Proceedings of the Annual Meeting of the Cognitive Science Society

Title

Eye-tracking Investigation of Visual Search Strategies When Mediated by Language

Permalink

https://escholarship.org/uc/item/6311g331

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 36(36)

ISSN

1069-7977

Authors

Chiu, Eric
Rigoli, Lillian
Spivey, Michael

Publication Date

2014

Eye-tracking Investigation of Visual Search Strategies When Mediated by Language

Eric M. Chiu (echiu@ucmerced.edu)
Lillian M. Rigoli (lrigoli@ucmerced.edu)
Michael J. Spivey (spivey@ucmerced.edu)
Cognitive and Information Sciences, 5200 North Lake Road,
Merced, CA 95343 USA

Abstract

Traditional parallel and serial descriptions of the visual search process are often inadequate when describing recent findings. Accordingly, literature and computational models have evolved from a dichotomous parallel and serial explanation to an account of search efficiency that is graded and continuous. In our current experiment, we replicate findings showing concurrent incremental information processing, via auditory spoken language, mediates visual search and improves search efficiency (Spivey et al., 2001; Reali et al., 2006; Chiu & Spivey, 2012). Novel to this study is the use of eye-tracking to investigate the role of language in mediating and improving strategies for visual search. We find evidence that search is best described as a purely parallel mechanism that immediately and rapidly integrates linguistic and visual information. This finding supports an interactive account of visual attention and spoken language.

Keywords: visual search, language, eye-tracking, dense-sampling, conjunction

Introduction

Humans are inherently limited capacity creatures and as a result crossmodal interactions bestow considerable behavioral advantages. At any given time, only a small amount of the existing information on the retina can be processed and mapped onto motor output. Giving attention to any one stimulus leaves less processing for any others because of the selectivity of attention. The ability to filter out unwanted information allows for awareness of attended stimuli but generally results in unawareness of unattended ones.

Treisman and Gelade's (1980) Feature Integration Theory distinguishes between two perspectives for processing visual search arrays. First is the initial *parallel processing* perspective, which institutes a *single feature* array and accounts for the majority of parallel processing observations. Parallel processing responses are based on a single map of partially active representations of objects simultaneously contending for probabilistic mapping onto motor output. These single feature arrays often induce what is called a perceptual "pop-out" effect, where the unique target object that differs from distractor objects by the only feature (e.g., color, orientation, etc.) in the array appears to pop-out (Treisman & Gelade, 1980; Treisman & Gormican, 1988).

A *conjunction search* uses multiple features, thus multiple maps are needed to identify the presence and subsequently map the location of each feature in a visual field. According to Feature Integration Theory, the mechanism used on conjunction-search arrays is referred to as a *serial search process*, which claims that observers allocate complete attentional resources discretely and wholly to individual objects one at a time (Treisman & Gelade, 1980; Treisman & Gormican, 1988).

Recent findings have demonstrated that instead of two apparently dichotomous perspectives, parallel and serial, attention in visual search may be better described as a single process of graded enhancement of feature salience. This is supported by observed gradual improvements of efficiency in visual search tasks (Olds, Cowan, & Jolicoeur, 2000a; 2000b; 2000c). In a series of experiments, Olds et al. observed facilitatory effects as a result of very brief presentations (less than 100 ms in some conditions) of displays with only single-feature distractors before transitioning to conjunction displays. Although observers' responses were not as fast as with pure pop-out displays, the data produced a graded improvement of efficiency.

To account for findings like those Maioli, Benaglio, Siri, Sosta, and Cappa (2001) argue for a time-limited competitive model of attention in visual search, in which both parallel and serial processing mechanisms are integrated. This perspective is supported in part by neural mechanisms in extrastriate visual cortex that exhibit a form of "biased competition" between multiple object representations that are partially active in parallel (Desimone & Duncan, 1995; Desimone, 1998; Reynolds & Desimone, 2001). Findings like Olds et al.'s (2000a, b, c) "search assistance," along with various other studies (Eckstein, 1998; Wolfe, 1998; Maioli et al., 2001; Watson, Brennan, Kingstone, & Enns, 2010) have largely shifted the description of visual search phenomena from a serial-parallel dichotomy to a graded and continuous account of visual search efficiency (Nakayama & Joseph, 1998).

