Vergence responses to vertical binocular disparity during lexical identification

Nikolova, M.[1], Jainta, S.[2], Blythe, H.I.[3], Jones, M.O.[4] and Liversedge, S.P.[5]

Affiliations:

[1:] School of Psychology, University of Southampton, Highfield Campus, Southampton, SO17 1BJ, UK. Tel: 0044-23-80-595078, Email: M.Nikolova@soton.ac.uk

[2] : Leibniz Research Centre for Working Environment and Human Factors, Ardeystrasse 67, D-44139 Dortmund, Germany. Tel: 0049-1084-272, Fax: 0049-1084401, Email: jainta@ifado.de

[3] : School of Psychology, University of Southampton, Highfiled Campus, Southampton SO17 1BJ, UK. Tel: 0044-23-8059-9399, Fax: 0044-23-8059-2606, Email: hib@soton.ac.uk

[5]: School of Psychology, University of Southampton, Highfiled Campus, Southampton SO17 1BJ, UK. Tel: 0044-23-8059-2917, Fax: 0044-23-8059-2606, Email: M.O.Jones@soton.ac.uk

[4] : School of Psychology, University of Southampton, Highfiled Campus, Southampton SO17 1BJ, UK. Tel: 0044-23-8059-2917, Fax: 0044-23-8059-2606, Email: S.P.Liversedge@soton.ac.uk

* Please address all correspondence to Mirela Nikolova, School of Psychology, University of Southampton, Highfield Capus, Southampton, SO17 1BJ, Tel: 0044-23-80-595078, Email: M.Nikolova@soton.ac.uk

Abstract

Humans typically make use of both eyes during reading, which necessitates precise binocular coordination in order to achieve a unified perceptual representation of written text. A number of studies have explored the magnitude and effects of naturally occurring and induced horizontal fixation disparity during reading and non-reading tasks. However, the literature concerning the processing of disparities in different dimensions, particularly in the context of reading, is considerably limited. We therefore investigated vertical vergence in response to stereoscopically presented linguistic stimuli with varying levels of vertical offset. A lexical decision task was used to explore the ability of participants to fuse binocular image disparity in the vertical direction during word identification. Additionally, a lexical frequency manipulation explored the potential interplay between visual fusion processes and linguistic processes. Results indicated that no significant motor fusional responses were made in the vertical dimension (all *p*s > .11), though that did not hinder successful lexical identification. In contrast, horizontal vergence movements were consistently observed on all fixations in the absence of a horizontal disparity manipulation. These findings add to the growing understanding of binocularity and its role in written language processing, and fit neatly with previous literature regarding binocular coordination in non-reading tasks.

Keywords: *vertical vergence, fixation disparity, binocular fusion, reading*

1. Introduction

Humans sample their visual environment by continuously orienting their eyes towards objects of interest in a sequence of saccades and fixations. Saccades are rapid ballistic movements of the eyes in the same direction that serve to redirect the visual axes to a new location. They are interspersed with brief periods of relative stillness, known as fixations, during which visual information is encoded (see Rayner, 1998 for review). Even though we sample visual information with two frontally placed and horizontally separated eyes, we perceive a single unified representation of the visual environment. This single percept is achieved via the sophisticated mechanisms of binocular fusion, which have been made functionally possible by the development of a vergence system that allows us to coherently merge the visual input received by each eye (Howard & Rogers, 1995; Schor & Ciuffreda, 1983).

Binocular coordination is required for efficiently performing a variety of tasks, including reading, which does not call for stereopsis, or large eye movements in depth (van Leeuwen et al., 1999). Since humans typically make use of both eyes during reading, it is important to understand how binocular coordination might impact on contributing processes involved in written language comprehension. It is relatively recent that research has begun to focus on the detailed investigation of binocular coordination during reading. A number of studies have revealed that during text processing, the two visual axes are often slightly misaligned, resulting in small vergence errors (i.e. fixation disparities) of more than 1 character space in a significant proportion of fixations (Blythe, Liversedge, & Findlay, 2010; Blythe et al., 2006; Juhasz et al., 2006; Liversedge et al., 2006a; Liversedge et al., 2006b; Nuthmann & Kliegl, 2009; Vernet & Kapula, 2009).

It has been established that because the stimulus in reading necessitates predominantly horizontal yoked eye movements, some transient divergence occurs during saccades, followed by horizontal misalignment on fixation onset (Collewijn et al. 1988; Hendriks, 1996; Yang & Kapoula 2003; Zee, Fitzgibbon, & Optican, 1993). Fine-grained oculomotor adjustments are then made during fixations in order to maximize the degree of correspondence between the two disparate retinal input (Jainta et al., 2010; Jainta & Jaschinski, 2012; Leigh & Zee, 2006). Generally, in every task – including reading - high-precision binocular vision is attained via the process of fusion, which incorporates two integral components: motor and sensory fusion (Partt-Johnson & Tillson, 2001; Schor & Tyler, 1981). Sensory fusion is a neurophysiological and psychological process whereby two independent representations are combined in the visual cortex into a single unified percept as a basic step for further processing (Howard & Rogers, 1995; Worth, 1921). Sensory fusion is only possible within a limited range of retinal disparities known as Panum's fusional area (Schor, Heckmann, & Tyler, 1989; Steinman, Steinman, & Garzia, 2000). Larger disparities typically trigger a motor fusional response, or cause diplopia. Motor fusion comprises of the aforementioned physiological mechanisms of vergence. That is, in subjects with normal binocular vision, slow disconjugate eye movements mainly triggered by retinal disparity are made in order to adjust the angle between the two visual axes (Schor, 1979).

To summarise, during reading, the visual system is primarily faced with horizontal disparities, which might be the reason why research in written language processing has focused mainly on horizontal binocular coordination (Blythe et al., 2010, Liversedge et al., 2009, see Kirkby et al., 2008 for review). Indeed, few studies so far have systematically investigated misalignments in reading in other dimensions, a limitation to the comprehensive understanding of binocular coordination that the current work aimed to address.

When conceptualizing the visual system's response to binocular misalignment, it is important to note that binocular motor fusion is characterised by horizontal vergence (along a plane containing the interocular axis), vertical vergence (along a plane orthogonal to the interocular axis) and cyclovergence (in opposite directions along the two visual axes, Boman & Kertesz, 1981; Howard & Rogers, 2012). While a significant body of work has investigated vergence movements driven by horizontal misalignments, the literature concerning responses to vertical and torsial disparities is considerably limited, particularly in the context of lexical processing. Although Nuthmann and Kliegl (2009) recently reported the presence of vertical misalignments in each reading fixation, their findings regarding vertical disparity were purely descriptive and no claims were made about any potential vertical vergence adjustments during fixations. In addition, Jainta, Blythe, Nikolova, Jones and Liversedge (2014) recently conducted a detailed investigation of disparities occurring during natural sentence reading. They reported that vertical disparities were of much smaller magnitude than horizontal disparities, and suggested that the limited activation of the vertical vergence system during reading could be due to functional differences between horizontal and vertical disparities and disparity reducing mechanisms in relation to maintaining a single unified perception of the written text. Aside from the two abovementioned accounts, no studies so far have systematically investigated the motor fusional response to stereoscopically imposed vertical disparities during lexical processing. Nevertheless, existing studies in non-reading tasks indicate that while serving complementary functions, horizontal and vertical vergence are considered as two different mechanisms (Howard & Rogers, 2012; Stevenson, Lott, & Yang, 1997). Research investigating the characteristics of vertical vergence revealed that when compared to its horizontal counterpart, it is limited in both amplitude and speed (Bharadvaj et al., 2007; Kertesz, 1981). Furthermore, Panum's fusion area has been shown to be elliptical in shape, that is, sensory fusion is possible over a larger range of horizontal disparities than vertical disparities (Fender & Julesz, 1967;Howard & Rogers, 1995; Jainta et al., 2014; Schor & Tyler, 1981). Interestingly, a recent study by Dysli, Vogel and Abegg (2014) investigated the assumption that latent heterophoria may be causally involved in reading problems. The authors changed the vergence tone of participants without reading difficulties using prisms that induced exophoria, esophoria and vertical phoria. It was found that none of the prism conditions affected reading speed, average fixation duration or saccadic amplitudes during paragraph reading. However, it is as yet unclear whether induced vertical disparity in written linguistic stimuli would affect the efficiency of lexical processing, or indeed what vergence adjustments would be made to compensate for any vertical misalignments.

