Language & Cognitive Dynamics Laboratory

Technical Report 2013.01
Version 1.0: December 2013

Preliminary validation of eye tracking and the "Visual World Paradigm" for participants with aphasia and limb apraxia

Dan Mirman
Kristen Graziano

Abstract

Eye tracking is a powerful method for studying language and cognitive processing. The viability of using eye tracking methods to study cognitive processing in individuals with aphasia and limb apraxia was examined by testing impaired and unimpaired individuals in several visual search tasks. The results recommend the use of perimetry and color discrimination pretests, chin rest or re-calibration to assure track quality, and a study design and/or data analysis plan for handling substantial differences in response times. With these challenges addressed, eye movements offer a powerful tool for examining the effects of brain lesions on language and cognitive function.

Author note

This is a summary of preliminary research conducted when we began using eye tracking to study language and cognitive processing in aphasia and limb apraxia. It was not intended as a comprehensive methodological assessment; rather, our goal was to determine what kinds of issues we should consider as we move forward with this research program using a particular eye tracker and configuring with a particular left hemisphere stroke population. We are sharing our findings for the benefit of other researchers who may be using similar set-ups. This research was supported by Albert Einstein Society grant 09-13 to DM and by the Moss Rehabilitation Research Institute. We thank Adelyn Brecher for her help with participant recruitment and Myrna Schwartz, John Whyte, and Laurel Buxbaum for their helpful suggestions.

Preliminary validation of eye tracking and the "Visual World Paradigm" for participants with aphasia and limb apraxia

Eye movements are among the most frequent of all human movements, with large ballistic scanning movements, called "saccades", typically occurring 3-4 times per second. Because the human eye monitors a visual field of about 200°, but receives detailed information from only about 2°, eye movements are fundamental to the operation of the visual system. Furthermore, due to their close relation to attentional mechanisms (e.g., Corbetta, 1998), eye movements can provide insight into a wide range of cognitive processes, including language comprehension, conceptual knowledge, memory, mental imagery, attention, and even social cognition. Eye tracking has emerged as a powerful method for studying cognitive processes, particularly since the re-discovery of the "Visual World Paradigm" (Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995; cf. Cooper, 1974).

In a typical Visual World Paradigm (VWP) experiment, several objects are shown in a display and participants are instructed to point to or click on one of the objects. As participants listen to the spoken phrase that specifies the target object, their eye movements are recorded. The proportion of fixations to a given object maps very closely and with high temporal precision onto the mental activation of the word or concept corresponding to that object, providing unique insights into the time course of cognitive processing. This experimental paradigm has become a critical tool in cognitive research for three main reasons.

First, the VWP task is relatively natural and places minimal additional demands on cognitive processing. Traditional experimental tasks require participants to make abstract judgments (e.g., evaluate whether a sentence is grammatical or whether a word refers to concrete object) and the analyses rely specifically on these overt responses. In contrast, in the Visual World Paradigm, the critical measure is eye fixations and participants simply point to or click on the named object, or even just look at the display (e.g., Altmann & Kamide, 1999; Ben-David et al., 2011; Mirman & Graziano, 2012a, 2012b). This advantage may be particularly helpful for testing neurologically impaired participants, who may have difficulty with complex task demands but relatively spared oculomotor control. Furthermore, neurologically impaired participants often differ in their strategic approach to overt responses (for example, some individuals with aphasia prefer not to respond rather than make an incorrect response), but such strategies are much less likely to affect eye movements.

Second, the VWP is much more sensitive than traditional cognitive experimental methods. For example, some theories of semantic knowledge predicted that concepts should partially activate all related concepts, even distantly related concepts (e.g., lion – beaver). Studies testing for this partial activation using semantic priming failed to show an effect, but Mirman and Magnuson (2009) found that listeners were more likely to look at such distantly related distractor objects than at completely unrelated objects (e.g., lion – hammer), indicating that they were indeed partially activated. Similarly, Allopenna, Magnuson, and Tanenhaus (1998) demonstrated more looks to rhyme distractors (e.g., beaker – speaker) than unrelated distractors, in contrast to previous failures to find rhyme-based priming. The high sensitivity of the VWP may be particularly important for understanding individual differences, especially when combined with Growth Curve Analysis: a statistical technique specifically adapted to quantifying individual differences in VWP experiments (Mirman, 2014; Mirman, Dixon, & Magnuson, 2008; see also Mirman, Yee, Blumstein, & Magnuson, 2011).

Third, by recording eye movements over the entire course of a trial, the VWP provides information about the time course of cognitive processing. This is particularly important for cognitive processes that evolve quickly in time and has played an important role in elucidating numerous aspects of healthy cognitive processing. For example, Tanenhaus et al. (1995) found that listeners use visual context to resolve syntactic ambiguities on-line, even at the earliest moments of linguistic processing. Similarly, Altmann and Kamide (1999) found that listeners integrate linguistic and visual context to anticipate upcoming information, and Magnuson, Tanenhaus, and Aslin (2008) found immediate effects of syntactic expectations and pragmatic constraints on word recognition processes.

Because of these advantages, eye tracking methods in general, and the VWP in particular, have been applied to a broad range of cognitive domains. In the domain of spoken language comprehension, the VWP has revealed important aspects of processing from sub-phonemic levels (e.g., McMurray, Tanenhaus, & Aslin, 2002; Salverda, Dahan, & McQueen, 2003) to syntactic (e.g., Tanenhaus et al., 1995) and pragmatic (e.g., Hanna, Tanenhaus, & Trueswell, 2003; Magnuson et al., 2008) levels. In the domain of conceptual processing, the VWP has revealed important aspects of the time course of activation of conceptual knowledge (e.g., Mirman & Magnuson, 2009; Yee, Huffstetler, & Thompson-Schill, 2011; Kalenine, Mirman, Middleton, & Buxbaum, 2012). Eye movements have also shed light on memory processes and representations (e.g., Richardson & Spivey, 2000) and on social/emotional cognition (e.g., Crosby, Monin, & Richardson, 2008).

