City Research Online

City, University of London Institutional Repository

Effects of word frequency and contextual predictability on sentence reading in aphasia: An eye movement analysis

Anneline Huck[1], Robin L. Thompson[2], Madeline Cruice[1], and Jane Marshall [1]

[1]*Language and Communication Science, School of Health Sciences, City, University of London, London, UK*

[2]*School of Psychology, University of Birmingham, Birmingham, UK*

Corresponding author: Anneline Huck: Anneline.Huck.1@city.ac.uk

Acknowledgments

We would like to express our gratitude to all individuals who took part in this study. We also want to thank Penny Roy for her advice in data analysis, Catherine Prior for reading aloud comprehension questions, and an anonymous reviewer for helpful comments. This work was supported by a PhD studentship from City, University of London, granted to the first author.

Abstract

Background: Mild reading difficulties are a pervasive symptom of aphasia. Whilst much research in aphasia has been devoted to the study of single word reading, little is known about the process of (silent) sentence reading. Reading research in the non-brain damaged population has benefited from the use of eye tracking methodology, allowing inferences on cognitive processing without participants making an articulatory response. This body of research identified two factors, which strongly influence reading at the sentence level: word frequency and contextual predictability (influence of context).
Aims: The main aim of this study was to investigate whether word frequency and contextual predictability influence sentence reading by people with

aphasia, in parallel to that of neurologically healthy individuals. A second aim was to examine whether readers with aphasia show individual differences in the effects, and whether these are related to their underlying language profile.

Methods & Procedures: Seventeen people with aphasia (PWA) and associated mild reading difficulties and twenty neurologically healthy individuals (NHI) took part in this study. Individuals with aphasia completed a range of language assessments. For the eye tracking experiment, participants silently read sentences that included target words varying in word frequency and predictability whilst their eye movements were recorded. Comprehension accuracy, fixation durations and the probability of first-pass fixations and first-pass regressions were measured.

Outcomes & Results: Eye movements by both groups were significantly influenced by word frequency and predictability, but the predictability effect was stronger for the people with aphasia than the neurologically healthy participants. Additionally, effects of word frequency and predictability were independent for the neurologically healthy individuals, but the individuals with aphasia showed a more interactive pattern. Correlational analyses revealed i) a significant relationship between lexical-semantic impairments and the word frequency effect score, and ii) a marginally significant association between the sentence comprehension skills and the predictability effect score.

Conclusions: Consistent with compensatory processing theories, these findings indicate that decreased reading efficiency may trigger a more interactive reading strategy that aims to compensate for poorer reading by putting more emphasis on a sentence context, particularly for low frequency words. For those individuals who have difficulties applying the strategy automatically, using a sentence context could be a beneficial strategy to focus on in reading intervention.

Keywords: reading, aphasia, eye movements, word frequency, predictability

Introduction

Mild reading difficulties are a pervasive symptom of aphasia (e.g. Cocks, Pritchard, Cornish, Johnson, & Cruice, 2013; Coelho, 2005; Meteyard, Bruce, Edmundson, & Oakhill, 2015). The reasons for reading difficulties can be myriad, as successful reading is based on the accurate and timely interplay of visual, linguistic and cognitive processes. Whilst a substantial amount of research has been devoted to the study of word-level impairments in oral reading in aphasia and acquired dyslexia (e.g. Cherney, 2004; Dickerson & Johnson, 2004; Patterson, 1994, 2000; Rapcsak et al., 2009; Warrington & Crutch, 2007), much less is known about reading at the sentence and text level, particularly with respect to the process of silent reading, i.e. reading for comprehension without articulation. However, it is essential to investigate silent sentence reading in more detail, as it comprises most of our natural everyday functional reading activities.

A promising method to study the process of silent sentence reading is eye-tracking, which has successfully informed reading research in the non-brain damaged population (for reviews, see Radach & Kennedy, 2004, 2013; Rayner, 1998). In the eye tracking while reading paradigm, a camera films a reader's eye gaze by tracking both their pupil and corneal reflection while they are reading from a computer monitor. Eye tracking is based on the assumption that there is an association between eye movements and cognitive processing such that eye movements allow us to make inferences about cognitive processes during reading (Rayner, Pollatsek, Ashby, & Clifton, 2012). Processing difficulties are detected by prolonged gaze durations or by a greater than usual number of regressions, fixations that return to earlier positions in the sentence (Boland, 2004). Eye tracking studies of healthy sentence reading have identified two factors that strongly influence reading at the sentence level: word frequency and contextual predictability. High frequency words attract shorter fixation durations than low frequency words, signalling decreases in processing load (Ashby, Rayner, & Clifton, 2005; Calvo & Meseguer, 2002; Inhoff & Rayner, 1986; Juhasz & Rayner, 2006; Kennedy, Pynte, Murray, & Paul, 2013; Rayner, Ashby, Pollatsek, & Reichle, 2004; Rayner, Reichle, Stroud, Williams, & Pollatsek, 2006). Similarly to word frequency effects, words that are highly predictable in a sentence context receive shorter fixation durations than those that are unpredictable (Calvo & Meseguer, 2002; Kennedy et al., 2013; Kliegl, Grabner, Rolfs, & Engbert, 2004; Rayner et al., 2004; Rayner, Slattery, Drieghe, & Liversedge, 2011; Zola, 1984). Both word frequency and predictability relate to our experience with language, and effects suggest that words that are more likely to occur on probabilistic grounds are easier to process than words that are less likely to occur.

The analyses of eye movements have recently also been applied to the diagnosis and treatment of reading impairment after brain damage, such as acquired central dyslexia (Ablinger, von Heyden, et al., 2014; Ablinger, Huber, & Radach, 2014; Kim & Lemke, 2016; Schattka, Radach, & Huber, 2010)[1]. There is also an emerging interest in using eye tracking to examine silent reading and sentence/text comprehension in aphasia (Chesneau, Joanette, & Ska, 2007; Kim & Bolger, 2012; Knilans & DeDe, 2015). One of these studies investigated the influence of semantic context on eye movements, comparing a group of ten people with aphasia to a group of eight control participants (Kim & Bolger,

[1] Notably, some research involving eye tracking in aphasia has been published earlier, but these studies were limited to more global parameters of eye movements such as saccade behaviour and number of fixations (e.g. Huber, Lüer, & Lass, 1983; Klingelhöfer & Conrad, 1984).

2012). The context had a significant effect on eye movements in the aphasia group, as evidenced by shorter fixation durations and by a smaller number of regressions on predictable words as compared to unpredictable words. The control group on the other hand was not affected by the manipulated context. This result is inconsistent with the eye tracking literature as summarised above, which generally reports context effects in skilled reading. A reason for the difference could be that the difference in predictability was not robust enough to affect reading in the control group. These recent studies indicate that eye tracking can serve as a valid method to investigate the process of sentence reading in aphasia. However, even though data are suggestive of a larger effect of context on reading in aphasia compared to healthy reading, it is not known how potential context effects relate to factors such as word frequency, and whether potential effects are associated with the type and/or severity of the underlying language impairment. Investigating variables that are known to affect healthy reading is an important starting point to understand the process of silent reading in aphasia. Hence, the purpose of the present study is to systematically investigate the influence of word frequency and contextual predictability on eye movements during sentence reading by people with aphasia in comparison to that of neurologically healthy readers. It is hoped that results of this study will contribute further to our understanding of the process of silent sentence reading in aphasia, specifically regarding the question of whether the language system is able to compensate for compromised reading efficiency. In the following, a summary on word frequency and context effects in aphasia will be provided, before the aims of this study are explained in more detail.

There is reason to believe that both word frequency and predictability are factors that influence reading and eye movements in aphasia. Much of the evidence for word frequency effects in aphasia stems from single word production and judgement tasks including naming, repetition, word reading and visual lexical decision (Bose & Buchanan, 2007; Bose, Lieshout, & Square, 2007; Bub, Cancelliere, & Kertesz, 1985; Cherney, 2004; Gerratt & Jones, 1987; Goodlass, Hyde, & Blumstein, 1969; Kittredge, Dell, Verkuilen, & Schwartz, 2008; Nozari & Dell, 2009; Schattka et al., 2010; Zingeser & Berndt, 1988). Schattka and colleagues (2010) showed an influence of word familiarity/word frequency on reading aloud in an eye tracking study of acquired dyslexia. Further, there is also evidence that word frequency effects are not limited to the word-level, but influence sentence comprehension in aphasia, as shown by a trend effect on self-paced reading (DeDe, 2012).

Context effects in aphasia have been described in relation to word retrieval (Mayer & Murray, 2003; Pashek & Tompkins, 2002; Pierce, 1991; Zingeser & Berndt, 1988) as well as to sentence comprehension (Germani & Pierce, 1992; Hough, Pierce, & Cannito, 1989; Pierce, 1988, 1991; Pierce & Wagner, 1985). For some individuals who present with word finding difficulties it is easier to produce a word if a relevant sentence frame is given, or if word retrieval is linked to connected speech rather than to confrontation naming.

In comparison to healthy speakers, frequency and context effects can be exaggerated in aphasia (DeDe, 2012; Kim & Bolger, 2012; Martin, 2013). Lexical impairments can lead to words having weaker lexical representations (DeDe, 2012). This implies that words generally need more activation in order to be accessed. Activation may suffice for high frequency words that have lower activation thresholds, but may not be strong enough to access low frequency words. A context provides probabilistic constraints that may increase the activation level of the target word(s) while reducing the activation level of other unrelated words and meanings. Finding a magnified context effect by readers with aphasia could be explained within the framework of interactive compensatory processing (Stanovich, 1980, 1986). If bottom-up processing is deficient, the processing system compensates and relies more on other sources of knowledge such as context (Stanovich, 1986). Good readers are efficient processors and thus they need less cognitive resources such as attention, working memory and concentration to process the visual information, and do not need extra input. Readers with a compromised language system on the other hand are less efficient processors and need more cognitive resources; these can be facilitated by extra input from a sentence context.

The main aim of this study is to find out whether both neurologically healthy individuals and individuals with aphasia show an influence of word frequency and contextual predictability when they read sentences silently, and whether these effects are equivalent or different. A further aim is to examine whether people with aphasia show any individual differences in the effects found, and if so, how effects relate to their underlying language impairments. Regarding influences of word frequency and predictability, we predict effects for people with aphasia as well as neurologically healthy individuals. However, we expect that both word frequency and predictability will show larger effects in the aphasia group in comparison to the NHI group, because aphasia is associated with weaker lexical and semantic representations of words. It is further expected that effects of frequency and predictability for people with aphasia will vary depending on the severity of their semantic and lexical impairments, with stronger effects being associated with larger impairments.

Methods

Ethical approval for this study was obtained from the School of Health Sciences Research Ethics Committee, City, University of London. All participants gave informed consent before the study commenced.

Participants

Seventeen[2] people with aphasia (PWA) and twenty neurologically healthy individuals (NHI) with no reported speech/language disorder or reading difficulty participated in the study. The PWA (ten women) were between 22 and 80 years old at the time of testing (mean age 58.76 years). They all had a single left hemisphere stroke and were between 10 months and 15 years/4 months post onset (mean post onset 5 years and 6 months).

