. Beyond the parafoveal view is the peripheral view which has even poorer visual acuity. As the eye perceives visual information with differing acuity, a reader must bring information to the fovea where it can be seen clearly and processed efficiently for linguistic information. The parafoveal view is able to provide preview information of upcoming words through collecting letter length, orthographical, and phonological information (Schotter et al. 2012).

The perceptual span. The functional area within vision, known as the perceptual span, covers the foveal and parafoveal views. It is asymmetrical and affected by the language of the reader. In English and other alphabetical languages, readers can perceive 3-4 letters to the left of fixation and 14-15 letters to the right (Rayner, 1998). In English readers, this spans roughly three words: the word fixated in the fovea and two words to the right in the parafovea (Rayner, Castelhano, & Yang, 2009; Schotter et al., 2012). The perceptual span is dynamic and is affected by a reader's foveal processing. The more effort a reader places on processing in the fovea, the less information that is processed in the parafovea, resulting in a smaller perceptual span.

Saccades and fixations. Reading occurs as a series of *saccades* and *fixations* in the direction of the text to bring new information into the perceptual span. Average durations of these measures have been well-established in healthy readers (Rayner, 1998; Rayner, 2009). Saccades are the eye-movements that bring new information into fovea, typically moving at the length of 7-9 letters. Not all eye-movements in reading are forward moving; about 10-15% of saccades are regressions back into text. Fixations are the moments between saccades, when the eyes pause to take in information. Words are typically fixated once or twice and for quick durations around 225-250 ms. Initial fixations typically land slightly left of word center, which is known as the *preferred viewing location* (Rayner, 1998).

Variables affecting reading. Saccades and fixations can be affected by several word level and text level factors (Rayner, 1998; Rayner, 2009). Variables at the word level include frequency, length, and grammatical category. Higher frequency words are more likely to be skipped or have shorter fixation durations. Length effects have been demonstrated by Rayner and McConkie (1976). In their study, they showed shorter words of 2-3 letters are likely to be skipped while longer words of 8 letters or more have a greater chance of being fixated. The effects of grammatical category are evident in that function words are skipped more than content words, which may be partially attributed to the fact that most function words are also short words. Other factors, such as concreteness and age of acquisition have also been investigated and found to be predictors of fixation duration (Juhasz & Rayner, 2003).

Saccades and fixations can also be affected by text level factors. For instance, Rayner and colleagues (Rayner, Chace, Slattery, & Ashby, 2006) demonstrated how global text difficulty affects eye-movements. Within their study, participants read passages of varying topics and difficulty. A separate group subjectively rated these passages based on how easy or difficult the text was to understand. An increase in text difficulty was correlated with an increase in the number of fixations, average fixation duration, and total fixation duration. Variables within the text that were associated with an increase of difficulty were complex sentences (e.g., relative clauses) and inconsistencies in text (e.g., inconsistent anaphors).

The contextual environment created by a sentence can mediate reading. Contextual constraint (otherwise, known as the context effect) refers to the degree to which syntax and context narrows the plausibility of an upcoming word or structure (Federmeier & Kutas, 1999; Federmeier, Wlotko, De Ochoa-Dewald, & Kutas, 2009). As the reader proceeds through a sentence, he or she uses informational cues to create an expectation of what words and structures may appear. When these expectations are correct, reading is facilitated. In the following example, the sentence is constrained so that "ferry" is the likely response:

> "Getting himself and his car to work on the neighboring island was time consuming. Every morning he drove for a few minutes and then boarded the..." (Federmeier & Kutas, 1999: 269)

Similar to frequency effects, words with high contextual constraint are more likely to be skipped or have shorter fixation durations (Rayner, Slattery, Drieghe, & Liversedge, 2011).

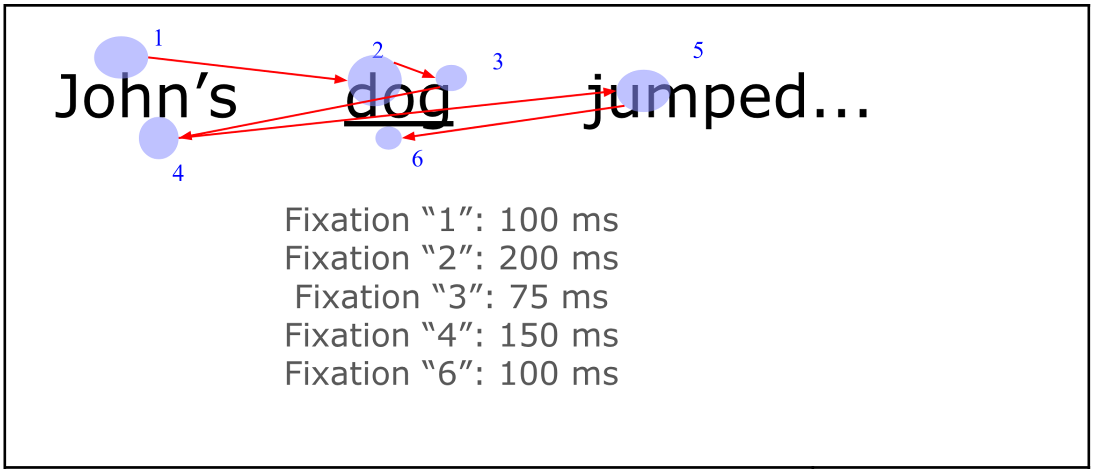
Eye-tracking measures of reading comprehension. When looking at reading behaviour, temporal and spatial measures are used to infer *lexical* or *language processing*. This can be captured through numerous measures, including fixation durations (i.e, the time spent fixating on a word), saccade amplitudes (i.e., how far the eyes travel), and landing positions (i.e., where the eyes land on a word) (Rayner 1998; Rayner 2009). When these measures are taken on single words, this is referred to as *lexical processing*. Interpretations of these same measures become more involved during sentence reading. Lexical processing of the fixated word still occurs, however, there is also the integration of syntax and context. This is now beyond the level of word access and involves higher level *language processing*. For sake of simplicity, I acknowledge that both are occurring during reading and will simply use the term *processing* to refer to lexical and language processing.

Numerous measures have been developed in an attempt to capture early and late stages of processing. Examples of these measures are provided in Figure 2.2. The early stage is considered the *first reading pass*. This is the first time the eyes encounter the word, before moving on to other words. The late stage is captured through more global measures that consider multiple reading passes, including regressions back to the word.

Two such measures that attempt to capture early processing in normal readers are *first fixation duration* and *first gaze duration* (Rayner, 1998). First fixation duration is simply the length of time of the first fixation on a word. First gaze duration, also known as the 'first reading pass', is the sum of all fixations on a word before moving forward (or backwards) in the passage. Both of these measures attempt to capture the initial processing of a word as soon as it is encountered, without including regressions. Though these measurements are obtained differently, Rayner (1998) argues that they provide similar results. A specific measure that looks at the difference between first fixation duration and first gaze duration is the *refixation time.* This includes the time spent fixating on a word after the first fixation in the first reading pass. Another measure of early processing is *mean fixation duration in the first gaze.* This represents the mean of all fixations during the first reading pass. Rayner (1998) has argued that averaging fixations will underestimate processing time.

A typical measure of later processing of a word is *total fixation duration* (also known as *total reading time*), which measures all fixation durations on a word, including regressions. Regressions, which occur about 10-15% of the time, are done to reread words and reflect difficulty of integrating information (Rayner, 2009). Thus, total fixation duration encompasses the first reading pass and, if they occur, subsequent reading passes. *Rereading time* is a specific measure that looks at the difference between the first gaze duration and the total fixation duration. Similar to total fixation duration, this measure may inform about reading difficulty.

The number of fixations and regressions are informative of reading comprehension. Particularly, they reflect the ease or difficulty of what is being read. More difficult words tend to have more fixations, while easier words have fewer fixations or can be skipped (Rayner, 1998). Regressions are thought to reflect a difficulty in the integration of what is being read (Rayner, 2009). As such, *total number of fixations* and *total number of regressions* are considered global measures of the reading difficulty.



Eye-tracking measure	**Total**
First Fixation Duration (fixation 2)	200 ms
Refixation duration (fixation 3)	75 ms
Gaze Duration (fixations 2 and 3)	275 ms
Mean Fixation Duration in the First Gaze (the mean of fixation 2 and 3)	137.5 ms
Total Fixation Duration (fixations 2, 3, and 6)	375 ms
Go Past Time (fixations 1, 2, 3, and 4)	525 ms
Total Number of Fixations (fixations 2, 3, and 6)	3 fixations
Total Number of Regressions (fixation 6)	1 regression
Rereading time (fixation 6)	100 ms

Figure 2.2: Examples of different eye-tracking measures describing the processing time on the target word "dog". Saccades are represented by arrows. Fixations and order are represented by the numbered circles, where the size of the circle displays the magnitude of the fixation duration.

The *amplitude of saccades* and *distribution of landing positions* also reflect reading processes. Generally, as reading gets more difficult, saccades are shorter in length (Rayner, 1998; Rayner, 2009). This in turn is expected to affect landing positions as shorter saccades can result in more landing positions near the beginning of the word. Saccades that are landing further away from the preferred viewing location are reflective of reading difficulty.

The measures described above have typically been used for single word reading. These measures can also be applied to sentence reading, where they can inform about lexical processing and integration of other sources of linguistic information. Sentence reading is naturally more complicated than single word reading and involves more complexity in describing eye-movements. A few measures have been created to capture this complexity, and reflect properties of the word itself, as well as properties of the surrounding context. For example, *skipping rates* (i.e., the rate at which a word is skipped when reading) and *probability of first fixation* (i.e., whether or not the upcoming word is fixated) are inversely related. As both these measures reflect word skipping, these are both sensitive to word frequency and sentence predictability effects (Rayner, Ashby, Pollatsek, & Reichle, 2004). *Probability of first-pass regression* measures if the next eye-movement immediately after entering a target word is a regression into previous parts of text (Ashby, Rayner, & Clifton, 2005). As this event occurs in the first pass of a word, it reflects early comprehension difficulty of a target word in relationship to previous parts of text. An additional measure for sentence reading is *go-past time*. This is a measure of time in an interest area before proceeding to the right of it (for English). It can be conceptualized like a checkpoint in a sentence that measures how long it took the reader's eyes to pass a certain point on the first reading pass. Go-past time has been sensitive to effects of sentence complexity, with more complex sentences (e.g., object cleft sentences) having longer go past times than simpler sentences (Staub, 2010).

No single measure informs about the entirety of how people read. Reading is a complex function of eye-movements with temporal and spatial dimensions. A single measure only describes a small part of how a reader reacted. As such, researchers often utilize a combination of measures to describe the effects they are seeking.

Eye-tracking methodology in aphasia studies. Eye-tracking methodology has seen limited use in the investigation of language processes in PWA. It may prove to be beneficial as it only requires eye gaze to inform about language processes, as described through Just and Carpenter's eye-mind and immediacy assumptions.

Eye-tracking paradigms have been used in PWA in the study of auditory comprehension (e.g., Dickey, Choy, & Thompson, 2007; Meyer, Mack, & Thompson, 2012; Thompson & Choy, 2009), language production (Thompson, Dickey, Cho, Lee, & Griffin, 2007), and cognitive processing (Heuer & Hallowell, 2015; Ivanova & Hallowell, 2012). Very few have focused on reading itself, despite the fact that reading difficulties often co-occur with aphasia. Eye-tracking has been used in PWA to describe overall reading profiles (Klingelhöfer & Conrad, 1984), single word reading (Schattka, Radach, & Huber, 2010), sentence reading (DeDe, 2017; Huck et al. 2017a; Huck et al. 201b; Knilans & DeDe, 2015), and reading treatment outcomes (Ablinger et al., 2014; Ablinger, Huber, & Radach, 2014; Ablinger & Radach, 2016; Kim & Lemke, 2016). Each of these studies focuses on different linguistic strengths and challenges that readers with aphasia encounter. A discussion about the similarities and differences between PWA and healthy readers may inform clinicians and researchers about the nature of aphasic reading difficulties.

Research Questions

Text reading is a complex task that can be seen as a combination of bottom-up and top-down factors providing cues to process the meaning of what is read. Bottom-up factors are cues at the single word level (e.g., orthographic, phonological, and semantic information). Top-down factors are the cues received from syntax and context. Studies of text treatments have argued that top-down factors that facilitate reading are an important component in the remediation of reading. However, it is unclear *what* top-down factors are supportive in their reading and *how* they utilize this information.

Eye-tracking studies have well described how healthy individuals process sentence and word level information when reading. This methodology is potentially useful in informing how individuals with acquired language impairments process what they are reading. However, the few studies of eye-tracking during reading in PWA are greatly variable in terms of their overall purposes and methods. Therefore, the purpose of this thesis is twofold: The first is to synthesize findings from existing studies of reading in PWA, and the second is to analyze an existing dataset designed to examine top-down and bottom-up factors in reading in PWA. Specifically, this thesis will address the following two research questions:

1. How are reading processes in PWA reflected in their eye-movements? This question is answered through a systematic review of studies that compare the eye-movements during reading of PWA to healthy readers.

2. What can eye-tracking measures tell us about the reading behaviour of PWA when reading is facilitated by word frequency (a bottom-up factor) and sentence predictability (a top-down factor)? This question was explored through a retroanalysis of Kim and Bolger (2012) using linear mixed effects modeling.

Chapter 3: Systematic Review

Reading difficulties commonly co-occur with spoken language difficulties in PWA (Brookshire et al., 2014; Leff & Starrfelt, 2013; Webb & Love, 1983). This may be the result of damage to the central language processes affecting multiple modalities. Reading may be prolonged and effortful, making activities of daily living difficult and recreational activities involving reading unenjoyable. Pure forms of central alexia resulting from damage to processes along the dual reading routes (Coltheart et al., 2001) have been described. However, individuals with PWA are more likely to present with a mixed alexia profile reflecting multiple levels of impairment (Leff & Starrfelt, 2013). Each PWA will present with a unique profile of symptoms and have their own strengths challenges during reading. It is important for clinicians to understand how the central language impairment is reflected in overall reading behaviour of PWA.

Reading behaviour of normal readers is well understood through the use of eye-tracking methodology (Rayner, 1998; Rayner, 2009). These methods have only recently been applied to PWA in single word (e.g., Schattka et al., 2010) and sentence reading (e.g., Huck et al., 2017b) studies. Each study forms a part of a puzzle that describe reading behaviour in PWA. They vary in the questions asked and the stimuli utilized. To build the foundation of understanding reading behaviour in PWA, a synthesis and discussion of these studies is necessary. To that end, a systematic review of eye-tracking reading studies that compared the eye-movements of PWA to healthy controls was carried out. Within this review, I discuss how a central language impairment is reflected in the overall reading profiles of PWA.

Methods

A systematic review of eye-tracking reading studies involving people with aphasia was conducted in April of 2017 and updated in November of 2018. With the assistance of a research librarian, the databases ComDisDome, CINAHL, MEDLINE, and PsycINFO were searched for peer-reviewed, English articles using the terms: (aphasi* or alexia or acquired dyslexia) AND reading AND (eye track* or eye movement*).

Inclusionary criteria for the search were: participants with a diagnosis of either an aphasia subtype or alexia subtype, control participants who were neurologically healthy, a research question into the reading behaviour of single-words or text, and the use of eye-tracking measures. Exclusionary criteria were participants with a reading impairment due to peripheral alexia (e.g., pure alexia), visual defect (e.g., hemianopsia), neglect, or dementia (e.g., Alzheimer's disease).

Screening. The screening process for inclusionary and exclusionary criteria was performed in two stages. First, studies were screened by reading titles and abstracts. Second, the remaining studies were further analyzed by reading through their contents. This process is illustrated in Figure 3.1 through a PRISMA flow diagram (Preferred Reporting Items for Systematic Reviews and Meta-Analyses: Moher, Liberati, Tetzlaff, Altman, and The PRISMA Group, 2009).

Prior to analysis, eleven studies were added by hand-searching the bibliographies of relevant articles. In total, 82 studies were identified by the search. After duplicates were removed, 53 studies were subjected to the initial screening process. This initial screening excluded 31 studies due to: investigations of visual defects, developmental dyslexia, or not involving reading. An in-depth analysis of the remaining studies excluded an additional 16 studies due to investigations of pure alexia, dementia, or problem solving. One article (Klingelhöfer & Conrad, 1983) met inclusionary and exclusionary criteria, but was too different from the other studies within this analysis because of its exploratory nature and its primary focus on qualitative description of eye-movements. For this reason, it was excluded from the analysis of eye-tracking measures. However, it was a foundational study describing the eye-movements of readers with aphasia and its approach can be appreciated. Characteristics of the study are described at the beginning of the results section. In total, 6 studies were included in this review.