Further support for this trend comes from work by Spivey, Tyler, Eberhard, and Tanenhaus (2001) that discovered another type of "search assistance" phenomenon. Observers in an *Audio/Visual-Concurrent* (A/V-concurrent) condition, where the conjunction-search display is presented concurrently with target identity delivery via auditory linguistic queries (e.g. "Is there a red vertical?"), exhibited dramatically improved search efficiency. By contrast, in an *Auditory-First* control condition, where the same spoken query of target identity was provided prior to visual display onset, visual search was notably inefficient. The findings suggest that in A/V-concurrent trials, upon hearing the first-mentioned adjective in the spoken query, visual attention is able to begin the search with only that feature, thus initiating the process more efficiently in a single-feature-like search. Then after hearing the second adjective, several hundred milliseconds later, observers can quickly identify the target among the now smaller and more salient subset of objects. Moreover, when target identity is delivered in a more traditional non-linguistic visual method for a conjunction search of this type (Spivey et al., 2001: Experiment 3; Chiu & Spivey, 2012) overall reaction time (e.g., y-intercept and slope) are nearly identical to the auditory-first linguistically mediated visual search condition, which supports a facilitory effect of concurrent visual and auditory linguistic delivery. This finding has been repeatedly reproduced and extended (Reali, Spivey, Tyler, & Terranova, 2006; Chiu & Spivey, 2012). Interestingly, Gibson, Eberhard, and Bryant (2005) found that with faster speech (4.8 vs. 3.0 syllables/second) the A/V-concurrent condition no longer provided an enhanced efficiency effect on conjunction-search tasks, indicating that linguistic mediation of visual search is sensitive to speech rate.

More recently, experiments by Jones, Kaschak, and Boot (2011) used eye-tracking to examine an alternative perspective to one that proposes search efficiency is increased due to language enhancing perceptual processing. Jones et al. (2011) observed eye movement patterns that suggest previously observed improvements in search efficiency with concurrent speech is not likely due to linguistic enhancement of perceptual processes but rather from delaying the onset of target-seeking eye movements. They suggest that the findings by Gibson et al. (2005) are better explained by this “preview” of search display because slower speech provides observers with additional search display viewing time, which affords additional information about potential target locations independently of the information conveyed by auditory linguistic speech stream.

With new advances in eye-tracking techniques researchers can now construct and quantify robust illustrations of real-time cognitive processes such as identify fixation rich regions over a time period. We use this method to investigate differences in eye-movement and -fixations during a linguistically mediated conjunction search task.

Experiment

In this experiment we observe, using eye-tracking methods, the mechanisms of visual search during a conjunction search task mediated by language.

Method

We utilized a mixed design. The search displays used were the same for all participants but presented in random order. Trials were split evenly between the A/V-concurrent and auditory-first conditions. Participants were randomly assigned to one of two groups. Participants in the first group, *Group A*, received half of the search displays in one of the two conditions (A/V-concurrent or auditory-first) and the other half of the search displays in the remaining condition. Participants in the second group, *Group B*, received the same search displays but had them presented in the opposite condition as the participants in the Group A, such that any given display was presented as both conditions across both groups. This allows for the between-subject comparison of search strategies among conditions for any given search display. Target-present and –absent trials along with the four set sizes (5, 10, 15, & 20) appeared randomly and equally. The two conditions were presented randomly and intermixed. While performing in the conjunction search task observers’ eye-movements were recorded for all trials using an Eye-Link II head mounted eye-tracker (SR Research Ltd., Mississauga, Ontario, Canada).

Participants Sixty-eight undergraduate students from the University of California, Merced received course credit for participation in this experiment. All of the participants had normal, non-corrected, vision as well as normal color perception. Those participants who did not reach 80% accuracy were omitted from the analysis. Three participants did not perform to these standards.