---

[2] Nineteen PWA took part in the study and completed the background assessments, but one participant's eye tracking file was corrupt, and another participant represented as an outlier in the eye tracking experiment. Hence, both participants were excluded from the whole dataset, and are not reported.

The NHI (13 women) were between 22 and 76 years old (mean age 53.60 years). All NHI participants had cognitive functioning commensurate with their age at the time of testing. This was established by administering the *Mini-Mental State Examination*, 2nd Edition Standard Version (Folstein, Folstein, White, & Messer, 2010) with a mean score of 29.3 and no score below 27/30.

All participants were (premorbidly) right-handed. They either spoke English as a first language or as their primary language since adulthood. None of the participants had developmental dyslexia, cognitive impairments such as dementia, or any evidence of visual (-spatial) impairment such as a cataract, glaucoma, visual neglect, or severe visual field impairment. Each individual's level of education was identified by categorising the education level from 1 (no formal) to 7 (doctoral degree), and calculating the average. There were no statistical differences between the two groups in terms of age, $t(35) = -1.06$, $p > .25$, nor in terms of education, $t(35) = 1.44$, $p > .15$. Table 1 presents an overview of the demographic information of both groups. This includes information about the PWA's stroke aetiology, which is based on medical reports that most participants were able to provide.

Table 1. Demographic information for participants.

Group	**ID**	**Gender**	**Age**	**Years.months post onset**	**Aetiology (all left hemisphere)**	**Education (Group)[a]**
NHI	1	f	71	n.a.	n.a.	Diploma (4)
NHI	2	m	44	n.a.	n.a.	Doctoral (7)
NHI	3	m	40	n.a.	n.a.	Bachelor's (5)
NHI	4	m	41	n.a.	n.a.	Bachelor's (5)
NHI	5	f	59	n.a.	n.a.	Master's (6)
NHI	6	f	53	n.a.	n.a.	GCSE (2)
NHI	7	f	53	n.a.	n.a.	Bachelor's (5)
NHI	8	f	22	n.a.	n.a.	Bachelor's (5)
NHI	9	f	50	n.a.	n.a.	Master's (6)
NHI	10	f	70	n.a.	n.a.	Master's (6)
NHI	11	f	69	n.a.	n.a.	Diploma (4)
NHI	12	m	38	n.a.	n.a.	Master's (6)
NHI	13	f	76	n.a.	n.a.	Bachelor's (5)
NHI	14	m	68	n.a.	n.a.	Bachelor's (5)
NHI	15	f	73	n.a.	n.a.	no formal (1)
NHI	16	f	51	n.a.	n.a.	Bachelor's (5)
NHI	17	f	54	n.a.	n.a.	Bachelor's (5)
NHI	18	m	51	n.a.	n.a.	A levels (3)
NHI	19	f	55	n.a.	n.a.	A levels (3)
NHI	20	m	34	n.a.	n.a.	Master's (6)
mean	n.a.	n.a.	53.6	n.a.	n.a.	4.7
PWA	1	f	75	11.6	CVA	Bachelor's (5)
PWA	2	f	61	13.5	CVA, ischemic	Master's (6)
PWA	3	f	46	1.4	CVA, MCA infarct	PhD (7)

PWA	4	f	40	4	CVA, post central left parietal lobe	GCSE + other (2)
PWA	5	f	54	4.3	CVA, large insular infarct with small area of hemorrhagic transformation	Master's (6)
PWA	6	m	70	7.2	CVA	Diploma (4)
PWA	7	m	74	8.6	CVA, MCA infarct, subdural hematoma	Master's (6)
PWA	8	m	57	2.1	CVA, subarachnoid hemorrhage and MCA infarct	GCSE (2)
PWA	9	f	80	4.3	CVA, left posterior putamen, insular cortex & corona radiata	Diploma (4)
PWA	10	m	65	4.8	CVA, MCA infarct with probable near occlusion of left ICA (pre and post central gyrus, middle and inferior frontal gyri, posterior insula and the underlying white matter of the centrum semiovale and corona radiata)	No formal (1)
PWA	11	f	22	3.9	CVA, lesion anterior and temporo-parietal	A levels (3)
PWA	12	f	53	15.4	CVA, MCA infarct, left carotid dissection leading to stroke	Bachelor's (5)
PWA	13	m	46	8	CVA, secondary hemorrhage, left frontal parietal craniotomy performed	Diploma (4)
PWA	14	m	68	1.1	CVA, ischemic changes in the left MCA territory	College (4)
PWA	15	m	73	2.6	CVA, MCA infarct, frontal lobe, thrombolysed. Developed left parietal bleed.	Apprenticeship (3)
PWA	16	f	64	0.10	CVA, MCA infarct	GCSE (2)
PWA	17	f	51	1.5	CVA, parietal infarct	Apprenticeship (3)
Mean	n.a.	n.a.	58.8	5.55	n.a.	3.94

Note: f = female; m = male; n.a. = not applicable;
[a]Education groups: (1) no formal, (2) GCSE, (3) A levels/Apprenticeship, (4) Diploma/College Degree, (5) Bachelor's Degree, (6) Master's Degree, (7) Doctoral Degree

Language assessments individuals with aphasia

The people with aphasia completed a range of different language assessments. Detailed results of the assessments are presented in Appendix A. As revealed by *The Western Aphasia Battery-Revised* (WAB-R, Kertesz, 2007), thirteen participants presented with a mild, and four participants with a moderate aphasia. Their mean AQ was 82.33 (range: 64.1–93.9). Types of aphasia were mixed with about half showing Anomic aphasia (see Table A.1 for an overview of individual WAB-R scores). We calculated composite scores by taking a straight average of scores from a number of subtests. The PWA had a mean lexical-semantics composite score of 0.88 (range: 0.80–0.96). This score was derived from the average outcome of the visual lexical decision task from the *Psycholinguistic Assessments of Language Processing in Aphasia* (PALPA) (Kay, Lesser, & Coltheart, 1997), the word to picture matching task from the PALPA, the object naming task from the PALPA, and the action naming test of the *Verb and Sentence Test* (VAST) (Bastiaanse, Edwards, & Rispens, 2002). Individual scores are presented in Table A.2. PWA were somewhat more compromised in sentence level tasks, with a sentence comprehension composite score of 0.84 (range: 0.62–0.96), see Table A.3. This comprised the mean score of the written PALPA sentence picture matching task and the mean score of the VAST sentence comprehension test. The latter is a test on auditory processing, but was changed into a written version for the purpose of this study. Whereas the PALPA test investigates a range of sentence structures, the VAST focuses on two canonical and two non-canonical sentence structures. Participants differed significantly between the comprehension of canonical and non-canonical sentences (z= -3.28, p= .001), indicating grammatical impairments (DeDe, 2013).

Material

Materials consisted of fifty-six sentences including fourteen pairs of high and low frequency words. Each target word appeared in a predictable and an unpredictable sentence context. All stimuli sentences are presented in Appendix B, Table B.1. High frequency nouns had a mean frequency of 186.86 occurrences per million[3], and low frequency nouns had a mean frequency of 14.79 occurrences per million. The difference in mean frequency between the low and the high frequency words was significant, $U = 6$, $z = -4.23$, $p < .001$, r = .80. Target words were between 4 and 8 letters long (mean: 6 letters) and word length was matched between the frequent and infrequent pair (+/- one letter[4]). Two sentence frames were constructed for each word pair (example sentences are shown in Table 2), making four sentence conditions: 1) a high frequency word in a predictable

[3] Written word frequencies were obtained from the SUBTLEX database (Brysbaert & New, 2009, http://www.ugent.be/pp/experimentele-psychologie/en/research/documents/ Accessed 07/07/2016). These frequency norms were shown to predict human processing latencies much better than existing norms so far (Brysbaert & New, 2009). Naming and lexical decision latencies based on British English are available through the *British Lexicon Project* (Keuleers et al., 2012).

[4] There was one exception: the word pair *doctor/explorer* differed in length by two letters.

context (HF P); 2) a high frequency word in an unpredictable context (HF U); 3) a low frequency word in a predictable context (LF P); and 4) a low frequency word in an unpredictable context (LF U). None of the sentences made the target words implausible or anomalous. The sentences ended in a further clause in order to include a region subsequent to the critical word; a region in the end of a sentence may attract different eye movements than regions in the middle of the sentence. For an example of a word pair in the four sentence conditions, see Table 2.

Table 2. Experimental target words in a predictable and unpredictable sentence context.

Condition	Frequency	Context	Example sentence
HF P	high frequency	predictable	Anna was able to get a reduced ticket for the show because she is a *student* working there.
HF U		unpredictable	Claire loves flowers and wants to be a *student* learning how to make nice bouquets.
LF P	low frequency	predictable	Claire loves flowers and wants to be a *florist* learning how to make nice bouquets.
LF U		unpredictable	Anna was able to get a reduced ticket for the show because she is a *florist* working there.

Note: Target words are printed in italics. HF = high frequency; P = predictable; LF = low frequency; U = unpredictable.

In order to determine predictability of the target words in their sentence contexts, two norming studies were conducted online. These involved participants without brain damage and a different group to the one taking part in the eye tracking study. In the first norming study (n = 67, mean age = 29.71, range = 18–69) participants were given the potential experimental sentences, up to, but not including the target word, and were asked to generate three different possible sentence endings. Overall, predictable items were offered as a potential closure 84% of the time, and unpredictable items less than 1% of the time. In the second norming study participants (n = 50, mean age = 26.69, range = 18 – 53) were asked to rate the fit of the potential target words, generated in the first norming study, on a scale of 1-7 (1 = very low; 7 = very high). This resulted in the following predictability ratings: 6.73 for HF P, 6.69 for LF P, 2.68 for HF U, and 3.14 for LF U. There was no statistical difference between the HF P and LF P condition, nor between the HF U and LF U condition. There was a significant difference between the HF P and HF U ($t(26) = 18.21$, $p < .0001$) condition, and between the LF P and LF U condition ($t(26) = 14.01$, $p < .0001$).

The four conditions were counterbalanced across two lists. One list comprised half of the items in a predictable context, and the other half in an unpredictable context. The other list included the reverse sentence contexts. The stimuli were mixed with 31 filler sentences in each list. Filler sentences comprised simple transitive and intransitive sentence structures as well as a range of different grammatical structures such as passives and comparative sentences. *Yes/no*-comprehension questions were developed to monitor comprehension of readers (e.g. Was Anna able to get a reduced ticket?). Questions were presented both auditorily and visually. The auditory version was implemented to ease comprehension for the individuals with aphasia. A female native speaker of British English who was blind to the answer of the questions read the questions with a consistent question intonation but otherwise monotonously. Recordings were made in a soundproofed room with a standard voice recorder, and were tailored using Praat (Boersma & Weenink, 2013). In summary, each list comprised 64 trials, with 5 practice trials given at the start and 59 experimental trials. Each of these trials consisted of a sentence to read, and a following yes/no-comprehension question. All participants read both lists in two separate sessions with a minimum of seven days in between. The presentation of the lists was counterbalanced across the participants, and sentences were randomised for individual participants.

