Incidental and context-responsive activation of structure- and function-based action features during object identification

Chia-lin Lee, **Erica Middleton**, **Daniel Mirman**, **Solène Kalénine**, and **Laurel J. Buxbaum**
Moss Rehabilitation Research Institute, Philadelphia, PA

Abstract

Previous studies suggest that action representations are activated during object processing, even when task-irrelevant. In addition, there is evidence that lexical-semantic context may affect such activation during object processing. Finally, prior work from our laboratory and others indicates that function-based ("use") and structure-based ("move") action subtypes may differ in their activation characteristics. Most studies assessing such effects, however, have required manual object-relevant motor responses, thereby plausibly influencing the activation of action representations. The present work utilizes eyetracking and a Visual World Paradigm task without object-relevant actions to assess the time course of activation of action representations, as well as their responsiveness to lexical-semantic context. In two experiments, participants heard a target word and selected its referent from an array of four objects. Gaze fixations on non-target objects signal activation of features shared between targets and non-targets. The experiments assessed activation of structure-based (Experiment 1) or function-based (Experiment 2) distractors, using neutral sentences ("S/he saw the …") or sentences with a relevant action verb (Experiment 1: "S/he picked up the……"; Experiment 2: "S/he used the…."). We observed task-irrelevant activations of action information in both experiments. In neutral contexts, structure-based activation was relatively faster-rising but more transient than function-based activation. Additionally, action verb contexts reliably modified patterns of activation in both Experiments. These data provide fine-grained information about the dynamics of activation of function-based and structure-based actions in neutral and action-relevant contexts, in support of the "Two Action System" model of object and action processing (e.g., Buxbaum & Kalénine, 2010).

Introduction

There is growing evidence suggesting that object-associated actions play an important role in object processing. For example, judgments about the categories, orientations, or sizes of object images or names are faster when signaled with a response gesture compatible with the object (e.g., precision grip for key and power grip for hammer; Tucker & Ellis, 1998, 2001,

Corresponding author: Chia-lin Lee Moss Rehabilitation Research Institute Medical Arts Building, 50 Township Line Road, Elkins Park PA, USA, 19027 phone: (+001) 215-663-6311 leechial@einstein.edu.

2004). These and other related findings have been taken as support for task-incidental activation of action attributes during object processing. The logic of this conclusion rests on the assumption that the response action is facilitated by the object. However, many such studies have required participants to prepare and execute manual actions as a means of responding, thereby plausibly influencing the degree to which action-related object attributes may be activated. It is well known that preparation of a specific grasping gesture can, on the one hand, facilitate visual detection of objects that are congruent with the planned grip (Craighero, Bello, Fadiga, & Rizzolatti, 2002; Craighero, Fadiga, Rizzolatti, & Umiltà, 1999; Müsseler, Steininger, & Wühr, 2001; also see Symes, Tucker, Ellis, Vainio, & Ottoboni, 2008; Vainio, Symes, Ellis, Tucker, & Ottoboni, 2008), and on the other, increase or reduce interference from nearby objects with congruent or incongruent action features (Bekkering & Neggers, 2002; Botvinick, Buxbaum, Bylsma, & Jax, 2009; Pavese & Buxbaum, 2002). These and other similar studies have demonstrated that preparing to act on an object produces faster processing of objects congruent with the planned movement ("*Motor-visual priming*"). From this perspective, it is uncertain whether activation of action features may be accurately characterized as task-incidental when an object-related response is required.

However, there are at least some circumstances in which action attributes are activated during object processing even in the absence of motor planning. These circumstances all appear to entail the performance of a lexical-semantic task. Thus, for example, matching an object picture to a word is less accurate if the target object is shown concurrently with another object sharing similar manipulation features (e.g. target: pincers, distracter: nutcracker) than when shown with an object not sharing these features (e.g. target: pincers, distracter: candle; Campanella & Shallice, 2011). Similarly, when cued by an auditory word to identify a target object (e.g., `typewriter') among distractors, participants looked at distractors that can be moved or used with similar actions (e.g., `piano') more than non-action-related objects (e.g., `couch') (Myung, Blumstein, & Sedivy, 2006). Recent studies have also shown that verbal context may draw participants' attention to action features (Kalénine, Mirman, Middleton, & Buxbaum, under review; Kamide, Altmann, & Haywood, 2003). For example, objects are identified faster following a sentence containing an action verb (`grasp') than an observation verb (`looked at') (Borghi & Riggio, 2009). Together, these findings suggest that activation of object-associated action attributes may occur in some lexical-semantic contexts even without overt object-oriented motor preparation.

An additional area of uncertainty concerns the types of actions that may be activated during object processing. Many manipulable objects are associated with several manual actions (e.g., a computer keyboard can be poked to use or clenched to move). One question is whether all or only some of these object-associated actions are activated incidentally when the task is being performed. Additionally, little is known about how these action types may differ in their patterns of activation.