Stimuli and Procedure Identical pre-generated search displays were used for each observer. The stimulus bars subtended 2.8° X 0.4° of visual angle and neighboring bars were separated from one another by an average of 2.0° of visual angle. The green and red bars had the same luminance of 13.4 cd/m^2. Appearance of the target object in quadrants (top-left & -right, bottom-left & -right) as well as the type of target (e.g., green horizontal), and set sizes of objects (5, 10, 15, & 20) appeared equally. Observers were randomly assigned to participate in one of two groups (A or B). The two groups were indistinguishable but differed in that identical search displays were presented in an auditory-first trial for one group and an A/V-concurrent trial for the other group. In half of the trials, a spoken query (e.g., “Is there a red vertical?”) informed participants of the targets’ identity prior to display onset (auditory-first), and for the other half, the first adjective of the spoken query coincided with the appearance of the visual display (A/V-concurrent condition). An identical 1000 ms prelude recording (“Is there a…”) was used with two target-identifying adjectives (color & orientation), together averaging 1500 ms (fig. 1).

An Eyelink II head mounted video-based eye-tracker with a temporal resolution of 250 Hz and a spatial resolution of 0.025° recorded eye movements by tracking pupil and corneal reflection. The video-based eye-tracker used two infrared LEDs mounted on the headband to illuminate each eye. Tracking was monocular although viewing was binocular. The eye-tracker classified an eye movement as a saccade when its distance exceeded 0.2° and its velocity reached 30°/second or when its distance exceeded 0.2° and its acceleration reached 9500°/second^2. The displays were generated using Mathworks MATLAB software and the experiment was designed using SR Research Experiment Builder. Stimuli were presented on a 22” ThinkVision LCD monitor with 1280 x 1024 resolution. The prerecorded speech queries, all from the same female speaker, are identical to Spivey et al. (2001) and were presented through Harmon Kardon HK206 desktop computer speakers.

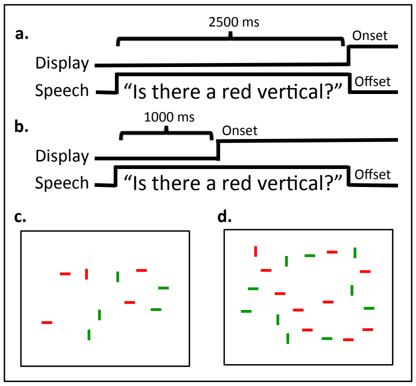

Figure 1: Examples of the auditory and visual stimuli. In the auditory-first control condition (a) the onset of the visual display coincided with the offset of the spoken target query. In the audiovisual-concurrent (A/V-concurrent) condition (b), the onset of the visual display coincided with the onset of the first target-feature word in the spoken query. The example displays show target-present trials with a set size of 10 (c) & 20 (d) where the target is a red vertical bar, which is accompanied by green vertical and horizontal as well as red horizontal distractor bars.

The Eyelink eye-tracker was calibrated using the standard nine-point calibration method for each participant. Calibration was followed by 16 practice trials to familiarize participants with the task and the eye-tracker. The experiment consisted of 128 trials. Observers were instructed to keep their fingers resting on the marked response keys and to respond as quickly and accurately as possible by pressing “YES” and “NO” if the target was present or absent, respectively. Before each trial, participants were required to fixate their gaze on a fixation cross in the center of the screen; this was also used as a “drift correct,” which verified that the initial calibration remained valid. Participants initiated each trail by pressing the space bar while fixating on the fixation cross. It was very rare for an observer to have an invalid drift correct, which required the experimenter to recalibrate. The entire experiment lasted approximately 30 minutes.

Results and Discussion

A hierarchal linear model (HLM) was used for this analysis because it accounts for the unbalanced number of observations by condition, due to our data culling process and repeated measures design. To fulfill the assumption of distribution normality the inferential statistics were performed on log-transformed RTs, as RT response data are bound on the left but not the right, and thus is naturally positively skewed (Luce, 1986). However, descriptive statistics (slopes and intercepts of RTs in milliseconds) continue to be reported from an untransformed HLM. Participants’ RTs as well as eye-movement and –fixation data were recorded from display onset. All trials with incorrect responses were removed from the analysis as well as trials with RTs greater than 2.5 interquartile ranges from the mean.