Apparatus and Set-up

An EyeLink 1000 video-based eye tracker (SR Research, Ottawa, Ontario, Canada) with low spatial and temporal noise was used to track eye gaze. Tracking was created via pupil and corneal reflection, and was monocular at a sampling frequency of 1000 Hz. The setup consisted of a Host PC for processing of the camera data (Eyelink computer), a laptop connected to a 24-inch widescreen monitor (display computer), a high-speed camera eye tracker, the desktop mount, a Microsoft sidewinder gamepad and the SR research chinrest. The gamepad was used to move between trials and to respond to the questions. In order to facilitate the use of the gamepad, all non-meaningful buttons were covered with a self-setting rubber. For individuals with a right hemiparesis, the gamepad was turned upside down which facilitated handling of the gamepad with one hand. The EyeLink 1000 desktop mount carried a 35mm lens and an IR illuminator, sat in front of the display monitor and 52cm away from the participant's eyes. The display monitor sat 92cm away from the participant's eyes. The stimuli sentences in lower and uppercase letters were displayed on a single line in the centre of the monitor. The sentences were written in black Arial 14p on a grey background, and the visual angle of a letter was 0.3°.

Procedure

Each eye tracking session started with an informal chat to make the participants feel comfortable in the room and with the setting. Participants were seated in front of the monitor and eye tracker. The chair was comfortable and adjustable in height. Participants were instructed to place their chin on the chinrest, and to lean their head against the forehead rest. A 9-point grid calibration was used aiming at an average error of less than 0.5° and a maximum error of less than 1°. These numbers indicate accuracy, that is, the correspondence between the calculated fixation location and the actual fixation location (Raney, Campbell, & Bovee, 2014). The visual angle of a letter was 0.3°. Hence, an error of 1° would mean that the fixations are shown about 3 letters away from their actual place. If this level of tracking accuracy was not successful, the set-up was changed to improve calibration. The calibration procedure was repeated when necessary, at least once halfway during the experiment.

The experiment started with a screen displaying instructions, which were read out aloud by the experimenter. Participants were instructed to read the test sentences for comprehension, and to answer a yes/no comprehension question after reading each sentence by pressing the left or right button on the gamepad. Trials started with a central dot on the screen to check accuracy of the eye gaze track. In order to direct eye gaze to the left side of the screen, a fixation cross was presented on the left side of the screen, followed by the sentence. Participants were instructed to press a large button on the gamepad when they had read the sentence. A comprehension question presented visually and via the loudspeakers followed each sentence. Participants could have a break whenever they needed one, usually once halfway through the session. The eye tracking procedure took approximately 60 minutes.

Data analysis procedures

In order to investigate whether the PWA differ from the NHI in terms of sentence comprehension, accuracy of response to the yes/no questions was used as offline measure. Only half of the comprehension questions targeted the critical word, hence, accuracy results did not necessarily reflect difficulties regarding the target word comprehension. Regarding sensitivity to word frequency and predictability, we examined four eye movement measurements on target words. All of these are conventionally used in eye-tracking reading research and all have been shown to reflect effects of word frequency and predictability in healthy reading. First, *gaze duration* was chosen to capture the initial processing of the text, i.e. first-pass reading (Rayner, 1998). Gaze duration sums up the duration of all fixations on the target word until a saccade is made to another area. Gaze durations are typically sensitive to influences of word frequency (Calvo & Meseguer, 2002; Rayner et al., 2004; Rayner, Binder, Ashby, & Pollatsek, 2001; Rayner et al., 2006), but can also reveal effects of predictability (Balota, Pollatsek, & Rayner, 1985; Rayner et al., 2004, 2001).[5] Second, *total fixation duration* (also referred to as total duration or total reading time) was used to capture more global and later processing stages of reading. This measurement includes all fixations on the target word, including those from first-pass and those from second-pass reading, i.e. re-reading (Rayner, 1998). A difference between gaze and total fixation durations indicates that the target word was re-read. Total durations have predominantly shown influences of predictability (Calvo & Meseguer, 2002; Kliegl et al., 2004; Rayner et al., 2004, 2006). However, effects on total fixation duration are also found for word frequency (Juhasz & Rayner, 2006; Kliegl et al., 2004; Rayner et al., 2004, 2006, 2011). Third, we analysed the *probability of a first-pass regression* out of a target

[5] Another measure that captures the earliest moment of processing is *first fixation duration*, which refers to the duration of the first fixation on the critical word. This measure was not chosen for the present analysis, because people with aphasia tend to make multiple fixations on a word even in first pass reading. Hence gaze duration was thought to be a more critical measure in this experiment.

word. This measure indicates whether the first fixation following fixation(s) on the target word was regressive relative to the target word or not. The probability of first pass regressions was calculated for all words that were fixated in first-pass reading. First-pass regressions have mostly been associated with predictability (Rayner et al., 2004). Words that are unpredictable within a sentence context create a small-scale garden path effect, which can initiate regressive eye movements and lead to re-analysis (Kliegl et al., 2004). This explains why predictability effects are often shown in re-reading measures. Lastly, the *probability of a first-pass fixation*, referring to whether a word was fixated as opposed to skipped, was analysed. Previous studies have demonstrated that readers are particularly likely to skip words if they are both high frequency and predictable (Kliegl et al., 2004; Rayner et al., 2004, 2011).

Following general practice, eye movement data were filtered with pre-determined cut-offs (Juhasz, Liversedge, White, & Rayner, 2006; Kliegl et al., 2004; Rayner et al., 2006, 2011; Schattka et al., 2010). Fixations that were shorter than 80ms, and which were within one character adjacent to another fixation (equal to a visual angle of 0.3° and equal to the size of one letter on the screen), were combined with that fixation. It is assumed that readers cannot extract information in fixations shorter than 80ms. Fixations shorter than 80ms with no near neighbour as well as fixations longer than 1200ms were excluded (Juhasz et al., 2006). Altogether this eliminated about 5% of the data.

Comprehension accuracy between groups was compared using the Mann Whitney U test as the data violated assumptions of a normal distribution and homogeneity of variance. The eye movement data were log-transformed and analysed via mixed multi-factorial Anovas, conducted with participants (F_1) as well as items (F_2) as random factors. The independent between-group variable was group (NHI vs. PWA), the within-group independent variables were frequency (high vs. low) and predictability (predictable vs. unpredictable). The dependent continuous variables were the fixation measurements as explained above. Mixed model Anovas were conducted in R (R Core Team, 2013) using the *ez* package (Lawrence, 2011). Pearson's correlation and the Spearman's test were carried out using the coin package, which implements permutation based tests (Hothorn, Hornik, Wiel, & Zeileis, 2008). Graphs were developed using the ggplot2 package (Wickham, 2009).

Results

Comprehension accuracy

Overall, the PWA (M = 85.22%; Mdn = 92.86%) were about 10% less accurate in answering the comprehension questions than the NHI (M = 95.63%; Mdn = 100%), U = 3966, z = 5.04, $p < .0001$, r = .83. The PWA also performed less accurately than the NHIs on three of the question subgroups: predictable HF (U = 253.5, z = 3.00, p = .002, r = .49); predictable LF (U = 237, z = 2.22, p = .03, r = .36) and unpredictable LF condition (U = 282, z = 3.57, $p < .001$, r = .59). There was no significant group difference in the unpredictable HF condition. All group comparisons except for the predictable low frequency condition remained significant after the Bonferroni correction that reduced the α level to .013.

Eye movements

Gaze durations

For gaze durations, all factors revealed main effects in the analysis by participants and by items (see Table 3)[6]. First, there was a main effect of group, with gaze durations significantly longer in the aphasia group[7] (*M* = 355.04ms) than in the neurologically healthy group (*M* = 249.57ms), $F_1(1,35) = 19.45$, $p < .0001$, $\eta_G^2 = .30$; $F_2(1,26) = 118.05$, $p < .0001$, $\eta_G^2 = .49$. A main effect of frequency was revealed by longer durations on low frequency words (*M* = 326.05ms) than on high frequency words (*M* = 270.01ms), $F_1(1,35) = 63.54$, $p < .0001$, $\eta_G^2 = .12$; $F_2(1,26) = 15.37$, $p < .001$, $\eta_G^2 = .24$. Finally, there was a main effect of predictability with longer gaze durations on unpredictable words (*M* = 318.49ms) than predictable words (*M* = 277.57ms), $F_1(1,35) = 35.93$, $p < .0001$, $\eta_G^2 = .07$; $F_2(1,26) = 26.86$, $p < .0001$, $\eta_G^2 = .13$. There was also a trend interaction between frequency and predictability, $F_1(1,35) = 3.71$; $p = .06$, $\eta_G^2 = .01$; $F_2(1,26) = 4.15$, $p = .05$, $\eta_G^2 = .02$, and a higher order trend interaction between group, frequency, and predictability, $F_1(1,35) = 3.36$; $p = .08$, $\eta_G^2 = .01$ (see Figure 1)[8]. An inspection of the graph for NHI in gaze durations suggests that word frequency and predictability effects occurred independently of each other. Specifically, the predictability effect was 27.37ms for high frequency words, and 30.14ms for low frequency words. The frequency effect was 41.15ms for predictable words, and 43.92ms for unpredictable words. In contrast, the graph for the PWA indicates greater interaction between frequency and predictability effects. The predictability effect was 19.51ms for high frequency words, and 90.96ms for low frequency words. The word frequency effect was 36.19ms for predictable items and 107.63ms for unpredictable items. However, since the higher interaction was only a trend effect, no post hoc tests were carried out, and the results have to be regarded with caution. Since the Anova on the critical word is carried out for four measurements, the Bonferroni correction was applied (α level = .013). Only the main effects remained significant after this correction.

Table 3. Mean fixation durations (in ms), probability of a first-pass regression (in %) and probability of a first-pass fixation (in %) for target words.

Measurement		**Condition**			
	Group	**HF P**	**HF U**	**LF P**	**LF U**

[6] For the purpose of readability, Table 3 represents raw, i.e. untransformed data.

[7] Since PWA and NHI spanned a large age range, the contributory role of age was checked with an additional analysis, which is presented in supplementary materials.

[8] Figures represent transformed data to match the results from data analysis.

Gaze duration	NHI	214.62	241.99	255.77	285.91
	PWA	309.32	328.84	345.51	436.47
Total duration	NHI	312.37	396.55	340.97	460.57
	PWA	752.18	1027.63	794.79	1331.88
Probability of first-pass regression	NHI	14.18	21.57	14.82	18.83
	PWA	25.85	35.31	33.39	36.89
Probability of first-pass fixation	NHI	78.93	80.36	80.71	87.50
	PWA	86.55	88.66	87.39	94.12

Note. HF P = high frequency predictable words; HF U = high frequency unpredictable words; LF P = low frequency predictable words; LF U= low frequency unpredictable words.