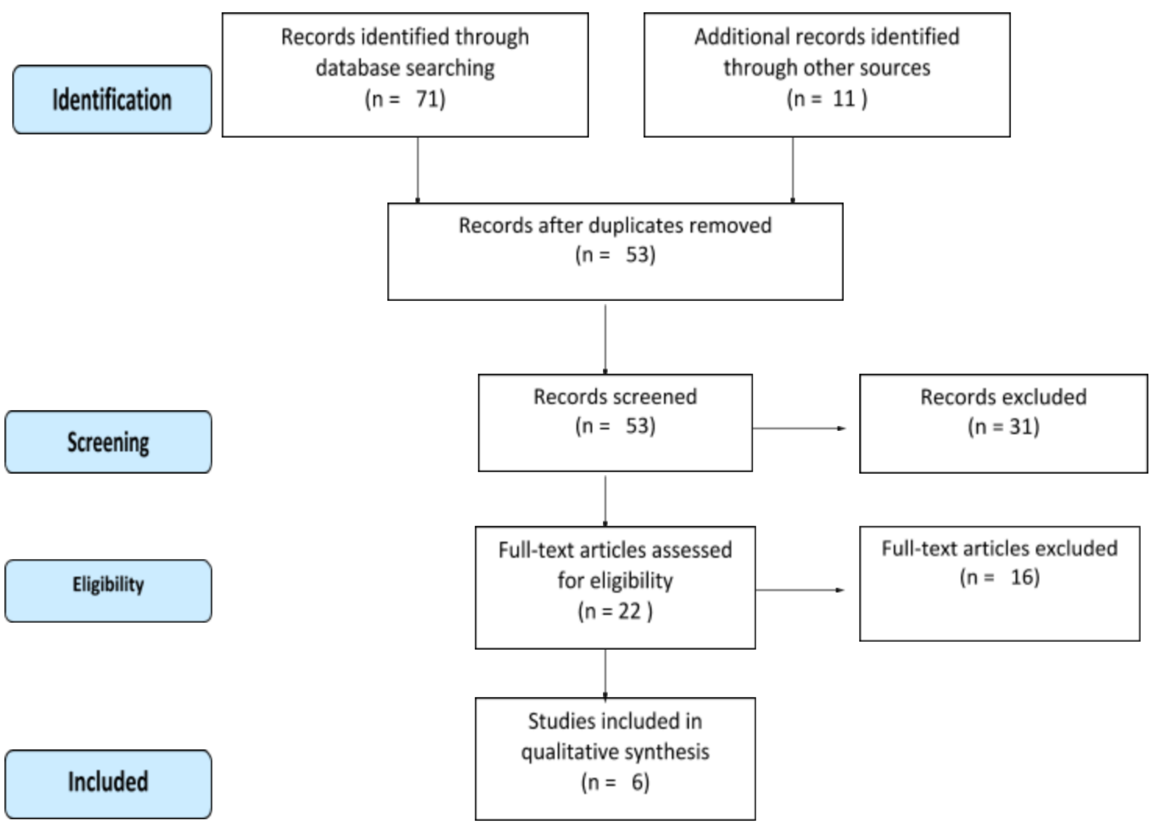


Figure 3.1: PRISMA flow diagram of the identification of reading studies with PWA employing eye-tracking methodology.

Data extraction and quality assessment. Demographic information/ participant characteristics (e.g., number of PWA and controls, aphasia type, age, time post-onset), study characteristics (e.g., stimuli, experimental variables, eye-tracking measures), and information relevant for quality assessment were extracted from the studies.

To assess the quality of each study, the STROBE Statement (STrengthening the Reporting of OBservational studies in Epidemiology: von Elm et al., 2007) checklist for case-control studies was used. The STROBE Statement is a checklist of 22 items throughout the entirety of the paper, including title, abstract, introduction, results, and discussion, that are recommended in the reporting of observational studies (Vandenbroucke et al., 2007). The inclusion of these items are to ensure transparent reporting of results.

Analysis and Results

This review included 6 studies, including 5 observational studies and 1 treatment study, with a total of 60 participants. All studies used video-based eye-tracking technology.

Quality assessments can be viewed in Appendix A. The review revealed several strengths in the reporting of these studies, namely: the inclusion of informative backgrounds, detailed methodologies, descriptions of participant characteristics, definitions of variables and outcome measures, and detailed interpretations of results. Weaknesses in reporting are that only a few studies included discussions about the generalizability of their sample and the limitations they encountered.

Potential methodological biases identified within these studies included the small sample sizes and varied characteristics of control participants. Typical of aphasia studies, the number of participants for each study was small, ranging from 6 to 17 PWA. Matching of groups was inconsistent and only considered one or two variables (e.g., age, education).

Table 3.1: Characteristics of studies included in systematic review.

Authors	**Research Question**	**Variables**	**Stimuli**
Treatment			
Ablinger, von Heyden, Vorstius, Halm, Huber, & Radach (2014)	How does a combined lexical and segmental reading treatment affect reading strategies in PWA?	Type of treatment (Lexical, segmental reading treatment)	Single words arranged as text
Observation			
Huck, Thompson, Cruice, & Marshall (2017a)	How do PWA respond to the effects of word frequency and sentence predictability in sentence reading?	Word frequency, sentence predictability	Sentences
Huck, Thompson, Cruice, & Marshall (2017b)	Do PWA use context cues when processing structurally ambiguous sentences?	Context bias, sentence ambiguity	Sentences
DeDe (2017)	How do PWA respond to lexical factors in sentence reading?	Word frequency, word length, word class	Sentences
Knilans & DeDe (2015)	How do PWA respond to structural frequency and sentence complexity?	Sentence type (object cleft, subject cleft)	Sentences
Schattka, Radach, & Huber (2010)	How are the effects of word length and word frequency reflected in the eye-movements of PWA?	Word frequency, word length	Single words arranged as text

Shown in Table 3.1, each study had a different experimental design and addressed different research questions about language or lexical processing. This review discusses the overall reading profile of PWA described by these studies and their response to linguistic variables.

It should be noted that there was an overlap in participants throughout studies. The same participants were used in the studies by DeDe (2017) and Knilans and DeDe (2015). Another overlap was apparent within the studies by Huck and colleagues (Huck et al., 2017a; Huck et al., 2017b), however, the exact number of participants who were common to both studies could not be identified. For this review, I report the total number of participants and acknowledge that these numbers do not necessarily reflect unique participants.

Participant characteristics. Characteristics of participants can be viewed in Table 3.2. PWA were either English or German speakers, and comprised varying fluent and non-fluent aphasia subtypes. Nearly all PWA were premorbidly right-handed, except for a single subject from Schattka et al. (2010) who was ambidextrous. Etiologies of aphasia were cerebrovascular accidents for nearly all participants, except for 3 with craniocerebral injuries. Four of six studies reported nearly all lesions in the left hemisphere; only 2 participants were reported with an etiology of right hemisphere damage (Huck et al., 2017a, Schattka et al., 2010). The lateralization of the etiology was not reported in the studies by Knilans and DeDe (2015) and DeDe (2017).

There was a considerably wide range in the time post-onset of aphasia. Most of the participants were in the chronic stage of recovery with a time post-onset of greater than one year. The earliest reported time post-onset was 4 months.

Table 3.2: Participant information from eye-tracking readings studies.

	Participants with Aphasia					**Control Participants**	
Authors	Number	Language Assessment	Subtype	Time Post-Onset (months)	Age Range (years)	Number	Matching Factor
Ablinger, von Heyden, Vorstius, Halm, Huber, & Radach (2014)	8	AAT	Anomic (1), Broca's (3), Wernicke's (2), Non-classifiable (2)	5-36 (mean= 14)	32-69 (mean= 52)	8	Age: yes Education: cannot be determined
Huck, Thompson, Cruice, & Marshall (2017a)	17[a]	WAB-R	Anomic (8), Anomia (3), Conduction (3), Broca's (1), Transcortical motor (2)	10-184 (mean= 66)	22-80 (mean= 59)	10	Age: Yes Education: Yes
Huck, Thompson, Cruice, & Marshall (2017b)	11[a]	WAB-R	Anomic (8), Broca's (1), Conduction (2)	13-198 (mean= 71)	41 to 71 (mean= 56)	11	Age: Yes Education: Yes
DeDe (2017)	9[b]	WAB-R / BDAE	Anomic (4), Broca's (4), Conduction (1)	At least 12 months	34-69 (mean= 55)	8	Age: Yes Education: No
Knilans & DeDe (2015)	9[b]	WAB-R / BDAE	Anomic (4), Broca's (4), Conduction (1)	At least 12 months	34-69 (mean= 55)	8	Age: Yes Education: No
Schattka, Radach, & Huber (2010)	6	AAT	Anomic (1), Broca's (3), Wernicke's (1), Residual (1)	4-53 (mean= 20)	35-61 (mean= 51)	11	Age: No Education: cannot be determined

Notes: a) There is a possible overlap with participants in the studies by Huck et al. (2017a) and (2017b). b) The studies by DeDe (2017) and Knilans & DeDe (2015) are comprised of the same participants. AAT= Aachen Aphasia Test; BDAE= Boston Diagnostic Aphasia Examination–Short Form; WAB-R = Western Aphasia Battery-Revised.

PWA had intact or corrected vision and no evidence of visual neglect or impairment as determined by visual screening (e.g., letter cancellation tasks); the visual screening methods used for Ablinger et al. (2014) and Schattka et al. (2010) were not reported.

Control participants were neurologically healthy adults. Cognitive screenings were performed using the *Mini-Mental State Examination*; the studies by Schattka et al. (2010) and Ablinger et al. (2014) did not report how healthy controls were screened. Most of the studies matched control participants with PWA by age. Only two studies additionally matched participants by education (Huck et al., 2017a; Huck et al., 2017b).

Stimuli. Two types of stimuli were used in these studies. The first kind of stimuli were single words arranged as text (Ablinger et al., 2014; Schattka et al., 2010). This was a single word reading task where unrelated nouns were arranged in a line to simulate text. This arrangement preserved the forward moving eye-movements that occur during sentence reading, while minimizing the cognitive load of higher order language processing that occurs with sentential contexts (Schattka et al., 2010). To ensure that each word was read, longer words (e.g., 6 letters or more) were used to avoid the skipping that typically occurs within reading.

The second kind of stimuli were coherent sentences (DeDe, 2017; Huck et al., 2017a; Huck et al., 2017b; Knilans, & DeDe, 2015) and more closely resembled what is encountered during real world reading. These stimuli were designed to look at how syntax and context affected the reading of target words (e.g., Huck et al., 2017a) or sections of the text (e.g., Knilans & DeDe, 2015). Unlike single word reading, there was opportunity for the eyes to perform more natural movements, such as skipping words and returning to previous parts of text.

Eye-movement measures were collected from interest areas in the stimuli. With the reading of single words arranged as text, the interest areas were the individual words. With sentence reading, the interest areas were either target words or sections of the text.

Eye-tracking measures. Numerous measures have been utilized within the studies reviewed to capture the initial and global processing of words and sentences (Table 3.3). The measures used across these studies were variable. However, using the categories described above, generalizations can be made as they all infer the processing of language.

The dependent measures in these studies have been categorized into the following: duration, amplitude, landing position, and number of events (numerosity). Duration measures count the length of time the eye has been fixated in the target area. Amplitude measures track the distance traveled by a saccade. Landing position measures indicate where a saccade has landed. Lastly, numerosity measures count the number of times an event has occurred through a total count or probability. These measures have been further divided into categories of early processing (i.e., occurring within the first reading pass) and late processing (i.e., including subsequent reading passes).

All studies reported a combination of early and late processing eye-movement measures. The most utilized measures of early processing were first fixation duration, first gaze duration, and probability of first pass regression. The most utilized measures of late processing were total fixation duration and rereading time. All studies reported measures of duration, while amplitudes and landing positions were only measured in Ablinger et al. (2014) and Schattka et al (2010). Five studies included a measure that counted the number of fixation events (e.g., probability of first pass regression) to further describe the reading behaviour.

Table 3.3: Eye-tracking measures used to compare PWA to controls and the number of studies reporting each measure.

	Measure	Number of PWA	Number of studies
DURATION			
	Early processing		
	First fixation duration	26	3[b, d, f]
	Mean fixation duration in the first gaze	6	1[f]
	Refixation time in the first gaze	6	1[f]
	First gaze duration	52	5[b, c, d, e, f]
	Go-Past Time	9	1[e]
	Late processing		
	Total fixation duration	51	5[a, c, d, e, f]
	Rereading time (all gazes)	24	3[b, e, f]
AMPLITUDE			
	Early processing		
	Amplitude of initial progressive saccade	6	1[f]
	Amplitude of all saccades in the first gaze	6	1[f]
	Late processing		
	Amplitude of all saccades in all gazes	6	1[f]
	Progressive Intra-word Saccades	6	1[a]
	Progressive Interword Saccades	6	1[a]
LANDING POSITION			
	Early processing		
	Landing position of initial progressive saccades	6	1[f]
	Landing position of all saccades in the first gaze	6	1[f]
	Late processing		
	Landing position of all saccades in all gazes	6	1[f]
NUMEROSITY			
	Early processing		
	Skipping Rate	9	1[b]
	Probability of First Pass Regression	28	2[c, d]
	Probability of First Fixation	17	1[c]
	Late processing		
	Total number of fixations in the first gaze	6	1[f]
	Total number of fixations in all gazes	14	2[a, f]
	Total number of gazes	6	1[f]

Note: a) Ablinger et al., 2014; b) DeDe, 2017; c) Huck et al., 2017a; d)Huck et al., 2017b; e) Knilans & DeDe, 2015; f) Schattka, Radach, & Huber, 2010

Klingelhöfer, J., & Conrad, B. (1984). This review begins with a short summary of Klingelhöfer and Conrad's (1984) seminal study investigating the reading profiles of PWA through eye-movements. The purpose of their study was to observe the type and rate of disturbance in the eye-movements of PWA. This study was different from the other studies within this analysis because of the exploratory nature of its experimental design and its use of global measures to form a qualitative description.

Klingelhöfer and Conrad compared a group of 21 PWA to a group of 40 healthy participants matched by age and education. Aphasia subtype was classified using the *Aachen Aphasia Test* and resulted in subdivision into those with Wernicke's aphasia (12), Broca's aphasia (5), and anomic aphasia (4). Using CT scans, lesions were identified as lateralized to the left hemisphere in 20 of the participants who were right-handed. The sole left-handed participant had a lesion lateralized to the right side. Reading ability of the control participants was categorized as "skilled" or "unskilled" based on a survey about their education, occupation, and reading habits.

Participants were tasked to perform silent and oral readings of two sets of stimuli while their eye-movements were tracked through the use of electrooculography. This method utilized electrodes placed near the eye to obtain evoked potentials, which can be used to identify eye position. The following global measures were obtained: total number of fixations, total number of regressions, and total fixation duration.

The stimuli presented in this study was unique compared to that of other studies included in this review. As a pilot study, Klingelhöfer and Conrad's focus was to describe the basic pattern of eye-movements in single word and text reading rather than provide an in-depth analysis. The first type of stimulus used was a 7-line passage of coherent text. This was comprised of 5 simple sentences with common words. The second stimulus type was a 7-line 'text' passage comprised of single words arranged as text. The first line began with 1 syllable words and increased in syllable length and morphological complexity in the proceeding lines. Line 4 contained a single word of 32 letters and 10 syllables. The final lines were 2 syllable nouns arranged as text. The exact contents of the stimuli could not be identified as examples were not provided within the study.

A qualitative analysis was performed by analyzing the pattern of saccades and fixation durations. These were used to make generalizations of the reading profiles of each Aphasia subgroup. Included were electroculargraphs of prototypical examples which illustrated their descriptions. It should be noted that they did not provide summary statistics of their data or perform inferential statistics. Skilled readers were observed to have a regular pattern of forward moving fixations and saccades with no regressions, which resembled a "staircase" pattern in the electroculargraphs; this was used as their reference *saccadic pattern*. Patients with anomic aphasia and unskilled readers in the control group were described to maintain a regular forward pattern of saccades and fixations, but with an increase in "irregularities" resulting from more fixations, regressions, and longer reading times.

The profile of eye-movements displayed by people with Wernicke's and Broca's aphasia was described as "pathologic". Some participants with Wernicke's aphasia were observed to maintain the saccadic pattern when reading coherent text. Their saccadic pattern was more irregular in reading single words arranged as text. Amplitudes of saccades were varied from word to word and had a tendency to continue decreasing in amplitude as they read the line (i.e., saccades getting smaller and smaller). This resulted in several mini-saccades to read a word and ultimately read the line. Even though they read with smaller saccades, longer saccades were possible as evident with return sweeps to the following line.