One study to experimentally increased disparity within written linguistic stimuli during lexical processing was conducted by Blythe et al. (2010). Using dichoptic presentations of single words with varying levels of horizontal offset, they estimated that the size of Panum's fusional area for linguistic stimuli was equal to approximately one character space for both children and adults. However, they did not include a frequency manipulation in their stimuli, focusing instead on the differences between the two participant groups. Another study to use dichoptic visual presentations during reading was conducted by Liversedge et al. (2006a) and explored binocular saccadic targeting. The authors found that conjugate eye movements in reading appear to be programmed on the basis of a combined signal sent to both eyes and that saccades in reading were targeted on the basis of a fused percept attained at an early processing stage. Both studies raise interesting questions regarding the response of the vergence and saccadic targeting system to stereoscopically presented vertical disparities during lexical processing.

We therefore set out to conduct a detailed investigation of vertical motor fusion in response to symmetric vertical offset during a lexical decision task. There were several aims to the study. Firstly, we were interested in the vertical vergence response to binocular image misalignment and its effect on lexical identification processes. Secondly, we investigated the sensitivity of saccade targeting mechanisms to vertical disparity in the parafovea. Finally, as a more specific exploration, we aimed to investigate the influence of the vertical stereoscopic disparity manipulation on a well-established finding in reading research: the frequency effect, or the increased efficiency of lexical processing for commonly occurring words (Inhoff & Rayner, 1986; Rayner, 1998; White, 2008). The theoretical motivation for this investigation is discussed in the context of the Interactive Activation (IA) model of word recognition (McClelland & Rumelhart, 1981). It is possible that the fusion of binocular inputs, both motor and sensory, is achieved at an earlier and separate stage of processing than lexical identification, prior to the feature extraction stage of the IA model. If that were the case, adding a level of complexity at the fusion stage of processing in the form of a disparity manipulation would cause an equal global increase in total reaction times (RTs) for both high-frequency (HF) and low-frequency (LF) words. However, it is also possible that visual fusion interferes with the feature extraction stage of processing, as fusion is central for attaining high quality binocular visual information. Therefore, making feature extraction more difficult by imposing vertical disparity in the stimuli would initially slow down the processing of both HF and LF words, but at the following (letter and word) stages of lexical identification, HF words would be processed faster. In other words, there might be an interaction between the two factors, such that the cost of adding complexity at the visual fusion stage would be larger for LF than for HF words. An alternative possibility would be that when presented with induced disparities within the range of those observed in normal reading, the vertical vergence system would remain inactive, which would indicate that vergence responses to this type of disparity are quite different to those associated with horizontal disparities. This in turn would be consistent with the claims of Jainta et al. (2014), who argue that vertical disparities provide much less useful information for stereopsis than do horizontal disparities given the horizontal alignment of the two eyes in the human visual system.

Based on previous findings, we made several predictions. We expected that, similar to Blythe et al. (2010), there would be a time cost associated with attaining a stable unified percept of the disparate dichoptic stimuli, which would be reflected in RTs on the lexical decision task. Furthermore, if participants found it impossible to fuse the imposed vertical disparities due to the vertical vergence limitations of the visual system, they might be unable to perform the lexical decision task, as it would be extremely difficult to distinguish the words from the non-words (Fig. 1; see also Blythe et al., 2010). Although we only attempted to actively drive vertical vergence, we expected that a small amount of horizontal vergence would likely be observed following a horizontal saccade. More critically, if the vertical disparity presentation triggered a vertical vergence response we would likely observe additional changes in horizontal fusional responses that typically occur in reading. In terms of saccadic programming, we expected that if saccades to dichoptically presented parafoveal targets were programmed on the basis of the individual input received by each eye, then that would be reflected in the direction and magnitude of the resulting fixation disparity. Finally, in terms of lexical processing, any potential interaction between the vertical disparity presentation and the lexical frequency manipulation would be informative as to the degree of interdependence between visual processes related to fusion and linguistic processes related to lexical identification. Such an interaction was observed in a recent study by Jainta, Blythe and Liversedge (2014), who found that the efficiency of lexical processing was diminished in monocular reading conditions. On the other hand, previous findings (Blythe et al., 2006; Juhasz et al., 2006) reported no influence of lexical frequency and orthographic manipulations on horizontal binocular disparity. Therefore, we explored whether vertical binocular disparity would interact with lexical processing, or if it would have an additive effect on total processing times for both high-frequency (HF) and low-frequency (LF) words.

2. Method

2.1. Participants

Participants were 8 native English speakers from the University of Southampton, who took part in the experiment in exchange for Psychology course credits, or payment at the rate of £6 per hour. All participants had normal or corrected to normal vision (with soft contact lenses) and no diagnosed reading difficulties. Testing their visual acuity with a Landolt C acuity chart confirmed that there were no considerable differences in acuity between the two eyes (best-corrected acuity in each eye in decimal units was 1.00 or higher). Additionally, a Titmus Stereotest indicated that all participants had functional stereopsis (minimal stereoacuity of 40 seconds of arc).

2.2. Apparatus

Binocular eye movements were measured using two Fourward Technologies Dual Purkinje Image (DPI) eye trackers, which recorded the position of both eyes every millisecond (sampling rate of 1000 Hz, spatial resolution < 1 min arc). Stereoscopic presentation of the target items was achieved through use of Cambridge Research Systems FE1 shutter goggles, which blocked the visual input received by each eye alternatively every 8.33 ms (corresponding to a 120 Hz refresh rate). The shutter goggles were synchronized with the eye trackers and interfaced with a Pentium 4 computer and a Philips 21B582BH 21” monitor. The experimental equipment made it possible to simultaneously track binocular eye movements whilst manipulating the unique visual input received by each eye. The monitor was situated at a viewing distance of 100 cm. To minimize head movements, participants leaned against two cushioned forehead rests and bit on an individually prepared bite bar.