Total durations

For total durations, all factors were main effects, in the same direction as above and as predicted (see Table 3). The aphasia group showed longer total durations (M = 976.62ms) than the neurologically healthy group (M = 377.61ms), $F_1(1,35) = 73.82$, $p < .0001$, η_G^2 = .63; $F_2(1,26) = 1389.98$, $p < .0001$, η_G^2 = .79. Second, there was a main effect of frequency for the analysis by participants, $F_1(1,35) = 15.68$, $p < .001$, η_G^2 = .03. Total durations were longer on low frequency words (M = 705.19ms) than on high frequency words (M = 600.47ms). Third, there was a main effect of predictability, $F_1(1,35) = 140.08$, p < .0001, η_G^2 = .19; $F_2(1,26) = 70.40$, $p < .0001$, η_G^2 = .33. Total fixation durations were longer on unpredictable words (M=773.70ms) than on predictable words (M = 531.96ms). Further, there was a significant interaction between group and predictability, $F_1(1,35) = 6.60$, p = .01, η_G^2 = .01; $F_2(1,26) = 7.44$, p = .01, η_G^2 = .02. Post hoc tests using dependent and independent t-tests revealed a significant group difference for predictable items, $t(59.17) = -10.43$, $p < .0001$, r = .81, as well as for unpredictable items, $t(52.81) = -10.75$, p <. 0001, r = .83. The effect of predictability was significant for the NHI, $t(39)= -8.53$, $p < .0001$, r = .81, and equally significant for the PWA, $t(33) = -8.04$, $p < .0001$, r = .81. These effects all remained significant after the Bonferroni correction (α level at .013). The predictability effect was 101.88ms for the NHI, and 406.28ms for the PWA. A t-test showed that the predictability difference score was significantly larger for the PWA than the NHI, $t(55.35) = -2.45$, p = .02, r = .31.

Finally, there was a marginally significant interaction of group, frequency, and predictability, $F_1(1,35) = 4.08$, p = .05, η_G^2 = .006; this just missed significance by items, $F_2(1,26) = 3.13$, p= .09, η_G^2 = .01. This interaction (see Figure 2) shows the same pattern as the three-way interaction for gaze durations. Further analyses with post hoc tests (α level at .013) revealed that for the NHI, there was a predictability effect for high frequency words ($t(19) = -5.78$, $p < .0001$, r= .80) as well as for low frequency words ($t(19) = -6.20$, p < .0001, r = .82). There was no statistical difference in the magnitude of this effect for high compared to low frequency items ($t(19) = 0.30$, p = .76, r = .07). There was a word frequency effect for predictable words ($t(19) = -2.12$, p = .05, r = .44), and the word frequency effect for unpredictable words reached the level of a trend (t(19 = -1.88, p = .08,

r = .40). Only the predictability effects remained significant after correcting for multiple analyses.

For the PWA, there was a predictability effect for the high frequency ($t(16) = -4.74$, $p < .001$, $r = .76$) as well as the low frequency words ($t(16) = -7.10$, $p < .0001$, $r = .87$). The predictability effect was stronger for the low frequency items than the high frequency items ($t(16) = -2.12$, $p = .05$, $r = .47$). Further, there was a word frequency effect for the unpredictable items ($t(16) = -3.08$, $p = .007$, $r = .61$), but there was no word frequency effect for the predictable items ($t(16) = -0.69$, $p = .50$, $r = .17$). Results from the post hoc tests were significant (α level at .013). In summary, results from the analyses of total fixation durations yielded main effects that remained significant after the Bonferroni correction, the interaction between group and predictability was at the corrected α level, and the three-way interaction was not significant after correction for multiple analyses.

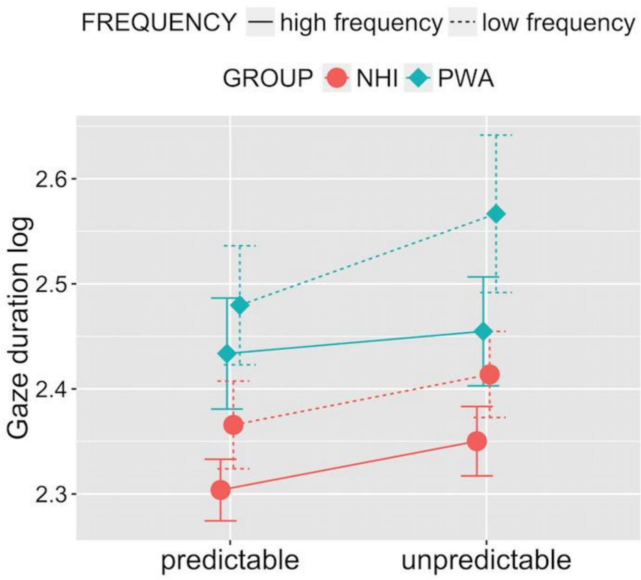

Figure 1. Effects of word frequency and predictability for NHI and PWA for gaze duration.

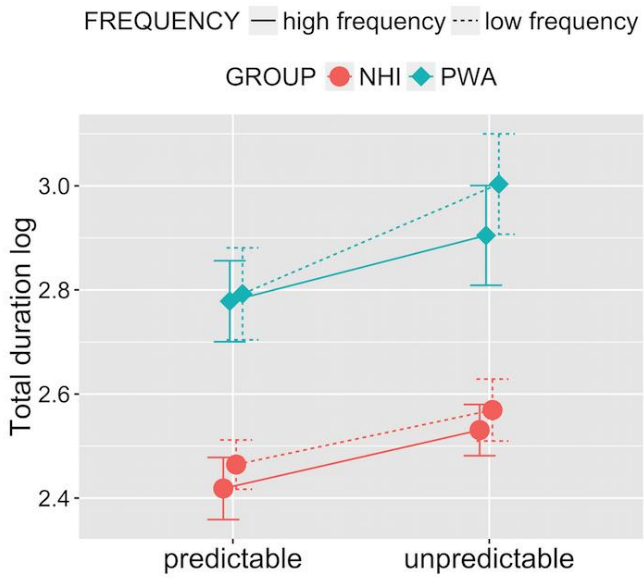

Figure 2. Effects of word frequency and predictability for NHI and PWA for total duration.

Probability of first-pass regression

For first-pass regressions, analyses of variance showed a main effect of group and predictability (see Table 3). PWA made more regressions (*M* = 32.86%) than NHI (*M* = 17.35%), $F_1(1,35) = 16.11$, $p < .001$, $\eta_G^2 = .23$; $F_2(1,26) = 63.67$, $p < .0001$, $\eta_G^2 = .30$. Readers from both groups regressed more out of unpredictable items (*M* = 27.51%) than out of predictable items (*M* = 21.45%), $F_1(1,35) = 13.30$, $p < .001$, $\eta_G^2 = .04$; $F_2(1,26) = 7.06$, $p = .01$, $\eta_G^2 = .07$. There was an interaction between group and frequency in the analysis by participants, $F_1(1,35) = 5.40$, $p < .05$, $\eta_G^2 = .01$. Post hoc analyses showed that PWA made more regressions than the NHI in both the high frequency ($t(71.28) = -3.62$, $p < .001$, $r = .39$) and the low frequency conditions ($t(70.38) = -5.66$, $p < .0001$, $r = .56$). The group difference was stronger for the low frequency conditions. Results (see Figure 3) suggest that the groups differed in their pattern of regression behaviour. The NHI made more regressions out of high frequency words than low frequency words, and the PWA regressed more out of low frequency than high frequency words. However, paired t-tests showed that the effect of word frequency on regressions was neither significant for the NHI ($t(39) = 0.48$, $p = .64$, $r = .08$) nor for the PWA ($t(33) = -1.69$, $p = .10$, $r = .28$). Results that remained significant after the Bonferroni correction were the main effect of group and predictability, but not the interaction between group and frequency (α level at .013).

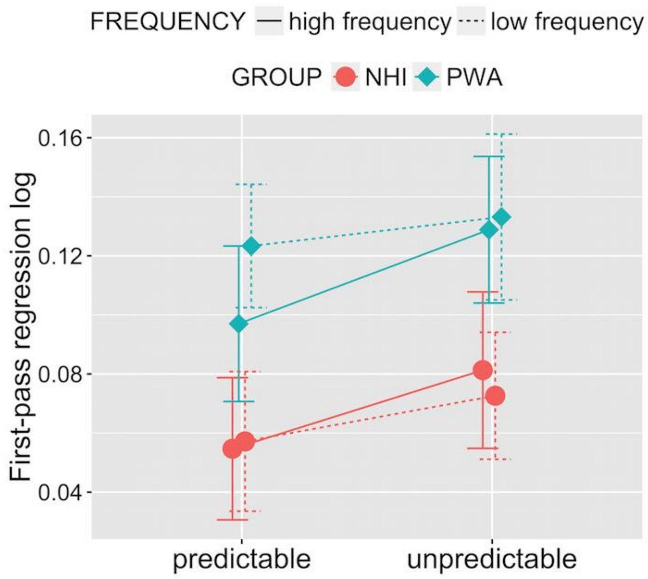

Figure 3. Effects of word frequency and predictability for NHI and PWA for the probability of first-pass regressions.

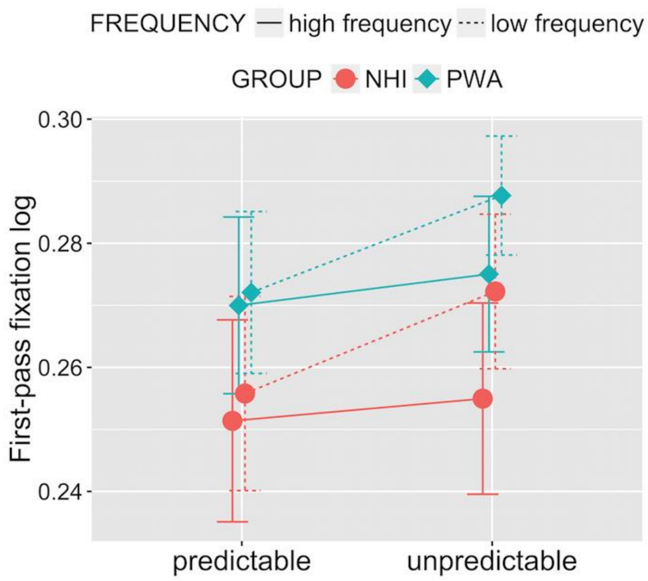

Figure 4. Effects of word frequency and predictability for NHI and PWA for the probability of first-pass fixations.

Probability of first-pass fixation

The analyses of the probability of first-pass fixations demonstrated a main effect of group, frequency and predictability (see Table 3). PWA were more likely to fixate a word in first-pass reading (M = 89.18%) than the NHI (M = 81.88%), $F_1(1,35) = 6.56$, $p < .05$, $\eta_G^2 = .09$, $F_2(1,26) = 15.24$, $p < .001$, $\eta_G^2 = .12$. In other words, PWA were less likely to skip words than NHI. Readers from both groups were also more likely to fixate a low frequency word (M = 87.16%) than a high frequency word (M = 83.30%), $F_1(1,35) = 5.70$, $p < .05$, $\eta_G^2 = .03$; however, the word frequency effect was not significant for the analysis by items. Finally, individuals from both groups demonstrated a larger probability of first-pass fixations if words were unpredictable (M = 87.36%) compared to predictable (M = 83.11%), $F_1(1,35) = 8.03$, $p < .01$, $\eta_G^2 = .09$, $F_2(1,26) = 5.40$, $p < .05$, $\eta_G^2 = .04$. There were no interactions. After Bonferroni correction, only the main effect of predictability remained significant for the analysis of participants. Results from first-pass fixations are pictured in Figure 4.