Patients with Broca's aphasia were described to be similar to unskilled readers in silent reading. However, oral reading seemed to exacerbate their reading difficulties, resulting in even more fixations, longer reading time, and more regressions when compared their silent reading.

Klingelhöfer and Conrad's (1984) quantitative measures showed that most of the PWA were similar to the control participants in the time spent reading. Where PWA differed from the control participants were in the number of eye-movements they made during the task. The patients with Wernicke's and Broca's aphasia had a greater number of fixations and regressions than the healthy participants. Some of the participants with anomic aphasia shared similarities with the unskilled readers in the number of fixations during reading.

The observations by Klingelhöfer and Conrad demonstrated that readers with aphasia had different eye-movement profiles than healthy readers and this may be related to their aphasia subtype. Since the time of their study, it would be more than twenty years until researchers would begin to investigate eye-movements in readers with aphasia again. Eye-tracking methodology and theories of language processing have advanced during this time. Current day studies involve a fine grained analysis of early and late processing in readers with aphasia. While earlier research has described how linguistic variables affect single word reading (Ablinger et al., 2014; Schattka et al., 2010), more recent studies have focused on how PWA respond to linguistic variables within sentences, including lexical, syntactic, and contextual information (e.g., DeDe, 2017; Knilans & DeDe, 2015; Huck et al., 2017a; Huck et al., 2017b).

Evidence from single words arranged as text. Two studies (Ablinger et al., 2014; Schattka et al., 2010) with 14 participants investigated the eye-movements when readers read single words arranged as text. Schattka et al. (2010) utilized a wide variety of temporal and spatial measures to assess lexical processing when readers read words varying in length and frequency. Within their study, 6 PWA were compared individually to a control group. What is reported in this section is their overall reading behaviour collapsed across length and frequency; the effects of these variables are discussed in the following sections. Ablinger and colleagues (Ablinger et al., 2014) performed a treatment study focusing on eye-movement specific training for whole-word and phonological reading for 8 PWA. While the details of this treatment are outside of the scope of this review, baseline temporal and spatial measures of their participants are reported on here. In reporting these results, the number of participants that differ significantly from HC, and the studies from which these results came from are reported in parentheses.

Duration measures. PWA demonstrated longer reading times than controls, particularly on measures of late processing. All PWA had significantly longer total fixation durations (n=14: Ablinger et al., 2014; Schattka, et al., 2010). Schattka et al. reported all PWA in their study had significantly longer rereading times on medium words and 5 of 6 had longer rereading times on long words. Taken together, these results suggest that not only do PWA spend more time reading a word, but much of this time is spent rereading it.

Measures of duration time in early processing have shown mixed results. Mean fixation duration in the first gaze did not differentiate between PWA and controls (0 participants, n=6) in Schattka et al.'s (2010) study, but between 3 to 5 PWA (n=6) were significantly different than controls on measures of first fixation duration, refixation time, and gaze duration. These results suggest that PWA demonstrate variability in early processing time. Although average fixation durations in their first reading pass may resemble that of normal readers, differences may arise on other measures of early processing, reflecting a difficulty in accessing word meaning during early processing.

Spatial measures. PWA had similar spatial eye-movements to healthy readers, which were shown in measures of saccade amplitudes and landing positions across early and late processing. Schattka et al. (2010) investigated amplitudes and landing positions in progressive saccades in first gaze, all saccades in the first gaze, and all saccades in all gazes, and reported no significant differences between any of their participants (n=6) and HC. Similarly, Ablinger et al. (2014) investigated the amplitudes of progressive interword and intraword saccades and also reported no significant differences between PWA and HC. Across all measures of amplitudes and landing positions, PWA showed greater variability but did not differentiate from healthy readers.

Numerosity measures. Counting the number of fixation events showed PWA made more eye-movements during reading of single words than healthy readers. These differences were not as evident in the first reading pass. The number of fixations in the first gaze did not differentiate between PWA and healthy readers (1 to 3 participants, n= 6: Schattka et al., 2010). Differences between groups emerged in later reading passes. PWA had a greater number of fixations in all gazes (n=14: Ablinger et al., 2014; Schattka et al., 2010). In Schattka et al's (2010) study, differences between PWA and healthy readers on the measure of total number of gazes was affected by word length. Five of the 6 PWA had significantly greater total number of gazes in medium words. In long words, this measure did not differentiate PWA from healthy readers. These results suggest that PWA produce a similar number of fixation events in the first reading pass as normal readers and use more eye-movements to reread words.

Word length and word frequency. Schattka et al. (2010) demonstrated length and frequency effects in several PWA (n=6) in duration measures representing early (e.g., refixation time) and late processing (e.g., total fixation duration). The earliest measure of lexical processing, first fixation duration, did not differ based on word length or word frequency.

Length was categorized into medium words (7-8 letters) and long words (11-12 letters). Not surprisingly, participants demonstrated significantly longer refixation time (3 participants) and total fixation duration (4 participants) for long words than medium words.

Frequency was categorized as low frequency words (average German CELEX rating of 0.62 occurences per million) and high frequency words (average German CELEX rating of 54 occurences per million). Similarly, the reading of low frequency words was significantly longer than high frequency words in measures of refixation time (3 participants) and total fixation duration (5 participants).

Summary of evidence. In these two studies utilizing single words arranged as lines of text, PWA had longer late processing measures and more total fixation events when reading. This appeared to be a result of more time spent rereading words, which was reflected in longer rereading times and longer total fixation durations. The eye-movements of PWA are similar to healthy readers in how far they move per saccade and where they land. In a study of lexical effects, PWA show the same pattern of response to word length and word frequency as control participants. These effects are likely to influence both early and late lexical processing.

Evidence from sentence reading. Four studies with 46 participants used sentence stimuli to investigate reading of PWA (DeDe, 2017; Huck et al., 2017a; Huck et al., 2017b; Knilans & DeDe, 2015). Across interest areas over the entire sentence, PWA had longer processing times than healthy readers in the first fixation (n=20; DeDe, 2017; Huck et al., 2017b), in the first gaze (n=37: DeDe, 2017; Huck et al., 2017a; Huck et al., 2017b), and in total fixation duration (n=37: DeDe, 2017; Huck et al., 2017a; Huck et al., 2017b). Additionally, PWA were more likely fixate on an upcoming word (n= 9: DeDe, 2017), more likely to regress to previous parts of text (n=28: Huck et al., 2017a; Huck et al., 2017b), and spent more time rereading (n=9: DeDe, 2017). Knilans and DeDe (2015) found similar results: PWA (n=9) had longer gaze durations, total fixation durations, and rereading times, but formal analyses of these differences were not performed due to statistical violations.

With the overall differences laid out, this section will now focus on how PWA process different parts of the sentence when compared to healthy readers. Sentence reading is a more complex task than single word reading. There are more linguistic variables to be considered and different interest areas that can be analyzed. For this reason, findings from various manipulations of stimuli and different interest areas will now be described in greater detail.

Lexical variables. Two studies investigated the effect of the lexical variables in sentence reading in PWA (DeDe, 2017; Huck et al. 2017a). These studies had a simple experimental design. Sentences were embedded with target words varying in word frequency, word length, and/or word class. The interest areas were the target words themselves. An example of stimuli from Huck et al. (2017a) is provided in Table 3.4.

Word frequency was categorized as high (140-190 occurrences per million) or low (1-15 occurrences per million) based on CELEX or Subtlex ratings (DeDe, 2017; Huck et al. 2017a). PWA (n=26: DeDe, 2017; Huck et al., 2017a) were responsive to the effect of word frequency. High frequency words were processed faster (i.e., shorter fixation durations) and were more likely to be skipped than low frequency words.

Word length was categorized using words that were 3, 4, 5, 6, 7, or 8 letters in length (DeDe, 2017). Similar to healthy readers, PWA (n=9: DeDe, 2017) showed a relatively intact response to word length. Short words were processed faster and were more likely to be skipped than long words.

Both effects were evident in duration measures of early and late processing. In the study by DeDe (2017), length and frequency effects were observed in first fixation durations, first gaze durations, rereading times, and skipping rates. Huck et al.'s (2017a) participants also demonstrated a frequency effect in first gaze duration and total fixation duration.

PWA (n=9: DeDe, 2017) were affected by word class when reading, but to a lesser degree than healthy readers. DeDe (2017) compared the skipping rates for short functional words ("the") against high frequency three letter nouns (e.g., "cat", "dog", "boy") and found an interaction between group and word class. The eye-movements of PWA were similar to healthy readers in that they skipped function words more than short content words. However, this skipping occurred at a lesser rate than in controls.

Sentence predictability. Huck et al. (2017a) investigated the effects of word frequency and sentence predictability in sentence reading in PWA (n=17). Their results of word frequency were discussed above. Sentence predictability, as discussed in Chapter 2, refers to the constraint of word meaning from its surrounding syntactic and semantic context. Words that are expected to occur within a specific context are more likely to be easier to process. A sample of the stimuli from Huck et al. (2017a) is provided in Table 3.4. The interest area was the target word itself.

Sentence predictability was categorized as predictable or unpredictable. This was normed through a cloze sentence task performed on a separate group of healthy participants. Predictable sentences had target words that were predicted about 84% of the time as determined by a clozed sentence task in a norming study. While unpredictable sentences had target words that were predicted less than 1% of the time. Target words were analyzed for first gaze duration, probability of first fixation, probability of first pass regression, and total fixation duration.

When compared to healthy readers, PWA had longer reading times and more difficulty reading the target words. This was indicated through longer first gaze durations, longer total fixation durations, and more regressions to previous parts of the sentence.

PWA were similar to healthy readers in their response to predictability. Target words in predictable sentences had shorter first gaze durations, shorter total fixation durations, less regressions, and were more likely to be skipped than target words in unpredictable sentences.

Table 3.4: Example stimuli of targets words of varying frequency (high frequency, low frequency) embedded in sentences of varying predictability (predictable, unpredictable) from Huck et al. (2017a).

Word Frequency	**Sentence Predictability**	**Sample Sentence**
High Frequency	Predictable	After the accident they rushed to the **hospital** to get the injury cleaned.
High Frequency	Unpredictable	The friends carry their tents to the **hospital** where they want to sleep.
Low Frequency	Predictable	The friends carry their tents to the **campsite** where they want to sleep.
Low Frequency	Unpredictable	After the accident they rushed to the **campsite** to get the injury cleaned.

Sentence complexity. Knilans and DeDe (2015) investigated how syntactic frequency and complexity affected the eye-movements in PWA (n=9). In their experiment, PWA read sentences with subject clefts and sentences with object clefts. Subject clefts are simpler in sentence structure. These sentences follow the canonical word order (i.e., in English: subject-verb-object) and the relationship between the verb phrase and its arguments are in close vicinity. Object clefts are more complex sentences with non-canonical word order (i.e., object-subject-verb). Moreover, the distance from the verb and its arguments is farther, making it more difficult for the reader to process the relationship between these items.

These sentence structures are also subject to the effect of frequency. The structure of subject clefts are more frequent than object clefts. As such, there is a bias to expecting the more frequent subject cleft over the less frequent object cleft. An example of the stimuli is provided in Table 3.5.

Table 3.5: Example stimuli of object and subject cleft sentences from Knilans and DeDe (2015).

	Sample Sentence
Subject cleft	It was the father that entertained the baby during the party last week.
Object cleft	It was the baby that the father entertained during the party last week.
Interest Area	[It was]= beginning of sentence [the baby]= first noun phrase [that]= relative clause **[the father]= second noun phrase** **[entertained]= verb phrase** [during…]= end of sentence

The interest areas observed were the first noun phrase ("the baby"), the relative clause ("that"), the second noun phrase ("the father"), the verb phrase ("entertained"), and the end of sentence ("during…"). Effects of *structural frequency* were expected to occur when viewing the second noun phrase ("the father" in the object cleft sentence; "the baby" in the subject cleft sentence). This was when the reader first encountered that the sentence structure may be different than what was expected. More expected structures should be processed quicker, while less expected structures should be processed slower. As such, readers are expected to spend more time reading the second noun phrase in the less frequent sentence structure (i.e., O-**S**-V in the object cleft sentence) than in more frequent structures (i.e., S-V-**O** in the subject cleft sentence).

Effects of *structural complexity* were expected to occur when reading the verb phrase ("entertained"). This was the point in the sentence where the reader established the relationship between the verb and its noun phrases. Readers are expected to have longer processing of the verb phrase in more complex sentences (i.e., object cleft sentences) than in simpler sentences (i.e., subject cleft sentences). Healthy controls were expected to respond to both effects of structural frequency and structural complexity (Staub, 2010).

Healthy readers performed as expected and were affected by both structural frequency and complexity. The effects of structural frequency were observed only in measures of late processing; HC demonstrated longer total fixation durations and rereading times when reading the second noun phrase in object cleft sentences than subject cleft sentences. The effects of structural complexity were observed in measures of early and late processing; HC demonstrated longer first gaze durations, go past times, and total fixation durations when reading the verb phrase of object cleft sentences than subject cleft sentences.

PWA showed similarities to HC in that they were responsive to both the effects of structural frequency and structural complexity. The effect of structural frequency was observed in PWA through longer go-past times of the second noun phrase when reading object cleft sentences. The structural complexity effect occurred in measures of late processing as observed by longer rereading times and total fixation durations of the verb phrase when reading object cleft sentences. The results of this experiment suggest the sensitivity to structural frequency and sentence complexity remains relatively intact in PWA, however, the effects of sentence complexity may occur in later reading passes than it does in healthy readers.

Context and sentence ambiguity. Huck et al. (2017b) investigated how PWA (n= 11) utilize context when interpreting ambiguous and non-ambiguous sentences. This was done was by manipulating a reader's expectation of the upcoming sentence structure. Lexical bias refers to the relationship between verb meaning and sentence structure (Gahl, 2002; Gahl et al., 2003). This relationship is evident in polysemous verbs. For example, the verb "recognize" can mean "to be fully aware or cognizant of". When it takes on this meaning, it most frequently takes on a subject complement as its argument. Recognize can also mean "detect with senses" where it frequently uses a direct object as its argument. Examples of these are provided in Table 3.6.

Table 3.6: Examples of direct object and subject complement sentence structures for the polysemous verb "recognize" from Huck et al. (2017b).

"recognize"		
Meaning	Sentence Structure	Sample Sentence
detect with senses	Direct object	At the library, I **recognized** the student sitting at the table.
to be fully aware or cognizant of	Sentence complement	At the library, I **recognized** (that) the student needed help finding a book.

Verb meanings are constructed within context. As such, contexts can be created to make a "sense" of a specific meaning and create a bias towards that meaning's most frequent sentence structure. Within their study, a polysemous verb was embedded into a target sentence. These target sentences were preceded by a context sentence that was constructed to bias a certain meaning and structure. An example of the stimuli is provided in Table 3.7. The first context sentence creates a direct object bias through the sense of "imagine" for the verb "project". The second context sentence creates the expectation of a subject complement through the sense of "project on a screen".

The target sentences all contained subject complement structures and were either ambiguous or non-ambiguous in their interpretation. The non-ambiguous sentences included the complementizer ("that") after the verb ("projected"), allowing a clearer interpretation that the noun phrase ("the film") is a subject complement. The ambiguous sentences excluded a complementizer following the verb, allowing the interpretation of the noun phrase to be temporarily ambiguous between a subject complement or a direct object complement.

Table 3.7: Example stimuli of structure biasing context sentences and (non)ambiguous sentences for the verb “projected” from Huck et al. (2017b).

	Sample Sentences	
Direct object bias: The journalist asked the filmmaker whether he expected the production would be a success**.** (Sense: imagine)		
	Ambiguous Sentence	He **projected** the film would be very popular with teenagers.
	Non-Ambiguous Sentence	He **projected** that the film would be very popular with teenagers.
Subject complement bias: At the meeting William wanted to show the video he made recently. (Sense: project on a screen)		
	Ambiguous Sentence	He **projected** the film would be very popular with nature lovers.
	Non-Ambiguous Sentence	He **projected** that the film would be very popular with nature lovers.
Interest areas	[projected] = verb region [that] = complementizer not analyzed [the film]= noun phrase region **[would be] = disambiguation region** [very popular with] = post disambiguation region	

The sentence was divided into several interest areas. The one in particular interest to the research question was the disambiguation region. When proceeding through the sentence, this was the first region where the reader was able to detect whether or not their expectation created by bias was correct. This region was analyzed for ambiguity and context effects. The *ambiguity effect* refers to how sentence ambiguity impacts the processing time of this region. This area was expected to be processed quicker in the non-ambiguous sentences than in the ambiguous sentences.