2.3. Materials and Design

All participants viewed 208 trials, each consisting of a single 6-letter item. The item was either one of 52 high-frequency (HF) words (e.g., *summer*), one of 52 low-frequency (LF) words (e.g., *acumen*) or one of 104 non-words (e.g., *worzer*). Non-words were formed in a similar fashion to Blythe et al. (2010) by substituting a single letter in the center of a word and creating an obvious misspelling (e.g., *summer* to *sumxer*). The 52 HF items had an average frequency of 118.48 counts per million (ranging from 18 to 850) and the 52 LF items had an average frequency of 2.58 counts per million (ranging from 0 to 9), as indexed in the English language CELEX lexical database (Baayen, Piepenbrock, & Gulikers, 1995). A *t*-test confirmed that HF words were significantly more frequent than LF words, t (51) = 5.31, $p < .001$.

All items were presented in red 20pt Courier New font on a black background. At the viewing distance of 100 cm, each letter height extended to 0.32 deg of visual angle. Each of the items was viewed by participants in one of four dichoptic presentation conditions: (1) aligned, where the two images were centered on the display monitor; (2) offset vertically by a total of 0.05 deg, (3) offset vertically by a total of 0.11 deg; (4) offset vertically by a total of 0.16 deg. The disparity presentation was symmetrical, i.e., the monocular images were offset by an equal amount in opposite directions in each eye. Conditions were counterbalanced such that every word appeared in each of the experimental conditions across participants. Additionally, whether the image presented to the left eye appeared above or below the image presented to the right eye was randomized and counterbalanced across conditions.

---------------- Insert Figure 1 about here -----------------

2.4. Procedure

The experimental procedure was approved by the University of Southampton Ethics and Research Governance Office and followed the conventions of the Declaration of Helsinki. Informed written consent was obtained from each participant after explanation of the procedure of the experiment.

Each trial consisted of a fixation point appearing on the left-hand side of the screen for 1 second, followed by the item (word or non-word) presented in the centre of the screen. The distance between the fixation point and the left edge of each stimulus/item was 2.54 deg visual angle.

Participants were instructed to look at the fixation point before looking at the word presented on the screen. They were asked press a button to indicate as quickly and accurately as possible whether the stimulus was a word or a non-word. They were not told that some of the stimuli would be presented with varying degrees of disparity. There were four practice trials to help participants become familiar and comfortable with the task.

Calibration was monocular (i.e., the left eye was occluded by the shutter goggle during calibration of the right eye, and vice versa). For calibration, participants were instructed to look at each of nine points in a 3x3 grid in a set sequence from the top left to the bottom right. Horizontal separation of the calibration points was 3.44 deg and the vertical separation was 1.26 deg relative to screen centre. Afterwards, the calibration was checked for accuracy and repeated if necessary. Once both eyes had been calibrated successfully, the experiment began. Calibration was checked for accuracy following every four trials and, if the drift in eye position was more than 0.06 degrees, the eye trackers were recalibrated.

2.5. Analyses

Custom-designed software was used for the data analyses. Saccades and fixations were manually identified in order to avoid contamination by dynamic overshoots (Deubel & Bridgeman, 1995) or artefacts due to blinks. From the separate signals of the two eyes, we calculated the horizontal and vertical conjugate eye component [(left eye + right eye)/2; i.e., the version signal] and the horizontal and vertical disconjugate eye component [left eye – right eye; i.e., the vergence signal]. Several parameters of binocular coordination were calculated for each fixation period: (1) vertical fixation disparity at the start and end of fixations, where a value of 0 represents alignment of the two eyes at eye height; positive values represent left-hyper fixations and negative values represent right-hyper fixations; (2) horizontal fixation disparity at the start and end of fixation; a value of 0 represents alignment of the two eyes at the depth of the screen, positive values represent crossed fixations, where the point of fixation is in front of the screen, and negative values represent uncrossed fixations, where the point of fixation is behind the screen; (3) net vertical and horizontal drift in vergence (Jainta et al., 2010; Liversedge, White, et al., 2006; Nuthmann & Kliegl, 2009; Vernet & Kapoula, 2009), which is the change in fixation disparity between the beginning and the end of the fixation period and (4) total change in vertical and horizontal eye position between the beginning of the first fixation and the end of the final fixation on each item. In addition, we calculated total reaction time (RT) and total number of fixations for each item.

For data analyses, we used linear mixed-effects models (lmer from package lme4 (Pinheiro & Bates, 2000) in R (R Development Core Team, 2009). P-values were estimated using posterior distributions for the model parameters obtained by Markov Chain Monte Carlo sampling, which include a typical sample size of 10000 (Baayen, Davidson, & Bates, 2008). The model was applied to the non-aggregated data and participants and items were treated as random effects, while lexical frequency (HF vs. LF) and binocular image disparity (0 deg, 0.05 deg, 0.11 deg or 0.16 deg) were treated as fixed effects.

3. Results

In the following sections, we report a variety of analyses based on different eye movement measures. Approximately 1.5% of the data were excluded due to tracker loss, resulting in a total of 5657 fixations on which the following analyses are based. We begin by giving a short descriptive account of the overall findings from the lexical decision task (3.1), then focus on the initial reaction of the vergence system to our disparity manipulation (3.2), changes in disparity throughout the duration of an entire multiple fixation trial (3.3.) and cases where only one fixation was made per trial (3.4.). Before reporting further results regarding eye movement measures, it is important to clarify certain terms that will be used throughout the following sections. Binocular image disparity refers to the induced offset between the dichoptic images presented on the screen. Binocular fixation disparity refers to the differences in position between the left and the right eye in degrees of visual angle, as measured by the eye trackers.

3.1. Lexical decision accuracy, reaction times (RTs) and number of fixations

The overall response accuracy in this experiment was 96%. Correct responses during lexical identification were taken as the behavioural indication that participants were able to successfully fuse the binocularly misaligned images. Table 1 contains information about participants' accuracy at the lexical decision task, mean RTs, fixation durations and number of fixations in all of the frequency and disparity conditions. Evidently, participants responded faster, made fewer fixations and were more accurate when identifying HF words than LF words and non-words.

---------------- Insert Table 1 about here -----------------

To further explore the frequency effect, an LME analysis was applied to the log-transformed RT values with participants and items as random effects and frequency and binocular image disparity as fixed effects. The results revealed a significant effect of frequency: participants were faster at identifying HF words than LF words ($t = 4.24$, $p < .001$) and non-words ($t = 5.13$, $p < .001$), with no significant difference between the latter two ($t < 1$). The size of the frequency effect was approximately 145 ms on average. There was no effect of binocular image disparity ($t = 1.13$, $p = .24$) and the interaction between the two fixed effects was not close to significant ($t < 1$). These results are also summarised in Figure 2. Clearly, participants were able to perform the lexical decision task without any interference from the vertical disparity manipulation. The following sections explore this by focusing on vergence responses during fixations.

----------------- Insert Figure 2 about here -----------------

3.2. Initial reaction to vertical disparity

With regard to binocular landing positions, *Figure 3* represents the distribution of disparities at the start of the first fixation on each item, plotted onto a Cartesian coordinate system. Positive values on the x-axis denote crossed disparities, and positive values on the y-axis represent left hyper-vertical disparities (where the left eye is fixating above the right eye). Negative values on the x-axis correspond to uncrossed disparities, and negative values on the x-axis represent right hyper-vertical disparities (where the right eye is fixating above the left eye). The data clearly indicate that horizontal disparities were predominantly uncrossed, while vertical disparities were predominantly left-hyper.