Based in part on the observation that patients with action deficits (apraxias) respond normally to objects' structural characteristics (shape, size, and volume) in the face of substantial deficits in knowledge-based use actions, the Two Action Systems (2AS) model hypothesizes a distinction between (1) grasp-to-move (power and precision grip) actions driven by object structural attributes, and (2) skilled use actions reliant upon knowledge of the identity and function of objects (e.g., Buxbaum, 2001; Buxbaum & Coslett, 1998; Buxbaum, Sirigu, Schwartz, & Klatzky, 2003; and see Jeannerod, 1997; Johnson-Frey, 2004; Pisella, Binkofski, Lasek, Toni, & Rossetti, 2006; Rizzolatti & Matelli, 2003; Vingerhoets, Acke, Vandemaele, & Achten, 2009). Furthermore, based on patient lesion and functional neuroimaging data (Buxbaum, Kyle, Grossman, & Coslett, 2007; Buxbaum et al., 2003; Sirigu et al., 1996), these `*structure-based*' (grasp-to-move) and `*function-based*'

(skilled use) actions are proposed to have distinct temporal processing characteristics (see Buxbaum & Kalénine, 2010 for a review). Structure-based action features are hypothesized to become active rapidly upon sight of an object, but only for a transient period of time. Function-based actions require more time to access but remain available for longer, an activation pattern that is characteristic of semantic memory (c.f. Campanella & Shallice, 2011).

The hypothesized differences in the processing characteristics of the two action systems are supported by a recent study that measured participants' initiation times to act on objects that are picked up and used with different actions (`conflict objects', e.g., calculator) or objects that are picked up and used with the same actions (`non-conflict objects', e.g., cup) (Jax & Buxbaum, 2010). Initiation times for function-based actions were slower for conflict objects than non-conflict objects, implicating interference from structure-based action attributes. For example, initiating movement for using a calculator with a "poking" action was slowed by the task-irrelevant activation of the clench action required to grasp the calculator (within-object grasp-on-use interference). In contrast, initiation times for structure-based actions were only slower for conflict- than non-conflict objects when participants had performed function-based actions on the same objects in earlier blocks. In other words, interference from function-based actions upon structure-based actions occurred only when the function-based actions had been activated previously, suggesting a comparatively slower pattern of activation and decay. Thus, the two types of object-related actions differ significantly in their patterns of temporal activation. Critically, however, the study design made it impossible to ascertain whether the observed activations occurred, at least in part, as a consequence of the preparation of object-related actions resulting in motor-visual priming. In addition, the method permitted only rough characterization of the temporal activation characteristics of the two action types.

In view of the outstanding issues, the goals of the current study were 1) to assess whether both function-based and structure-based action features are activated incidentally in a word-picture matching task without target object-related actions, 2) to extend the findings of Jax and Buxbaum (2010)by assessing the temporal dynamics of activation of function-based and structure-based action attributes over time, and 3) given the evidence reviewed earlier that verbal context may facilitate action features of objects, assess whether verbal context may differentially facilitate these two action types. To this end, two eye tracking experiments were conducted using the Visual World Paradigm.

In a typical Visual World Paradigm (VWP) study, participants' eye movements are recorded while they point to or click on an auditorily-cued target picture shown as part of a visual display. A related distractor ("competitor") that shares attributes of interest with the target is typically also displayed, along with unrelated distractor pictures that do not share these attributes. For example, for a given target object such as `typewriter', the distractors might include an object sharing action attributes with the target (the related distractor, e.g., `piano') as well as objects completely unrelated to the target (the unrelated distractor, e.g.,`couch'; examples taken from Myung, Blumstein, & Sedivy, 2006). As the related distractor and the unrelated distractors in the same array are typically matched on a number of features (e.g., visual complexity, familiarity, etc) and differ only in whether they share critical attributes with the target, more fixations on the related relative to unrelated distractor can be used to infer whether the critical attributes are incidentally activated. The competition effects can thus be thought of as analogous to priming (Huettig & Altmann, 2005). Moreover, by comparing gaze fixations on the competitors and unrelated items in the same display, one can infer the activation time course of the object attributes in question. This approach has been used to demonstrate task-incidental activation of relevant object attributes along dimensions such as phonology (e.g., Allopenna, Magnuson, & Tanenhaus, 1998; Dahan,

Magnuson, Tanenhaus, & Hogan, 2001), semantics (e.g., Huettig & Altmann, 2005; Mirman & Magnuson, 2009; Yee & Sedivy, 2006), and manipulation actions (Myung et al., 2006, 2010). Moreover, it has been shown that eye movement patterns closely track the unfolding of auditory instructions (Allopenna et al., 1998; Tanenhaus & Spivey-Knowlton, 1996; Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995). Such fine-grained implicit and continuous measurement of the word-to-picture matching process is thus ideal for the purposes of the current study.

To examine the time courses of structure-based and function-based action attributes as well as their potential modification by verbal contextual information, we manipulated (1) the action attributes shared between each target and its competitor and (2) the type of linguistic context. Targets and corresponding competitors shared structure-based actions in Experiment 1 (e.g., stapler and hammer) and function-based actions in Experiment 2 (e.g., remote control and key fob). In both experiments, sentence contexts included a neutral context (`S/he saw the ….' in both Experiments) and an action verb context (`S/he picked up the …' in Experiment 1 vs. `S/he used the ….' in Experiment 2).