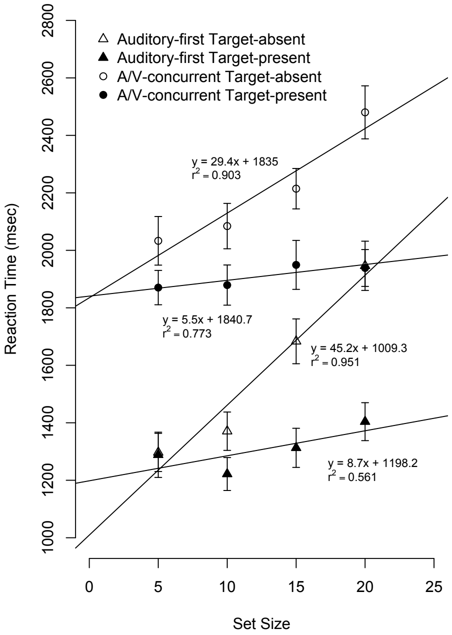

Figure 2: Experiment results. Shown separately for target-present (filled symbols) and –absent trials (open symbols) for cue-first (triangles) and cue-concurrent (circles) conditions. Each line is accompanied by the best-fit linear equation and the proportion of variance accounted for (r^2). Error bars indicate standard error of the mean.

In this experiment, we replicated previous findings demonstrated by Spivey et al. (2001) and Reali et al. (2006). The RT-by-set-size functions are highly linear for the auditory-first condition in both target-present, $r^2 = .561$, and -absent, $r^2 = .951$, trials as well as for the A/V-concurrent condition for target-present, $r^2 = .773$, and -absent, $r^2 = .903$, trials, which is typical of standard conjunction search tasks. Mean accuracy across all trials is 95.0%, which is consistent with previous studies (Spivey et al., 2001; Reali et al., 2006, Chiu & Spivey, 2012). As expected, the slopes of the RT-by-set-size functions reveal that the A/V-concurrent condition produces more efficient visual search (shallower slope) when compared with the auditory-first condition (fig. 2). The HLM analysis revealed significantly shallower slopes for the A/V-concurrent condition compared to the auditory-first condition in target-present (5.5 vs. 8.7 ms/item), $t(64) = -3.23$, $p < .001$, and -absent (29.4 vs. 45.2 ms/item), $t(64) = -10.24$, $p < .001$, trials, as previously observed by Spivey et al. (2001), Reali et al. (2006), and Chiu and Spivey (2012). Overall mean RT, as well as y-intercepts, were significantly slower in A/V-concurrent conditions because complete delivery of target identity was delayed by approximately 1500 ms relative to the auditory-first condition for both target-present, $t(64) = 184.79$, $p < .001$, and -absent, $t(64) = 250.27$, $p < .001$, trials.

We see the results of this experiment continue to show observers were able to find the target object in a way that was substantially less affected by the number of distractors, simply by adjusting the timing of spoken query so the two target-feature words are heard while viewing the visual display. It appears that the incremental nature of speech processing allows the visual search process to begin when only a single feature of the target identity has been heard. When the initial feature is identified the search proceeds in an efficient nearly-parallel fashion so when the second adjective is heard, a substantial amount of the target identification process has been completed – and thus the presence of multiple distractors is less disruptive.

Table 1: Number of fixations for target-present trials.

Set Size	A/V-concurrent		Auditory-first	
5	*M* = 12.1	*SD* = 5.1	*M* = 13.4	*SD* = 7.3
10	*M* = 12.7	*SD* = 4.9	*M* = 14.0	*SD* = 5.7
15	*M* = 14.2	*SD* = 10.0	*M* = 14.5	*SD* = 6.2
20	*M* = 13.2	*SD* = 5.9	*M* = 14.5	*SD* = 5.4

Of primary and novel interest in this experiment is the analysis of eye-movement patterns during the well replicated linguistically mediated conjunction search task (Spivey et al., 2001; Reali et al., 2006; Chiu & Spivey, 2011). Figure 3 shows a 100 ms time slice of target-present trials with fixations; the four set sizes are shown separately for A/V-concurrent and auditory-first. One can clearly see that for the same time-slice, fixations in the A/V-concurrent trials are primarily focused on color-matched objects, while fixations in the auditory-first trials do not appear to exhibit any pattern. This phenomenon is consistent across all of the trials. The following analyses investigate the claim that A/V-concurrent target-feature delivery does indeed elicit a different and more efficient oculomotor search of the display than the auditory-first condition.