The *context effect* refers to how the context bias affects the processing of the disambiguation region. In the subject complement bias, the disambiguation region should be processed quicker as the reader’s expectation correctly matches the sentence type. In the direct object bias, the reader discovers that he/she has made a misanalysis on the expected sentence type, resulting in a longer processing time. Healthy readers have been shown to respond to both ambiguity effects (Trueswell, Tanenhaus, & Kello, 1993) and context bias (Hare, McRae, & Elman, 2003).

Sentences were analyzed for measures of early processing (first fixation duration, first gaze duration, probability of first pass regression) and late processing (total fixation duration). Group differences were evident throughout all measures. Overall, sentence reading of PWA was described as having more eye-movements than controls participants in all regions of the sentence. PWA had prolonged early processing (first fixation duration, first gaze duration), prolonged late processing (total fixation duration), and more regressions.

In the analysis of the disambiguation region, they found that PWA behaved similarly to healthy readers. An interaction of context and ambiguity was found in first gaze duration. This suggested that ambiguous sentences were more difficult to process for both groups in the first reading pass when the context provided the incorrect verb meaning-structure bias. In the measures of total fixation duration and probability of first pass regression, there was a main effect of context. Both groups processed the disambiguation region quicker and had fewer regressions when provided a context bias that was consistent with the actual sentence structure. These results indicated that PWA have a relatively intact ability to predict syntactic structures based on informational cues from context and syntax.

Summary of evidence. Studies of sentence reading show an overall difference between groups. PWA have longer reading times, evident in measures of early and late processing when considering reading across the whole sentence. In describing fixation events, PWA are less likely to skip words and more likely return to previous parts of text (i.e., regress) than controls. These studies investigated how linguistic effects impact the sentence reading of PWA. PWA, similarly to controls, were observed to utilize lexical, syntactic, and contextual cues to facilitate the processing of target words and syntax. Word length, word frequency, and context (both sentence predictability and context bias) were observed in early and late processing. This suggests that PWA are sensitive to linguistic and contextual effects in the first reading pass and these effects persist into later reading passes. As PWA have characteristically long late processing, later reading passes may be used to integrate this information.

The effect of structural complexity was observed only in late processing in the study by Knilans and DeDe (2015). In their study, they suggest that lexical information is collected during the first reading pass and syntactic information is collected in later reading passes. Indeed, this hypothesis was supported by the studies analyzing lexical variables. DeDe (2017) and Huck et al. (2017a) observed lexical effects to emerge during the first reading pass. Further studies examining syntactic information processed in later reading passes would help elucidate this hypothesis further.

Discussion

A systematic review of eye-tracking reading studies identified 6 studies that compared the eye-movements of PWA to healthy readers. Quality assessment of these studies revealed several reporting strengths, especially in the description of background, participant characteristics, and variables, and the interpretation of results. Weaknesses in reporting were the lack of discussions about the generalization of the studies or the limitations encountered. Potential methodological biases were identified in the small sample size of PWA and the matching of control participants. These biases may impact the reporting of individual results.

Numerous eye-tracking measures were utilized throughout these studies to describe processing time. A commonality that was shared was that each study utilized duration measures of early and late processing. Some studies counted the number of fixation events which was used to detail specific reading behaviours such as regressions and word skipping.

Relative to healthy readers, PWA had longer durations for both early and late processing measures, were less likely to skip words, were more likely to regress to previous parts of text, and spent more time rereading. Difficulties were evident at the word level and persisted into sentence reading. Overall longer processing times reflect that PWA are less efficient in integrating linguistic information.

Numerous linguistic effects were shown to mediate reading in PWA. PWA, similar to healthy readers, were able to utilize lexical, syntactic, and contextual cues throughout a sentence. Constraint based approaches describe language processing as an interaction of multiple and concurrent informational cues from linguistic information and language knowledge (Seidenberg & MacDonald, 1999). Following this approach, PWA may be using their language knowledge to integrate informational cues across the sentence to assist weakened reading routes.

Prolonged reading times were largely due to excessive rereading. Rereading may be a strategy to cope with the less efficient processing in the first reading pass. As evident in the studies investigating linguistic effects, late processing in PWA is mediated by linguistic variables such as context and sentence complexity. It is possible PWA are utilizing later reading passes to integrate informational cues in the sentence.

The results of this systematic review are consistent with Klingelhöfer and Conrad's (1984) findings that PWA make more fixations overall when reading. Their subgroup analysis revealed a pattern of microsaccades in Wernicke's aphasia and longer oral reading times in Broca's aphasia. The results of this systematic review did not reveal that saccadic patterns in PWA differed from healthy readers (Schattka et al., 2010; Ablinger et al., 2014). Moreover, longer processing was evident for PWA as a group, not just for a single subtype. It should be noted however, that unlike Klingelhöfer and Conrad, aphasia subtypes were not provided throughout all studies, so similar analysis could not be performed.

Chapter 4: Retroanalysis

Traditional treatments of acquired alexia have focused on remediating at the level of impairment, which is typically at the word level (Cherney, 2004). These have resulted in the improvement of single word reading, but generalization to text reading has been limited. Two text treatments developed are Multiple Oral Re-reading (MOR; Moyer, 1979) and Oral Reading of Language for Aphasia (ORLA; Cherney, Merbitz, & Grip, 1986). These have shown to increase oral accuracy and reading rate across a variety of aphasia subtypes (e.g., Cherney, Merbitz, & Grip, 1986; Cherney, 2010) and alexia subtypes (e.g., Beeson & Insalaco, 1998; Cherney, 2004). Researchers have proposed that top-down processing facilitates text reading through the use of syntax and contextual cues (Beeson & Insalaco, 1998; Tuomainen & Laine, 1991). Studies investigating the context effect have little consensus on the underlying mechanisms (Silverberg et al., 1998; Mitchum et al., 2005), which may be due to their use of oral reading as an outcome measure.

Just and Carpenter's (1980) eye-mind assumption and immediacy assumption posit that language processing during reading is reflected in eye-movements. Indeed, eye-tracking methodology has been successfully used to describe the reading processes in healthy readers (Rayner, 1998). Subtleties in eye-movements have shown that reading is mediated by properties of the word ("bottom up factors" including word length, word frequency, and word class) and properties of the sentence ("top-down factors" including grammatical complexity, word order, and context) (Rayner, 1998). Recent reading studies have shown that PWA are similar to healthy readers in their response to these factors, including lexical, syntactic, and contextual variables (DeDe, 2017; Huck et al., 2017a; Huck et al., 2017b; Knilans & DeDe, 2015). Eye-tracking methodology may be an advantageous tool in elucidating what factors of text facilitate reading in PWA.

In this section, I present the results of a retroanalysis of data from a conference paper by Kim and Bolger (2012). The goal of performing this analysis was to examine how bottom-up and top-down factors are reflected in the eye-movements of PWA. In their study, Kim and Bolger compared eye-movement patterns of PWA to healthy control participants during sentence reading. Stimuli were manipulated to examine the effects of sentence predictability (a top-down factor) and word frequency (a bottom-up factor). In the original study, Kim and Bolger used factorial analyses (ANOVAs) to investigate group differences. PWA are a variable group and this traditional method of analysis may not accurately explain subtleties in their reading behaviour. Similarly, the context effect involves a combination of syntactic, lexical, and semantic cues which may not be controlled by traditional analyses. To attribute for the variability within participants and stimuli, I analyzed the data with linear mixed effects modeling (LMEM). LMEM allowed participants and stimuli to contribute to the dependent variable as random effects. The characteristics of the original study and retroanalysis are described below.

Kim & Bolger (2012)

Procedure. People with aphasia and/or acquired alexia and healthy controls completed a behavioural assessment battery and eye-tracking assessments over 2 to 4 sessions. Behavioural assessments comprised various cognitive domains including general language ability, general cognitive ability, single word reading, text reading production, and text reading comprehension.

Participants. Since its publication, more participants were added to the Kim and Bolger (2012) study. This retroanalysis utilized that updated data. Participants comprised 12 people with aphasia and/or alexia (five female and seven male) and 10 age-matched healthy controls. Table 4.1 presents participants' demographic information.

Etiology for 11 of the participants was a left hemisphere ischemic stroke; 1 participant suffered a left hemisphere hemorrhagic stroke after an arteriovenous malformation burst. Mean age of subjects was 56.7 years, ranging from 22 to 83 years old. Mean years of education was 11.8 years, ranging from 8 to 14 years. Mean Aphasia Quotient scores (WAB AQ) from the *Western Aphasia Battery-Revised* was 66.3 (out of 100), ranging from 21.3 to 94.7.

PWA were compared to a group of healthy control participants (HC), comprising 10 people with no history of neurological damage or reading impairment. Mean age of control participants was 56.9 years old, ranging from 40 to 73 years old. Mean years of education was 15.2 years, ranging from 12 to 18 years. Normal cognitive status of controls was confirmed with the Mini Mental State Exam.

The two groups were matched for age, $t(20) = -0.04$, $p = 0.97$. However, the control group had a higher average years of education than the group of subjects, $t(19) = -3.46$, $p < 0.01$.

All participants passed screenings for hearing, vision, and nonverbal problem solving. Hearing was screened bilaterally through a minimal pair discrimination task or pure tone air-conduction screening at 0.5, 1, 2 and 4k Hz. Vision was screened through the Rosenbaum pocket vision screener, finger perimetry, and letter cancellation. Hearing and vision for all participants was either normal or corrected.

Normal nonverbal problem solving abilities was confirmed by Raven's Coloured Progressive Matrices (Raven's CPM); all participants scored greater than the 5th percentile for their age/education (Smits, Smit, van den Heuvel, & Jonker, 1997).

Table 4.1: Demographic information of participants.

Participants with Aphasia and/or Acquired Alexia						
	Age	Education	Gender	Raven's CPM (out of 37)	WAB AQ (out of 100)	Aphasia Type
P1	61	9	M	24	60.4	Broca's
P2	42	12	M	20	83.2	Anomic
P3	56	14	M	35	76.7	Conduction
P4	67	11	F	17	44.6	Broca's
P5	52	14	F	14	21.3	Broca's
P6	67	8	M	25	65.9	Conduction
P7	83	13	F	26	73.3	Anomic
P8	41	12	F	34	53.3	Broca's
P9	66	8	M	19	79.5	Anomic
P10	68	12	F	30	71.4	Anomic
P11	22	14	M	34	71.7	Anomic
P12	55	14	M	34	94.7	Non-Aphasic
Average	**56.7**	**11.8**	**5F/7M**	**26.0**	**66.3**	
Control Participants						
	Age	Education	Gender	Raven's CPM (out of 37)	MMSE (out of 30)	
C1	54	16	M	34	28	
C2	52	14	F	36	30	
C3	58	12	F	37	30	
C4	59	18	M	33	28	
C5	66	18	F	33	30	
C6	75	16	M	34	28	
C7	40	12	M	35	30	
C8	43	18	M	33	26	
C9	49	14	F	29	28	
C10	73	14	F	34	28	
Average	**56.9**	**15.2**	**5F/5M**	**33.8**	**28.6**	

Reading behaviour assessment. Reading profiles were characterized with the *Arizona Battery for Reading and Spelling.* One participant was unable to complete the assessment due to severe apraxia of speech. PWA had significantly more errors than HC, $t(19) = -6.54$, $p < 0.0001$. Seven of the PWA were described with phonological alexia, characterized by a large lexicality effect (i.e., better word reading than non-word reading), and 4 PWA were described with mixed alexia, characterized by lexicality, frequency, and regularity effects. In comparing performances across groups, PWA had the most difficulty reading non-words while HC performed near ceiling on regular words, irregular words, and non-words.

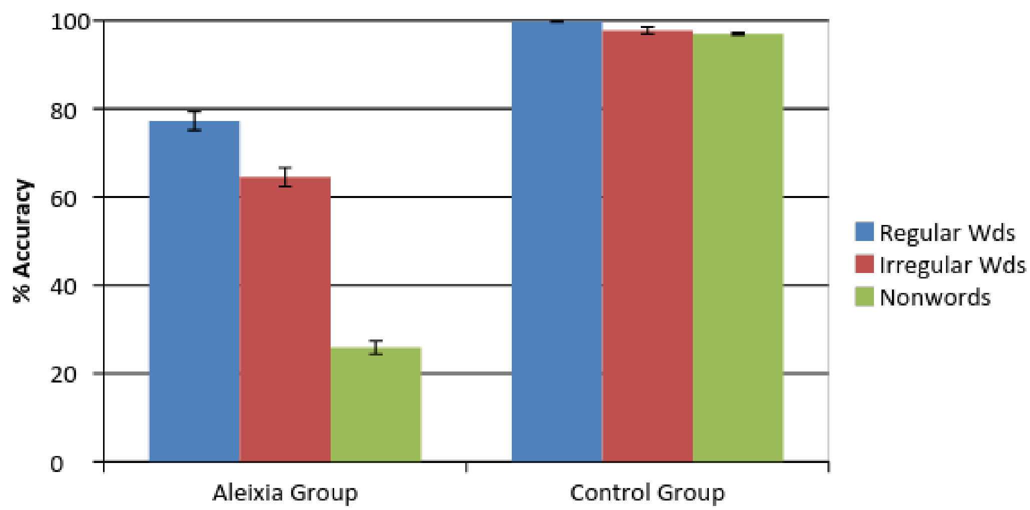


Figure 4.1: Single-word reading performance on *Arizona Battery for Reading and Spelling* for PWA and HC.

Text-reading fluency was assessed with the *Gray Oral Reading Test-4.* Two PWA were unable to be assessed due oral speech impairments. This test was comprised of 13 passages with increasing length and difficulty. PWA read between 2 to 8 passages (M= 5) while HC read all 13. Measures of average reading rate (words per minute), average reading errors (per 100 words), and reading comprehension scores (percent of multiple choice questions answered correctly) were collected from each participant. PWA differed significantly from HC on all measures: rate, t(18) = -11.71, $p= 0.000$, errors, t(18) = 3.43, $p< 0.01$, and comprehension, t(18) = -2.37, $p< 0.05$. Individual analysis was performed by converting into measures into z-scores. All PWA had significantly slower reading rate and more errors ($p < 0.001$). Only 3 PWA had significantly worse reading comprehension when compared to HC ($p < 0.05$).

Table 4.2: *Gray Oral Reading Test-4* performance and z-scores for PWA compared to HC group average

	# of passages read	Rate words/min	Errors per 100 wds	Comprehension % correct
P1	2	23.8 (-5.4)	23.6 (53.2)	50 (-3.3)*
P2	8	66.6 (-3.9)	14.4 (31.8)	75 (-.8)
P3	5	27.2 (-5.3)	11.5 (25.0)	80 (-.2)
P4	2	24.5 (-5.4)	63.8 (147.4)	70 (-1.3)
P6	6	73.1 (-3.7)	6.0 (12.1)	56.7 (-2.6)*
P7	3	22.5 (-5.5)	36.6 (83.7)	66.7 (-1.6)
P8	4	19.5 (-5.6)	18.3 (40.8)	85 (.3)
P9	6	59 (-4.2)	8.2 (17.3)	56.7 (-2.6)*
P11	4	38.6 (-4.9)	18.3 (40.9)	75 (-.7)
P12	8	72.6 (-3.7)	3.3 (5.9)	87.5 (.6)
Controls	10	175.5	0.8	82.2

Note: *z*-scores are in parentheses; all participants demonstrated significantly impaired reading rate and more errors relative to controls ($p < 0.001$). Participants with significantly impaired reading comprehension relative to controls ($p < 0.05$) indicated by *. P5 and P10 did not complete text-reading assessment.

Methods. Eye-tracking data was collected from the participant's left-eye using the SR Research EyeLink 1000 with desktop mounted camera. Participants were seated with their head on a chinrest that was positioned 60 cm away from an LCD computer screen. Stimuli was presented in 18 pt. Courier font in white against a black background. Calibration and validation procedures preceded each trial. Participants initiated the trials by gazing at a fixation cross on the left side of the screen and ended the trials by a button push.

Stimuli. Stimuli was comprised of 20 high frequency and 20 low frequency target nouns that were embedded within pairs of sentences of varying predictability. High frequency words had an average frequency of 215 occurences per million words based on the Kucera-Francis values. Low frequency words had an average frequency of 10 occurrences per million words. Target nouns were never embedded at the beginning or the end of the sentence because these regions are known to impact eye-tracking validity (Rayner, 1998). Table 4. presents examples of the stimuli.

Table 4.3: Example of sentence stimuli from Kim and Bolger (2012).

		Frequency	
		High	Low
Context	High	The couple was married by the minister of the church.	The boy kissed his mother on the cheek and said goodbye.
	Low	The woman was praised by the minister for her actions.	The mosquito bit the boy on his cheek and flew away.

Sentence predictability was normed by 38 undergraduate students in a cloze sentence task. The targets of sentences were produced 74% of the time for sentences with high predictability and 4% of the time for sentences with low predictability. These values are consistent with the literature descriptions of high/low sentence predictability (Balota, Pollatsek, & Rayner, 1985).