There were several predictions. First, if there is indeed task-incidental activation of both function-based and structure-based action features during object processing, then we would expect to see more fixations on action-related competitors than unrelated items in the neutral contexts of both experiments. Second, if the patterns observed by Jax and Buxbaum (2010) reflect differences in the activation time courses of the two action types during object processing, rather than byproducts of motor planning, we would expect to see similar temporal differences in the current study, although with a much finer temporal resolution. Specifically, in the neutral context we predicted a relatively fast-rising but transient structure-based competition effect and a slower rising but longer-lasting function-based competition effect. We also predicted modulation by action verb context. Action verbs and action-related linguistic materials have been shown to trigger neural activations in primary motor and/or pre-motor regions (Buccino et al., 2005; Pulvermuller, 2005; Raposo, Moss, Stamatakis, & Tyler, 2009) and facilitate compatible motor gestures (Glenberg & Kaschak, 2002; Taylor & Zwaan, 2008). We therefore predicted facilitated object identification and competition effects in the action verb contexts as compared to the neutral context. Finally, we speculated that we might observe differences in the degree to which action verb context facilitates structure-based versus function-based actions. Specifically, if function-based actions have relatively close links to the lexical-semantic system, as has been proposed (e.g., Buxbaum & Kalénine, 2010) it might be possible that "used" sentences are relatively facilitatory of function-based activations. Similarly, if "picked up" actions are relatively more strongly dependent on structural object attributes, then the "picked up" verb context may have only a weak effect on structure-based actions.

Experiment 1

Method

Participants—Twenty healthy older adults (5 males; mean age: 68.8 years; range 57–78 years; SD: 6.1 years) participated in the study[1]. All participants were recruited from a subject database maintained by Moss Rehabilitation Research Institute. All participants gave informed consent to participate in accordance with the IRB guidelines of the Albert Einstein Healthcare Network and were remunerated with cash for their participation. All participants

---

[1]Older adults were selected for the purpose of future comparison with the stroke patients frequently run in our laboratory. Numerous publications from our laboratory assessing action and object representations have pursued this strategy (e.g., Botvinick et al., 2009; Pavese & Buxbaum, 2002). In one such study, results from two experiments were highly comparable whether younger or older adults were tested (Botvinick et al., 2009).

were right handed, had normal or corrected-to-normal vision, normal hearing, no history of neurological/psychiatric disorders or brain damage based on self-report, and scored within the normal range on the Mini Mental State Examination (MMSE; Folstein, Folstein, & McHugh, 1975) (mean score: 29.0/30, SD: 1.0, range: 27–30). On average, these participants had a mean education level of 16 years (SD=1.8 years; range 12–18 years).

Materials

Visual stimuli: Twenty-two arrays of color object images were created for critical trials. Each array included a target, a competitor, and two unrelated images. All target images were objects involving distinct function-based (skilled use) and structure-based (grasp-to-move) action gestures (e.g. a TV remote control). Competitors shared structure-based, but not function-based, action features with their corresponding targets (e.g. a blackboard eraser). For unrelated items, care was taken to assure that no unrelated items shared any action features with the corresponding target or the competitor. This yielded a strong constraint on the choice of items such that even after efforts to match the images in each array for visual and action-related characteristics, the unrelated items tended to be less visually similar to the target than the competitor or less manipulable. In view of possible influences of these differences in stimuli characteristics, we collected norming values for the stimuli used in Experiment 1 and 2, on (1) the visual similarity between distracter items and target items and (2) the degree to which the critical images fit expectations generated by the verb context. These norming results were incorporated into the subsequent analyses in order to control for their contribution to the effects of interest. None of the participants in either norming session participated in the eye tracking experiments.

Visual similarity: Each of the distracter images from both experiments was paired with their corresponding target image, totaling 157 pairs. Images in a pair were arranged side by side, with target images always being on the left. Sixteen participants were asked to judge, on a 1 to 7 point scale, the visual similarity of the two images in each pair (7 = highly similar, 1 = not at all similar). Specifically, participants were asked to rate the similarity of the images in appearance, but not the objects they represented. Means and standard deviations of the rating scores for each object type are provided in Table 1.

Context fitness: Each image from the critical arrays was shown individually below the sentence contexts in which it occurred in the experiments (`S/he picked up/used the _____.' and `S/he saw the _____.'). A different group of 17 participants were asked to rate, on a 1–7 point scale, the extent to which the displayed object fit expectations generated by the sentence context (7 = perfect fit to context, 1 = no fit to context). Means and standard deviations of the rating scores for each object type are provided in Table 2.

Auditory Stimuli—Auditory stimuli included each of the target names (e.g., `the stapler') as well as the sentence contexts: two carrier phrases for action verb contexts (e.g., `She/He picked up') and two for neutral contexts (e.g., `She/He saw'). All stimuli were produced by a female native speaker of American English. Each auditory target word was recorded separately and then appended to the end of each auditory sentence context using the open source sound-editing program *Audacity*. Onset of the target words started 1400ms after the beginning of the sound file.

Apparatus

Gaze position was recorded using an EyeLink 1000 desktop eyetracker at 250 Hz and parsed into fixations using the built-in algorithm with default settings. Stimulus presentations and response recording were conducted by E-Prime software (Psychological Software Tools, Pittsburgh, PA).