----------------- Insert Figure 3 about here -----------------

Furthermore, we were interested in the sensitivity of the saccade programming system to vertical disparity in the parafovea. More specifically, we explored the relationship between the nature of the dichoptic presentation (left-hyper or right-hyper) and the resulting disparity at the start of each trial. Close correspondence between the two categorical variables would indicate that during the initial saccade onto the stimulus, each of the eyes targeted the monocular image presented to it separately via the shutter goggles. Recall that 75% of our stimuli were presented with some degree of vertical misalignment. Regardless of presentation condition, 38% of vertical disparities at the start of the initial fixation were right-hyper and 62% were left-hyper. A Chi-square test revealed that right-hyper and left-hyper disparities did not closely correspond to presentation conditions. In fact, left-hyper disparities were the predominant case, regardless of the binocular image manipulation (X^2 (1) = 15.10, $p < .001$).

The following part of the analyses focuses on how the vergence system responded when presented with a disparate image upon initial fixation on the target on trials with multiple fixations. We only included fixation disparities and fixation durations within 2SD of each participant's mean, which resulted in exclusion of approximately 3% of the data. 1375 fixations in total were analysed. *Figure 4* illustrates the distribution of horizontal and vertical disparities at the start (a) and at the end (b) of the initial fixation. The distribution of vertical disparities in both cases is clearly more leptokurtic, indicating that vertical disparities are generally smaller in magnitude than horizontal disparities. Fixation disparity at the start was not significantly affected by either lexical frequency, disparity manipulation or the interaction between the two (*t*s < 1, n.s.). Disparities at the end of fixations were also not influenced by either manipulation (all *p*s > .13). The average vertical disparity was 0.12 deg (*SD* = 0.09) at the start of the initial fixation and 0.11 deg (*SD* = 0.10) at the end of the fixation. A *t*-test revealed no difference in the drift in vertical fixation disparity throughout the fixation ($t < 1$). These results indicate that no considerable vertical vergence movements were made during the initial fixation on the target.

----------------- Insert Figure 4 about here -----------------

Interestingly, however, we observed a small but significant change in horizontal disparity during the initial fixation. Horizontal disparities at the start of the fixation had an average magnitude of 0.18 deg (*SD* = 0.15), which was reduced to 0.16 deg (*SD* = 0.14) by the end of the fixation, $t = 2.98$, $p < .01$. A tendency for disparity-reducing vergence movements seemed to emerge as early as the first fixation on an item, even in the absence of any horizontal stereoscopic manipulation. This was not affected by either the frequency or the binocular image manipulation (*t*s < 1). Furthermore, there was no significant correlation between the magnitude of horizontal and vertical fixation disparities at the start or at the end of the initial fixation (*p*s > .19, n.s.). In other words, we observed a rapid horizontal vergence response during the first fixation on each item, following the horizontal saccade onto the stimulus, but no vertical vergence response to our disparity manipulation. The following sections further explore this pattern across all fixations made during a trial.

3.3. Reaction to vertical disparity throughout an entire trial

The previous sections demonstrate that the vertical and horizontal vergence system seem to make very different initial responses to parafoveal stereoscopic targets. Recall, however, that participants typically made more than one fixation on each item, hinting at the possibility that vergence movements occurred after the initial fixation. Therefore, a comparison was made between the start of the first fixation and the end of the final fixation on each multiple fixation item in order to capture any change in vergence throughout the duration of each trial. There was no significant difference in vertical fixation disparity between the two measures ($t < 1$, $p > .16$). In addition, an LME analysis investigated the magnitude of change in vertical fixation disparity between the start of the initial fixation and the end of the final fixation. There was no significant effect of lexical frequency, binocular image disparity, or the interaction between the two fixed factors (*t*s < 1). The average magnitude of vertical fixation disparity at the end of the final fixation on each item was 0.16 deg (*SD* = .13). Considering that the last fixation of each trial was the one during which participants pressed the button to indicate their lexical decision, this mean magnitude could be taken as an approximation of the amount of vertical fixation disparity which the visual system could easily tolerate in order to successfully process lexical information. Note, however, that no disturbances in fusion were reported by any of the participants, suggesting that the vertical limits of Panum's fusional area are likely larger than the reported value.

As for horizontal disparity, a consistent vergence response was observed throughout the duration of each trial: participants displayed a tendency for disparity-reducing vergence movements and a transition from uncrossed to aligned binocular disparities. This effect was significant ($t = 4.12$, $p < .001$), despite the absence of a stimulus that was intended to actively drive horizontal vergence.

Horizontal disparities at the end of the final fixation on each trial were on average 0.02 deg smaller than they were at the start. In addition, an LME analysis confirmed that horizontal vergence measures at the end of each trial were not affected by the vertical disparity manipulation ($t < 1$), nor were they correlated with the magnitude of vertical fixation disparity, $r\ (1373) = .03$, $p = .29$.

3.4. Single fixation trials

As a final step in our investigation, we explored cases in which only one fixation was made per trial. It was important to include single fixation trials in the analyses, as they would undoubtedly provide insight into any potential interactions between low-level visual processes involved in disparity processing and high-level lexical identification processing. In addition, cases in which vertical disparity was dealt with in a single fixation would enable us to closely monitor any potential vergence responses to our vertical manipulation. Note, however, that single fixations were made on only 17% of trials. Data were included in the analyses if fixation duration, horizontal and vertical disparities at the start and the end of fixations fell within 2 *SD* of each participant's mean. This resulted in 5% data loss – 240 fixations in total were analysed.

Reaction time data are presented in Table 1. A significant lexical frequency effect of 119ms was observed in single fixation trials, $t = 3.38$, $p < .001$. There was no significant effect of binocular image disparity or the interaction between the two fixed effects (*t*s < 1). Therefore, it appears that single fixation trials did not differ significantly from multiple fixation trials in terms of participants' responses during the lexical decision task. Again, it is evident from these results that although a robust frequency effect was observed in the data, it was not affected by the visual disparity manipulation.

As for disparity measures, the mean magnitude of vertical disparity was 0.12 deg ($SD = 0.10$) at the start and 0.11 deg ($SD = 0.09$) at the end of single fixation trials. No significant vertical vergence movements were observed throughout the fixation ($t < 1$). In addition, LME analyses revealed that vertical disparities at the start and the end of the fixations were not affected by the frequency manipulation ($t_{start} = 1.55$, $p = .12$; $t_{end} < 1$), the disparity manipulation ($t_{start} = 1.63$, $t_{end} = 1.54$, $ps = .11$) or the interaction between the two fixed effects (*t*s < 1). However, once again we observed a consistent disparity-reducing vergence response in the horizontal dimension. A tendency emerged for horizontal disparities to move from uncrossed to aligned throughout a fixation. The mean magnitude of disparity was 0.15 deg (SD = 0.13) at the start and 0.12 (SD = 0.10) at the end of the trial. Disparity was reduced by an average of 0.03 deg throughout the duration of the fixation, $t = 3.97$, $p < .001$. There was no significant correlation between the magnitude of horizontal and vertical disparity at the end of fixations ($r\ (238) = .09$, $p = .19$), and an LME analysis revealed no effect of the vertical disparity manipulation on horizontal disparity measures ($t = 1.31$, $p = .19$).