The relation between language skills and word frequency and predictability effects

Individual effects of word frequency and predictability were calculated and correlated with background language assessments. We calculated frequency and predictability effect scores as proportions[9] for each participant in the sample. For the word frequency effect score, fixation durations in the low frequency conditions were divided by the fixation durations in the high frequency conditions. For the predictability effect score, fixation durations in the unpredictable conditions were divided by the fixation durations in the predictable conditions. From both effect scores, 1 was subtracted to gain a proportional effect score, and to show the difference in percentages. As an example, an effect score of 1.5 would mean that fixation durations in one condition are 1.5 times longer than in the other conditions. Having subtracted the 1, 0.5 means that fixation durations are 50% longer in one as compared to the other condition. Subtracting 1 led to negative values for the non-predicted effects. The average proportional effect score for the neurologically healthy group for gaze duration was 0.19 for the word frequency effect and 0.13 for the predictability effect. For total duration, the average effect score was 0.14 for the word frequency effect and 0.33 for the predictability effect.

[9] Simple difference scores (calculated by subtracting fixation durations in the high frequency items from fixation durations in low frequency items) ran the risk of confounding any group differences.

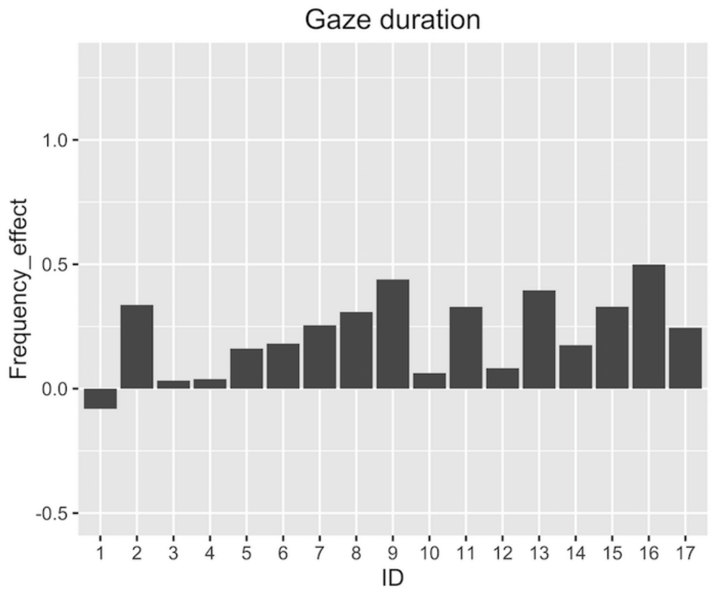

Figure 5. Proportional word frequency effect scores for gaze duration by the PWA
Note. Bars represent proportional word frequency effect scores that were calculated by dividing gaze durations in the low frequency conditions by gaze durations in the high frequency conditions minus 1. Positive values refer to effects in the predicted direction, and negative values refer to effects in the non-predicted direction.

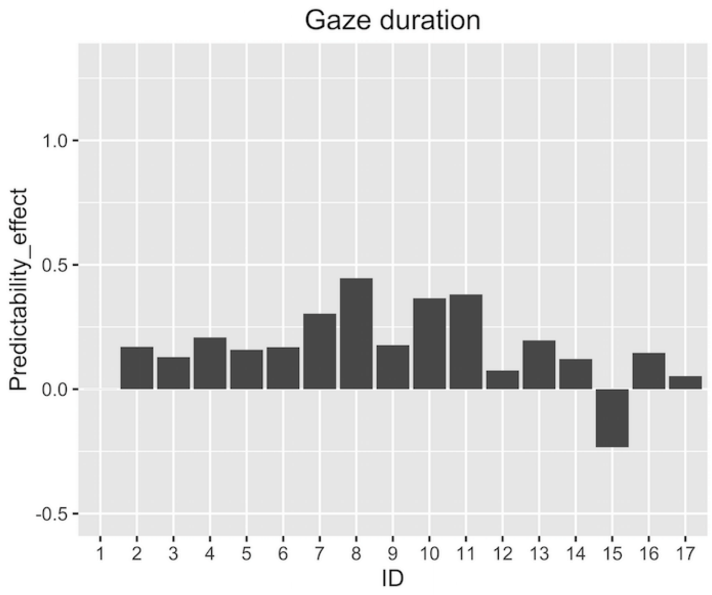

Figure 6. Proportional predictability effect scores for gaze duration by the PWA
Note. Bars represent proportional predictability effect scores that were calculated by dividing total durations in the unpredictable conditions by total durations in the predictable conditions minus 1. Positive values refer to effects in the predicted direction, and negative values refer to effects in the non-predicted direction.
ID 1 has a predictability effect score .0002 for gaze duration, hence the effect is too small to be visualised in the graph.

Figures 5 – 8 illustrate proportional effect scores for the seventeen participants with aphasia. This demonstrates that the word frequency and predictability effects for gaze duration and total fixation duration were in the predicted direction for most individuals with aphasia. However, there was individual variability regarding the size of the effects, particularly for the predictability effect for total durations. Further, some individuals with aphasia did not reveal the predicted effects, or demonstrated effects in the non-predicted direction. Individual 15, for example, showed longer gaze durations on predictable compared to unpredictable words, leading to a negative value of the effect.

The Aphasia Quotient and the two composite scores were correlated with the word frequency and predictability effect scores for both gaze and total duration. Analyses revealed a trend correlation between the Aphasia Quotient and the word frequency effect score for total durations, $r = -.44$, $p = .08$. The higher the AQ score, the smaller the word frequency effect. Further, the lexical-semantics composite score correlated with the word frequency effect score for total durations, $r = .54$, $p < .05$. The higher the participants with aphasia scored in the lexical-semantics tasks, the smaller was their effect of word frequency in the eye tracking experiment. Additionally, there was a relationship between the sentence comprehension composite score and the predictability effect score for total durations, which was marginally significant, $r_s = .48$, $p=.05$. Interestingly, the better the readers performed in sentence comprehension tasks, the larger was their context effect in the experiment. None of the correlations involving word frequency and predictability effects for gaze durations were significant.

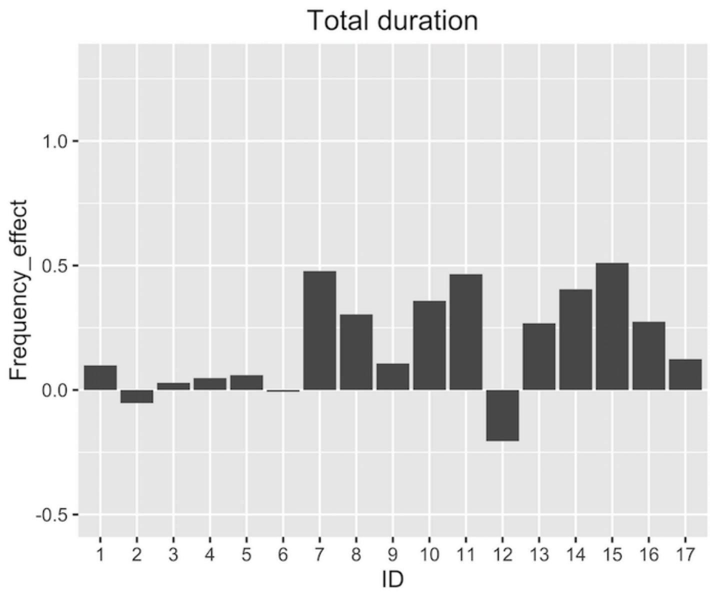

Figure 7. Proportional word frequency effect scores for total duration by the PWA

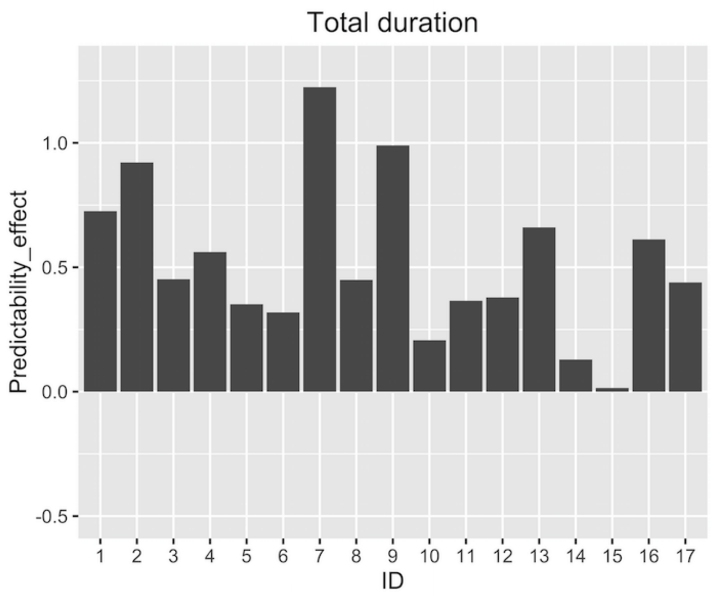

Figure 8. Proportional predictability effect scores for total duration by the PWA

Discussion

This experiment investigated whether eye movements by neurologically healthy individuals and individuals with aphasia are influenced by word frequency and contextual predictability when they read sentences silently. The presence of aphasia mildly affected comprehension accuracy, and strongly affected all measures of eye movements. PWA showed prolonged gaze and total durations, regressed more frequently out of target words, and demonstrated a larger probability to fixate a word than neurologically healthy participants. Increased durations in first-pass reading were probably due to a delay in the time course of lexical processing, rather than lexical accuracy, which was only mildly impaired in this group. Slowed lexical processing has also been demonstrated in eye-tracking-listening studies (Choy & Thompson, 2010; Meyer, Mack, & Thompson, 2012). PWA's prolonged total fixation durations suggest that they engage in more re-reading than the neurologically healthy individuals. Re-reading, in turn, is associated with high demands in post lexical integration processes (Ashby et al., 2005). An eye movement analysis of acquired dyslexia, Schattka et al. (2010) also suggests that prolonged re-reading times can reflect general linguistic processing difficulties as well as self-monitoring. Finally, group differences in first-pass regressions and in word skipping behaviour could be an indicator that groups differ in reading strategy (Kliegl et al., 2004; Rayner et al., 2006). Results suggest that readers with aphasia adopt a more careful reading strategy than the healthy readers, as demonstrated by a higher probability of fixating words and a higher probability of regressing to earlier parts of the sentence.

As expected, eye movement measures by both groups were influenced by word frequency and predictability. Readers inspected low frequency items for longer than high frequency items. Predictability was the most robust factor and influenced both groups in all four measurements. Both gaze and total durations were longer if the word was unpredictable as compared to when it was predictable, and first-pass fixation and first-pass regression probabilities were more pronounced on unpredictable than predictable words. The finding of word frequency and predictability effects during reading in the healthy population supports a large number of studies showing these influences on eye movements (Ashby et al., 2005; Calvo & Meseguer, 2002; Juhasz & Rayner, 2006; Kennedy et al., 2013; Kliegl et al., 2004; Rayner et al., 2004, 2006). Regarding silent reading in aphasia, effects are consistent with findings of these effects in other tasks (Hough et al., 1989; Jescheniak & Levelt, 1994; Kittredge et al., 2008; Nozari, Kittredge, Dell, & Schwartz, 2010; Zingeser & Berndt, 1988), and extend previous findings from a self-paced reading study that demonstrated that word frequency effects can be revealed at the sentence level (DeDe, 2012). Further, results of the context effect are consistent with findings by Kim and Bolger (2010). Establishing word frequency and predictability effects using eye tracking further emphasises their integral part in language processing, as eye movements are known to be strongly related to cognitive processes. The eye movement analysis revealed that the effects are shown in largely automatic reading that is free of meta-strategies and not influenced by having to focus on an additional task.