Given the strengths and broad applicability of eye tracking, it is potentially a very powerful technique for studying cognitive processing in neurologically impaired populations. Indeed, there have been a few recent efforts to apply this method to the study of aphasia and limb apraxia (Dickey, Choy, and Thompson, 2007; Kalenine, Mirman, & Buxbaum, 2012; Mirman & Graziano, 2012a; Mirman et al., 2011; Myung et al., 2010; Yee, Blumstein, and Sedivy, 2008). Although each of these studies report interesting and important results, each one only tested a relatively small number of individuals with aphasia, all of whom had relatively mild impairments. As a result, it is not clear how applicable those findings are to the broad aphasic population, particularly more severely impaired individuals. Because of its minimal task demands, the VWP may be an effective method for testing relatively severely impaired individuals, who are unable to perform many experimental tasks. In addition, those studies excluded individuals with frank visual or oculomotor deficits. However, it is unclear what effect these impairments would have on performance in this task and whether relatively subtle visual impairments may have impacted performance. In fact, Hallowell, Douglas, Wertz, and Kim (2004) pointed out that the routine failure to evaluate visual function in research on aphasia may lead to invalid data collection and interpretation.

This study explored to what extent subtle and not-so-subtle visual and oculomotor impairments interfere with the use of eye tracking methods to study cognitive processing in individuals with aphasia and limb apraxia.

Methods

Participants

Participants with aphasia were recruited from the Neuro-Cognitive Rehabilitation Research Patient Registry at the Moss Rehabilitation Research Institute (Schwartz et al., 2005). The inclusion criteria were (1) an acute diagnosis of aphasia secondary to left hemisphere cerebrovascular accident, (2) currently in the chronic phase (>6 months post onset, actual range: 28-266 months post onset), (3) right-handed (Edinburgh Handedness Inventory; Oldfield, 1971) and native English speakers (to match typical selection criteria for studies on language processing), (4) intact vision and hearing (HHIE; Ventry & Weinstein, 1983), and (5) no major psychiatric or neurologic co-morbidities. These individuals were selected to be relatively diverse with respect to aphasia subtype and severity, and lesion location. However, due to selecting for an acute diagnosis of aphasia and the nature of the Research Registry, most of the individuals with aphasia had had middle cerebral artery strokes and were in the mild-to-moderate aphasia severity range, though we tested at least one participant with severe aphasia (WAB

AQ: 25.2) and several in the severe-to-moderate impairment range (WAB AQ: 50-70). The participants with aphasia had mean age of 59 (range: 41-74) and mean years of education of 16 (range: 10-22).

The participants with aphasia had previously participated in a multi-session language assessment, which included the Western Aphasia Battery (WAB; Kertesz, 1982, 2006), and MRI or CT imaging to determine the precise location of their lesion. At the time of the experiment, these participants also completed the 10-item *Transitive Gesture to Sight Test* of limb apraxia (Buxbaum, Giovannetti, & Libon, 2000). This test assessed their ability to correctly produce common transitive gestures (e.g., "show me how to wind a watch"), while imagining they are holding and using the specified item with their left hand. Items were in view while the gesture was produced. Gesture productions were scored on five components: content, hand posture, arm posture, amplitude, and timing. Mean performance was 43 out of 50, with scores ranging from severe limb apraxia (25) to no apraxia (48). Detailed information about the participants with aphasia is provided in Supplemental Table S1.

Unimpaired control participants were recruited to be matched to the aphasic group on mean age (M = 53, range: 34-65) and mean years of education (M = 15, range: 12-20) and to have no history of major psychiatric or neurological conditions. As with the aphasic group, only right-handed native English speakers with intact vision and hearing were recruited for the study. A total of 40 participants completed the study and an additional 5 participants (3 aphasic, 2 control) were excluded due to failure to obtain a reliable eye-tracker calibration. The final sample included twenty-two participants with aphasia (38% females; 48% African American) and eighteen neurologically intact control participants (50% female, 33% African American).

All participants gave informed consent to participate in accordance with guidelines of Albert Einstein Healthcare Network and were paid for their participation and reimbursed for travel and related expenses. All were living in the community at time of testing.

Apparatus

Participants were seated approximately 24 inches away from a 17-inch monitor with screen resolution set to 1024x768 dpi. Stimuli were presented using E-Prime Professional 2.0 experimental design software (Psychology Software Tools, Inc.). Responses were recorded by the experimenter or using a Serial Response Box (Psychology Software Tools, Inc.) or mouse (see procedure for details). During the testing session, a remote EyeLink 1000 eye tracker was used to record participants' left eye gaze position at 250 Hz. A remote eye tracker is less invasive and more comfortable for participants, but these benefits are typically associated with lower resolution (spatial and/or temporal) and the possibility of track loss due to excessive head motion. The latter occurs if the participant moves her or his head out of the eye camera's field of view. Although not very common, this can occur due to substantial shifts in sitting position. We will return to this issue in the discussion.

Procedure

Testing was completed in a single session lasting approximately 60 minutes. The test session was composed of 4 subtests: (1) a simple perimetry test to assess visual impairments such as scotomas, visual field cuts, neglect, or extinction; (2) single-feature and conjunction letter search tasks (Treisman & Gelade, 1980) to test basic visual search processes; (3) a difficult naturalistic visual search task using "Where's Waldo?" pictures (e.g., Klein & MacInnes, 1999) to test basic visual search under more difficult conditions; (4) a "silent" version of a typical visual world paradigm task (Tanenhaus et al., 1995) to test the kind of highly constrained visual search involved in typical VWP experiments without requiring online language processing. The tasks were presented in a fixed order progressively approximating the VWP task: perimetry, single-feature search, conjunction search, naturalistic search, silent VWP. The following subsections describe each task in detail.

Perimetry task. Participants were asked to fixate on a plus sign (+) in the center of the screen and if they saw a circle appear, to indicate whether the circle was on the left, right, or both sides of the screen. Individuals with aphasia responded either verbally, or by pointing to printed arrows that corresponded to the left, right, or both sides. A blue circle (175 pixel diameter, approximately 2.3in) appeared for 300ms in one of six locations around the perimeter of the screen (upper left, middle left, lower left, upper right, middle right, and lower right), once individually and once with another circle on the opposite side of the screen. Time between trials and location of circle were randomized to prevent guessing.