Procedure

Twenty-two critical arrays were presented once in the action verb context and once with the neutral context. In addition, 88 arrays were created as filler trials from the 22 critical arrays according to the following scheme (also illustrated in Table 3). To make the target images in the critical trials unpredictable and reduce the prominence of the action verb context, each critical array was presented on a third trial in the neutral context, but with one of the original unrelated items as the target (e.g. Filler 1 in Table 3). To make the relation between the targets and competitors less noticeable and to again reduce the predictability of conflict objects being the targets, each of the distracter items in the original critical arrays was mixed with three other new images to form new arrays and served as targets in these new arrays (e.g. Filler 2–4 in Table 3). Among these new images, two were occurrences of a target from another critical array (e.g. *Corkscrew* in Filler 2 and 3 was the target image from another critical array). Half of these 66 new arrays were presented with the neutral context and the other half with the action verb context.

Overall, each participant saw 132 trials of which 44 were experimental trials with action related objects. Of the 132 trials in total, 77 trials had a neutral verb context and 55 had an action verb context.

Participants were seated with their eyes approximately 27 inches from a 17-inch screen (resolution 1024×768 pixels). Each trial started with the participant clicking on a central fixation cross. Four images were presented simultaneously subsequent to the mouse click; each image was presented near one of the screen corners with a maximum size of 200×200 pixels (each picture subtended about 3.5° of visual angle). The location of target, related, and unrelated distracters was randomized on each trial. After a 1 second preview to allow for initial fixations driven by random factors or visual salience (as opposed to concept processing), participants heard the auditory stimuli through speakers. They were instructed to click on the image corresponding to the word at the end of the sentence as fast as possible. Upon the mouse-click response, the visual array disappeared and was replaced by two text boxes presented side-by-side on the screen, each containing one verb (`saw' or `picked up'). Participants were instructed to click on the verb mentioned in the sentence they just heard. This was to ensure that participants were paying attention to the sentence context during the experiment. The text boxes disappeared upon participants' mouse-click response, terminating the trial.

All participants used their left hand to respond for the purpose of future comparison with left-hemisphere stroke patients who may not be able to use their contralesional hand. In the beginning of the session, participants were given a familiarization session to ensure that they were familiar with the labels for each image. In the familiarization session, each image was presented at the center of the screen with its label presented visually below the image as well as auditorily through the speakers. Participants pressed the space bar to advance to the next image. Prior to the experiment, participants were given a 30-trial practice session to orient them to the task.

Eye movement recording and data analysis

Eye movements were recorded from the beginning of each trial until the mouse-click response on the images. Four areas of interest (AOI) associated with the displayed pictures were defined as 400×300 pixel quadrants situated in the 4 corners of the computer screen. Fixations were counted toward each object type (Target, Competitor, and Unrelated items) when falling into the corresponding AOI.

To reduce noise in the time course estimates of the fixations and to facilitate statistical model fitting (described in the next section), for each object type—Target, Competitor, and

Unrelated, fixation proportions were calculated over every 40ms time bin. For all time bins, the total number of trials (for by-subject analysis) or the total number of subjects (for by-item analysis) was used as the denominator to avoid the selection bias introduced by varying trial-termination times (c.f. Kukona, Fang, Aicher, Chen, & Magnuson, 2011; Mirman & Magnuson, 2009; Mirman, Strauss, Dixon, & Magnuson, 2010). Only trials on which both the target image and the verb were correctly identified were included in the fixation analyses.

Two sets of data analysis were carried out on the fixation data within each experiment, including: (1) Target Fixation Analysis, which focused on the comparisons between target fixations across the neutral and the action verb contexts and (2) Distracter Fixation Analysis, which focused on the comparisons between fixations on competitors relative to unrelated images across the neutral and the action verb contexts.

Growth curve analysis (GCA) with orthogonal polynomials was used to quantify fixation differences across conditions during target identification (see Mirman, Dixon, & Magnuson, 2008 for a detailed description of this approach). Briefly, GCA uses hierarchically related sub-models to capture the data pattern. The first sub-model, usually called *Level 1*, captures the effect of time on fixation proportions using fifth-order orthogonal polynomials[2]. Specifically, the intercept term reflects average overall fixation proportion, the linear term reflects a monotonic change in fixation proportion (similar to a linear regression of fixation proportion as a function of time), the quadratic term reflects the symmetric rise and fall rate around a central inflexion point, and the higher terms similarly reflect the steepness of the curve around inflexion points and capture additional deflections in the curves. The *Level 2* models then capture the effects of experimental manipulations (as well as differences between participants or items) on the time terms. Specifically, in the following analyses, the *Level 2* models for Target Fixation Analysis included the factor *Context Verb* (Action verb vs. Neutral), and the *Level 2* models for Distracter Fixation Analysis included *Action Relatedness* (Related vs. Unrelated), *Context Verb* (Action verb vs. Neutral), and the interaction between *Action Relatedness* and *Context Verb* in incremental order. In the Level 2 models in the by item analyses, before adding the effects of experimental manipulations, we added the effect of *Context Fitness* for Target Fixation Analysis, and the effects of *Visual Similarity* and *Context Fitness* for Distracter Fixation Analysis. By so doing, we can account for the influence of these factors and demonstrate the effects of interest (e.g., *Context Verb* and *Action Relatedness*) over and above these possible confounds.