Measures and analyses. Dependent variables used on target words were indicators of late processing in PWA (Johnson & Rayner, 2007): *total fixation duration* on the target word, *total number of fixations* on the target word, and *total number of regressions* throughout the sentence. Kim and Bolger's original analysis was performed using a repeated measures ANCOVA with context as the within-subjects factor and education as a covariate.

Results. Their results revealed a significant main effect of group and a significant interaction of group and context condition across both measures of total fixation duration and total number of fixations. PWA had more fixations and longer reading times of the target word than HC. In addition, PWA were affected by the predictability more than controls. Low predictability sentences resulted in longer fixations and a greater number of fixations in PWA relative to HC.

For the measure of total number of regressions, there was a significant main effect of context. This indicated that both PWA and HC demonstrated more regressions for low predictability sentences.

Linear Mixed Effects Modeling

The present analysis refines Kim and Bolger's analyses through the use of linear mixed-effects modeling (LMEM). Traditional analyses (e.g., ANOVA) aggregate data to formulate means which are compared to each other and usually do not control for random variables. LMEM utilizes the entire data set throughout the analysis and is able to control for variability in participants and stimuli. This may be advantageous to use for populations who exhibit a wide variability in symptoms, such as aphasia.

LMEM considers the dependent variable as a function of both fixed effects and random effects (Winter, 2013). Fixed effects are the systematic part of the experiment that we measure and control. For this experiment, word frequency and sentence predictability are both fixed effects. Both of these are manipulated variables that are controlled by being categorized as "high" or "low". Through these categories, they encompass the totality of what is being measured.

Random effects are the part of the experiment that are not within our control but affect the dependent variable. Baayen, Davidson and Bates (2008) described participants and stimuli as random effects, which I considered as well for this study. Both of these are considered random effects because they are sampled from a larger population and what is represented within the study does not exhaust all possibilities that are available. Each individual participant introduces randomness to the experiment through unique personal, biological, and cultural backgrounds which may influence his or her responses. Specifically, PWA have their own unique story and background which have given them different pre- and post-onset language skills. Similarly, stimuli introduce a sense of randomness as they are composed of numerous linguistic properties that may influence responses. Within this study, some of this was controlled within the experimental design (e.g., word frequency, word predictability), but other properties were not as tightly controlled. For example, the sentence lengths of the stimuli were only controlled across pairs of sentences. There is also the consideration of several linguistic factors, including sentence type, word order, and the meanings of the words included. Each of these factors may contribute to the contextual effect, so it is important to consider the impact of the stimuli within the analysis. The strength of LMEM was that it considers that participants and stimuli are made up of different factors that may influence the dependent variable, where traditional analyses tend to ignore these factors.

Analysis was performed in R (R Core Team, 2016) with the packages "lme4" (Bates, Maechler, Bolker, and Walker, 2015) and "lmerTest" (Kuznetsova, Brockhoff, & Christensen, 2017).

Assumptions of LMEM. In this section, I will summarize the assumptions of LMEM. According to Winter (2013), these follow the same assumptions as linear regression models. Assumption 1 is the *assumption of linearity*. This means that the relationship between the response variable and the explainable variables should be linear.

Assumption 2 is the *absence of collinearity* which means that predictors should not be correlated with each other. This can occur in over-explained models where the predictors are very similar to each other. This becomes problematic as it becomes difficult to interpret which predictor is actually contributing to the model.

Assumption 3 is the *homoscedasticity of residuals.* Residuals are the difference between a model's predicted value and the actual value. It is expected that a good model has equal variances of residuals, meaning that it is reliable at explaining the the response variable at various values of the dependent variable.

Assumption 4 is the *normality of residuals*. Residuals should have a normal distribution, resembling that of a bell-curve.

Assumption 5 is the *absence of influential data points*. Outliers within the data can drastically affect a model's ability to predict.

Assumption 6 is the most important and it is the *assumption of independence*. It assumes that all data points have come from different participants. This is not the case for many experiments as multiple responses are usually collected from participants. LMEM resolves this assumption by using subjects as a random effect. The model assumes that each subject is different and estimates each one as a different baseline value (i.e., random intercept). This means that multiple responses from a single subject are affected by their own unique baseline, which distinguish them from other participants.

Table 4.4: Assumptions of linear mixed effects modeling (Winter, 2013).

Assumption	**Description**
1) Linearity	The relationship between the dependant variable and explanatory variables should be linear.
2) Absence of Collinearity	Fixed effects should not be correlated with each other.
3) Homoscedasticity of Residuals	Variances of the predicted values should be approximately equal.
4) Normality of Residuals	The distribution of residuals should resemble a normal distribution.
5) Absence of Influential Data Points	There should be no outliers in the dataset.
6) Independence	Each data point should come from a different subject.

Response variables. The measures used in this retroanalysis were *first fixation duration* (FFD), *first gaze duration* (FGD), and *total fixation duration* (TFD) on target words. Differing from the original study, these response variables were chosen as they provided information on different stages of lexical processing and had continuous distributions that would likely fit LMEM. First fixation duration is the duration of the first fixation on the word. First gaze duration is the duration of all fixations on a word before the eyes move to the left or right of it. Both are measures of early processing as they capture the reading of a word during the first encounter. Though these measures may seem redundant, first gaze duration may capture difficulty processing a word in the first reading pass if more than one fixation occurs, whereas first fixation duration would miss it. Total fixation duration provided a more global measure of lexical processing and represents all fixations on a word in all reading passes. PWA have been described as taking longer in late stages of processing than HC (e.g., Schattka et al., 2010). As such, any differences are most likely to be reflected in total fixation duration. Based on the results of the sentence reading studies by DeDe (2017) and Huck et al. (2017a; 2017b), I hypothesized frequency and contextual effects to emerge in both early and late stages of processing.

Data omission. A standard method of data omission for eye-tracking studies was used. Measures were omitted if they were <70 ms or >2000 ms, or 3 SD from the mean (Inhoff & Radach, 1998). This was to avoid accidental eye-slips on to the target word or off-task gazing from being included in the analysis. This also removed any potential outliers, ensuring that the assumption of absence of influential data points is followed. Including missing data, a total of 9.8% of the data was omitted from the analysis.

Model fitting. Data was fitted using a top-down approach described by Zuur and colleagues (Zuur, Walker, Saveliev, & Smith, 2009). It began with a model called a *beyond optimal mode*l. This model over-explained the response variable by containing all fixed and random effects that most likely contributed to it. The model was optimized using a form of hypothesis testing that determined which variables significantly contributed to the model. Using likelihood ratio tests, variables of interest were tested by comparing a full model to a nested model which excluded the variable of interest. If there was a statistically significant difference (alpha= 0.05) between the models, then the variable of interest significantly contributed to the model. The final model included only the variables that significantly contributed to it. The protocol is outlined below:

1. A beyond optimal model was built using all explainable fixed and random effects.
2. The random effects structure was optimized through a series of likelihood ratio tests using restricted maximum likelihood (REML) estimation. This step included a comparison to a linear regression model (i.e., no random effects), which was used to justify if a random effects structure was necessary.
3. The fixed effects structure was optimized through a series of likelihood ratio tests using maximum likelihood (ML) estimation.
4. The results of the final model was presented with REML estimation.

Results and Analysis

Data description. An overview of eye-measurements for both PWA and HC are provided in Table 4.5. Patterns within the data were immediately evident.

Throughout all measures, PWA demonstrated greater means and standard deviations than HC. Thus, PWA not only had longer fixations and more fixations than HC, but they also demonstrated more variability in their eye-movements.

Table 4.5: Mean eye-tracking measures of PWA and HC when reading target words of varying frequency (HF = high frequency, LF = low frequency frequency) embedded in varying context conditions (HP = high predictability, LP = low predictability).

	Control (n =10)					**PWA (n = 12)**				
	Overall	HP	LP	HF	LF	Overall	HP	LP	HF	LF
First Fixation Duration (ms)	222 ± 73	220	224	220	225	287 ± 97	281	292	284	289
First Gaze Duration (ms)	246 ± 102	241	252	239	254	392 ± 193	380	403	386	397
Total Fixation Duration (ms)	342 ± 199	320	363	332	353	774 ± 425	705	846	764	785
Total Number of Fixations	1.2 ±.91	1.1	1.3	1.2	1.2	2.8 ± 1.7	2.5	3.2	2.8	2.8

Histograms of the data were created using the package “ggplot2” (Wickham, 2016) and arranged with “gridExtra” (Auguie & Antonov, 2017). Figure 4.2 (top) shows that the data was positively skewed for all response variables. An assumption of homogeneity of variances is not included in the assumptions for LMEM (Winter, 2013). However, when the models were created, the residual plots for each model violated the assumption of homoscedasticity of residuals. To improve the residuals for each model, all response variables were log transformed. The bottom of Figure 4.2 shows that this transformation reduced skewness in all response variables.

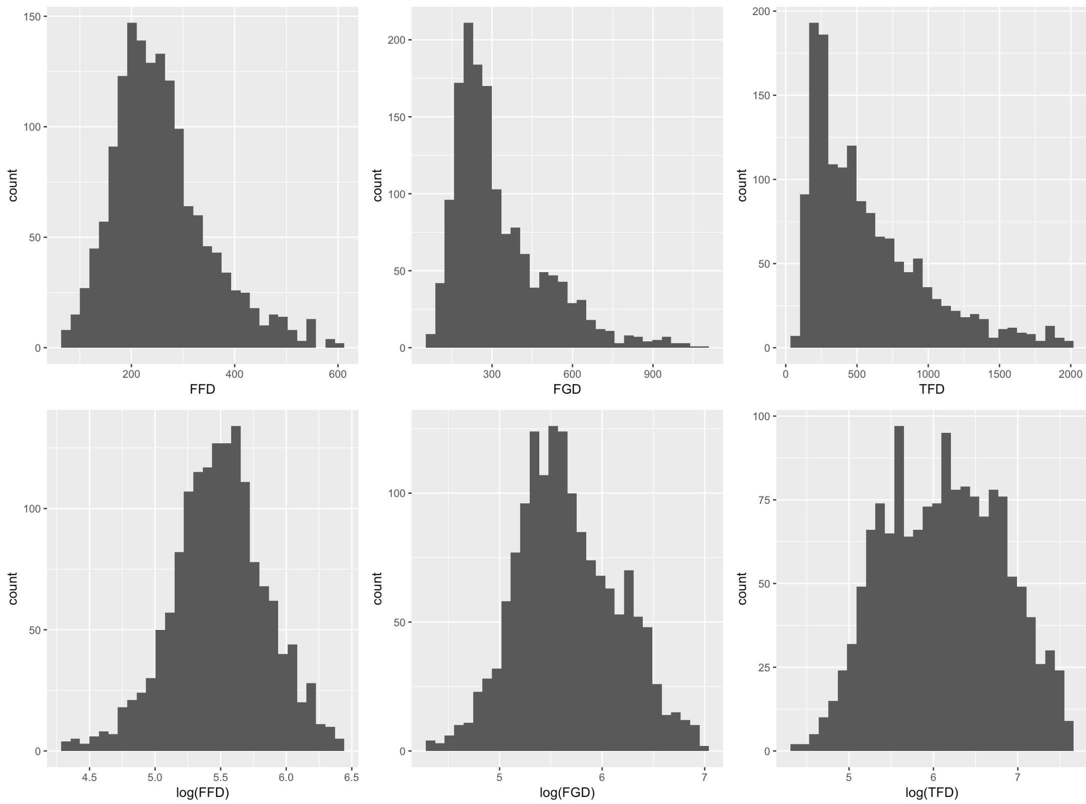


Figure 4.2: Distributions of non-transformed (top) and transformed (bottom) data for the response variables *first fixation duration*, *first gaze duration*, and *total fixation duration*.

Beyond optimal model. The beyond optimal model was formed by including all the explainable variables. The fixed effects chosen were the experimental variables of Group (HC, PWA), Context (High, Low), and Frequency (High, Low). An interaction term of Group and Context was included based on Kim and Bolger's (2012) finding that PWA had a greater response to predictability effects than controls. The random effect terms were Subject and Stimulus. A random slope for Education was included within the by-subject term, as it was logical that subject responses to linguistic information may have differed based on their years of education. Each of the response variables began with the same beyond optimal model, which is presented in Table 4.6:

Table 4.6: Terms and descriptions for a linear mixed effect model.

Beyond optimal model: response variable ~ Group * Context + Frequency + (1+Education|Subject) + (1|Stimulus)	
Term	**Description**
Group	Fixed effect
Context	Fixed effect
Frequency	Fixed effect
"*"	Main effect AND interaction
":"	Interaction only
Education	Random slope
(1|Subject)	Random intercept
(1|Stimulus)	Random intercept

Beta values of main effects can be interpreted in respect to their reference variable. The reference variable for Group was the healthy control participants; the reference variable for Context was high predictability sentences; the reference variable for Frequency was high frequency words.

Total fixation duration. The optimization of the random effects structure is presented in Table 4.7. First, model (a), which was a linear regression model (i.e., no random effects), was compared to the beyond optimal model. There was a significant difference between these models, justifying the use of LMEM. After, the goal was to identify which terms contributed to the random effects structure. Model (b), which excluded the random slope, was compared to the beyond optimal model. There was no difference between models, suggesting that the simpler intercept-only model was appropriate for the data. Next, model (c), which excluded the random effect of Subject, was compared to the intercept-only model. This variable was found to significantly contribute to the model. Similarly, this process was repeated for model (d), which excluded the random effect of Stimulus, and was also found to be significant. The resulting random effects structure was comprised of (1|Subject) + (1|Stimulus). This structure was kept stable for the following step.

The fixed effects structure was optimized in the same manner as the random effects structure. Model (e), which excluded the interaction term, was compared to semi-optimized model (1). There was no significant difference, suggesting that the interaction term could be excluded. Models (f), (g), and (h) excluded the terms of Group, Context, and Frequency, respectively. These models were compared to semi-optimized model (2). Only the terms Group and Context condition were found to be significant.

The result was a final model for TFD which was comprised of the fixed effects of Group and Context and the random effects of Subject and Stimulus. It had the significant main effects for Group (β= 0.84, SE= 0.11, t= 7.86, $p < 0.001$) and Context (β= 0.18, SE = 0.04, t= 4.58, $p < 0.001$). Total fixation durations were longer in PWA (M= 774 ms) than HC (M= 342 ms). Likewise, total fixation durations were longer in low predictability sentences (M= 620 ms) than high predictability sentences (M= 534 ms).

Table 4.7: The model selection process for *total fixation duration* using likelihood ratio tests.

	Likelihood Ratio Test		
Dropped Term	X^2	**df**	***p***
Beyond Optimal Model: log(TFD) ~ Group * Context + Frequency + (1+Education|Subject) + (1|Stimulus)			
a) (1+Education|Subject) + (1|Stimulus)	223.46	4	< 0.001*
b) Random slope (Education)	1.88	2	0.39
Semi-Optimized Model 1: log(TFD) ~ Group * Context + Frequency + (1|Subject) + (1|Stimulus)			
c) (1|Subject)	210.95	1	< 0.001*
d) (1|Stimulus)	29.12	1	< 0.001*
e) Group:Context	3.10	1	0.08
Semi-Optimized Model 2: log(TFD) ~ Group + Context + Frequency + (1|Subject) + (1|Stimulus)			
f) Group	30.37	1	< 0.001*
g) Context	19.04	1	< 0.001*
h) Frequency	0.87	1	0.35
Final Model: log(TFD)~ Group + Context + (1|Subject) + (1|Stimulus)			

Model validation. The model was validated by fitting residuals against fitted values. Residuals are the difference between the model’s predicted value and actual value. This is compared against the fitted values (i.e., the zero line) to get an understanding of how the model behaves. Ideally, a cluster of residuals equally surrounding the zero line would mean that the model follows the assumption of homoscedasticity of residuals. Patterns that arise that violate this assumption represent “fanning of data”. Residuals would have higher variance at higher fitted values, spreading out like a fan. This would reflect that the model has difficulty predicting data at certain values.

The residual plot is also checked for the assumption linearity. A cluster of residuals shows that the relationship between the explainable variables and the response variable is linear. A pattern that indicates a violation of this assumption is curvature (e.g., the pattern of residuals forms a "u-shape").

The model was judged to follow all assumptions of LMEM. The residual plot in Figure 4.3 showed no patterning. I inferred that it followed the assumption of homoscedasticity of residuals and assumption of linearity. A histogram of the residuals (Appendix B) showed the model followed the assumption of normality of residuals. The model summary (Appendix C) showed no collinearity between the predictors. The assumption of independence was resolved by the inclusion of Subject as a random effect.