4. Discussion

Binocular coordination is critical to successfully attaining a fused stable representation of the visual environment, which is essential for performing a variety of tasks, including reading. Recent findings have begun to explore the role of binocularity in reading, the way it affects language processing and the relative importance of various binocular visual processes for written language comprehension (see Kirkby et al., 2008 for review). The present study adds to that growing literature by making an exploration of the role of vertical binocular disparities in lexical processing. We focused on investigating the motor fusional response to induced vertical misalignments in the parafovea and, upon fixation, its potential influence on horizontal vergence movements that typically occur following saccades in reading, and its effect on lexical identification.

Our findings revealed that when participants made a horizontal saccade onto a centrally presented stimulus with induced vertical disparity, no change was observed in the vertical vergence system. That is, participants did not make significant disparity reducing vertical vergence movements during the initial saccade, nor when first fixating on the target or even throughout the duration of a trial. Importantly, there was a clear dissociation between the presentation on the monitor and the perceptual experience of our participants. Their subjective reports did not indicate any experience of diplopia or visual disturbances, or any awareness of our manipulation. This was further evidenced by the high lexical decision accuracy in all disparity conditions, as well as the robust frequency effect we observed across single and multiple fixation trials. The present findings are in direct contrast to the vergence responses to words presented with a horizontal disparity observed by Blythe at al. (2010), who used dichoptic presentation of single words with equal amount of horizontal offset (from 0 to 0.74 deg in total, or up to 2 character spaces), and reported that the measured vergence responses were rapid and direction-appropriate.

Furthermore, we were interested in the sensitivity of the saccadic targeting system to disparity in the parafovea. Previous findings regarding horizontal disparity have revealed that the vergence system reacts actively to disparity from fixation onset, but makes no adjustments during saccades when stimuli are presented stereoscopically (Blythe et al., 2012; but see Kapoula, Eggert & Bucci, 1995 for an alternative account). Furthermore, Blythe et al. (2010) observed that when making a horizontal saccade onto a stereoscopic stimulus, participants targeted the preferred viewing location (O'Regan, 1981; Rayner, 1979) for an unfused letter string with a length equal to the combined length of the two monocular images. For instance, if a 6-letter word was presented independently to each eye with 2 character spaces of stereoscopic disparity, then the resulting letter string appeared 8 characters long on the screen. This is an important point to consider: it appears that when inducing horizontal disparity within single words, the disparate images were combined, but not fused prior to fixation. In addition, Liversedge et al. (2006a) presented different parts of a target word within a sentence individually to each eye (e.g. for the word cowboy, cowb was only seen by one eye and wboy was only seen by the other eye). They found that saccades were targeted to stereoscopic lexical stimuli based on a combined percept, regardless of which constituent of the target word was available to which eye monocularly. Recall that the majority of the stimuli in the present study were presented with some degree of vertical disparity and we explored the relationship between the direction of the visual offset in the stimuli and the resulting fixation disparity, as measured by the eye-trackers. Evidence for close correspondence between the two categorical variables would hint at the possibility of independent monocular saccade targeting, as outlined by Liversedge et al. (2006a), which would in turn violate Hering's law of equal innervation. Our findings, however, indicated otherwise. Similar to Liversedge et al. (2006a), the present results indicated that landing positions on the vertically disparate stimuli were not affected by the direction of the visual presentation. Vertical disparities at the start of the initial fixation on the target were predominantly left-hyper, regardless of whether the left monocular image appeared above or below the right monocular image. The left-hyper predominance was also observed in trials where the dichoptic images were presented without disparity. In addition, the magnitude of vertical fixation disparity at the start or at the end of each trial was not affected by the magnitude of binocular image disparity present on the screen. Indeed, vertical disparities larger than 1 character space were only measured on less than 10% of fixations, regardless of the fact that in 75% of trials the vertically disparate stimuli exceeded the height of one character by up to 50%. Therefore, it appears that when presented with a relatively small magnitude of vertical disparity in the parafovea, participants performed parallel saccades in both eyes, regardless of the vertical disparity in the stimulus. That is to say, the two monocular dichoptic images on the screen did not appear to have been used as separate saccade targets for each eye.

Interestingly, while participants made no vertical vergence movements in response to the vertical disparity manipulation, we observed significant systematic horizontal vergence movements as early as the first fixation on the stimulus, even in the absence of a horizontal disparity manipulation. In other words, the horizontal motor fusional system was automatically activated following a horizontal saccade, as is typically observed in normal reading, whereas the vertical system showed no significant activation. Importantly, we found no correlation between the magnitude of horizontal and vertical disparity and drift measures, indicating that in the current study, the two systems did not interact during lexical processing. Furthermore, the LME analyses found no effect of the vertical disparity manipulation on horizontal disparity magnitude and drift measures. All these findings suggest that the horizontal and vertical vergence systems react differently to imposed vertical disparities, which is compatible with early studies investigating horizontal and vertical vergence responses to symmetric disparity presentations (Perlmutter & Kertesz, 1978). Future studies would ideally investigate the interaction between horizontal and vertical vergence when disparities are induced in both dimensions simultaneously, as well as the degree of automaticity in horizontal vergence during sentence reading.

Another potential direction for future research would be to explore the natural vertical limitations of fusion, that is, the thresholds for vertical vergence in reading. Such research could show how vertical fusion limits could impact on reading processes and more specifically, potentially interact with reading difficulties. However, recent work by Dysli et al. (2014) demonstrated that inducing vertical phoria (vertical disparity) of up to 2 prism diopters (approximately 1 degree of visual angle) had no effect on reading performance while participants read a paragraph of text aloud. Note though that differences exist between eye movements during silent reading and reading aloud (e.g. longer fixation durations in the latter condition, see Rayner, 1998 for review).

When contrasting our findings about vertical disparity patterns with those reported by Nuthmann et al. (2009), several points become immediately apparent. Firstly, we observed a larger proportion of exo (uncrossed) than eso (crossed) horizontal disparities, while Nuthmann and colleagues reported the opposite pattern. These differences in the direction of horizontal disparities, as reported in different studies, have been discussed in detail elsewhere (Kirkby et al., 2013). Importantly, it has been suggested that viewing conditions associated with different data acquisition techniques (e.g., light text over dark background or vice versa) amongst a variety of other factors, such as font colour and viewing distance, might affect the pattern of horizontal disparities in reading. As for vertical disparities, we observed the same left-hyper predominance in all induced vertical disparity conditions as Nuthmann et al. (2009) observed during sentence reading. It appears, therefore, that vertical disparities that occur during language processing are much less sensitive to viewing conditions than their horizontal counterpart. Furthermore, our findings regarding the range of horizontal and vertical disparities over which fusion is possible are compatible with the notion of an elliptical pattern for Panum's fusional area, indicating that fusion operates over a limited range of vertical disparities and a larger range of horizontal disparities. More critically, the vertical motor fusional mechanisms showed limited activation, even in the presence of a disparity manipulation designed to elicit a vergence response. While these findings differ from Nuthmann et al.'s (2009) report of approximately equal magnitude of horizontal and vertical disparity in reading, the present data fit neatly with studies in non-reading tasks, which suggest that the vertical limitations in Panum's area are caused in part by the visual system's diminished capacity to compensate for vertical misalignments with disparity reducing vergence movements (Houtman, Roze, & Scheper, 1977; Steinmann et al., 2000). This is also consistent with Jainta et al.'s (2014) accounts of vertical disparity in normal reading, and suggests that the difference in activation between the two oculomotor systems may be due to the separate but complementary functions that they serve.