Models were fit using Maximum Likelihood Estimation and compared using the - 2LL deviance statistic (minus 2 times the log-likelihood), which is distributed as χ^2 with k degrees of freedom corresponding to the k parameters added (Mirman et al., 2008). In the current study, step-wise factor-level comparisons were used to evaluate the overall effects of factors in incremental order (i.e., *Action Relatedness*, *Context Verb*, and *Action Relatedness* × *Context Verb* interaction). In addition, tests on individual parameter estimates were used to evaluate specific condition differences on individual orthogonal time terms.

If our predictions are correct, we expect to see more fixations on competitors then unrelated items, reflected on the intercept term (the overall competition effect). In addition, we expect the action verb context to modulate the fixations on targets and competitors containing the

---

[2]Previous studies using GCA with eye tracking data have frequently used fourth order orthogonal time terms for Level 1 models. We used fifth order orthogonal time terms in the current study in order to better capture the curves present in the data. Note, however, that because the time terms are orthogonal, they are independent of each other, so the additional fifth-order terms would not affect the results of the other time terms, and more importantly, would not change our interpretation of the results of the lower time terms that we focus on in the current study.

action features of interest. Specifically, we expect that the action verb phrase will drive earlier activation of action-related objects (as reflected in fixation probabilities).

Results

Behavioral performance calculated from critical trials showed that participants were highly accurate in identifying the target images from the visual arrays (94.8%) and identifying the verbs (95.5%). Among the correct trials, mean mouse click reaction times from target word onset were 2358 ms and 2177 ms in the neutral and action verb context, respectively. Despite the numeric trend of better performance in the action verb context than in the neutral context, these differences were not statistically reliable (Accuracy: t(19)=0.44, p=0.74; Log-transformed reaction times: t(19)=1.61, p=0.12).

Figure 1 shows the average fixation proportions to targets, competitors, and unrelated items in both contexts for about 1.5 seconds starting from the onset of the target words. Visual inspection of the data shows that, in the beginning of the time window in the `saw' context, slightly less fixations were allocated to competitors than targets and unrelated items. However, crucially, in both contexts, following increasing fixation proportions for targets, more fixations appear to be allocated to competitors than to unrelated items, with the effect emerging earlier in the `picked up' context than in the `saw' context.

To quantify these results, fixation proportions on targets and distracter items were subjected to GCA separately, and are reported next.

Target Fixation Analysis—Given that the location of the target image is unknown prior to presentation of the target word, the analysis window began 200ms after target word onset to allow for saccade initiation (Altmann & Kamide, 2004). The analysis time window continued to 4 seconds after the trial onset (i.e., 1.6 seconds after target word onset), at which point overall average target fixation proportions were no longer increasing. For all fixation analyses reported in this paper, only correct-response trials were included.

The results of model fit comparisons showed no effect of *Context Verb* in either by subject or by item analysis (*ps* 0.2). That is, targets were identified equally quickly in both contexts.

Distracter Fixation Analysis—Fixations on competitors and unrelated images were analyzed in the same time window. The results of model comparisons showed a significant overall effect of *Action Relatedness* (by subject: $\chi^2(6)$ =83.51, p <0. 001; by item: $\chi^2(6)$ =79.65, p <.001), reflecting more fixations on the competitors than the unrelated items. There were no overall effects of *Context Verb (ps 0.3)*. However, there was a reliable interaction between *Action Relatedness* and *Context Verb* (by subject: $\chi^2(6)$ =62.85, p <0. 001; by item: $\chi^2(6)$ =63.65, p <0. 001), indicating that the action verb context modulated action relatedness competition. All effects remained significant after incorporating the norming values for visual similarity and context fitness into the by-item analysis (*Action Relatedness*:$\chi^2(6)$ = 55.25, p < 0.001; *Context Verb*:$\chi^2(6)$ = 15.21, p <0.05; *Action Relatedness* by *Context Verb* interaction:$\chi^2(6)$ = 55.66, p <0.001).

Significance tests were then carried out on the *Action Relatedness* by *Context Verb* interaction effects for the time terms of interest (intercept, linear and quadratic terms) by removing each term from the model and testing the decrement in model fit (using the -2 Log Likelihood deviance statistic, c.f. Mirman et al., 2008). The results revealed a reliable difference on the intercept and linear terms (Table 4). The positive effect on the intercept indicates more overall activation of structure-based relations in the `picked up' context. In addition, compared to the `saw' context, the slope of the competition effect in the `picked up'

context was more negative, reflecting that it occurred closer to the beginning of the time window rather than the middle (for both contexts, the average competitor fixation peak latencies were 512 ms after target onset[3]). In other words, consistent with our prediction, the competition effect emerged earlier and was larger in the `picked up' context than in the `saw' context.

In summary, this experiment shows that competitors sharing structure-based actions with targets (e.g. target - stapler; competitor - hammer) attracted more looks (the competition effect) than did unrelated items in the same visual array[4]. The occurrence of this competition effect did not hinge on the presence of an action verb. However, the action verb context exaggerated the effect, such that the competition effect emerged earlier and was greater in amplitude in the `picked up' than in the `saw' context. These results suggest that structure-based action knowledge is incidentally activated during spoken word comprehension and that an action verb context facilitates this activation.