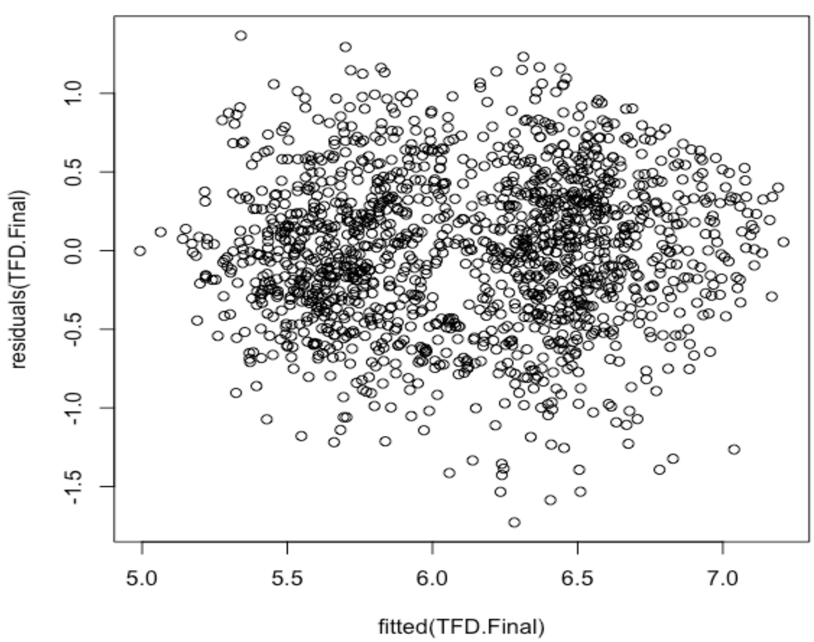


Figure 4.3: Residual plot for *total fixation duration*.

First fixation duration. The same protocol was performed with FFD and shown in Table 4.8. Resulting from this process was a final model comprised of the fixed effect of Group and the random effect of Subject. The main effect of Group was significant (β= 0.26, SE = 0.07, t= 3.72, *p*=0.001). It showed that first fixation durations were longer in PWA (M= 287 ms) than HC (M= 222 ms).

Table 4.8: The model selection process for *first fixation duration* using likelihood ratio tests.

	Likelihood Ratio Test		
Dropped Term	X^2	**df**	***p***
Beyond Optimal Model: log(FFD) ~ Group * Context + Frequency + (1+Education|Subject) + (1|Stimulus)			
(1+Education|Subject) + (1|Stimulus)	271.81	4	< 0.001*
Random slope (Education)	0.11	2	0.95
Semi-Optimized Model 2: log(FFD) ~ Group * Context + Frequency + (1|Subject) + (1|Stimulus)			
(1|Subject)	276.56	1	< 0.001*
(1|Stimulus)	0.12	1	0.73
Semi-Optimized Model 2: log(FFD) ~ Group * Context + Frequency + (1|Subject)			
Group:Context	0.03	1	0.86
Semi-Optimized Model 3: log(FFD) ~ Group + Context + Frequency + (1|Subject)			
Group	11.49	1	< 0.001*
Context	2.9	1	0.09
Frequency	1.79	1	0.18
Final Model: log(FFD)~ Group + (1|Subject)			

Model validation. The residual plot, shown in Figure 4.4, revealed vertical banding. This may be the result of the model missing an explainable variable to explain the outcome. It was difficult to identify if the model followed the assumption of homoscedasticity of residuals or linearity. No other patterns, such as fanning or curvature, were identified so it is possible that these assumptions are followed.

However, I acknowledge that it is possible that the vertical bands may mask any patterns in the data. The model followed the assumption of normality of residuals (Appendix B) and the assumption of absence of collinearity (Appendix C).

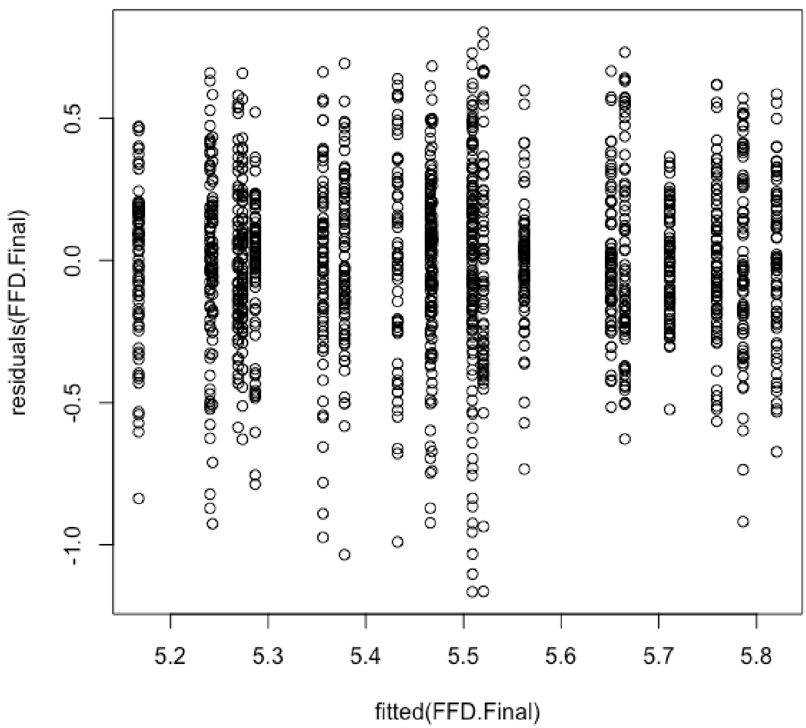


Figure 4.4: Residual plot for *first fixation duration*.

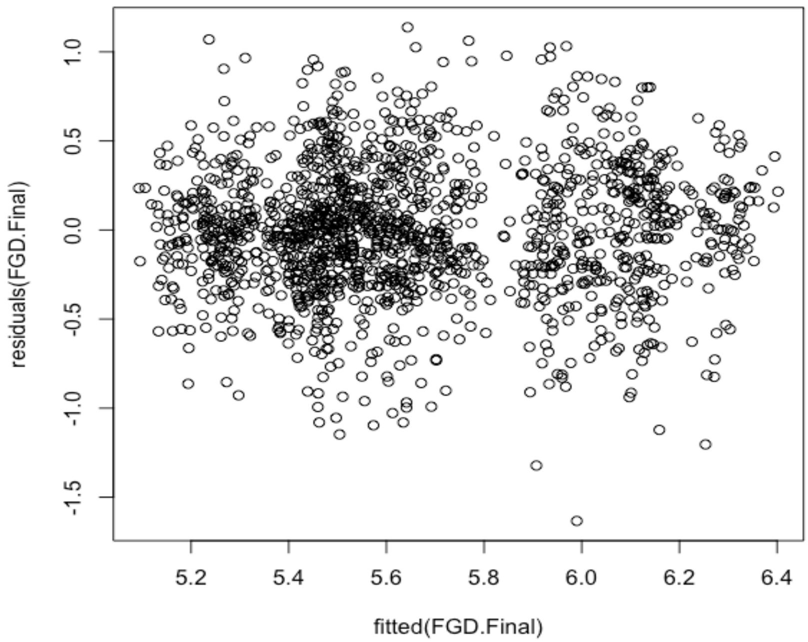
First gaze duration. The protocol for model fitting for FGD can be followed in Table 4.9. The final model was comprised of the fixed effect of Group and the random effects of Subject and Stimuli. The main effect of Group was significant (β= 0.4345, SE= 0.11, t= 4.05, *p*< 0.001). First gaze durations were longer in PWA (M= 392 ms) than HC (M= 246 ms).

Table 4.9: The model selection process for *first gaze duration* using likelihood ratio tests.

	Likelihood Ratio Test		
Dropped Term	X^2	**df**	***p***
Beyond Optimal Model: log(FGD) ~ Group * Context + Frequency + (1+Education|Subject) + (1|Stimulus)			
(1+Education|Subject) + (1|Stimulus)	432.96	4	< 0.001*
Random slope (Education)	0.67	2	0.71
Semi-Optimized Model 1: log(FGD) ~ Group * Context + Frequency + (1|Subject) + (1|Stimulus)			
(1|Subject)	437.77	1	< 0.001*
(1|Stimulus)	7.02	1	< 0.01*
Group:Context	0.01	1	0.93
Semi-Optimized Model 2: log(FGD) ~ Group + Context + Frequency + (1|Subject) + (1|Stimulus)			
Group	13.07	1	< 0.001*
Context	3.31	1	0.07
Frequency	2.26	1	0.13
Final Model: log(FGD)~ Group + (1|Subject) + (1|Stimulus)			

Model validation. The resulting residual plot can be viewed in Figure 4.5. The model was judged to follow the assumption of homoscedasticity of residuals and assumption of linearity. Noticeable in the plot was clumps of data, which may reflect the difference in responses between groups. The model followed the assumption of normality of residuals (Appendix B) and the assumption of absence of collinearity (Appendix C).

Figure 4.5: Residual plot for *first gaze duration*.



Summary. Group differences were seen in the models for all three variables. PWA had consistently longer reading times, shown in measures of first fixation duration, first gaze duration, and total fixation duration, reflecting an overall difficulty in language processing. The main effect of context was observed only in the measure of late processing (TFD), but not in measures of early processing (FFD and FGD). Total fixation durations showed that PWA, similar to HC, responded to changes in context. Both groups showed an increase in total fixation duration when viewing words embedded in low predictability sentences relative to high predictability sentences.

Discussion

Text treatments in PWA are relatively few and have resulted in an increase in oral reading accuracy and rate (Cherney, 2004; Cherney, Merbitz, & Grip, 1986; Moyer, 1979). Investigators have proposed that that text may facilitate reading as it provides an opportunity to utilize top-down processes to integrate semantic and contextual information (context effect) (Beeson & Insalaco, 1998; Tuomainen & Laine, 1991. Previous studies have utilized oral reading as an outcome measure and have not come to a consensus on the factors underlying the contextual effect (Silverberg et al., 1998; Mitchum et al., 2005). Eye-tracking methodology may provide an insight on this phenomenon.

This study was a retroanalysis of a conference paper by Kim and Bolger. Linear mixed effects modeling (LMEM) was used to analyze how sentence predictability (High, Low) and word frequency (High, Low) affect the sentence reading of PWA relative to healthy readers. LMEM is considered advantageous as it utilizes the entire dataset and is able to control for the random effects of subject and stimuli. To capture the entire reading event, measures of early processing (first fixation duration, first gaze duration) and late processing (total fixation duration) were used as the response variables.

The results of this retroanalysis found that PWA had overall longer early and late processing durations than healthy readers. This is in line with evidence of prolonged processing in other eye-tracking reading studies examining sentence reading (DeDe, 2017; Huck et al., 2017a; Huck et al., 2017b) and single word reading (Ablinger et a., 2014; Schattka et al., 2010). These results are consistent in the hypothesis that PWA require more time to process words than healthy readers.

The context effect was found in total fixation duration, which is a measure of late processing. PWA were similar to healthy readers in their response to sentence predictability as both groups were negatively affected by low predictability. Both groups had longer total fixation durations when reading words in low predictability sentences than in high predictability sentences. This is consistent with the study by

Huck et al. (2017a), which showed that the context effect can appear in late processing. Unlike their study, this was not found in any measure of early processing. Differences in results between this study and similar sentence studies investigating bottom-up and top-down processes may be due to small sample sizes and variability of PWA. With only 12 PWA in the current study, it is possible that participant characteristics (i.e., reading difficulty, education) throughout these studies are variable.

Unexpectedly, there was no effect of frequency for either group. DeDe (2017) and Huck et al. (2017a) demonstrated that frequency effects can appear in early and late processing in PWA. The reason why this did not occur in this study is unclear. Cut-offs for high and low frequency conditions were comparable to other sentence reading studies by DeDe (2017) and Huck et al. (2017a). A possibility may be due to the design of the sentence frames, which had less control than Huck et al. (2017). In their study, two sentence frames were matched with each of target words. In this current study, sentences frames were matched with a single target word. A benefit of utilizing LMEM is that it can control for the random effects within stimuli. Regardless, the frequency effect was not observed within this retroanalysis.

Similar to the results of Kim and Bolger (2012), re-analyzing the data using LMEM provides support for the finding that PWA utilize sentence context to influence their processing of word meanings. Kim and Bolger suggested that PWA have a greater response to context than healthy readers as indicated in an interaction between group and context condition in their findings. This was not supported within this retroanalysis as no interaction between group and context condition was found in any models. The results of the retroanalysis suggest that PWA and HC respond rather similarly to the context effect. However, I note that this interpretation of the results may be tentative. During the model fitting process, the interaction term of group and context approached significance for the TFD model. It is possible if more participants were collected, this interaction term may significantly contribute to the model.

Differences in results between the current study and the original study may be due to differences in data preparation and analyses. This current analysis used log transformed data to achieve a better model fit and utilized LMEM as method of analyses. The original study utilized ANCOVA to analyze non-transformed data. Arguably, the use of LMEM is more accurate than traditional methods of analyses as it considered the entire data set and was able to control for the random effects of subject and stimuli.

In the systematic review, PWA were found to have more instances of rereading. The hypothesis from that analysis was that PWA are utilizing rereading as a strategy to integrate syntactic and contextual information to facilitate access to word meaning. The results from this retroanalysis are consistent with this hypothesis as late processing was shown to be mediated by sentence predictability. This study provides support that sentence predictability may be one of the contributing factors to the context effect seen in text treatments.

Chapter 5: General Discussion

This thesis explored how eye-movements reflect the reading processes in PWA and how they can be mediated by properties of text. This investigation occurred in two parts. The first was a systematic review of the extant literature that compared the eye-movements of PWA to healthy readers. An analysis of these six studies described that PWA had a reading profile of prolonged early processing (i.e., the first reading pass), prolonged late processing (i.e., including subsequent reading passes), less word skipping, and a disproportionate amount of time spent rereading. These eye-movements reflect an overall difficulty in the processing and integration of information during reading.

Interestingly, the eye-movements of PWA were similar to healthy readers in terms of spatial measures. Saccade amplitude and landing positions did not distinguish between PWA and healthy readers. Reading difficulties are associated with shorter saccadic amplitudes, which may affect landing position distributions (Rayner, 1998). Patterns of short saccades are observable in cases of peripheral alexia (e.g., Ablinger, Huber, Schattka & Radach, 2013), where clients utilize a letter-by-letter reading strategy. Though reading processes in PWA are disturbed, shorter amplitudes were not observed. It should be noted that this observation should be taken with caution. Only two studies of single word reading included comparisons of saccade amplitudes and landing positions to healthy readers, limiting the interpretation of these results. Moreover, PWA were similar to healthy readers in that they used lexical, syntactic, and contextual information while reading single words and sentences.

The second part of this thesis was retroanalysis of a conference paper by Kim and Bolger (2012) that investigated how top-down and bottom-up factors were reflected in the eye-movements of PWA. The analysis was performed using linear mixed effects modeling, which considers the response variable as a function of fixed and random effects. LMEM is advantageous over traditional analyses (e.g., ANOVA) as it considered the entire data set in the analysis and allowed the random variables of subject and stimuli to contribute to the response variable. Sentence reading was manipulated with the experimental conditions of word frequency and sentence predictability.

Results showed that, relative to healthy readers, PWA had prolonged early and late processing as measured by first fixation duration, first gaze duration, and total fixation duration. This observation of prolonged processing time was consistent with other sentence reading studies (DeDe, 2017; Huck et al., 2017a; Huck et al., 2017b). Similar to Kim and Bolger's original study, the context effect was found in the late processing measure of total fixation duration for both PWA and healthy readers. PWA had longer total reading times for words in low predictability sentences than words in high predictability sentences. This finding is consistent with Huck et al. (2017a). No frequency effect was observed for either group.

Unlike Kim and Bolger's original study, no interaction of group and context condition was found. It is likely that the differences in findings are due to the differences in data preparation and analyses. LMEM is arguably more accurate as it provided a fine-grained analyses of the entire dataset while controlling for random variables.

Eye-movements in PWA

Similarities and differences between groups may be explained in how the eyes process linguistic information. It is possible that an impairment to central linguistic processes (e.g., phonological, semantic, and lexical components) leads to less efficient processing of linguistic information in the foveal and parafoveal views. The foveal view has the visual acuity to efficiently collect phonological, lexical, and semantic information. A central language impairment may result in a word form not being activated in the first fixation, requiring the need for multiple fixations, multiple gazes, and/or regressions back to the word once the reader becomes aware of a lack of comprehension. The parafoveal view is responsible for collecting preview information, including letter length, phonological and orthographical information (Schotter et al., 2012), which guides the eyes to the next landing position. Higher rates of fixation for upcoming words may be the result of inefficient processing of preview information.