Letter search task (based on Treisman & Gelade, 1980). Participants completed two kinds of letter search tasks: a single-feature search task and a conjunction search task. In the single-feature search task, they were asked to indicate whether the display contained the letter "O", which appeared among "N" and "X" distractors. In the conjunction search task, they were asked to indicate whether the display contained a green letter "N", which appeared among brown "N" and green "X" distractors. The set size of the trials varied (1, 5, 15, 30), with six trials at each level (3 target-present trials and 3 target-absent trials). Participants responded by pressing one of two defined buttons on the Serial Response Box. Stimuli appeared in a random order with a 1000ms inter-trial interval during which a blank fixation screen was displayed.

Naturalistic visual search task ("Where's Waldo?"). Participants viewed digitized pictures from the "Where's Waldo" series of puzzle books and were asked to indicate whether each picture contained the character Waldo (a curly-haired man with a walking stick, red-and-white striped shirt, and blue slacks) or the Wizard (a long-robed, white-haired man with a striped staff). There were 20 trials in this task, with Waldo occurring in 10 of the trials, and the Wizard occurring in the other 10 trials. Trials were presented in a random order with a 1000ms inter-trial interval during which a blank fixation screen was displayed. Across trials, the target (Waldo or Wizard) occurred in all regions on the screen.

"Silent" Visual World Paradigm task. Participants were instructed to click on the picture of the animal in each display. This variation on the typical VWP task removed the need for on-line language processing while maintaining the essential visual search properties of the task. On each trial the participant saw 4 images, one near each of the 4 corners of the display. Exactly one of those images was an animal. There were 30 trials in this task, with the target animals presented in a random order and location. Participants controlled the mouse with their preferred hand, which was generally the right hand except for a few participants with aphasia who used their left hand due to a right-sided hemiparesis/hemiplegia.

Results

Overview

This exploratory study was designed evaluate to what extent visual and oculomotor impairments interfere with the use of eye tracking methods in studies with individuals with aphasia and limb apraxia. Because of its exploratory nature, there was a wide variety of measures that we analyzed. We begin by considering standard behavioral measures (reaction time and accuracy) for each task. We then consider track quality (i.e., track loss) and basic measures of fixation and saccade properties (duration, amplitude, etc.). Group means by task and t-tests for differences between groups on behavioral performance, track quality, and eye movement measures are shown in Table 1. In addition, we examine differences among individuals with aphasia by testing correlations between demographic and clinical measures and eye movement, track quality, and behavioral performance measures. Overall, the patterns of results for control participants and individuals with aphasia were quite similar, so we focus on documenting and explaining (so far as possible) differences or individual deviations from the overall pattern.

Table 1. Group differences by task: Means (SD in parenthesis) for each group for each task on eye movement, track quality, and behavioral performance measures and results of t-tests of the difference between the two groups.

	Number of Fixations	Fixation Duration	Number Saccades	Saccade Duration	Saccade Amplitude	Track Loss Proportion	Off Screen Fix. Prop.	Reaction Time	Accuracy
Simple Letter Search									
Control	2.85 (0.85)	283.1 (88.0)	2.94 (0.82)	212.7 (660.2)	5.15 (2.01)	0.09 (0.22)	0.03 (0.05)	1031.0 (293.6)	98.6 (2.5)
Aphasic	5.41 (2.24)	234.2 (49.3)	5.53 (2.06)	291.8 (529.5)	8.75 (4.53)	0.23 (0.25)	0.10 (0.09)	1985.0 (857.3)	92.8 (13.0)
t	4.64	2.07	5.08	0.40	2.83	1.73	3.15	4.58	.92
p<	0.001	0.05	0.0001	n.s.	0.01	0.1	0.01	0.001	n.s.
Conjunction Letter Search									
Control	5.85 (2.03)	224.9 (37.3)	6.03 (2.13)	80.7 (125.7)	6.55 (2.09)	0.06 (0.12)	0.04 (0.05)	1632.6 (498.0)	89.1 (13.1)
Aphasic	9.35 (4.07)	239.3 (51.0)	9.35 (4.01)	152.8 (145.8)	7.62 (4.16)	0.22 (0.21)	0.10 (0.14)	3018.9 (1483.8)	81.4 (17.7)
t	3.54	1.02	3.35	1.68	1.05	3.0	2.02	4.11	1.57
p<	0.01	n.s.	0.01	n.s.	n.s.	0.01	0.01	0.001	n.s.
"Where's Waldo"									
Control	49.1 (12.8)	261.5 (47.6)	49.0 (12.7)	120.0 (126.2)	5.82 (2.68)	0.17 (0.19)	0.13 (0.10)	-	60.0 (18.2)
Aphasic	45.3 (16.8)	265.5 (67.5)	45.0 (16.7)	417.6 (684.8)	6.41 (2.94)	0.35 (0.27)	0.20 (0.19)	-	56.4 (17.5)
t	0.82	0.22	0.85	2.0	0.66	2.56	1.46	-	0.64
p<	n.s.	n.s.	n.s.	0.1	n.s.	0.05	n.s.	-	n.s.
Silent VWP									
Control	4.11 (1.09)	304.7 (63.2)	4.16 (1.09)	46.4 (12.3)	9.00 (1.52)	0.02 (0.03)	0.06 (0.12)	1443.8 (317.8)	100.0 (0.0)
Aphasic	7.49 (2.88)	276.0 (76.0)	7.52 (2.82)	196.9 (460.0)	11.2 (8.68)	0.16 (0.20)	0.18 (0.20)	2867.7 (1174.7)	99.7 (1.0)
t	5.08	1.31	5.12	1.53	1.14	3.31	2.30	5.45	1.45
p<	0.0001	n.s.	0.0001	n.s.	n.s.	0.01	0.05	0.0001	n.s.

We then turn to two sets of analyses that are particularly relevant for the use of the VWP with participants with aphasia and limb apraxia: spatial analyses and target fixation time course analysis. The spatial analyses examined whether the control and impaired participants differed in the spatial distribution of their fixations by considering the proportion of fixations to the top half vs. bottom half and the left half vs. right half of the display. The VWP target fixation time course analysis mimicked standard analysis of VWP data. Accordingly, the likelihood of fixating the target picture (the only animal in each display) was computed for every 40ms time bin from display onset. The resulting target fixation proportion curves were fit using multilevel regression with fourth-order orthogonal polynomials. Group differences were modeled as effects on the parameters of the target fixation curves (for a detailed description of growth curve analysis see Mirman, 2014).