Experiment 2

Extending Experiment 1, Experiment 2 investigated the activation dynamics of action attributes relevant to functional use of objects. The same paradigm and manipulations used in Experiment 1 were applied to Experiment 2 except that, in this case, targets and corresponding competitors shared function-based (but not structure-based) actions. For example, a key fob served as a competitor for the remote control target; both objects involve a poke gesture to use but different gestures to grasp (pinch for the key fob and clench for the remote control). Sentence contexts again included a neutral context (`S/he saw the ....') as well as an action verb context (`S/he used the ....').

We expected to see similar competition effects elicited by the function-based competitors, i.e. more looks to the competitors compared to the unrelated items, as reflected on the intercept term. We also predicted that eye movement patterns would be modulated by the action verb context. Similar to Experiment 1, we predicted that the action verb phrase would lead to earlier activation of function-based action features, reflecting in the fixation probabilities of action-related objects (targets and/or competitors).

Method

Materials

Visual stimuli: Experiment 2 used the same 22 target images as Experiment 1. In Experiment 2, targets (e.g. a calculator) were paired with competitors that shared `use', but not `grasp', action features with their corresponding targets (e.g. a doorbell). Again, care was taken to assure that unrelated items shared no action features with the corresponding target or the competitor. Norming values for visual similarity between targets and each of their distractors in the same array can be found in Table 1, and norming values for how each image fits the context in Table 2. Following the same scheme illustrated in Table 3, 88 arrays were generated as fillers.

Auditory Stimuli: Auditory stimuli were produced following the same guidelines as in Experiment 1. Both target names and the two sentence contexts, including an action verb

---

[3]Peaks of the competitor fixations were calculated for each participant as the time point when the competitor fixations were greater than both the preceding and the following time point. If multiple time points were obtained, the time when the competitor fixation was of the greatest value was chosen to be the peak.

[4]It is not clear why the competitors attracted fewer fixations than did other images in the same array in the `saw' context in the very beginning of the time window. It is important to note, however, that this fixation difference precedes the time when fixation proportions for the target started to separate from the non-targets. Therefore, it may be more related to the residual fixation patterns from the preview session, and less related to the processing of the target word, which is of central interest in this study.

context (e.g., `S/he used') and a neutral context (e.g., `S/he saw'), were recorded by a female native speaker of English and conjoined using *Audacity*. In keeping with the stimuli used in Experiment 1, onset of the target words started at 1400ms from the beginning of the sound file.

Participants—A different group of twenty healthy older adults (6 males; mean age: 61.6 years; range 49–73 years; SD: 7.4 years) participated, all recruited from a subject database maintained by Moss Rehabilitation Research Institute. All participants gave informed consent to participate in accordance with the IRB guidelines of the Albert Einstein Healthcare Network and were remunerated with cash for their participation. All participants were right handed, had normal or corrected-to-normal vision, normal hearing, no history of neurological/psychiatric disorders or brain damage based on self-report, and scored in the normal range on the MMSE (Folstein et al., 1975) conducted prior to the experiment (mean score: 28.8/30, SD: 1.1, range: 27–30). On average, these participants had a mean education level of 16 years (range = 12–21 years, SD=2.6 years).

The apparatus, experiment procedure, and analysis procedures in this experiment were identical to those used in Experiment 1.

Results

Behavioral performance calculated from critical trials showed that participants were again highly accurate in identifying the target images from the visual array and the verbs, with percentages of accurately choosing both the target images and verbs at 96.6% and 98.6% in the "saw" and "used" contexts, respectively. Among the correct trials, mean mouse click reaction times from target word onset were 2949ms and 2714ms in the "saw" and action "used" context, respectively. In keeping with the trend shown in Experiment 1, participants appeared to perform better in the action verb context than in the neutral "saw" context. The results of paired t-test revealed statistically reliable differences in the log-transformed reaction times (t(19)=3.37, p<0.005), but not accuracy (t(19)=7.75 p=0.08).

Figure 2 shows the average fixation proportions to targets, competitors and unrelated items in both contexts, starting from the onset of the target words to about 1.5 second afterwards. On visual inspection, target fixation proportions appeared to separate from the distractors (particularly the unrelated distractors) at an earlier point in the "used" than "saw" context. Consistent with the results obtained in Experiment 1, there were more overall fixations on competitors than on unrelated items, and the competition effects emerged earlier in the "used" than "saw" context.

To quantify these results, fixation proportions on targets and distracter items were subjected to GCA separately, and are reported next. Only trials on which both the target image and the verb were correctly identified were included in these analyses.

Target Fixation Analysis—Target fixations were measured using the same time window as in Experiment 1, i.e. from 200ms until about 1.5 seconds after the target onset. The results of model fit comparisons showed a reliable effect of *Context Verb* (by subject: $\chi^2(6)$ =87.55, p <.001; by item: $\chi^2(6)$ =88.09, p< .001), which remained reliable after context fitness norming values were incorporated into the by item analysis ($\chi^2(6)$ =36.64, p <.001).