Patterns in eye-movements of PWA can be further explained using the E-Z Reader model. The E-Z Reader is a computational model that describes how reading is an the interaction of oculomotor and linguistic processes (Reichle et al., 2006; Reichle, Warren, & McConnell, 2009); a schematic of these interactions are presented in Figure 5.1. The model assumes that lexical processing of text occurs in a serial manner (i.e., word-by-word) which is guided by attention. Attention is treated like a spotlight, which can be focused on words in the foveal and parafoveal view. Upon fixation, the attention focuses on word *n*, hypothetically in the foveal view, and analyzes it for visual properties ("V"). Next, the word undergoes lexical processing, which is characterized by early ("L1") and late ("L2") stages. It is important to note that in this thesis, the terms "early processing" and "late processing" were used differentiate between first and later reading passes. These terms are used differently within the E-Z Reader model as they refer to different stages of lexical processing *within a single fixation*. On the condition that oculomotor processing (completion of "M1") has not been completed, the lexical processing stages are able to influence when attention can shift within a fixation and/or when the eyes can move. If the early lexical processing stage completes, a command for the eyes to move to the next word (*n+1)* is sent to the oculomotor system ("M1" and "M2") where it can be processed. If the later lexical processing stage completes, attention ("A") shifts to word *n+1* within the parafoveal view and pre-processing of the upcoming word begins. Integration ("I") occurs after lexical processing has finished and represents higher-level language processes that incorporate that word's meaning into previously collected syntax and context. When oculomotor processing completes, the next saccade is triggered and this entire interaction can begin again. In normal readers, these interactions allow the eyes to move forward when enough information has been collected from words. It also allows for words to be skipped if early lexical processing finishes on a word in the parafoveal view. When this occurs, a command is sent to the oculomotor system to move the eyes to word *n+2*, instead of word *n+1*.

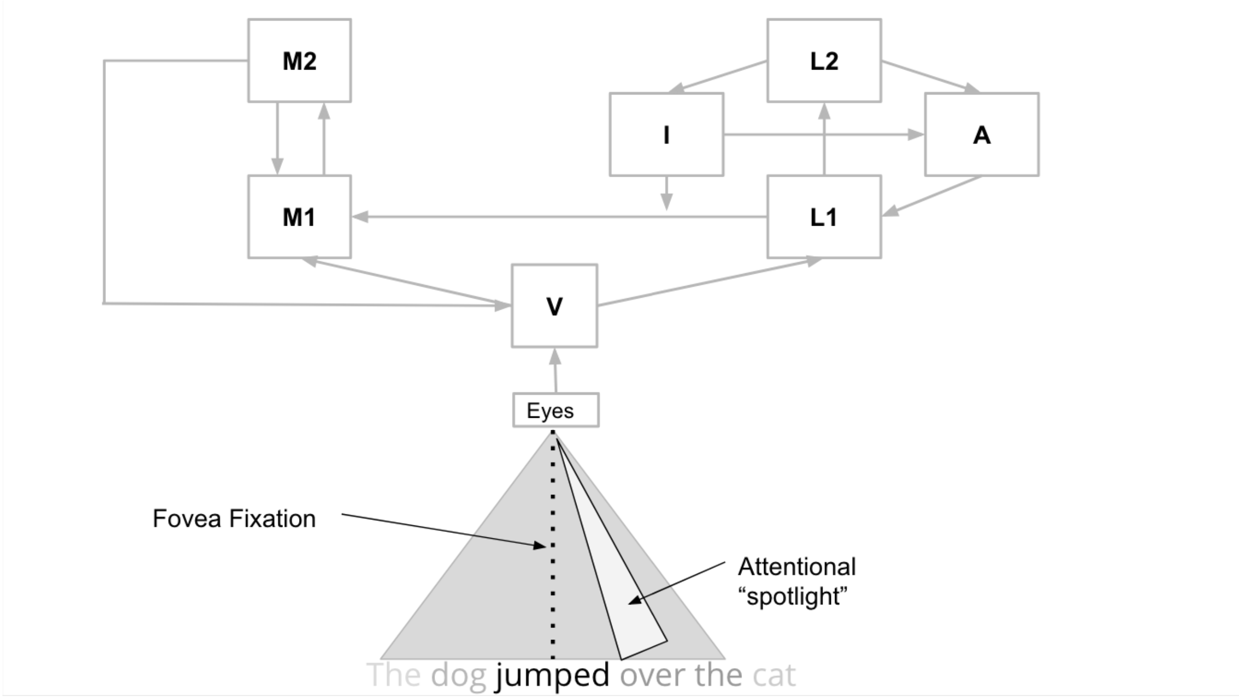


Figure 5.1: Schematic of E-Z Reader 10 model adapted from Reichle, Rayner, and Pollatsek (2003) and Reichle, Warren, and McConnell (2009).

The framework for the E-Z Reader model simplifies complex interactions through its components (e.g., "L1", "L2", "M1", "M2"), whose properties are only modestly described. The strength within this framework is that these components can be fit into deeper theories by researchers. For this thesis, I conceptualize the "L1" stage to be the processing of orthographical and phonological information and the "L2" to be the processing of semantic information. I also make the assumption that lexical processing and integration are less efficient in PWA due to the central language impairment. The model then describes the eye-movements in PWA as a consequence of slower lexical processing in relation to a (likely) intact oculomotor system. This means that the oculomotor system is more likely to finish processing before lexical processing completes. Refixations would be the result of slow early lexical processing ("L1") as the processing of this linguistic information is used to tell the eyes move to the next word. Failing to complete early lexical processing

("L1") before oculomotor processing ("M1") triggers a refixation on word *n*. Higher rates of fixations (i.e., reduced skipping rates) may result from less efficient pre-processing in the parafoveal view. This may be a result of one of two reasons. One possibility is that attentional resources are spent on processing word *n* (e.g., "L2" of word *n* does not reach completion), preventing an attentional shift to word *n+1,* and ultimately preventing the pre-processing of the upcoming word. Alternatively, pre-processing of word *n+1* has began, but the process is slow enough that early lexical processing of word *n+1* does not finish. Because the processing of word *n+1* is incomplete, the eyes will fixate on it on the next saccade. With this said, both refixations and the higher rates of fixation are related to each other. Words that are slower to process require more fixations; since more attentional resources are spent on that word, pre-processing of upcoming words is impaired, which will guarantee the need to fixate on word *n+1*.

Facilitated Reading in PWA

Text has been argued to facilitate the reading in PWA through allowing "top-down" processes to integrate syntactic and contextual cues (Beeson & Insalaco, 1998; Cherney, 2004). The findings in the systematic review and retroanalysis are supportive of this view. PWA are responsive to several linguistic variables within text, including word frequency, word length, sentence complexity, sentence frequency, sentence predictability, sentence ambiguity, and lexical bias. Text may be advantageous for PWA as they provide rich linguistic cues that may support weakened reading routes.

It is possible that PWA may be utilizing the lexical, syntactic, and contextual cues differently than healthy readers. The studies by DeDe (2017) and Knilans and

DeDe (2015) propose that lexical cues are processed early in the first reading pass, while syntactic cues are processed through later reading passes. In other words, the first reading pass is utilized to capture information of the words themselves, while later reading passes are used obtain information about syntax. Huck et al. (2017b) demonstrated that syntactic and contextual cues can affect early processing and persist into late processing. These both share the view that later reading passes are utilized to integrate information.

Research in reading development may provide an insight into the cognitive skills that allow readers to integrate knowledge in text. Hogan and colleagues (Hogan, Bridges, Justice, & Cain, 2011) used the Simple View of Reading model to argue that the higher level language skills that improve reading comprehension include inferencing, comprehension monitoring, and text structure knowledge. Inferencing allows a reader to bring in previous known knowledge to fill in gaps in comprehension. Comprehension monitoring is a reader's ability to evaluate their own understanding of what is being read. Text structure knowledge is the understanding of the schema of the text. A possibility is that the contextual effect is a result of a distributed network of cognitive skills managing information from the surrounding context and providing clues to facilitate access of lexical-semantic information of target words.

Clinical Significance

These results are clinically significant as they inform about the specific reading difficulties PWA encounter and how it may be mediated. The symptom of effortful reading in PWA may be primarily the result of more time spent rereading. This may be a sign of prolonged processing and difficulty integrating information across text. Moreover, processing in PWA can be mediated by factors within text. The text environment may provide lexical, syntactical, and contextual cues that can be utilized to facilitate the online processing of syntax and word meanings. It is possible that text based treatments are capitalizing on the client's ability integrate this information.

Limitations and Future Direction

This thesis encountered limitations in both sections. The systematic review collected a small number of studies and participants. This problem was complicated even further by each study using its own set of measures. Only one measure was consistently used across these studies (total fixation duration). However, many of the measures were used only in a few studies. The reliability of these findings would be bolstered if these measures were reported more consistently across studies.

There were two limitations identified in the retroanalysis. Typically in eye-tracking studies, data filtering is performed by merging nearby fixations under a cutoff (e.g., 80 ms) to nearby interest areas. It is unclear what cutoff was used for the original study as this data was not available. Secondly, the sentence frames in the experiment were not controlled for in length or sentence structure. One of the strengths of LMEM is its ability to manage the data by-stimulus and it appeared the variability in the sentence frames did not cause any issues with the analysis.

This thesis identified properties of text that can mediate reading in PWA. However, it is unclear which specific higher level cognitive processes are responsible for the integration of this information. It is recommended that future reading studies explore the cognitive processes that support text reading.

Chapter 6: Conclusion

Aphasia is caused by a central language impairment resulting from damage to the language-dominant hemisphere. Potentially, it can affect multiple modalities, including that of reading. Reading difficulties have been reported to commonly co-occur with aphasia. This may make life difficult for PWA as intact literacy skills are used to navigate the world. From books, to signs, to computer screens, a reader must decode single words and text to enjoy recreational activities, communicate with friends and family, and perform activities of daily living.

The ultimate goal of reading therapy is to being able to read text accurately and quickly. Rehabilitation of acquired alexia has typically focused on single word reading, but has seen limited generalization to text. Text treatments are few and have been able to remediate reading resulting in increased oral accuracy and rate. It has been proposed that top-down processes facilitate the reading of text through the use of syntactic and contextual constraints. However, the underlying mechanisms are unclear.

To elucidate the factors contributing to the context effect, this thesis investigated the reading behaviour of PWA utilizing eye-tracking methodology. A systematic review of eye-tracking reading studies indicated that PWA have a prolonged processing time, perform less word skipping, and spend more time rereading text relative to healthy readers. Importantly, this thesis provided support for the hypothesis that PWA utilize top-down processing in text reading by identifying lexical, syntactic, and contextual variables within text that mediate early and late processing in PWA. All of these variables were seen to impact late processing. Similarly, the retroanalysis confirmed the facilitative effect of context in late processing. As such, this thesis concludes that PWA may be utilizing a strategy of rereading to integrate lexical, syntactic, and contextual cues within text to cope for long processing times.

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.

Ablinger, I., Huber, W., & Radach, R. (2014). Eye movement analyses indicate the underlying reading strategy in the recovery of lexical readers. *Aphasiology*, *28*(6), 640-657.

Ablinger, I., Huber, W., Schattka, K. I., & Radach, R. (2013). Recovery in a letter-by-letter reader: More efficiency at the expense of normal reading strategy. *Neurocase, 19*(3), 236-255.

Ablinger, I., & Radach, R. (2016). Diverging receptive and expressive word processing mechanisms in a deep dyslexic reader. *Neuropsychologia*, *81*, 12-21.

Auguie, B., & Antonov, A. (2017). Miscellaneous Functions for "Grid"' Graphics. (Version 2.3) [Software].

Ashby, J., Rayner, K., & Clifton, C. (2005). Eye movements of highly skilled and average readers: Differential effects of frequency and predictability. *The Quarterly Journal of Experimental Psychology Section A*, *58*(6), 1065-1086.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390-412.

Balota, D. A., Pollatsek, A., & Rayner, K. (1985). The interaction of contextual constraints and parafoveal visual information in reading. *Cognitive Psychology*, *17*(3), 364-390.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*(1), 1-48. (Version 1.1-19) [Software].

Beeson, P. M. & Henry, M. L. (2008). Comprehension and production of written words. In R. Chapey (Ed.), *Language intervention strategies in aphasia and related neurogenic communication disorders*, 5th ed. (pp. 654-688). Baltimore, MD: Lippincott Williams & Wilkins.

Beeson, P. M. & Insalaco, D. (1998). Acquired alexia: Lessons from successful treatment. *Journal of the International Neuropsychological Society*, *4*, 621-635.

Beeson, P. M., Rising, K. & Rapcsak, S. Z. (2011). Reading and writing impairments. In L. L. LaPointe (Ed.), *Aphasia and related neurogenic language disorders*, 4th ed. (pp. 121-139). New York, NY: Thieme Medical Publishers, Inc.

Brookshire, C. E., Wilson, J. P., Nadeau, S. E., Gonzalez Rothi, L. J., & Kendall, D. L. (2014). Frequency, nature, and predictors of alexia in a convenience sample of individuals with chronic aphasia. *Aphasiology*, *28*(12), 1464-1480.

Cherney, L. (1995). Efficacy of oral reading in the treatment of two patients with chronic Broca's aphasia. *Topics In Stroke Rehabilitation*, *2*(1), 57-67.

Cherney, L. (2004). Aphasia, alexia, and oral reading. *Topics In Stroke Rehabilitation, 11*(1), 22-36.

Cherney, L. (2010). Oral Reading for Language in Aphasia: impact of aphasia severity on cross-modal outcomes in chronic nonfluent aphasia. *Seminars in Speech & Language, 31*(1), 42-51.

Cherney, L. , Merbitz, C. T., & Grip, J. C. (1986). Efficacy of oral reading in aphasia treatment outcome. *Rehabilitation Literature, 47*(5-6), 112-118.

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.

DeDe, G. (2017). Effects of Lexical Variables on Silent Reading Comprehension in Individuals With Aphasia: Evidence From Eye Tracking. *Journal of Speech, Language, and Hearing Research*, 60(9), 2589-2602.

Dickey, M. W., Choy, J. J., & Thompson, C. K. (2007). Real-time comprehension of wh-movement in aphasia: Evidence from eyetracking while listening. *Brain and Language*, *100*, 1-22.

von Elm, E., Altman, D. G., Egger, M., Pocock, S. J., Gøtzsche, P. C., Vandenbroucke, J. P., & Strobe Initiative. (2007). The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *PLoS medicine*, *4*(10), e296.

Federmeier, K. D., & Kutas, M. (1999). A rose by any other name: Long-term memory structure and sentence processing. *Journal of Memory and Language*, *41*(4), 469-495.

Federmeier, K. D., Wlotko, E. W., De Ochoa-Dewald, E., & Kutas, M. (2007). Multiple effects of sentential constraint on word processing. *Brain Research*, *1146*, 75-84.

Gahl, S. (2002). Lexical biases in aphasic sentence comprehension: An experimental and corpus linguistic study. *Aphasiology*, *16*(12), 1173-1198.

Gahl, S., Menn, L., Ramsberger, G., Jurafsky, D. S., Elder, E., Rewega, M., & Audrey, L. H. (2003). Syntactic frame and verb bias in aphasia: Plausibility judgments of undergoer-subject sentences. *Brain and Cognition*, *53*(2), 223-228.

Garrett, M. F., (1982). Production of speech: Observations from normal and pathological language use. In A. W. Ellis (Ed.), *Normality and pathology in cognitive functions* (pp. 19-76). London: Academic Press.

Hare, M., McRae, K., & Elman, J. L. (2003). Sense and structure: Meaning as a determinant of verb subcategorization preferences. *Journal of Memory and Language*, *48*(2), 281-303.

Harm, M. W., & Seidenberg, M. S. (1999). Phonology, reading acquisition, and dyslexia: insights from connectionist models. *Psychological Review*, *106*(3), 491.

Heuer, S., & Hallowell, B. (2015). A novel eye-tracking method to assess attention allocation in individuals with and without aphasia using a dual-task paradigm. *Journal of Communication Disorders*, *55*, 15-30.

Hogan, T. P., Bridges, M. S., Justice, L. M., & Cain, K. (2011). Increasing Higher Level Language Skills to Improve Reading Comprehension. *Focus on Exceptional Children*, *44*(3), 1–20.

Hoover, W. A., & Gough, P. B. (1990). The simple view of reading. Reading and writing, 2(2), 127-160.

Huck, A., Thompson, R. L., Cruice, M., & Marshall, J. (2017). Effects of word frequency and contextual predictability on sentence reading in aphasia: an eye movement analysis. *Aphasiology*, 1-26.