In addition, the functional differences between vertical and horizontal fusional mechanisms are particularly relevant to understanding of the interplay between visual and linguistic processes in the present experiment. Our findings indicated that word identification was not disturbed by the particular nature of the binocular presentation. We observed no interaction between lexical frequency and vertical disparity, but also found no additive effect of the disparity presentation on global processing times for HF and LF words. A robust significant frequency effect was observed, regardless of the magnitude of disparity present in the stimuli. These findings are different from those reported by Blythe et al. (2010), who found that increasing horizontal disparity also increased the time taken to make a lexical decision. Note, however, that the magnitude of disparity they introduced in their stimuli was larger that the present experiment. In addition, their study did not include a lexical frequency manipulation, and they only reported the effect of induced disparity on total trial viewing times. Jainta et al. (2014), on the other hand, observed that presenting text monocularly, rather than binocularly, significantly reduced the frequency effect for HF words. Although we are cautious when comparing data from natural reading and lexical decision experiments, what we can nevertheless glean from those findings is that in the present study, despite the disparity manipulation, participants were able to derive the benefits of binocular vision during word identification and display the well-documented increased efficiency of lexical processing for HF words. It is likely that a fused percept of our stimuli was obtained at an early stage of visual processing, possibly prior to the feature extraction stage of lexical identification. Furthermore, it may well be the case that induced vertical disparities of the magnitude typically observed in reading caused no disturbance in lexical processing because they are informative in a different way to horizontal disparities. As Jainta et al. (2014) suggested, this dissociation between the two oculomotor responses is very likely due to the physical arrangement of the visual system and the resulting effect on binocular coordination, the computation of depth and stereopsis.

In conclusion, the present study demonstrated that during lexical identification, the visual system responds differently to stereoscopic vertical disparity than it does to horizontal disparity. Our findings suggest that the visual system programs saccades to vertically misaligned lexical stimuli based on a fused percept attained at an early stage of processing, as indicated by the observed pattern of landing positions and the reported vergence and disparity measures. Further work is needed to investigate the response of the visual system to induced disparities in all directions during lexical processing in order to quantify the degree of interdependence between horizontal and vertical fusional mechanisms.

Word count: 7000 words

References

Blythe, H. I., S. P. Liversedge, et al. (2010). The effective fusional range for words in a natural viewing situation. *Vision Research, 50,* 1559-1570.

Blythe, H.I., Liversedge, S.P., Joseph, H.S.S.L., White, S.J., Findlay, J.M., & Rayner, K. (2006). The binocular co-ordination of eye movements during reading in children and adults. *Vision Research, 46,* 3898-3908.

Bharadwaj, S.R., Hoening, M.P., Sivaramakrishnan, V.C., Karthikeyan, B., Simonian, D., Mau, K., Rastani, S., & Schor, C.M. (2007). *Investig. Ophthalmol. & Vis. Sciences, 48* (4), 1592-15600.

Boman, D.K., & Kertesz, A.E. (1983). Interaction between horizontal and vertical fusional responses. *Perception and Psychophysics, 33*, 565–570.

Collewijn, H., Erkelens, C. J., & Steinman, R. M. (1988). Binocular coordination of human horizontal saccadic eye movements. *Journal of Physiology, 404,* 157–182

Deubel, H. & B. Bridgeman (1995). Fourth Purkinje image signals reveal lens deviations and retinal image distortions during saccadic eye movements. *Vision Research 35*, 529-538.

Dysli, M., Vogel, N., & Abegg, M. (2014). Reading performance is not affected by a prism induced increase in horizontal and vertical vergence demand. *Frontiers in Human Neuroscience, 8*, 431. doi: 10.3389/fnhum.2014.00431

Fender, D., & Julesz, B. (1967). Extension of Panum's fusional area in binocularly stabilized vision. *Journal of the Optical Society of America, 57,* 819–830.

Hendriks, A. (1996). Vergence movements during fixations in reading. *Acta Psychologica, 92*, 131–151.

Hering, E. (1868). The theory of binocular vision. In B. Bridgeman and L. Stark (Eds.). New York: Plenum Press.

Houtman, W. A., Roze, J. H., & Scheper, W. (1977). Vertical motor fusion. *Documenta Ophthalmologica, 44,* 179–185.

Howard, I. P., & Rogers, B. J. (1995). Binocular vision and stereopsis. New York: Oxford University Press.

Howard, I. P., & Rogers, B. J. (2012). *Perceiving in depth. Volume 2, Stereoscopic vision.* Oxford: Oxford University Press.

Inhoff, A.W., & Rayner, K. (1986). Parafoveal word processing during eye fixations in reading: Effects of word frequency. *Perception and Psychophysics, 40*, 431-439.

Jainta, S., Blythe, H.I., & Liversedge, S.P. (2014). Binocular advantages in reading. *Current Biology, 24,* 526-530.

Jainta, S, Blythe, H.I., Nikolova, M., Jones, M.O., & Liversedge, S.P. (2014). *A comparative analysis of vertical and horizontal fixation disparity in sentence reading.* Manuscript submitted for publication.

Jainta, S., Hoorman, J., Kloke, W. B., & Jaschinski, W. (2010). Binocularity during reading fixations: Properties of the minimum fixation disparity. *Vision Research, 50*, 1775–1785.

Jainta, S. & Jaschinski, W. (2012). Individual differences in binocular coordination are uncovered by directly comparing monocular and binocular reading conditions. Investigative Ophthalmology and Visual Sciences, 53, 5762-5769.

Juhasz, B.J., Liversedge, S.P., White, S.J., & Rayner, K. (2006). Binocular coordination of the eyes during reading: Word frequency and case alteration affect fixation duration but not fixation disparity. *Quarterly Journal of Experimental Psycgology, 59 (9),* 1614-1625.

Kapoula, Z., Eggert, T., & Bucci, M.P. (1995). Immediate saccade amplitude disconjugacy induced by unequal images. *Vision Research, 35*, 3529-3540.

Kertesz, A. E. (1981). Effect of stimulus size on fusion and vergence *Journal of the Optical Society of America, 71*, 289–293.

Kirkby, J.A., Blythe, H.I., Drieghe, D., Benson, V. and Liversedge, S.P., 2013. Investigating eye movement acquisition and analysis technologies as a causal factor in differential prevalence of crossed and uncrossed fixation disparity during reading and dot scanning. *Behavior Research Methods*, 45 (3), 664-678.

Kirkby, J. A., Webster, L. A. D., Blythe, H. I., & Liversedge, S. P. (2008). Binocular coordination during reading and non-reading tasks. *Psychological Bulletin, 134*, 742–763.

Leigh, R. J. & D. S. Zee (2006). The neurology of eye movements. New York, Oxford University Press.

Liversedge, S.P., Rayner, K., White, S.J., Findlay, J.M., & McSorley, E. (2006a). Binocular coordination of the eyes during reading. *Current Biology, 16,* 1726-1729.

Liversedge, S.P., White, S.J., Findlay, J.M., & Rayner, K. (2006b). Binocular coordination of eye movements during reading. *Vision Research, 46,* 2363-2374.