Since word frequency is known to be a robust factor in aphasia, it was expected that it would exert a stronger influence on reading in the aphasia group than in the control group. Contrary to these expectations, the size of the word frequency effect was similar between the groups. A group with more marked lexical impairments may have shown magnified frequency effects. Another reason for the absence of a group by frequency interaction could be a trade off between fixation durations and first-pass regressions. Results revealed an interaction between group and word frequency, showing that the group difference in the number of first-pass regressions was larger for low than for high frequency words. The PWA showed an increase in regressions when approaching a low frequency word. It has been argued that regressions can signal incomplete lexical access (Engbert, Nuthmann, Richter, & Kliegl, 2005). Instead of fixating low frequency words for longer, PWA may have regressed out of the low frequency word to reread the sentence from an earlier point, hereby gaining facilitation through the sentence context.

Consistent with our predictions, there was a larger effect of contextual predictability for the PWA than for the NHI, revealed for total fixation durations. Thus predictability influenced a global measure of eye movements that includes re-reading durations. The sentence context was particularly facilitative for the reading of low frequency words. As argued above, readers who had difficulties in accessing or integrating low frequency words may have regressed to earlier parts of the sentence to re-read with support from the sentence context. Thus, compared to the NHI, the PWA relied more on the sentence context during reading. This result supports interactive compensatory processing theories (Stanovich, 1980, 1986), which assume that the use of a sentence context to facilitate reading is dependent on reading efficiency. More precisely, the magnitude of a context effect is inversely related to word recognition abilities (Stanovich, 1986). If bottom-up processing is not fully efficient, the system can compensate by posing more demands on other information sources such as the sentence context. More evidence for an interactive pattern of results in the aphasia group is provided by the finding that the word frequency effect was significant for unpredictable but not for predictable words. Unpredictable words are more difficult to integrate into a sentence context, leading to an effect of frequency whereas predictable words are easier to integrate, making frequency differences obsolete. In contrast to the aphasia group, the neurologically healthy participants showed a word frequency effect independent of predictability, and an effect of predictability that was independent of word frequency, a finding that is largely consistent with results from previous eye tracking studies in psycholinguistics (Ashby et al., 2005; Kliegl et al., 2004; Rayner et al., 2004, 2006). The group difference demonstrated here could imply that reduced reading efficiency leads to a more interactive reading strategy (see Ashby et al., 2005 for a similar pattern of word frequency and predictability effects as established for skilled compared to average readers). The aphasia group showed a strong reduction in processing speed compared to the control group. The individuals with aphasia did, however, understand the experimental sentences most of the time; their overall comprehension accuracy was only 10% lower than that of the control group. This may suggest that people who have mild reading impairments associated with their aphasia incur more processing costs and expend more effort to read, but often accomplish comprehension through an over-reliance on the sentence context, through re-reading, and by employing a more dynamic processing system than the neurologically healthy individuals.

Results from correlational analyses between linguistic background assessments and experimental effects point to a relationship between lexical-semantic processing and the effect of word frequency in that compromised lexical-semantic processing was associated with an increased difference between the processing of low and high frequency words. This suggests that in the context of weaker lexical-semantic representations, activation is more likely to suffice for accessing high frequency words than low frequency words. The correlation supports the assumption that mild lexical impairments in this group had an influence on how frequency affected their reading and eye movement behaviour. The variability in lexical scores in this aphasia group was however limited with many participants scoring close to ceiling. This could explain why the correlation was only marginally significant, and why no significant interactions were found between group and frequency in the mixed model Anovas. Correlational analyses further revealed that the magnitude of the predictability effect varied depending on sentence comprehension skills. The higher PWA scored in sentence comprehension tests, the more did they seem to rely on the context in the experimental task. A tentative explanation for this relation could be that using a sentence context results in better sentence comprehension. A sentence context provides semantic, phonological, syntactic and probabilistic constraints (Pashek & Tompkins, 2002). These constraints can facilitate word recognition and may result in improved reading comprehension, because effortless decoding is needed to free cognitive resources for higher-level demands (Perfetti, 2007).

Finally, results from this study have implications for reading rehabilitation. First, finding a relation between the use of a sentence context and sentence comprehension skills suggests that “context reading” could be a beneficial strategy to target in reading treatment. This may prove particularly useful for people that have more severe reading difficulties and who struggle to read at the sentence level. Second, a more interactive reading strategy could be supported in reading treatment by targeting the reading of low frequency words in predictable sentence contexts before they are incorporated into less predictable sentence contexts.

In summary, this paper adds to a recent interest in studying sentence reading in aphasia using the analyses of eye movements. The outcome suggests that mild reading difficulties in aphasia may be associated with a larger reliance on the sentence context as this compensates for less efficient bottom-up processing. Hence, the context influence is part and parcel of the normal reading process, but can be magnified if reading is compromised, consistent with interactive compensatory theories.

References

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. http://doi.org/10.1080/02687038.2014.894960

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. http://doi.org/10.1080/09602011.2014.913530

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*, *58A*(6), 1065–1086. http://doi.org/10.1080/02724980443000476

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. http://doi.org/10.1016/0010-0285(85)90013-1

Bastiaanse, R., Edwards, S., & Rispens, J. (2002). *Verb and Sentence Test (VAST).* Bury St. Edwards, UK: Thames Valley Test Company Ltd.

Boersma, P., & Weenink, D. (2013). Praat: Doing phonetics by computer [Computer program]. Version 5.4.21, retrieved January 2013 from http://www.praat.org/.

Boland, J. (2004). Linking eye movements to sentence comprehension in reading and listening. In M. Carreiras & C. Clifton Jr (Eds.), *The on-line study of sentence comprehension: Eyetracking, ERPs, and beyond* (1st ed., pp. 51–76). Hove, UK: Psychology Press. http://doi.org/10.4324/9780203509050

Bose, A., & Buchanan, L. (2007). A cognitive and psycholinguistic investigation of neologisms, *21*, 726–738. http://doi.org/10.1080/02687030701192315

Bose, A., Lieshout, P. Van, & Square, P. A. (2007). Word frequency and bigram frequency effects on linguistic processing and speech motor performance in individuals with aphasia and normal speakers. *Journal of Neurolinguistics*, *20*(1), 65–88.

http://doi.org/10.1016/j.jneuroling.2006.05.001

Brysbaert, M., & New, B. (2009). Moving beyond Kucera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods*, *41*(4), 977–90. http://doi.org/10.3758/BRM.41.4.977

Bub, D., Cancelliere, A., & Kertesz, A. (1985). Whole-word and analytic translation of spelling to sound in a non-semantic reader. In K. E. Patterson, J. C. Marshall, & M. Coltheart (Eds.), *Surface dyslexia: Cognitive and neuropsychological studies of phonological reading*. Amsterdam: North-Holland Publishing.

Calvo, M. G., & Meseguer, E. (2002). Eye movements and processing stages in reading: Relative contribution of visual, lexical, and contextual factors. *The Spanish Journal of Psychology*, *5*(1), 66–77. http://doi.org/10.1017/S1138741600005849

Cherney, L. R. (2004). Aphasia, alexia, and oral reading. *Topics in Stroke Rehabilitation*, *11*(1), 22–36. http://doi.org/10.1080/02687038.2015.1052728

Chesneau, S., Joanette, Y., & Ska, B. (2007). Text comprehension and eye movements after aphasia recovery. *Brain and Language*, *103*(1), 232–233. http://doi.org/10.1016/j.bandl.2007.07.019

Choy, J. J., & Thompson, C. K. (2010). Binding in agrammatic aphasia: Processing to comprehension. *Aphasiology*, *24*(5), 551–579. http://doi.org/10.1080/02687030802634025

Cocks, N., Pritchard, M., Cornish, H., Johnson, N., & Cruice, M. (2013). A "novel" reading therapy programme for reading difficulties after a subarachnoid haemorrhage. *Aphasiology*, *27*(5), 509–531. http://doi.org/10.1080/02687038.2013.780283

Coelho, C. (2005). Direct attention training as a treatment for reading impairment in mild aphasia. *Aphasiology*, *19*(3–5), 275–283. http://doi.org/10.1080/02687030444000741

DeDe, G. (2012). Effects of word frequency and modality on sentence comprehension impariments in people with aphasia. *American Journal of Speech-Language Pathology*, *21*(2), 103–114. http://doi.org/10.1044/1058-0360(2012/11-0082).Effects

DeDe, G. (2013). Verb transitivity bias affects on-line sentence reading in people with aphasia. *Aphasiology*, *27*(3), 326–343. http://doi.org/10.1080/02687038.2012.725243

Dickerson, J., & Johnson, H. (2004). Sub-types of deep dyslexia: A case study of central deep dyslexia. *Neurocase*, *10*(1), 39–47. http://doi.org/10.1080/13554790490960477

Engbert, R., Nuthmann, A., Richter, E. M., & Kliegl, R. (2005). SWIFT: A dynamical model of saccade generation during reading. *Psychological Review*, *112*(4), 777–813.

http://doi.org/10.1037/0033-295X.112.4.777

Folstein, M. F., Folstein, S. F., White, T., & Messer, M. A. (2010). *MMSE-2 Mini-Mental State Examination*. Lutz, FL: PAR.

Germani, M. J., & Pierce, R. S. (1992). Contextual influences in reading comprehension in aphasia. *Brain and Language*, *42*, 308–319. http://doi.org/10.1016/0093-934X(92)90103-L

Gerratt, B. R., & Jones, D. (1987). Aphasic performance on a lexical decision task: Multiple meanings and word frequency. *Brain and Language*, *30*, 106–115. http://doi.org/10.1016/0093-934X(87)90031-9

Goodlass, H., Hyde, M. R., & Blumstein, S. (1969). Frequency, picturability and availability of nouns in aphasia. *Cortex*, *5*(2), 104–119.

Hothorn, T., Hornik, K., Wiel, M. A. van de, & Zeileis, A. (2008). Implementing a class of permutation tests: The coin package. *Journal of Statistical Software*, *28*(8), 1–23. http://doi.org/10.18637/jss.v028.i08

Hough, M. S., Pierce, R. S., & Cannito, M. P. (1989). Contextual influences in aphasia: Effects of predictive versus nonpredictive narratives. *Brain and Language*, *36*(2), 325–334. http://doi.org/10.1016/0093-934X(89)90069-2

Huber, W., Lüer, G., & Lass, U. (1983). Processing of sentences in conditions of aphasia as assessed by recording eye movements. In R. Groner, C. Menz, & R. Monty (Eds.), *Eye movements: An international perspective* (pp. 315–335). N.J.: Erlbaum, Hillsdale.