Behavioral Data

Perimetry task. Control participants correctly detected each target presentation and identified its location (i.e., 100% accuracy). Most participants with aphasia also performed essentially at ceiling (above 85% correct with no discernable pattern of errors), with a few notable exceptions. Participant MR0166 was much poorer at detecting targets on the right than on the left side (left: 100% correct; right: 15% correct) and only reported the left target when targets were presented on both sides of the screen, consistent with a severe right hemifield neglect or a right side field cut. Participant MR0913 performed substantially worse when two targets were presented (33% correct) than when only one was presented (left: 100% correct; right: 83% correct), with a slight tendency to report the target on the left, suggesting a possible case of extinction or simultagnosia. One participant (MR1619) completed an early version of the task in which the inter-trial delays were too short, and consequently missed some targets, though there was no spatial pattern to the misses.

Letter search task. Overall accuracy in both letter search tasks was very high for control participants and participants with aphasia (Single-feature search: >90% correct; Conjunction search: >80% correct) with no statistically reliable differences between the two groups.

Exceptions to this pattern include participants MR0166 and MR0913 who were significantly impaired in the single-feature search (50% and 67% correct, respectively). Their difficulty with this task was most likely a result of their diminished ability to attend to stimuli on the right side of the display. In the conjunction search, participants MR0206 and MR0913 responded incorrectly to exactly half of the stimuli (50% correct), correctly responding only to the target-absent trials. This is most likely a result of their (self-reported) inability to detect the green color of the target "N". In the conjunction search task, participant MR0166 again showed significant difficulty (50% correct), suggesting that the possible right hemifield neglect impaired performance. Participants MR1238 and MR0583 were also significantly impaired on the conjunction search task (58% and 50% correct, respectively), showing poorer performance as the stimuli display size increased. Performance was intact on all other tasks, suggesting inadequate attentional resources required to accurately complete this compound task.

Across both task types, the participants with aphasia took nearly twice as long as control participants to respond. In the single-feature search, both groups showed minimal differences in response times between the four display sizes (1, 5, 15, 30), although on average, the participants with aphasia were significantly slower to respond (Control: 1031ms; Aphasic: 1985ms). In the conjunction search, control participants showed the expected pattern: an approximately linear increase in reaction time as a function of set size (Treisman & Gelade, 1980). The participants with aphasia showed a similar, and greatly exaggerated pattern of response time on this task (Overall mean: Control: 1633ms; Aphasic: 3019ms). The mean reaction times for correct-response trials are plotted in Figure 1.

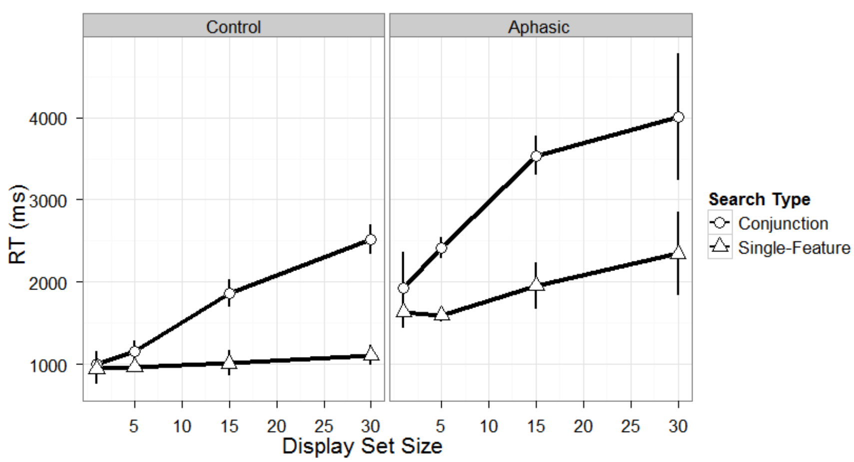

Figure 1. Mean single-feature and conjunction-feature letter search reaction times as a function of number of distractors (set size) for control (left panel) and aphasic (right panel) groups. Error bars indicate ±1SE.

Naturalistic visual search task ("Where's Waldo"). Overall, this task was difficult for both groups of participants, with both groups performing at approximately 60% correct. Both groups also tended to perform better on the Waldo trials (approximately 70% correct) than the Wizard trials (approximately 40% correct), possibly because the Waldo character is more familiar and occurs more frequently in the "Where's Waldo" picture books. Omission errors accounted for approximately 30% of the trails, with the other 10% being commission errors (e.g., responding "Waldo" on a "Wizard" trial). Participants were allotted up to 30 seconds to successfully complete each trial and reaction times were not recorded for this task.

"Silent" Visual World Paradigm task. Control participants and participants with aphasia performed near ceiling on this task with every participant performing above 95% correct. As in the letter search task, participants with aphasia took approximately twice as long to respond as control participants did (Controls: 1444ms; Patients: 2868ms).

Track Quality

Eye track quality varied greatly among participants, with a significant difference between the control and aphasic groups. Compared to control participants, participants with aphasia had a substantially higher degree of track loss as indicated by proportion of missing data samples (Control: 0.09; Patient: 0.24) across all four tasks ($t(37) = 3.8$, $p < 0.001$), although this was not correlated with impairment severity as measured by WAB AQ (all $r < 0.1$). The largest proportions of track loss occurred during the "Where's Waldo" task (Control: 0.17; Aphasic: 0.35), most likely because the difficulty of the task caused participants to shift position during trials, resulting in temporary track loss. Conversely, the lowest proportion of track loss occurred during the Silent VWP task (Control: 0.02; Aphasic: 0.16), most likely because of the constraints and ease of the task (only four fixed image locations, easily identifiable target).

Eye Movements

Across tasks, participants with aphasia made approximately twice as many fixations and saccades as control participants did. This is consistent with their reaction times, which were generally twice as slow. No consistent statistically reliable differences between groups were found for fixation durations, saccade durations, or saccade amplitudes.

In the single-feature letter search, participants with aphasia had smaller fixation durations and larger saccade amplitudes than control participants did. This is possibly due to the greater cognitive demand and resulting increased effort by the participants with aphasia with the start of this first "cognitive" task. These differences were not found in any of the later tasks. In the "Where's Waldo" task, participants with aphasia only differed from control participants in duration of saccades – participants with aphasia exhibited shorter saccade durations. This suggests that the complex nature of the "Where's Waldo" stimuli caused difficulty processing large portions of the images at once, leading to a tendency to make shorter saccades. Both groups made similar numbers of fixations and saccades, possibly because of the extensive search required to complete this task.