McCelland, J.L. & Rumelhard, D.E. (1981). An interactive activation model of context effects in letter perception: I. An account of basic findings. *Psychological Review, 88 (5),* 375-407.

Nuthmann, A. & Kliegl, R. (2009). An examination of binocular reading fixations based on sentence corpus data. *Journal of Vision, 9* (5), 1-28.

Perlmutter, A.L. & Kertesz, A.E. (1978). Measurement of human vertical fusional response. *Vision Research, 18* (2), 219-223.

Pratt-Johnson, J.A. & Tillson, G. (2001). Management of strabismus and amblyopia: A practical guide. New York: Thieme.

R-Development-Core-Team (2008). R: A Language and Environment for Statistical Computing, (< http: //www.R-project.org >).

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124* (3), 372-422.

Schor, C.M. (1979). The relationship between fusional vergence eye movements and fixation disparity. *Vision Research, 19* (12), 1359-1367.

Schor, C.M. & Tyler, C.W. (1981). Spatio-temporal properties of Panum's fusional area. *Vision Research, 21*(5), 683-692.

Schor, C.M. & Ciuffreda, K.J. (1983). Vergence eye movements: Basic and clinical aspects. Boston, MA: Butterworth.

Schor, C.M., Heckman, T., & Tyler, C.W. (1989). Binocular fusion limits are independent of contrast, luminance gradient and component phases. *Vision Research, 29*, 837-847.

Steinman, S. B., B. A. Steinman, & Garzia, R.P. (2000). Foundations of binocular vision: A clinical perspective. New York: The MacGraw-Hill Companies.

Stevenson, S.B., Lott, L.A., & Yang, J. (1997). The influence of subject instruction on horizontal and vertical vergence tracking. *Vision Research, 37,* 2891-2898.

van Leeuwen, A.F., Westen, M.J., van der Steen, J., de Faber, J.T.H.N., & Collewijn, H. (1999). Gaze-shift dynamics in subjects with and without symptoms of convergence insufficiency: Influence of monocular preference and the effect of training, *Vision Research, 39*, 3095–3107.

Vernet, M. & Z. Kapoula (2009). Binocular motor coordination during saccades and fixations while reading: A magnitude and time anaysis. *Journal of Vision, 9 (7)*, 1-13.

White, S.J. (2008). Eye movement control during reading: Effects of word frequency and orthographic familiarity. *Journal of Experimental Psychology: Human Perception and Performance, 34,* 205-223.

Worth, C. (1921). Squint. Its causes, pathology and treatment. Philadelphia: Blakiston.

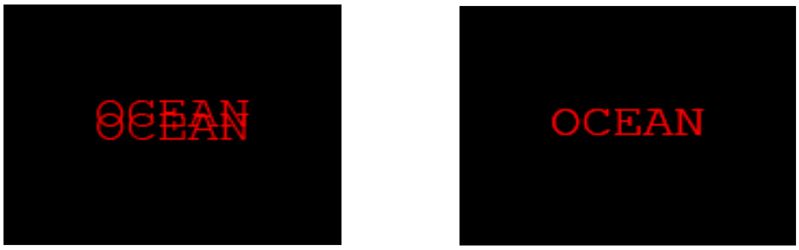
Yang, Q. & Z. Kapoula (2003). Binocular coordination of saccades at far and at near in children and in adults. *Journal of Vision 3*(8), 554-561.

Zee, D.S., Fitzgibbon, E.J., & Optican, L.M. (1992). Saccade-vergence interactions in humans. *Journal of Neurophysiology, 68 (5),* 1624-1641.

Acknowledgements

This research was funded by the University of Southampton. M. Nikolova was supported by a Psychology Jubilee Scholarship. We would like to thank the participants who volunteered to take part in this experiment. We would also like to thank two anonymous reviewers for their helpful comments on an earlier version of this manuscript.

Figure 1. Dichoptic presentation of the experimental stimuli without fusion (A) and with fusion (B)



A) B)

Figure 2. Mean differences in total reaction times (RTs) between high-frequency words (HF), low-frequency words (LF) and non-words (NW) in the different disparity conditions.

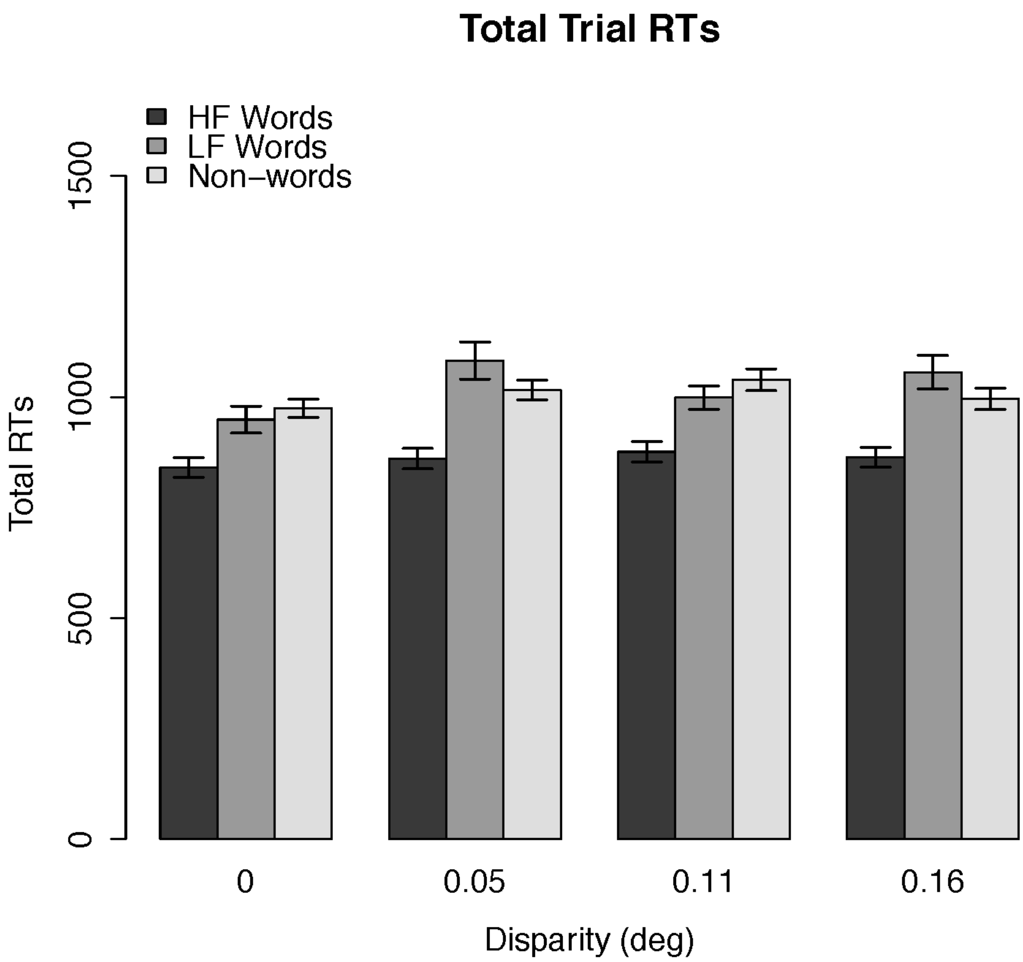


Figure 3. Horizontal and vertical disparities at the start of the initial fixation on each item plotted on a Cartesian coordinate system