First, we find significantly fewer fixations across each trial in the A/V-concurrent condition ($M = 13.08$, $SD = 6.89$) as compared to the in the auditory-first condition ($M = 17.23$, $SD = 23.76$), $t(64) = 17.18$, $p < .001$, which is consistent with the idea that upon hearing the first target-identifying adjective a rapid parallel-like search process weeds out conflicting colored distractors and increases the saliency of fitting objects. Further analysis finds that, across the four set sizes, the number of fixations is again significantly smaller for the A/V-concurrent condition, $f(64) = 116.7$, $p < .001$ (Table 1). It should be noted that the descriptive statistics reported here solely involve target-present trials because search strategies in target-absent trials have been found to differ from those of target-present and are notoriously difficult to simulate. Although many of the differences between an A/V-concurrent target-identity delivery and an auditory-first delivery are observed with target-absent trials as well, they are beyond the scope of this report and will not be discussed.

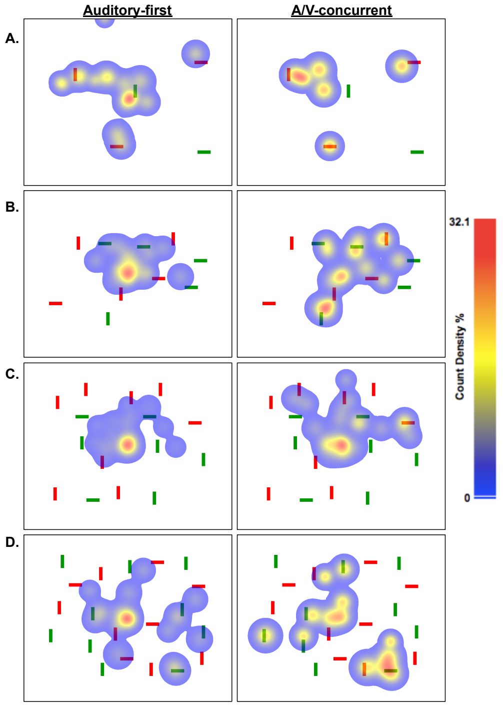

Figure 3: Eye-tracking results. Target-present trials are shown separately for auditory-first control and the A/V-concurrent trials. Search displays are overlapped with a heat map representing fixation activity (blue = low, yellow = medium, and red = high). A single 100 ms time period is depicted for each set size: 1600-1700 ms for 5 (A), 1900-2000 ms for 10 (B), 1700-1800 ms for 15 (C), and 1100-1200 ms for 20 (D). Targets for each trial are as follows: 5 = red vertical, 10 = green vertical, 15 = red horizontal, and 20 = green horizontal. Fixation patterns differed drastically between trials thus different time period were chosen for each set size but were the same between conditions.

Second, the average duration of each fixation is also significantly shorter for the A/V-concurrent condition, 335.7 ms ($SD = 395.0$), than for the auditory-first condition, 382.2 ms ($SD = 503.4$), $t(64) = 7.33$, $p < .001$. Since observers in the auditory-first condition receive both target-identifying adjectives before the onset of the search display, it is possible that they are judging each fixated object by both features at once in search of the unique target, which would explain the longer fixation durations when compared to A/V-concurrent trials. Since observers in the A/V-concurrent condition appear to have already isolated attention to objects that match the identified color feature, they would only have to judge each fixated object on the one remaining feature (orientation). Further analysis of this effect has found that fixation durations are significantly shorter for A/V-concurrent trials when compared to auditory-first trials across all four set sizes, *f*(64) = 39.1, $p < .001$ (Table 2).

Interestingly, across both conditions, eye-fixation duration decreased (392.6, 356.2, 349.0, and 345.8 ms) as set size increased; this main effect is significant, *f*(64) = 24.83, $p < .001$. This pattern may be the result of a desire to respond quickly, as instructed at the beginning of the experiment. Thus when observers see there are more objects, they may speed their search strategy, surprisingly with no significant affect on accuracy (5.1, 4.7, 5.0, & 6.2% errors), *f*(64) = 0.46, $p < .497$. This may be the result of including the first fixation (from display onset to the first saccade) in the analysis, which is longer than most fixations and may have drove up the average fixation duration. A later report of these findings will test this hypothesis.

Table 2: Fixation duration for target-present trials.