Inhoff, A. W., & Rayner, K. (1986). Parafoveal word processing during eye fixations in reading: Effects of word frequency. *Perception & Psychophysics*, *40*(6), 431–439. http://doi.org/10.3758/BF03208203

Jescheniak, J. D., & Levelt, W. J. M. (1994). Word frequency effects in speech production: Retrieval of syntactic information and of phonological form. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*(4), 824–843. http://doi.org/10.1037/0278-7393.20.4.824

Joseph, H. S. S. L., Nation, K., & Liversedge, S. P. (2013). Using eye movements to investigate word frequency effects in children's sentence reading. *School Psychology Review*, *42*(2), 207–222. Retrieved from http://www.scopus.com/inward/record.url?eid=2-s2.0-84880947572&partnerID=40&md5=67bcebe1a944a3f17f0cc9a03b33aab6

Juhasz, B. J., Liversedge, S. P., White, S. J., & Rayner, K. (2006). Binocular coordination of the eyes during reading: Word frequency and case alternation affect fixation duration but not fixation disparity. *Quarterly Journal of Experimental Psychology (2006)*, *59*(9), 1614–25. http://doi.org/10.1080/17470210500497722

Juhasz, B. J., & Rayner, K. (2006). The role of age of acquisition and word frequency in reading: Evidence from eye fixation durations. *Visual Cognition*, *13*(7–8), 846–863.

http://doi.org/10.1080/13506280544000075

Kay, J., Lesser, R., & Coltheart, M. (1997). *Psycholinguistic assessments of language processing in aphasia*. Hove, England: Psychology Press. http://doi.org/10.1080/02687039608248403

Kennedy, A., Pynte, J., Murray, W. S., & Paul, S. (2013). Frequency and predictability effects in the Dundee Corpus: An eye movement analysis. *The Quarterly Journal of Experimental Psychology*, *66*(3), 601–618. http://doi.org/10.1080/17470218.2012.676054

Kertesz, A. (2007). *Western Aphasia Battery-Revised.* San Antonio, Tx: Harcourt Assessment.

Keuleers, E., Lacey, P., Rastle, K., & Brysbaert, M. (2012). The British lexicon project: Lexical decision data for 28,730 monosyllabic and disyllabic English words. *Behavior Research Methods*, *44*(1), 287–304. http://doi.org/10.3758/s13428-011-0118-4

Kim, E. S., & Bolger, P. (2012). Examining the facilitative effect of semantic context on sentence reading in aphasia using eye-tracking. Proceedings from the 50th Academy of Aphasia Conference [Abstract]. *Procedia - Social and Behavioral Sciences*, *61*, 58–59. http://doi.org/10.1016/j.sbspro.2012.10.078

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. http://doi.org/10.1080/09602011.2014.999688

Kittredge, A. K., Dell, G. S., Verkuilen, J., & Schwartz, M. F. (2008). Where is the effect of frequency in word production? Insights from aphasic picture-naming errors. *Cognitive Neuropsychology*, *25*(4), 463–92. http://doi.org/10.1080/02643290701674851

Kliegl, R., Grabner, E., Rolfs, M., & Engbert, R. (2004). Length, frequency, and predictability effects of words on eye movements in reading. *European Journal of Cognitive Psychology*, *16*(1–2), 262–284. http://doi.org/10.1080/09541440340000213

Klingelhöfer, J., & Conrad, B. (1984). Eye movements during reading in aphasics. *European Archives of Psychiatry Neurological Sciences*, *234*, 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*(2), 961–973. http://doi.org/10.1044/2015

Lawrence, M. (2011). Ez: Easy analysis and visualization of factorial experiments. R package (version 3.0-0.). Available from http://CRAN.R-project.org/package=ez.

Martin, N. (2013). Disorders of word production. In I. Papathanasiou, P. Coppens, & C. Potagas (Eds.), *Aphasia and related neurogenic communication disorders* (1st ed., pp. 131–155). Burlington, MA: Jones & Bartlett Learning.

Mayer, J. F., & Murray, L. (2003). Functional measures of naming in aphasia: Word retrieval in confrontation naming versus connected speech. *Aphasiology*, *17*(5), 481–497. http://doi.org/10.1080/02687030344000148

Meteyard, L., Bruce, C., Edmundson, A., & Oakhill, J. (2015). Profiling text comprehension impairments in aphasia. *Aphasiology*, *29*(1), 1–28. http://doi.org/10.1080/02687038.2014.955388

Meyer, A. M., Mack, J. E., & Thompson, C. K. (2012). Tracking passive sentence comprehension in agrammatic aphasia. *Journal of Neurolinguistics*, *25*(1), 31–43. http://doi.org/10.1016/j.jneuroling.2011.08.001

Nozari, N., & Dell, G. S. (2009). More on lexical bias: How efficient can a “lexical editor” be? *Journal of Memory and Language*, *60*(2), 291–307. http://doi.org/10.1016/j.jml.2008.09.006.

Nozari, N., Kittredge, A. K., Dell, G. S., & Schwartz, M. F. (2010). Naming and repetition in aphasia: Steps, routes, and frequency effects. *Journal of Memory and Language*, *63*(4), 541–559. http://doi.org/10.1016/j.jml.2010.08.001

Pashek, G. V., & Tompkins, C. A. (2002). Context and word class influences on lexical retrieval in aphasia. *Aphasiology*, *16*(3), 261–286. http://doi.org/10.1080/02687040143000573

Patterson, K. E. (1994). Reading, writing and rehabilitation: A reckoning. In M. J. Riddoch & G. W. Humphreys (Eds.), *Cognitive Neuropsychology and Cognitive Rehabilitation*. Hove, UK: Erlbaum (UK) Taylor & Francis.

Patterson, K. E. (2000). Phonological alexia: The case of the singing detective. In E. Funnel (Ed.), *Case studies in the neuropsychology of reading* (1st ed., pp. 57–84). Hove, UK: Lawrence Erlbaum Assoc.

Perfetti, C. (2007). Reading Ability: Lexical Quality to Comprehension. *Scientific Studies of Reading*, *11*(4), 357–383. http://doi.org/10.1080/10888430701530730

Pierce, R. S. (1988). Influence of prior and subsequent context on comprehension in aphasia. *Aphasiology*, *2*(6), 577–582. http://doi.org/10.1080/02687038808248968

Pierce, R. S. (1991). Contextual influences during comprehension in aphasia. *Aphasiology*, *5*(4–5), 379–381. http://doi.org/10.1080/02687039108248539

Pierce, R. S., & Wagner, C. M. (1985). The role of context in facilitating syntactic decoding in aphasia. *Journal of Communication Disorders*, *18*(3), 203–213. http://doi.org/10.1016/0021-9924(85)90021-8

R Core Team. (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. (Version 3.2.1) [Software]. Available from URL http://www.R-project.org/.

Radach, R., & Kennedy, A. (2004). Theoretical perspectives on eye movements in reading: Past controversies, current issues, and an agenda for future research. *European Journal of Cognitive Psychology*, *16*(1–2), 3–26. http://doi.org/10.1080/09541440340000295

Radach, R., & Kennedy, A. (2013). Eye movements in reading: Some theoretical context. *Quarterly Journal of Experimental Psychology (2006)*, *66*(3), 429–52. http://doi.org/10.1080/17470218.2012.750676

Raney, G. E., Campbell, S. J., & Bovee, J. C. (2014). Using eye movements to evaluate the cognitive processes involved in text comprehension. *Journal of Visualized Experiments*, (83), e50780. http://doi.org/10.3791/50780

Rapcsak, S. Z., Beeson, P. M., Henry, M. L., Leyden, A., Kim, E., Rising, K., … Cho, H. (2009). Phonological dyslexia and dysgraphia: Cognitive mechanisms and neural substrates. *Cortex*, *45*, 575–591. http://doi.org/10.1016/j.cortex.2008.04.006

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, *124*(3), 372–422. http://doi.org/10.1037/0033-2909.124.3.372

Rayner, K., Ashby, J., Pollatsek, A., & Reichle, E. D. (2004). The effects of frequency and predictability on eye fixations in reading: Implications for the E-Z Reader model. *Journal of Experimental Psychology. Human Perception and Performance*, *30*(4), 720–32. http://doi.org/10.1037/0096-1523.30.4.720

Rayner, K., Binder, K. S., Ashby, J., & Pollatsek, A. (2001). Eye movement control in reading: Word predictability has little influence on initial landing positions in words. *Vision Research*, *41*(7), 943–54. http://doi.org/10.1016/S0042-6989(00)00310-2

Rayner, K., Pollatsek, A., Ashby, J., & Clifton, C. (2012). *Psychology of reading* (2nd ed.). New York: Psychology Press.

Rayner, K., Reichle, E. D., Stroud, M. J., Williams, C. C., & Pollatsek, A. (2006). The effect of word frequency, word predictability, and font difficulty on the eye movements of young and older readers. *Psychology and Aging*, *21*(3), 448–465. http://doi.org/10.1037/0882-7974.21.3.448

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), 514–28. http://doi.org/10.1037/a0020990

Schattka, K. I., Radach, R., & Huber, W. (2010). Eye movement correlates of acquired central dyslexia. *Neuropsychologia*, *48*(10), 2959–2973. http://doi.org/10.1016/j.neuropsychologia.2010.06.005

Stanovich, K. E. (1980). Toward and interactive-compensatory model of individual differences in the development of reading fluency. *Reading Research Quarterly*, *16*(1), 32–71.

Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly*, *21*(4), 360–407.

Warrington, E. K., & Crutch, S. J. (2007). Selective category and modality effects in deep dyslexia. *Neurocase*, *13*, 144–153. http://doi.org/10.1080/13554790701440462

Wickham, H. (2009). *Ggplot2: Elegant graphics for data analysis*. New York: Springer. http://doi.org/10.1007/978-0-387-98141-3

Zingeser, L. B., & Berndt, R. S. (1988). Grammatical class and context effects in a case of pure anomia: Implications for models of language production. *Cognitive Neuropsychology*, *5*(4), 473–516. http://doi.org/10.1080/02643298808253270

Zola, D. (1984). Redundancy and word perception during reading. *Perception & Psychophysics*, *36*(3), 277–284. http://doi.org/10.3758/BF03206369

Web reference

SUBTLEXUS frequency norms (2009). http://www.ugent.be/pp/experimentele-psychologie/en/research/documents/ Last Accessed 07/07/2016.

Appendix A. Individual language assessment scores for people with aphasia.

Table A.1. Individual (and mean) scores on the Western Aphasia Battery – Revised.

PWA ID	Spontaneous Speech (max=20)	Auditory Comprehension (max=10)	Repetition (max=10)	Naming (max=10)	Aphasia Quotient (max= 100)	WAB-R Subtype
1	17	9.35	8.4	9.1	87.7	Anomic
2	17	10	8.2	9.5	89.4	Anomic
3	17	10	9.6	8.7	90.6	Anomic
4	17	10	9.3	9.5	91.6	Anomic
5	15	9.4	9.2	8.5	84.2	Anomic[a]
6	19	9.95	9.1	8.9	93.9	Anomic
7	17	9.45	8.2	7	83.3	Anomic
8	16	8.5	5.4	7.3	74.4	Conduction
9	18	9.3	9	7	86.6	Anomic
10	12	8.65	9.1	7.5	74.5	Transcortical motor
11	13	9.35	3.4	6.3	64.1	Broca
12	14	8.7	6.4	9.2	76.6	Conduction[b]
13	17	9.9	9	9.1	90	Anomia
14	15	9.95	8.7	8.2	83.7	Anomia
15	13	9	8.2	8.1	76.6	Transcortical motor
16	18	9.15	7.2	7.1	82.9	Anomia
17	15	7.95	4.8	7	69.5	Conduction
Mean	15.88	9.33	7.84	8.12	82.33	n.a.