Eye movement patterns were not correlated with measures such as aphasia severity (WAB AQ), although the WAB comprehension component was negatively correlated with fixation duration and reaction time. This suggests that comprehension impairment may require the participant to fixate longer on stimuli to allow full understanding, and therefore take longer overall to respond to stimuli.

Differences among Participants with Aphasia

In addition to testing overall differences between participant groups, we examined whether differences within our diverse group of participants with aphasia and limb apraxia were associated with differences in any of the key measures. Correlations between demographic and clinical measures and behavioral performance, track quality, and eye movement measures are shown in Table 2. Overall accuracy was not correlated with any of the measures, which was expected since the tasks were not directly related to the language and action impairments that characterize aphasia and limb apraxia. Reaction times were positively correlated with months post onset, suggesting that individuals who have been living with aphasia and/or limb apraxia for a long time may be compensating by performing tasks more slowly (though note that all of these participants were in the chronic stage and were at least 2 years from the onset of their aphasia). Reaction times were also positively correlated with lesion volume and negatively correlated with Gesture-to-Sight scores, possibly because larger lesions and increasingly severe limb apraxia are likely to affect motor planning and execution, thereby slowing down responses. Comprehension scores were negatively correlated with reaction times, indicating faster responding by participants with better comprehension.

Table 2. Correlations between demographic and clinical measures and eye movement, track quality, and behavioral performance measures for participants with aphasia.

	Number of Fixations	Fixation Duration	Number of Saccades	Saccade Duration	Saccade Amplitude	Track Loss Prop.	Off Screen Fix. Prop.	Reaction Time	Accuracy
Age	0.061	0.226*	0.059	-0.188	-0.039	-0.148	0.019	0.015	-0.128
Education	0.144	0.023	0.141	-0.243*	0.012	-0.258*	-0.033	0.021	-0.121
Months Post Onset	0.001	0.081	-0.008	0.073	0.065	0.116	0.039	0.478**	-0.168
Aphasia Quotient	-0.007	-0.124	-0.01	0.107	0.04	0.011	0.158	-0.151	-0.021
Comprehension	-0.083	-0.217*	-0.082	0.05	0.104	0.011	0.162	-0.483**	0.092
Fluency	0.018	-0.102	0.017	0.122	0.022	-0.036	0.12	-0.166	-0.007
Gesture to Sight	-0.037	-0.164	-0.037	0.07	0.02	0.016	0.089	-0.468**	-0.001
Lesion Volume	0.013	-0.041	0.01	-0.027	0.122	0.058	0.091	0.398**	0.002

* $p < 0.05$, ** $p < 0.001$

Despite these substantial effects on reaction times, there were no consistent effects of demographic or clinical variables on eye movement measures. Track quality tended to be somewhat better (lower proportion of missing eye data samples) for individuals with higher education levels and higher education level was also associated with somewhat faster eye movements (shorter saccade durations), though neither of these patterns lends itself to a clear account. Fixation duration was positively correlated with age, suggesting that older participants may have spent a longer time examining and identifying individual display elements (though note that age was not correlated with reaction time). Fixation duration was also negatively correlated with comprehension scores, suggesting that participants with comprehension impairments may have used a more exploratory fixation strategy rather than a planned analytical strategy (though note that comprehension scores were not correlated with number of fixations). In sum, it appears that although demographic and clinical variables have strong effects on reaction time, the effects on eye movement measures are relatively weak. This suggests that – with appropriate matching for age and education – eye movements may provide a reliable measure of the effects of brain lesions on language and cognitive function.

Spatial Analysis

To validate the VWP method it was also important to examine whether control participants and participants with aphasia differed in the general spatial distribution of their fixations. To do this we examined the relative proportion of fixations and cumulative fixation duration to the top vs. bottom half of the display and left vs. right half of the display. Table 3 shows the group means (with standard deviations) for each task and the result of a logistic regression test for difference between groups. Not surprisingly, control participants exhibited a consistent top-bias and left-bias in their fixation behavior. That is, more than 50% of their fixations were directed to the top half and the left half of the display and the same pattern held for cumulative fixation duration (i.e., time spent looking at the display). Participants with aphasia exhibited a substantially reduced top-bias and left-bias, for some tasks even exhibiting a slight bottom-half and/or right-side bias. However, this group difference was not nearly as pronounced when only the first fixation was considered (marginally significant group difference in top/bottom bias, no significant group difference in left/right bias), suggesting that all participants tended to start their visual search on the top and on the left side of the display, but participants with aphasia were simply more conservative in their decision-making and tended to look in all places on the screen before making their response. This interpretation is also consistent with their much slower response times.

Table 3. Mean (SD in parenthesis) proportions of fixations, fixation duration, and first fixations on the top half and left half of the display during each task for each group.

	Single Feature Search		Conjunction Search		"Where's Waldo"		Silent VWP		Group effect	
	Control	Aphasic	Control	Aphasic	Control	Aphasic	Control	Aphasic	Estimate	*p* <
Top Half										
Proportion of Fixations	0.621 (0.103)	0.508 (0.128)	0.562 (0.081)	0.537 (0.104)	0.479 (0.101)	0.504 (0.100)	0.565 (0.083)	0.496 (0.175)	-0.3804	0.01
Proportion of Fixation Duration	0.634 (0.117)	0.518 (0.148)	0.572 (0.074)	0.522 (0.123)	0.472 (0.104)	0.509 (0.127)	0.533 (0.083)	0.484 (0.175)	-0.4135	0.001
Proportion of First Fixations	0.712 (0.171)	0.595 (0.175)	0.718 (0.154)	0.668 (0.207)	0.665 (0.171)	0.652 (0.207)	0.693 (0.223)	0.707 (0.292)	-0.4743	0.1
Left Half										
Proportion of Fixations	0.614 (0.080)	0.432 (0.162)	0.618 (0.066)	0.507 (0.142)	0.496 (0.038)	0.472 (0.091)	0.549 (0.102)	0.460 (0.129)	-0.8382	0.0001
Proportion of Fixation Duration	0.556 (0.109)	0.445 (0.151)	0.606 (0.076)	0.503 (0.149)	0.496 (0.042)	0.487 (0.093)	0.524 (0.098)	0.469 (0.123)	-0.5366	0.0001
Proportion of First Fixations	0.680 (0.122)	0.573 (0.210)	0.720 (0.137)	0.613 (0.240)	0.658 (0.195)	0.609 (0.234)	0.657 (0.190)	0.618 (0.303)	-0.3889	0.2

The two participants whose perimetry test results suggested visual deficits affecting the right hemifield (MR0166: neglect or field cut; MR0913: extinction) were more likely overall to fixate the right side of the display (approx. 56% of their fixations were on the right side), though their first fixations were still more likely to be on the left side of the display (approx. 55% of fixations were on the left side). This pattern suggests that these individuals were using compensatory strategies to make up for right hemifield deficits. However, it is important to note that neither their overall pattern nor their pattern of first fixations was substantially different from the overall aphasic group mean.