Huck, A., Thompson, R. L., Cruice, M., & Marshall, J. (2017). The influence of sense-contingent argument structure frequencies on ambiguity resolution in aphasia. *Neuropsychologia*, *100*, 171-194.

Inhoff, A. W., & Radach, R. (1998). Chapter 2: Definition and Computation of Oculomotor Measures in the Study of Cognitive Processes. In G. Underwood (Ed.), *Eye Guidance in Reading and Scene Perception* (pp. 29–53). Oxford: Elsevier Science Ltd.

Ivanova, M. V., & Hallowell, B. (2012). Validity of an eye-tracking method to index working memory in people with and without aphasia. *Aphasiology*, *26*(3-4), 556-578.

Johnson, R. L., & Rayner, K. (2007). Top-down and bottom-up effects in pure alexia: Evidence from eye movements. *Neuropsychologia*, *45*(10), 2246-2257.

Juhasz, B. J., & Rayner, K. (2003). Investigating the effects of a set of intercorrelated variables on eye fixation durations in reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*(6), 1312.

Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological review*, *87*(4), 329.

Kim, E. S., & Bolger, P. (2012). Examining the facilitative effect of semantic context on sentence reading in aphasia using eye-tracking. *Procedia-Social and Behavioral Sciences*, *61*, 58-59.

Kim, E. S., & Lemke, S. F. (2016). Behavioural and eye-movement outcomes in response to text-based reading treatment for acquired alexia. *Neuropsychological Rehabilitation*, *26*(1), 60-86.

Kim, M., & Russo, S. (2010). Multiple oral rereading (MOR) treatment: Who is it for. *Contemporary Issues in Communication Science and Disorders*, *37*, 58-68.

Kim, E. S., Rising, K., Rapcsak, S. Z., & Beeson, P. M. (2015). Treatment for Alexia

With Agraphia Following Left Ventral Occipito-Temporal Damage: Strengthening Orthographic Representations Common to Reading and Spelling. *Journal of Speech, Language & Hearing Research*, *58*(5), 1521-1537.

Klingelhöfer, J., & Conrad, B. (1984). Eye movements during reading in aphasics. *European archives of psychiatry and neurological sciences*, 234(3), 175-183.

Knilans, J., & DeDe, G. (2015). Online sentence reading in people with aphasia: Evidence from eye tracking. *American Journal of Speech-Language Pathology*, *24*(4), S961-S973.

Kuznetsova, A., Brockhoff, P.B., & Christensen R.H.B. (2017). "lmerTest Package: Tests in Linear Mixed Effects Models." *Journal of Statistical Software*, 82(13), pp. 1–26. (Version 3.0-1) [Software].

Leff, A., & Starrfelt, R. (2013). *Alexia: diagnosis, treatment and theory*. Springer Science & Business Media.

Meyer, A. M., Mack, J. E., & Thompson, C. K. (2012). Tracking passive sentence comprehension in agrammatic aphasia. *Journal of Neurolinguistics*, *25*(1), 31-43.

Mitchum, C., Haendiges, A., & Berndt, R. S. (2005). Oral reading of words and sentences: Investigating the source of context effects. *Aphasiology*, *19*(7), 615-631.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D.G., & The PRISMA Group. (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *Annals of Internal Medicine*, *151*(4), 264-269.

Moyer, S. B. (1979). Rehabilitation of alexia: A case study. *Cortex*, *15*(1), 139-144.

R Core Team (2016). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. (Version 3.3.2) [Software].

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, 124, 372-422.

Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *The Quarterly Journal of Experimental Psychology*, *62*(8), 1457-1506.

Rayner, K., Ashby, J., Pollatsek, A., & Reichle, E. D. (2004). The effects of frequency and predictability on eye fixations in reading: Implications for the EZ Reader model. *Journal of Experimental Psychology: Human Perception and Performance*, *30*(4), 720.

Rayner, K., Castelhano, M. S., & Yang, J. (2009). Eye movements and the perceptual span in older and younger readers. *Psychology and Aging*, *24*(3), 755.

Rayner, K., Chace, K. H., Slattery, T. J., & Ashby, J. (2006). Eye movements as reflections of comprehension processes in reading. *Scientific Studies of Reading*, *10*(3), 241-255.

Rayner, K., & McConkie, G. W. (1976). What guides a reader's eye movements?. *Vision Research*, *16*(8), 829-837.

Rayner, K., Slattery, T. J., Drieghe, D., & Liversedge, S. P. (2011). Eye movements and word skipping during reading: effects of word length and predictability. *Journal of Experimental Psychology: Human Perception and Performance*, *37*(2), 514.

Reichle, E. D., Pollatsek, A., & Rayner, K. (2006). E–Z Reader: A cognitive-control, serial-attention model of eye-movement behavior during reading. *Cognitive Systems Research*, *7*(1), 4-22.

Reichle, E. D., Warren, T., & McConnell, K. (2009). Using EZ Reader to model the effects of higher level language processing on eye movements during reading. *Psychonomic Bulletin & Review*, *16*(1), 1-21.

Riley, E. A., & Kendall, D. L. (2017). The acquired disorders of reading. In I. Papathanasiou, P. Coppens, & C. Potagas (Eds.), *Aphasia and related neurogenic communication disorders*, 2nd ed. (pp. 195–218). Burlington, MA: Jones & Bartlett Learning.

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

Schotter, E. R., Angele, B., & Rayner, K. (2012). Parafoveal processing in reading. *Attention, Perception, & Psychophysics*, *74*(1), 5-35.

Seidenberg, M. S., & MacDonald, M. C. (1999). A probabilistic constraints approach to language acquisition and processing. *Cognitive Science*, *23*(4), 569-588.

Silverberg, N., Vigliocco, G., Insalaco, D., & Garrett, M. (1998). When reading a sentence is easier than reading a little word: The role of production processes in deep dyslexics' reading aloud. *Aphasiology, 12*, 335-356.

Smits, C. H., Smit, J. H., van den Heuvel, N., & Jonker, C. (1997). Norms for an abbreviated Raven's Coloured Progressive Matrices in an older sample. *Journal of Clinical Psychology*, *53*(7), 687-697.

Staub, A. (2010). Eye movements and processing difficulty in object relative clauses. *Cognition*, *116*(1), 71-86.

Thompson, C. K. & Choy, J. J. (2009). Pronominal resolution and gap filling in agrammatic aphasia: Evidence from eye movements. *Journal of Psycholinguistic Research*, *38*, 255-283.

Thompson, C. K., Dickey, M. W., Cho, S., Lee, J., & Griffin, Z. (2007). Verb argument structure encoding during sentence production in agrammatic aphasic speakers: An eye-tracking study. *Brain and Language*, *103*(1-2), 24-26.

Trueswell, J., & Tanenhaus, M. (1994). Toward a lexical framework of constraint-based syntactic ambiguity resolution. Perspectives on sentence processing, 155-179.

Trueswell, J. C., Tanenhaus, M. K., & Kello, C. (1993). Verb-specific constraints in sentence processing: Separating effects of lexical preference from garden-paths. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*(3), 528-553.

Tuomainen, J., & Laine, M. (1991). Multiple oral rereading technique in rehabilitation of pure alexia. *Aphasiology, 5*(4-5), 401-409.

Vandenbroucke, J. P., Von Elm, E., Altman, D. G., Gøtzsche, P. C., Mulrow, C. D., Pocock, S. J., Poole, C., Schlesselman, J. J., Egger, M., & Strobe Initiative. (2007). Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): explanation and elaboration. *PLoS medicine*, *4*(10), e297.

Webb, W. G., & Love, R. J. (1983). Reading problems in chronic aphasia. *Journal of Speech and Hearing Disorders*, *48*(2), 164-171.

Wickham, H. (2016). ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York.

Winter, B. (2013). Linear models and linear mixed effects models in R with linguistic applications.

Zuur, A. F. (2009). *Mixed effects models and extensions in ecology with R.* New York: Springer.

Appendix A: Quality Assessment

Table A.1:Quality assessment of eye-tracking studies investigating reading behaviour in PWA using the STROBE Statement checklist for case-control studies

Item No	Recommendation	Ablinger et al. (2014)	Huck et al. (2017a)	Huck et al. (2017b)	DeDe (2017)	Knilans & DeDe (2015)	Schattka, Radach, & Huber (2010)
Title and abstract							
1	(a) Indicate the study's design with a commonly used term in the title or the abstract	No	No	No	No	No	No
	(b) Provide in the abstract an informative and balanced summary of what was done and what was found	Yes	Yes	Yes	Yes	Yes	Yes
Introduction							
2	Background/rationale Explain the scientific background and rationale for the investigation being reported	Yes	Yes	Yes	Yes	Yes	Yes
3	Objectives State specific objectives, including any prespecified hypotheses	Yes	Yes	Yes	Yes	Yes	Yes
Methods							
4	Study design Present key elements of study design early in the paper	No	No	No	Yes	No	No
5	Setting Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	Recruitment information provided for PWA	No	No	No	No	Recruitment information provided for PWA
6	Participants (a) Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls	PWA only; methods of ascertainment for controls not provided	Yes	Yes	Yes	Yes	PWA only; methods of ascertainment for controls not provided

	(b) For matched studies, give matching criteria and the number of controls per case	n/a	n/a	n/a	n/a	n/a	n/a
7	Variables Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	Yes	Yes	Yes	Yes	Yes	Yes
8	Data sources/ measurement For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	Yes	Yes	Yes	Yes	Yes	Yes
9	Bias Describe any efforts to address potential sources of bias	No	Yes	Yes	Yes	Yes	No
10	Study size Explain how the study size was arrived at	No	No	No	No	No	No
11	Quantitative variables Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	Yes	Yes	Yes	Yes	Yes	Yes
12	Statistical methods (a) Describe all statistical methods, including those used to control for confounding	Yes	Yes	Yes	Yes	Yes	Yes
	(b) Describe any methods used to examine subgroups and interactions	n/a	Yes	Yes	Yes	Yes	n/a
	(c) Explain how missing data were addressed	n/a	n/a	n/a	n/a	n/a	n/a
	(d) If applicable, explain how matching of cases and controls was addressed	n/a	n/a	n/a	n/a	n/a	n/a
	(e) Describe any sensitivity analyses	n/a	n/a	n/a	n/a	n/a	n/a
Results							

13	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	Yes	Yes	Yes	Yes	Yes	Yes
	(b) Give reasons for non-participation at each stage	n/a	Yes	Yes	Yes	Yes	n/a
	(c) Consider use of a flow diagram	n/a	n/a	n/a	n/a	n/a	n/a
14	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	Yes	Yes	Yes	Yes	Yes	Yes
	(b) Indicate number of participants with missing data for each variable of interest	Total missing data reported	Total missing data reported	Total missing data reported	Total missing data reported	Total missing data reported	Total missing data reported
15	Report numbers in each exposure category, or summary measures of exposure	n/a	n/a	n/a	n/a	n/a	n/a
16	(a) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	Yes	Yes	Yes	Yes	Yes	Yes
	(b) Report category boundaries when continuous variables were categorized	n/a	Yes	Yes	Yes	n/a	n/a
	(c) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	n/a	n/a	n/a	n/a	n/a	n/a
17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	n/a	Yes	Yes	Yes	Yes	Yes
				Discussion			
18	Summarise key results with reference to study objectives	Yes	Yes	Yes	Yes	Yes	Yes

19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	No	No	No	Yes	Yes	Yes
20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	Yes	No	Yes	Yes	Yes	No
21	Discuss the generalisability (external validity) of the study results	No	No	No	Yes	Yes	No
Other Information							
22	Funding Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	Yes	Yes	Yes	Yes	Yes	Yes

Appendix B: Histograms of Residuals

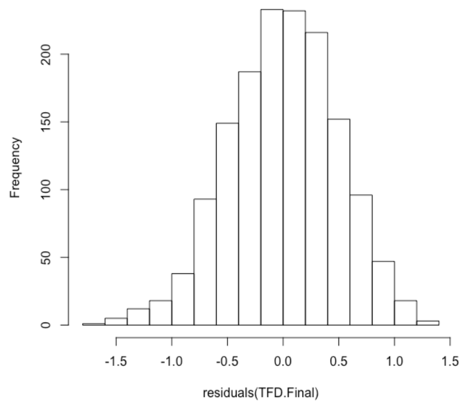


Figure B1: Histogram of residuals for *total fixation durations*

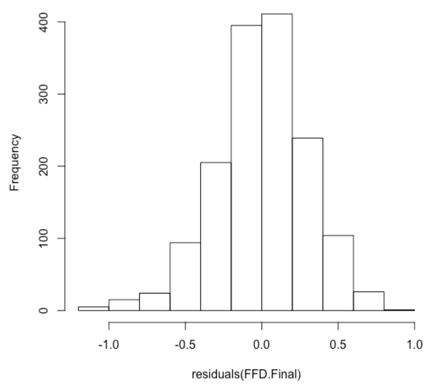


Figure B2: Histogram of residuals for *first fixation duration*

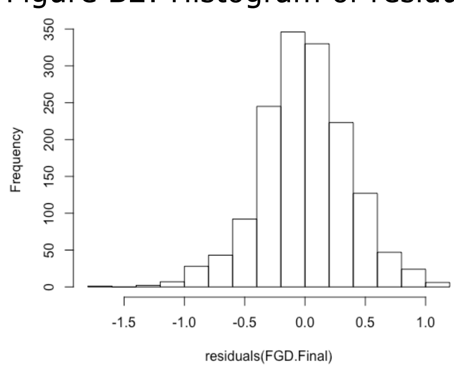


Figure B3: Histogram of residuals for *first gaze duration*

Appendix C: Model Summaries

```
>summary (TFD.Final)
Linear mixed model fit by REML. t-tests use Satterthwaite's method ['lmerModLmerTest']
Formula: log(TFD) ~ Group + Context + (1 | Subject) + (1 | Stimuli)
   Data: C

REML criterion at convergence: 2301.1

Scaled residuals:
    Min      1Q      Median      3Q     Max
-3.4631 -0.6466  0.0230  0.6679  2.7387

Random effects:
 Groups   Name        Variance Std.Dev.
 Stimuli  (Intercept) 0.01587  0.126
 Subject  (Intercept) 0.05570  0.236
 Residual             0.24901  0.499
Number of obs: 1500, groups:  Stimuli, 80; Subject, 21

Fixed effects:
             Estimate   Std. Error   df         t value   Pr(>|t|)
(Intercept)   5.57812   0.08073    22.78193  69.094    < 2e-16 ***
GroupPWA      0.83605   0.10636    18.90680   7.860     2.25e-07 ***
ContextLow    0.17508   0.03822    78.07453   4.581     1.73e-05 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Correlation of Fixed Effects:
            (Intr)  GrpPWA
GroupPWA   -0.693
ContextLow -0.240  0.005
```

Figure C1: The model summary for *total fixation duration.*

```
Linear mixed model fit by REML. t-tests use Satterthwaite's method [lmerModLmerTest]
Formula: log(FFD) ~ Group + (1 | Subject)
   Data: C

REML criterion at convergence: 742.7

Scaled residuals:
    Min      1Q  Median      3Q     Max
-3.8577 -0.5836  0.0329  0.6430  2.6555

Random effects:
 Groups    Name        Variance Std.Dev.
 Subject   (Intercept)  0.02387  0.1545
 Residual              0.09123  0.3020
Number of obs: 1519, groups:  Subject, 21

Fixed effects:
             Estimate  Std. Error        df  t value    Pr(>|t|)
(Intercept)   5.34482    0.05024   19.21007 106.382  < 2e-16 ***
GroupPWA      0.25772    0.06931   19.08933   3.718   0.00145 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Correlation of Fixed Effects:
           (Intr)
GroupPWA -0.725
```
Figure C2: The model summary for *first fixation duration*.

```
Linear mixed model fit by REML. t-tests use Satterthwaite's method ['lmerModLmerTest']
Formula: log(FGD) ~ Group + (1 | Subject) + (1 | Stimuli)
   Data: C

REML criterion at convergence: 1429.9

Scaled residuals:
    Min      1Q  Median      3Q     Max
-4.3772 -0.5973 -0.0022  0.6087  3.0494

Random effects:
 Groups   Name        Variance  Std.Dev.
 Stimuli  (Intercept) 0.004197  0.06478
 Subject  (Intercept) 0.058390  0.24164
 Residual             0.139208  0.37311
Number of obs: 1521, groups:  Stimuli, 80; Subject, 21

Fixed effects:
            Estimate  Std. Error       df  t value  Pr(>|t|)
(Intercept)   5.4204     0.0781   19.4741   69.404  < 2e-16 ***
GroupPWA      0.4345     0.1073   19.0695    4.048  0.000682 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Correlation of Fixed Effects:
         (Intr)
GroupPWA -0.721
```

Figure C3: The model summary for *first gaze duration*.