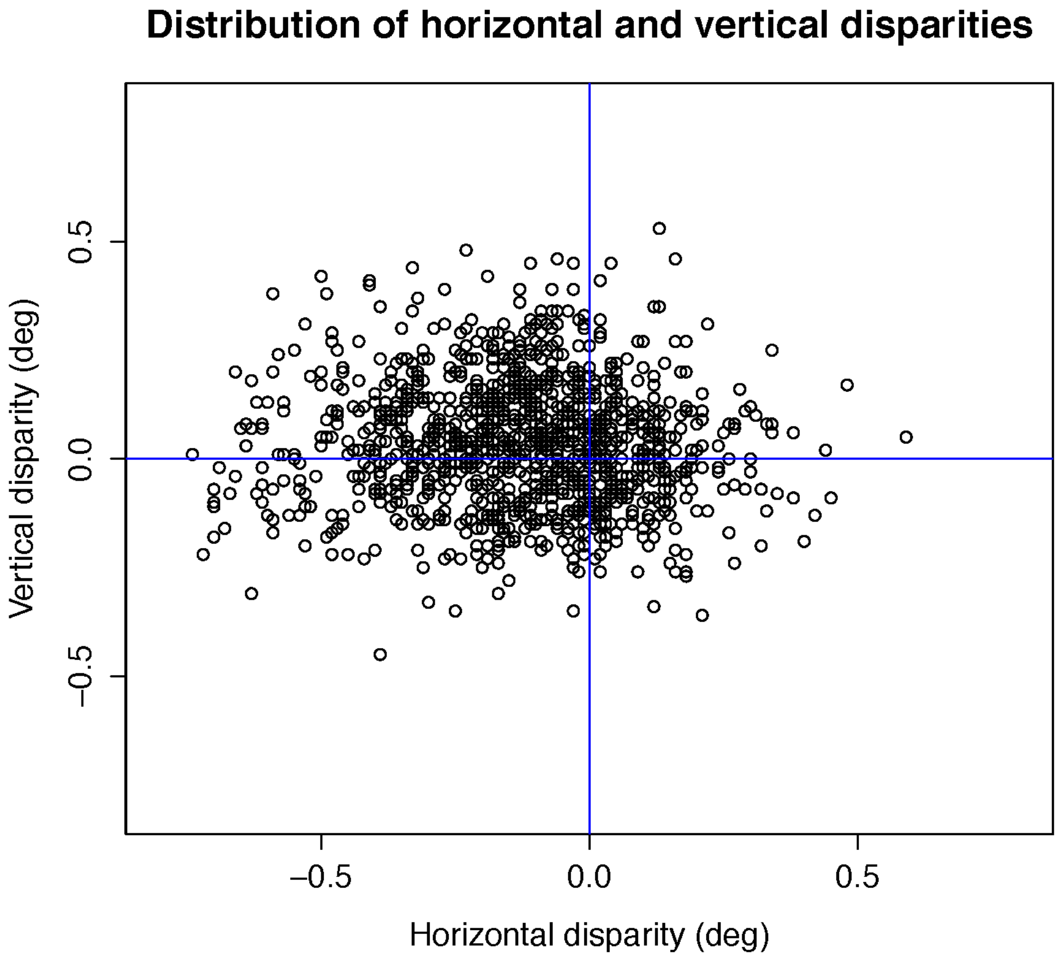


Figure 4. Distribution of horizontal (green) and vertical (blue) disparities between the start (A) and the end (B) of the initial fixation on each item

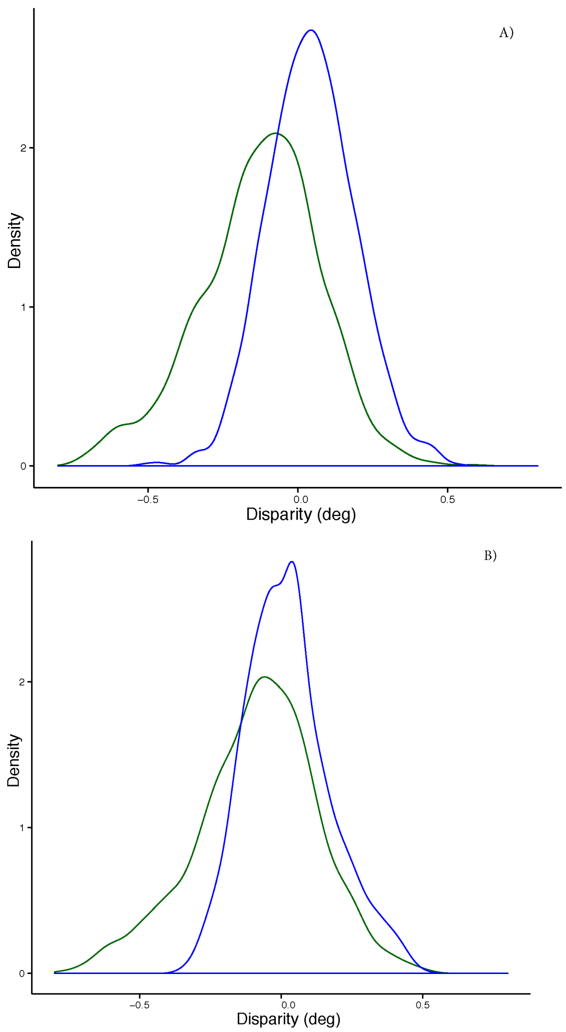


Table 1.

Descriptive data about lexical decision accuracy, total reaction times (RTs), first fixation duration (FFD), single fixation duration (SFD) and mean number of fixations per trial (*SD*s in parentheses) for high-frequency words (HF), low-frequency words (LF) and non-words (NW).

	Frequency	Disparity (deg)			
		0	0.5	0.11	0.16
Lexical decision accuracy	HF	100%	99%	99%	99%
	LF	92%	92%	93%	93%
	NW	97%	98%	98%	98%
Total RTs (ms)	HF	840.79 (22.23)	860.70 (23.16)	876.05 (23.11)	864.24 (22.48)
	LF	948.84 (30.05)	1082.37 (42.08)	998.53 (27.01)	1056.15 (37.91)
	NW	974.39 (20.27)	1015.50 (22.55)	1038.97 (24.49)	995.83 (24.13)
FFD (ms)	HF	407.32 (23.03)	393.01 (19.49)	414.57 (20.11)	371.73 (17.34)
	LF	374.74 (19.04)	369.03 (23.13)	369.03 (23.13)	378.81 (23.05)
	NW	399.41 (15.48)	380.71 (15.44)	401.20 (15.58)	362.49 (14)
SFD (ms)	HF	608.14 (31.35)	712.56 (65)	683.92 (46.73)	733.96 (48)
	LF	736.40 (170.27)	857.92 (98.16)	805.79 (66.16)	813.67 (75.12)
	NW	733.11 (54.89)	710.62 (69)	679.75 (55.50)	758.52 (60.64)
Number of fixations per trial	HF	2.41 (.07)	2.47 (.07)	2.42 (.07)	2.45 (.07)
	LF	2.78 (.10)	2.87 (.11)	2.66 (.09)	2.85 (.11)
	NW	2.71 (.06)	2.77 (.07)	2.78 (.07)	2.79 (.07)