VWP Target Fixation Time Course

The critical data in typical VWP studies is the time course of target (or distractor) fixation. Figure 2 shows the mean target fixation time courses for the control and aphasic groups. Consistent with the behavioral response time data, the eye movement data show substantially slower performance for participants with aphasia than controls. There was only a small difference in overall target fixations (i.e., area under the fixation proportion curves), indicating that both groups spent an approximately equal amount of time examining the target pictures (though the aphasic group was marginally higher: *Estimate* = 0.073, *SE* = 0.04, $p < 0.1$). In contrast, there were much stronger differences between participant groups in the time course. For example, the overall linear slope was negative (*Estimate* = -

1.49, *SE* = 0.28, $p < 0.0001$) reflecting that most target fixations occurred relatively early with target fixation proportion gradually decreasing over time. However, the participants with aphasia had a significantly more positive slope relative to controls (*Estimate* = 0.94, *SE* = 0.37, $p < 0.05$), that is, a slope closer to 0, reflecting much slower decrease in target fixations for participants with aphasia compared to control participants. There were also significant group differences on the quadratic, cubic, and quartic terms (all $p < 0.0001$), which reflect differences in the time course of target fixations. The time course difference was also confirmed by repeating this analysis on data just from the first 2000ms (where there was no linear slope effect for the control group: $p > 0.8$). In this time window there was no group difference in overall fixation proportion ($p > 0.7$), but there were very strong group differences on the linear, quadratic, and cubic terms (all $p < 0.0001$).

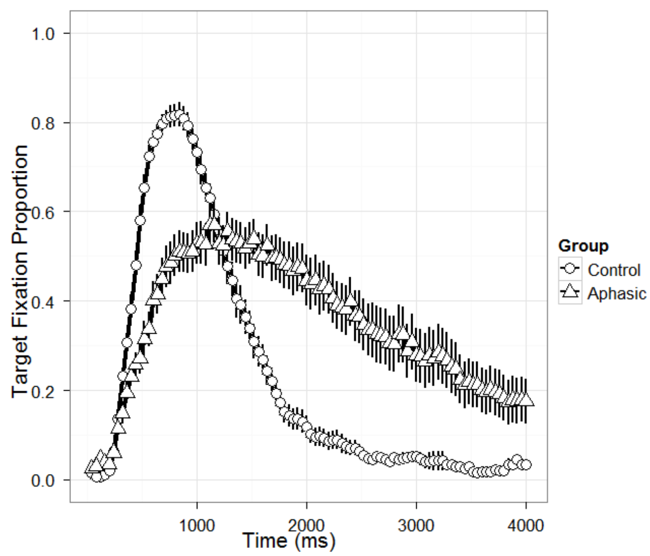

Figure 2. Mean target fixation time course in the silent VWP task for control and aphasic groups. Error bars indicate ±1SE.

Summary and Conclusions

Eye tracking has emerged as a powerful method for the study of language and cognitive processes and affords particular advantages for testing individuals with neurological impairments. However, in order to understand the applicability and limitations of using eye tracking methods with neurologically impaired populations, it is important to compare a relatively broad range of such participants with control participants in basic visual search tasks that are separate from the neurologically impaired participants' core deficits. We tested 18 neurologically healthy controls and 22 participants with aphasia and limb apraxia in a series of visual search tasks. The tasks were chosen to test visual field problems (perimetry test for field cuts, neglect, etc.), basic visual search processes (letter search, difficult naturalistic scene search), and the kind of highly constrained visual search involved in typical "visual world paradigm" experiments. The participants with aphasia and limb apraxia were selected to have varying degrees of severity, subtype, and lesion location, within the constraints of chronic aphasia secondary to left hemisphere (typically MCA) stroke and interest in participating in such studies (i.e., few participants with global aphasia or other very severe cases). Overall, the results validate the use of eye tracking with these populations, though they raise a few issues that researchers should consider when designing experiments and analyzing data.

In general, basic oculomotor aspects of fixations and saccades did not differ between controls and participants with aphasia and were not correlated with any demographic or clinical measures. Fixation durations, saccade durations, and saccade amplitudes were generally the same for control participants and participants with aphasia across tasks, validating the use of fixation-based measures of cognitive processing. These results are consistent with previous studies showing minimal age-related differences in oculomotor control between younger and older adults (e.g., Pratt, Dodd, & Welsh, 2006). However, track quality (proportion of missing data samples) was consistently poorer for participants with aphasia than for control participants. We used a remote (desktop) eye tracker with free head movement and a standard 9-point calibration at the start of the experiment. Researchers concerned about track quality may wish to consider a chin rest or other device to keep the participant's head in a more fixed location and/or periodic re-calibration during the course of their experiments.

Response accuracy was generally quite high and approximately equal for control participants and participants with aphasia, indicating that the visual search demands of eye tracking experiments are unlikely to cause substantial problems for participants with aphasia and limb apraxia. Two participants with aphasia did exhibit modest visual impairments in a perimetry test, but these impairments seemed only to affect performance in letter search tasks and not the naturalistic or VWP search tasks. Nevertheless, we suggest that a short perimetry test would be a useful control in studies using eye tracking with neurologically impaired individuals. More importantly, a few patients had substantial difficulty distinguishing moderate color differences (somewhat muted green and brown), so researchers should avoid visual stimuli that critically depend on color discriminations. Note also that we only tested individuals who had left hemisphere strokes, but neglect and related visual impairments are more common following right hemisphere damage (e.g., Corbetta & Shulman, 2011), so the prevalence and severity of such impairments may be greater in right hemisphere stroke populations. Furthermore, we selected participants with chronic aphasia (since the VWP has been primarily used to study language and related cognitive functions), which typically results from middle cerebral artery strokes; visual impairments may be more common and may have a bigger impact on tests of participants who had posterior cerebral artery strokes.