Set Size	A/V-concurrent		Auditory-first	
5	*M* = 361.6	*SD* = 414.4	*M* = 421.6	*SD* = 547.5
10	*M* = 334.6	*SD* = 376.2	*M* = 375.5	*SD* = 488.0
15	*M* = 327.2	*SD* = 397.6	*M* = 368.3	*SD* = 484.9
20	*M* = 320.8	*SD* = 391.2	*M* = 367.9	*SD* = 494.5

The average length of saccades, rapid ballistic movements of the eye between fixation points, referred to as amplitude and measured in degrees of visual angle, are significantly longer for A/V-concurrent, 5.24 (*SD* = 6.88), than for auditory-first, 4.73 (*SD* = 7.03), $t(64) = -4.95$, $p < .001$, trials. If it is the case that observers in an auditory-first trial are performing a traditional serial search process, where they attend to each object wholly and discretely to judge whether it is the target, then we can presume their attention would jump from one object to the next closest object to optimize their search strategy until the target was found. This scenario would describe why saccade amplitudes are shorter for auditory-first trials than for A/V-concurrent trials. Because half of the objects are effectively ruled out in A/V-concurrent trials, distance from one viable object to another viable color-matched object would tend to be longer than simply to the next closest object. Further analysis found that saccade amplitudes are significantly shorter for auditory-first trials when compared to A/V-concurrent trials across all four set sizes, *f*(64) = 14.21, $p < .001$ (Table 3). Furthermore, saccade amplitude appears to be decreasing as set size increases, which would support a serial like search process because the to-be-next-fixation object would be closer with a larger set size than with a smaller set size. We do not see the same pattern with A/V-concurrent trials.

Table 3: Saccade amplitude for target-present trials.

Set Size	A/V-concurrent		Auditory-first	
5	*M* = 5.33	*SD* = 7.00	*M* = 5.22	*SD* = 8.06
10	*M* = 4.45	*SD* = 6.14	*M* = 4.38	*SD* = 6.55
15	*M* = 5.32	*SD* = 6.91	*M* = 4.43	*SD* = 6.50
20	*M* = 5.84	*SD* = 7.36	*M* = 4.96	*SD* = 7.01

The strongest evidence supporting the claim that search patterns differ dramatically between A/V-concurrent and auditory-first trials comes from an analysis of a target-present trial with a set size of 20 where the target-object is a green horizontal bar. We see that the average amount of time spent fixating green bars is greater than red when presented as an A/V-concurrent trial (46.5 vs. 17.5 ms; difference of 29 ms) than when presented prior in an auditory-first trial (48.4 vs. 21.2 ms; difference of 27.2 ms), *f*(64) = 7.33, $p < .001$; given a range of 21.1-14.5 fixations per trial this difference would add up. Moreover, when we look at only the first 1000 ms of the trial, we see that observers continue to fixate significantly longer on green bars than red for A/V-concurrent trials (19.7 vs. 9.4 ms), $t(64) = -2.02$, $p = .044$, but not for auditory-first trials (12.6 vs. 12.0 ms), $t(64) = -0.13$, $p = .897$. Thus, after hearing the color feature and only part of the orientation feature, in the first 1000 ms of an A/V-concurrent trial, observers appear to be immediately using the concurrent audio-visual information to bias their eye fixations toward color-appropriate objects, thereby improving efficiency compared to when the spoken target query precedes onset of the visual display.

General Discussion

The findings here are consistent with the inferences made in prior linguistically mediated visual search reports (Spivey et al., 2001; Reali et al., 2006; Chiu & Spivey, 2012). The significantly fewer fixations, shorter fixation durations, and larger saccade amplitudes observed when auditory linguistic target features are delivered concurrent with display onset, in the A/V-concurrent condition, compared to when target features were delivered prior to display onset, in the auditory-first condition, provides further evidence supporting the notion that observers employ distinctive search strategies when display onset timing is altered in relation to feature delivery. Furthermore, the longer dwell time observed with color-matched objects than non-matched distractors throughout an A/V-concurrent trial, and especially in the first 1000 ms, supports the existence of a fast-acting efficient parallel search process that does not occur in an auditory-first trial.