Significance tests were then carried out on the *Context Verb* effects on each parameter estimate (with the neutral context as the reference level) (Table 5). The results revealed reliable differences on the intercept and linear time term, reflecting overall higher target fixation proportions and slower increase in target fixation proportions in the "used" context than in the "saw" context. This slower increase reflects the fact that, in the "used" context,

target fixation proportions start to rise earlier, but reach the same peak level at about the same time (the maximum fixation proportion was reached at about 1314ms post target onset in the "saw" context, and 1298ms in the "used" context). In other words, the reduced slope suggests that the "used" context facilitated the beginning of target word recognition.

Distracter Fixation Analysis—Fixations on competitors and unrelated items were measured using the same time window as the previous analyses. The results of model comparisons showed a significant overall effect of *Action Relatedness* (by subject: $\chi^2(6)$ =35.14, p <0.001; by item: $\chi^2(6)$ =18.85, p <0.01), reflecting more fixations on the competitors than the unrelated items. There was also an overall effect of *Context Verb* (by subject: $\chi^2(6)$ =47.49, p <0.001; by item: $\chi^2(6)$ =39.79, p <0.001), reflecting less overall distractor fixation in the `used' context. The two effects interacted (by subject: $\chi^2(6)$ =37.46, p <0.001; by item: $\chi^2(6)$ =31.43, p <0.001). The effects of *Action Relatedness* and *Context Verb* remained highly significant after the rating scores for visual similarity and context fitness were incorporated into the by-item analysis (*Action relatedness*: $\chi^2(6) = 21.89$, p =. 001, *Context Verb*: $\chi^2(6) = 42.08$, p <0.001). The interaction between *Action Relatedness* and *Context Verb*, however, was no longer statistically significant (p=0.9).

In summary, similar to Experiment 1, Experiment 2 showed that competitors sharing only function-based actions (i.e., not structure-based actions) with targets attracted fixations more than did unrelated items in the same display. Additionally, action verb context modulated the observed eye movement patterns. Unlike in Experiment 1, the context modulation was manifested on target fixations, instead of distracter fixations: the incremental increases on target fixations started earlier in the action verb contexts than in the neutral contexts. Thus, the action verb appeared to facilitate target detection, consistent with the response time analyses.

In order to directly contrast the activation time course of structure-based and function-based action attributes, we directly compared the competition effects elicited by the structure-based competitors (Experiment 1) and function-based competitors (Experiment 2) in the neutral contexts.

**Comparison of competition effects elicited by structure-based and function-based competitors in the neutral `saw' context**—Data from the neutral `saw' contexts in both experiments were directly compared in by-subject and by-item analyses. The results showed reliable effects of *Action Relatedness* (by subject: $\chi^2(6)$ =109.71, p <0.001; by item: $\chi^2(6)$ =66.96, p <.001) reflecting overall differences in fixating competitors versus unrelated items. There was also reliable effects of *Competitor Type* in the by item analysis (by subject: $\chi^2(6)$ =5.06, p=0.53; by item: $\chi^2(6)$ =45.63, p <0.001) reflecting more distracter fixations with function-based as compared to structure-based displays. In addition, there was a reliable *Action Relatedness * Competitor Type* interaction (by subject: $\chi^2(6)$ =63.11, p<.001; by item: $\chi^2(6)$ =85.12, p <0.001) reflecting differences in the time course of structure-based versus function-based competition. All effects remained reliable after norming values of visual similarity and context fitness were incorporated into the model (*Action Relatedness*: $\chi^2(6)$ =22.34, p =0.001; *Competitor Type*: $\chi^2(6)$ =36.35, p <0.001; *Action Relatedness * Competitor Type*: $\chi^2(3)$ =75.54, p <0.001).

Significance tests on parameter estimates in the *Action relatedness * Competitor type* interaction effect with the structure-based display as the baseline showed the following results (also see Table 6). There was overall more competition evoked by function-based competitors than by structure-based competitors (effect on intercept term). In addition, compared to the structure-based competition effect, the function-based competition effect

ramped up more slowly (effect on linear term) and was less steeply peaked (more wide spread in time; effect on quadratic term).

Figure 3 presents point-by-point differences in the competition effects from both competitor types (competitors – unrelated items) in the neutral contexts. In summary, function-based action features become active at a slower rate than do structure-based action features. However, function-based competitors elicit greater and more lasting competition effects than do structure-based competitors, consistent with the idea that function-based action features remain active for a longer time.

General Discussion

Using eye-movement recording, the present study investigated whether action attributes that are incidental to task demands may be activated during word-to-picture matching. Based on previous data (Jax & Buxbaum, 2010) and the Two Action System (2AS) model, we predicted that activation of both function-based and structure-based activation would be observed. In a neutral context in which the target word was presented with a verb not conveying manual action ("saw"), we predicted that structure-based competition would be faster-rising and more transient than function-based competition. Furthermore, we predicted an enhancement of competition effects in action verb contexts. Finally, we speculated that function-based actions might show relatively greater facilitation by verbal context than structure-based actions.