Response speed was the most consistent and largest difference between control participants and participants with aphasia. Participants with aphasia took approximately twice as long to respond across tasks compared to controls, which was accompanied by making approximately twice as many fixations and saccades, and much slower target fixation time courses in the VWP task. This seemed to be part of a generally conservative response strategy – control participants tended to search only until they found the target, but participants with aphasia tended to search the entire display in order to verify that their intended response was correct. This pattern is likely a compensatory strategy used by individuals who have adjusted to living with a cognitive impairment (it was particularly strong among individuals with visual field problems and was positively correlated with months post onset).

A number of study design and analysis options are available to researchers concerned that longer baseline response times may interfere with interpretability of results. Perhaps the simplest would be to allow participants to simply "look and listen" without requiring a response. We have used this approach in recent studies (e.g., Mirman & Graziano, 2012a; 2012b) and found little effect on the time course of target or distractor fixations (Mirman & Graziano, 2012b), though the lack of a response may obscure comprehension failures (i.e., if there is no response, it is hard to know whether the participants understood the target word or not). An alternative strategy is to encourage participants to respond as quickly as possible. Such instructions may produce data that are more directly comparable, but it is important to consider that time pressure may also affect the dynamics of cognitive processing (e.g., Kello, 2004; Dahan & Mirman, under review) or produce speed-accuracy trade-offs, which may have differential effects on control and neurologically impaired participants. Finally, researchers can analyze relative – rather than absolute – time course differences. For example, if the letter search response times are expressed as percent change relative to the no-distractor (i.e., Set Size = 1) condition, the participants with aphasia no longer show a bigger effect of set size than control participants do.

Our “silent” visual world paradigm task is most relevant to researchers considering using eye tracking with aphasic and/or apraxic participants, since this is the paradigm that has been used so widely and effectively to study language and cognitive processing. Control participants and participants with aphasia showed the least differences in this task, with both groups exhibiting very high accuracy and minimal differences on oculomotor measures. The strongest difference was the consistent slower response time for participants with aphasia compared to control participants, which had a very strong effect on target fixation time course. When the baseline time course is so different, it may be difficult to interpret the effects of experimental manipulations. As mentioned above, one could eliminate the need for a response or instruct participants to respond within a certain time window, though that may have consequences on the underlying dynamics of processing (e.g., Dahan & Mirman, under review). Alternatively, the data could be time-normalized to convert from raw trial time to proportion of trial duration (e.g., Spivey, Grosjean, & Knoblich, 2005) or researchers could examine the proportional change in parameter estimates provided by growth curve analysis. For example, researchers could use a standard method to define group-specific analysis windows, such as “from trial onset to a point when 90% of group’s trials have terminated”, and consider relative differences in condition effects on parameters for each group. In other words, this would mean re-expressing the condition effect parameter estimates in terms of percent change between conditions; much like the above example of re-expressing letter search response times in terms of percent change relative to a baseline condition.

In sum, aphasia and apraxia seem not to have substantial effects on oculomotor patterns in a range of visual search tasks. We recommend that researchers include brief perimetry and color discrimination pretests to assure adequate visual perceptual function and implement procedures to assure adequate track quality. In addition, researchers need to have a data analysis plan for fixation time course data that can deal with large differences in base response time between control participants and participants with aphasia and limb apraxia. If these potential challenges are addressed, we believe eye tracking methods hold tremendous potential as tools for experimental and clinical neuropsychology.

Supplemental Table S1. Demographic and clinical descriptions of participants with aphasia.

Participant ID	Age	Education	MPO	Aphasia subtype	WAB AQ	WAB Comp	WAB Fluency	Gesture to Sight	Lesion Volume (cc)	Lesion Location
MR0042	56	17	132.8	Anomic	87.1	7.85	9	41	224.4	Both
MR0044	58	12	117.2	Anomic	95.2	10	9	46	78.5	Posterior
MR0083	54	11	116.7	Anomic	95.1	9.85	9	46	51.0	Anterior
MR0186	49	10	173.3	Anomic	73.4	8.3	5	29	253.0	Both
MR0206	57	16	127.4	Anomic	92.3	9.95	9	42	103.9	Both
MR0419	41	12	96.8	Anomic	91.5	8.85	9	48	51.9	Anterior
MR0913	67	16	62.5	Anomic	89.0	9	9	48	64.5	Posterior
MR1088	47	18	60.3	Anomic	78.8	8.8	8	46	89.1	Posterior
MR1392	66	19	38.6	Anomic	90.2	8.7	9	47	84.9	Posterior
MR1687	72	22	29.3	Anomic	88.1	9.75	8	48	41.0	Anterior
MR0166	62	18	266.1	Broca’s	70.6	6.6	5	35	231.1	Both
MR0190	63	16	121	Broca’s	56.8	7	4	36	205.3	Both
MR0583	64	19	116	Broca’s	55.1	7.25	4	43	35.6	Both
MR1238	52	14	59.5	Broca’s	67.8	10	4	48	186.4	Both
MR1510	51	18	60.3	Broca’s	62.8	8	4	38	219.9	Both
MR1626	73	12	34.5	Broca’s	67.8	9.4	4	48	77.3	Both
MR2027	65	16	28.5	Broca’s	25.2	6.5	1	25	272.0	Both
MR0281	50	16	120	Conduction	82.9	9.15	9	41	151.3	Posterior
TU1449	57	12	54.2	Conduction	81.8	8.1	9	48	30.0	Posterior
MR1619	74	21	33.4	Conduction	56.1	7.95	5	-	8.0	Posterior
MR2038	51	14	27.5	Conduction	68.0	9.4	5	45	31.4	Posterior
XO2540	54	12	31.8	Conduction	65.3	7.05	6	47	57.6	Both

Note: MPO = Months Post Onset, WAB = Western Aphasia Battery, AQ = Aphasia Quotient, Comp = Comprehension.