The novel discoveries here further promote the claim that upon hearing the first-mentioned adjective in a spoken query, visual attention is able to begin the search with only that single feature. Thus, the process is initiated with a highly efficient single-feature search such that when the second adjective is delivered, several hundred milliseconds later, the target can be quickly found among the attended subset of objects. Conversely, trials presented in the auditory-first condition appear to exhibit a search strategy representative of a traditional serial search processes, by which each object in the search display is compared to the aforementioned target-object one at a time until the target-object is located in a target-present trial.

Conclusion

These results support a robust interactive account of visual perception that explains language mediation of visual search is chiefly due to the capacity to rapidly and immediately integrate incremental linguistic information with visual information. This study provides us with significant insight into the mechanisms of auditory language mediated visual search but also adds to the complexity of the dynamic relationship, which escalates the importance and compels the need for additional exploration with additional experimental tests such as this.

Acknowledgments

We are grateful to Greg Perlman from SR Research for help with experiment design. We are also thankful to Lydia Goes and Maria Vega for their help with data collection.

References

Chiu, E. M. & Spivey, M. J. (2012). The role of preview and incremental delivery on visual search. In N. Miyake, D. Peebles, & R. P. Cooper (Eds.), *Proceedings of the 34th Annual Conference of the Cognitive Science Society* (pp. 216-221). Austin, TX: Cognitive Science Society.

Desimone, R. (1998). Visual attention mediated by biased competition in extrastriate visual cortex. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *353*(1373), 1245.

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience*, *18*(1), 193–222.

Eckstein, M. P. (1998). The lower visual search efficiency for conjunctions is due to noise and not serial attention processing, *Psychological Science, 9*, 111-118.

Gibson, B. S., Eberhard, K. M., & Bryant, T. A. (2005). Linguistically mediated visual search: The critical role of speech rate. *Psychonomic Bulletin and Review*, *12*(2), 276.

Jones, J. J., Kaschak, M. P., & Boot, W. R. (2011). Language mediated visual search: The role of display preview. In N. Miyake, D. Peebles, & R. P. Cooper (Eds.), *Proceedings of the 34th Annual Conference of the Cognitive Science Society* (pp. 2739-2744). Austin, TX: Cognitive Science Society.

Luce, R. D. (1986). *Response Times: Their Role in Inferring Elementary Mental Organization3* (Vol. 8). Oxford University Press.

Maioli, C., Benaglio, I., Siri, S., Sosta, K., & Cappa, S. (2001). The integration of parallel and serial processing mechanisms in visual search: Evidence from eye movement recording. *European Journal of Neuroscience. 13*(2), 364-372.

Nakayama, K., & Joseph, J. S. (1998). Attention, pattern recognition, and pop-out in visual search. *The attentive brain*, 279–298.

Olds, E. S., Cowan, W. B., & Jolicoeur, P. (2000a). Partial orientation pop-out helps difficult search for orientation. *Perception & psychophysics*, *62*(7), 1341–1347.

Olds, E. S., Cowan, W. B., & Jolicoeur, P. (2000b). The time-course of pop-out search. *Vision Research*, *40*(8), 891–912.

Olds, E. S., Cowan, W. B., & Jolicoeur, P. (2000c). Tracking visual search over space and time. *Psychonomic Bulletin and Review*, *7*(2), 292–300.

Reali, F., Spivey, M. J., Tyler, M. J., & Terranova, J. (2006). Inefficient conjunction-search made efficient by concurrent spoken delivery of target identity. *Perception and Psychophysics*, *68*(6), 959.

Reynolds, J., & Desimone, R. (2001). Neural mechanisms of attentional selection. *Visual attention and cortical circuits (Braun J, Koch C, Davis JL, eds)*, 121–136.

Spivey, M. J., Tyler, M. J., Eberhard, K. M., & Tanenhaus, M. K. (2001). Linguistically mediated visual search. *Psychological Science*, *12*(4), 282-286.

Treisman, A., & Gormican, S. (1988). Feature analysis in early vision: Evidence from search asymmetries. *Psychological Review*, *95*(1), 15–48.

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive psychology*, *12*(1), 97–136.

Watson, M. R., Brennan, A. A., Kingstone, A., & Enns, J. T. (2010). Looking versus seeing: Strategies alter eye movements during visual search. *Psychonomic Bulletin & Review, 17*(4), 543-549.

Wolfe, J. M. (1998). What can 1 million trials tell us about visual search? *Psychological Science, 9*, 33-39.