[a] ID 5 showed a history of Broca's aphasia. At the time of testing her speech was non-fluent and effortful, but with good monitoring skills and few errors. Grammatical impairments persisted in sentence comprehension as shown by an advantage of canonical over non-canonical sentences (Table C.1.).

[b] ID 12 was classified as Conduction but presented symptoms of Broca's aphasia. Her non-fluent speech contained omissions of determiners as well as errors of verb inflection.

Table A.2. Individual (and mean) scores on lexical-semantic processing (all in proportions).

PWA ID	Word - picture matching (PALPA)	Visual lexical decision (PALPA)	Naming nouns (PALPA)	Naming verbs (VAST)	Lexical production (VAST)	Lexical-semantics composite score
1	1.00	1.00	0.93	0.91	0.92	0.96
2	1.00	1.00	0.95	0.79	0.87	0.93
3	0.98	0.94	0.95	0.80	0.88	0.92
4	0.98	0.89	1.00	0.98	0.99	0.96
5	0.93	0.74	0.85	0.71	0.78	0.81
6	0.93	0.99	0.93	0.80	0.87	0.91
7	0.98	0.82	0.90	0.63	0.76	0.83
8	0.98	0.99	0.90	0.80	0.85	0.92
9	1.00	0.98	0.95	0.78	0.86	0.93

10	0.98	0.81	0.77	0.70	0.73	0.81
11	1.00	0.97	0.78	0.43	0.61	0.80
12	1.00	1.00	0.95	0.65	0.80	0.90
13	1.00	0.93	1.00	0.91	0.95	0.96
14	1.00	0.99	0.90	0.68	0.79	0.89
15	0.95	0.68	0.82	0.78	0.80	0.81
16	0.98	0.87	0.92	0.65	0.78	0.85
17	0.95	0.93	0.83	0.68	0.76	0.85
Mean	0.98	0.91	0.90	0.74	0.82	0.88

Note: Column representing composite score is shaded in grey.

Table A.3. Individual scores on sentence comprehension (all in proportions).

PWA ID	Sentence-picture matching (PALPA)	Total canonical (VAST)	Total non-canonical (VAST)	Total sentence comprehension (VAST)	Sentence comprehension composite score
1	0.90	1.00	0.95	0.98	0.94
2	0.88	1.00	0.65	0.83	0.85
3	0.97	1.00	0.90	0.95	0.96
4	0.98	0.95	0.85	0.90	0.94
5	0.83	0.95	0.35	0.65	0.74
6	0.87	1.00	0.70	0.85	0.86
7	0.87	1.00	0.95	0.98	0.92
8	0.88	0.95	0.70	0.83	0.85
9	0.90	0.90	1.00	0.95	0.93
10	0.72	0.80	0.25	0.53	0.62
11	0.78	0.90	0.70	0.80	0.79
12	0.77	0.95	0.75	0.85	0.81
13	0.95	1.00	0.90	0.95	0.95
14	0.90	0.95	1.00	0.98	0.94
15	0.67	0.95	0.50	0.73	0.70
16	0.75	0.90	0.40	0.65	0.70
17	0.85	0.85	0.65	0.75	0.80
Mean	0.85	0.94	0.72	0.83	0.84

Note: Column representing composite score is shaded in grey.

Appendix B. Stimuli sentences used in the eye tracking experiment.

Table B.1. Stimuli sentences used in the experiment with word frequency and predictability information.

	Experimental sentence	Condition	Predictability rating	Occurrence per million (Subtlex)
1	The book describes a strong tie between the parent and the *child* living in the country.	HF P	6.80	157.65
1	Scooby-Doo is a great Dane but Lassie is a *child* who has performed in many movies.	HF U	1.47	157.65
1	Scooby-Doo is a great Dane, but Lassie is a *collie* who has performed in many movies.	LF P	6.55	1.04
1	The book describes a strong tie between the parent and the *collie* living in the country.	LF U	2.90	1.04
2	After a long day the children were hungry for *dinner,* which is healthy.	HF P	6.70	202.67
2	Popeye is strong because he likes to eat *dinner* in the evening.	HF U	2.90	202.67
2	Popeye is strong because he likes to eat *spinach i*n the evening.	LF P	6.53	2.00
2	After a long day the children were hungry for *spinach,* which was their favourite.	LF U	3.90	2.00
3	Anna was able to get a reduced ticket for the show because she is a *student* working there.	HF P	7.00	43.04
3	Claire loves flowers and wants to be a *student* learning how to make nice bouquets.	HF U	3.07	43.04
3	Claire loves flowers and wants to be a *florist* learning how to make nice bouquets.	LF P	6.95	2.41
3	Anna was able to get a reduced ticket for the show because she is a *florist* working there.	LF U	3.40	2.41
4	The young couple are saving to buy a *house* to refurbish.	HF P	6.77	514.00
4	The poor backpackers are staying in a *house* in New York.	HF U	2.90	514.00
4	The poor backpackers are staying in a *hostel* in New York.	LF P	6.67	0.57
4	The young couple are saving to buy a *hostel* to refurbish.	LF U	2.30	0.57

5	To find out about vaccinations they sought the advice of an experienced *doctor* before the trip.	HF P	6.80	263.94
5	Captain Scott was an Antarctic *doctor* who was not afraid of challenges.	HF U	2.40	263.94
5	Captain Scott was an Antarctic *explorer* who was not afraid of challenges.	LF P	6.75	1.90
5	To find out about vaccinations they sought the advice of an experienced *explorer* before the trip.	LF U	3.80	1.90
6	John withdraws money from the *bank* to go shopping.	HF P	6.70	84.98
6	Carla keeps her jewellery in a *bank* when she goes on holidays.	HF U	3.75	84.98
6	Carla keeps her jewellery in a *safe* when she goes on holidays.	LF P	6.33	143.00
6	John withdraws money from the *safe* to go shopping.	LF U	4.50	143.00
7	Although she is tired she reads another *chapter* before going to sleep.	HF P	6.75	11.84
7	Hannah has difficulties with the new computer and consults the *chapter* to find a solution.	HF U	2.67	11.84
7	Hannah has difficulties with the new computer and consults the *manual* to find a solution.	LF P	6.55	8.00
7	Although she is tired she reads another *manual* before going to sleep.	LF U	4.33	8.00
8	Thomas holds shares in a large *company* but wants to sell them.	HF P	6.73	147.00
8	Tim was interested in beer making and visited a *company* who explained all about it.	HF U	4.00	147.00
8	Tim was interested in beer making and visited a *brewery* who explained all about it.	LF P	6.70	1.80
8	Thomas holds shares in a large *brewery* but wants to sell them.	LF U	4.45	1.80
9	Before she goes to bed her mum reads her a short *story* written by her father.	HF P	6.90	220.78
9	Lisa does not like letters, but prefers to write a quick *story* to tell others about her news.	HF U	3.33	220.78
9	Lisa does not like letters, but prefers to write a quick *email* to	LF P	6.95	2.08

	tell others about her news.			
9	Before she goes to bed her mum reads her a short *email* written by her father.	LF U	3.03	2.08
10	After a long day at work she forgot her keys and had to go back to the *office* to get them.	HF P	6.57	203.90
10	After a long day the gambler went to play in the *office* near his home.	HF U	1.85	203.90
10	After a long day the gambler went to play in the *casino* near his home.	LF P	6.63	20.37
10	After a long day at work she forgot her keys and had to go back to the *casino* to get them.	LF U	2.70	20.37
11	Ryan loves old castles and is interested in their *history* and tales.	HF P	6.80	83.92
11	James wanted to know how rocks were formed so read a book about *history* on Sunday.	HF U	2.47	83.92
11	James wanted to know how rocks were formed so read a book about *geology* on Sunday.	LF P	6.90	0.92
11	Ryan loves old castles and is interested in their *geology* and tales.	LF U	2.53	0.92
12	Every day Liz picks up her 12 year-old from the *school* and goes home.	HF P	6.40	333.12
12	William needs a new custom made suit and goes to the *school* to get one.	HF U	1.45	333.12
12	William needs a new custom made suit and goes to the *tailor* to get one.	LF P	6.73	4.18
12	Every day Liz picks up her 12 year old from the *tailor* and goes home.	LF U	1.65	4.18
13	The athlete drinks lots of *water* at the weekend.	HF P	6.73	225.06
13	At the distillery in Scotland the man bought a bottle of *water* to take home.	HF U	3.45	225.06
13	At the distillery in Scotland the man bought a bottle of *whisky* to take home.	LF P	6.67	4.00
13	The athlete drinks lots of *whisky* at the weekend.	LF U	2.15	4.00
14	After the accident they rushed to the *hospital* to get the injury cleaned.	HF P	6.70	124.20
14	The friends carry their tents to the *hospital* where they want to sleep.	HF U	1.75	124.20

14	The friends carry their tents to the *campsite* where they want to sleep.	LF P	6.73	0.59
14	After the accident they rushed to the *campsite* to get the injury cleaned.	LF U	2.30	0.59

Footnotes

[1] Notably, some research involving eye tracking in aphasia has been published earlier, but these studies were limited to more global parameters of eye movements such as saccade behaviour and number of fixations (e.g. Huber, Lüer, & Lass, 1983; Klingelhöfer & Conrad, 1984).

[2] Nineteen PWA took part in the study and completed the background assessments, but one participant's eye tracking file was corrupt, and another participant represented as an outlier in the eye tracking experiment. Hence, both participants were excluded from the whole dataset, and are not reported.

[3] Written word frequencies were obtained from the SUBTLEX database (Brysbaert & New, 2009, http://www.ugent.be/pp/experimentele-psychologie/en/research/documents/ Accessed 07.07.2016). These frequency norms were shown to predict human processing latencies much better than existing norms so far (Brysbaert & New, 2009). Naming and lexical decision latencies based on British English are available through the *British Lexicon Project* (Keuleers, Lacey, Rastle, & Brysbaert, 2012).

[4] There was one exception: the word pair *doctor/explorer* differed in length by two letters.

[5] Another measure that captures the earliest moment of processing is *first fixation duration*, which refers to the duration of the first fixation on the critical word. This measure was not chosen for the present analysis, because people with aphasia tend to make multiple fixations on a word even in first pass reading. Hence gaze duration was thought to be a more critical measure in this experiment.

[6] For the purpose of readability, Table 3 represents raw, i.e. untransformed data.

[7] Since PWA and NHI spanned a large age range, the contributory role of age was checked with an additional analysis, which is presented in supplementary materials.

[8] Graphs represent transformed data to match the results from data analysis.

[9] Simple difference scores (calculated by subtracting fixation durations in the high frequency items from fixation durations in low frequency items) ran the risk of confounding any group differences.