References

Allopenna, P. D., Magnuson, J. S., & Tanenhaus, M. K. (1998). Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. *Journal of Memory & Language, 38*(4), 419-439.

Altmann, G. T. M., & Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition, 73*(3), 247-264.

Ben-David, B. M., Chambers, C. G., Daneman, M., Pichora-Fuller, M. K., Reingold, E. M., & Schneider, B. A. (2011). Effects of aging and noise on real-time spoken word recognition: Evidence from eye movements. *Journal of Speech, Language, and Hearing Research, 54*, 243-262.

Buxbaum, L. J., Giovannetti, T., & Libon, D. (2000). The Role of the Dynamic Body Schema in Praxis: Evidence from Primary Progressive Apraxia. *Brain and Cognition, 44,* 166-191.

Cooper, R. M. (1974). The control of eye fixation by the meaning of spoken language: A new methodology for the real-time investigation of speech perception, memory, and language processing. *Cognitive Psychology, 6*(1), 84-107.

Corbetta, M. (1998). Frontoparietal cortical networks for directing attention and the eye to visual locations: Identical, independent, or overlapping neural systems? *Proceedings of the National Academy of Sciences, 95*, 831-838.

Corbetta, M., & Shulman, G. L. (2011). Spatial neglect and attention networks. *Annual Review of Neuroscience*, *34*, 569–599. doi:10.1146/annurev-neuro-061010-113731.

Crosby, J. R., Monin, B., & Richardson, D. (2008). Where do we look during potentially offensive behavior? *Psychological Science, 19*(3), 226-228.

Dahan, D., & Mirman, D. (under review). *Effects of word frequency and time pressure on eye movements to visual referents during the recognition of spoken words*.

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

Hallowell, B., Douglas, N., Wertz, R. T., & Kim, S. (2004). Control and description of visual function in research on aphasia and related disorders. *Aphasiology, 18*(5/6/7), 611-623.

Hanna, J. E., Tanenhaus, M. K., & Trueswell, J. C. (2003). The effects of common ground and perspective on domains of referential interpretation. *Journal of Memory and Language, 49*(1), 43-61.

Kalenine, S., Mirman, D., Middleton, E. L., & Buxbaum, L. J. (under review). Temporal dynamics of activation of thematic and functional action knowledge during auditory comprehension of artifact words.

Kello, C. T. (2004). Control over the time course of cognition in the tempo-naming task. *Journal of Experimental Psychology: Human Perception and Performance, 30*(5), 942-955.

Kertesz, A. (1982). *Western aphasia battery*. New York: Grune & Stratton.

Kertesz, A. (2006). *Western aphasia battery – revised.* San Antonio: Harcourt Assessment, Inc.

Magnuson, J. S., Tanenhaus, M. K., & Aslin, R. N. (2008). Immediate effects of form-class constraints on spoken word recognition. *Cognition, 108*(3), 866-873.

McMurray, B., Tanenhaus, M. K., & Aslin, R. N. (2002). Gradient effects of within-category phonetic variation on lexical access. *Cognition, 86*(2), B33-B42.

Mirman, D. (2014). *Growth Curve Analysis and Visualization Using R*. Florida, USA: Chapman & Hall/CRC.

Mirman, D., Dixon, J. A., & Magnuson, J. S. (2008). Statistical and computational models of the visual world paradigm: Growth curves and individual differences. *Journal of Memory and Language, 59*(4), 475-494.

Mirman, D., & Graziano, K. M. (2012a). Damage to temporo-parietal cortex decreases incidental activation of thematic relations during spoken word comprehension. *Neuropsychologia*, *50*(8), 1990–1997.

Mirman, D., & Graziano, K. M. (2012b). Individual differences in the strength of taxonomic versus thematic relations. *Journal of Experimental Psychology: General*, *141*(4), 601–609. doi:10.1037/a0026451

Mirman, D., & Magnuson, J. S. (2009). Dynamics of activation of semantically similar concepts during spoken word recognition. *Memory & Cognition, 37*(7), 1026-1039.

Mirman, D., Yee, E., Blumstein, S. E., & Magnuson, J. S. (2011). Theories of spoken word recognition deficits in aphasia: Evidence from eye-tracking and computational modeling. *Brain and Language*, *117*(2), 53–68. doi:10.1016/j.bandl.2011.01.004

Myung, J.-Y., Blumstein, S. E., Yee, E., Sedivy, J. C., Thompson-Schill, S. L., & Buxbaum, L. J. (2010). Impaired access to manipulation features in apraxia: Evidence from eyetracking and semantic judgment tasks. *Brain & Language, 112*, 101-112.

Pratt, J., Dodd, M., & Welsh, T. (2006). Growing older does not always mean moving slower: Examining aging and the saccadic motor system. *Journal of Motor Behavior, 38*(5), 373-382.

Richardson, D. C., & Spivey, M. J. (2000). Representation, space and hollywood squares: Looking at things that aren't there anymore. *Cognition, 76*(3), 269-295.

Salverda, A. P., Dahan, D., & McQueen, J. M. (2003). The role of prosodic boundaries in the resolution of lexical embedding in speech comprehension. *Cognition, 90*(1), 51-89.

Schwartz, M. F., Brecher, A. R., Whyte, J., & Klein, M. G. (2005). A patient registry for cognitive rehabilitation research: A strategy for balancing patients' privacy rights with researchers' need for access. *Archives of Physical Medicine and Rehabilitation, 86*, 1807-1814.

Spivey, M. J., Grosjean, M., & Knoblich, G. (2005). Continuous attraction toward phonological competitors. *Proceedings of the National Academy of Sciences, 102*(29), 10393-10398.

Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science, 268*(5217), 632-634.

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology, 12*, 97-136.

Ventry, I.M., & Weinstein, B.E. (1983). Identification of Elderly people with hearing problems. *American Speech-Language-Hearing Association, 25*, 37-42.

Yee, E., Blumstein, S., & Sedivy, J. C. (2008). Lexical–semantic activation in Broca's and Wernicke's aphasia: Evidence from eye movements. *Journal of Cognitive Neuroscience, 20*(4), 1-21.

Yee, E., Huffstetler, S., & Thompson-Schill, S. L. (in press). Function follows form: Activation of shape and function features during word recognition. *Journal of Experimental Psychology: General*.