The findings were largely consistent with these predictions. When presented with neutral sentence contexts, participants were more likely to fixate on both function-based and structure-based competitors than on unrelated objects, suggesting task-incidental activation of both action types in the absence of an object-related action task. Importantly, the structure-based competition effect ramped up faster and peaked more steeply than did the function-based effect, indicating distinct activation dynamics consistent with the 2AS model. Finally, provision of a contextual sentence containing a manual action verb modulated the pattern of results, leading to more and earlier fixations on competitors containing the relevant action features (Experiment 1) as well as improved target identification (Experiment 2). These data suggest, contrary to our expectation, that activation of both types of action attributes may be facilitated by lexical-semantic contextual information, albeit in somewhat different ways. In the next sections, we will discuss each of these findings in turn.

Task-Incidental activation of both function-based and structure-based action features in a word-picture matching task—In two experiments designed to assess separately function-based and structure-based activations, participants were more likely to fixate on action-related distractors than unrelated items. Previous studies using the Visual World Paradigm have shown that participants' gaze is directed to competitors when their visual forms (Dahan & Tanenhaus, 2005; Huettig & Altmann, 2004, 2005, 2007; Yee, Huffstetler, & Thompson-Schill, 2011) or conceptual properties (Mirman & Magnuson, 2009; Myung et al., 2006, 2010; Yee & Sedivy, 2006) are similar to the targets. These competition effects tend to start as soon as 200ms after the onset of the target word, and in many cases well before the offset of the target word. Given that programming and executing an eye movement has been estimated to take at least 150–200ms (Hallett, 1986), the rapid onset of these language-mediated eye movements suggest that competition effects such as these reflect implicit and automatic partial activation of the distractors due to feature overlap with the target (see Salverda & Altmann, 2011 for more discussion on this issue). Altmann and Kamide (2007), for example, proposed that, prior to auditory instructions, participants' inspection of the visual array leads to pre-activation of conceptual features of the displayed

objects, leaving conceptually enriched episodic traces associated with each object. As the verbal instructions unfold, the conceptual features activated by the verbal input make contact with the features pre-activated from the visual array and effectively re-activate these episodic traces, which then leads to a shift in visual attention such that participants are more prone to make a saccadic eye movement towards the object with these features (see also Salverda & Altmann, 2011). Along these lines, the greater fixation proportions we observed to both structure-based and function-based distracters relative to the unrelated items could be thought of as reflections of overlapping action features that are incidentally activated by the target images, distracter images and the spoken words. Although previous investigations using the VWP have demonstrated that gaze is diverted to manipulation-related distractors (Myung et al., 2006, Exp2), this is the first study, to our knowledge, to demonstrate competition for visual attention based on two distinct subtypes of action features.

A related series of studies by Bub, Masson, and colleagues used combinations of pictorial and verbal materials to assess function-based and structure-based activations as measured by priming effects; specifically, the degree to which pictures or words facilitated congruent actions. Importantly, in all of these studies responses were signaled by manual object-related actions (grasping a manipulandum); thus, the intention to perform an action may have influenced target processing. In a series of studies presenting object names either in isolation or in sentences with non-manipulation verbs (`The young scientist looked at the stapler'; `Jane forgot the calculator'), these investigators found evidence of activation of use actions, but not grasp actions (Masson, Bub, & Newton-Taylor, 2008; Masson, Bub, & Warren, 2008). On the other hand, when function-based or structure-based verbs were present in the sentences, both function-based and structure-based actions were primed (Bub & Masson, 2010). Finally, several studies using real or pictured objects, rather than words, have reported task-incidental activation of both function-based and structure-based actions (e.g., Bub & Masson, 2006; Bub, Masson, & Cree, 2008; Jax & Buxbaum, 2010), though these studies also required manual object-relevant responses.

One might argue that the reach and mouse-click movements involved in the present study may be sufficient to potentiate task-irrelevant action features. However, this possibility is mitigated by data from Pavese and Buxbaum (2002) and Bub, Masson, and Cree (2008), both of which showed that although motor responses including object-relevant hand postures induced object-relevant action features, a simple reach to touch or reach to button-press movement did not. Therefore, we believe that it is unlikely that the action activations we observed in our data were induced by the reach and mouse-click movements.

Viewed together, the evidence suggests that function-based actions may be incidentally evoked either from object names or visual images, whereas incidental activations of structure-based actions require either an appropriate verb (e.g., "pick up") or presentation of the visual form of the object. Critically, our method extends prior work by enabling us to capture differences in the time course of activation of the two action types, which we discuss next.

Different time courses for the activation of structure-based and function-based features—To our knowledge, this study is the first to provide fine-grained information about the dynamics of activation of function-based and structure-based actions in a neutral context with a relatively naturalistic experimental paradigm. We suggest that the observed differences in activation profiles are likely to reflect differences in underlying functional neuroanatomic mechanisms. According to the 2AS model, structure-based actions are mediated largely by the dorso-dorsal visual processing stream, a bilateral system that is specialized for grasping and moving objects. The dorso-dorsal stream processes current visuo-spatial information, maintains information for milliseconds to seconds, and may in

some circumstances operate independent of long-term conceptual representations (Cant, Westwood, Valyear, & Goodale, 2005; Garofeanu, Kroliczak, Goodale, & Humphrey, 2004; Jax & Rosenbaum, 2009). In contrast, function-based actions are mediated by the ventral part of the dorsal stream in the left hemisphere, including the left inferior parietal and posterior temporal lobes. A number of lines of evidence suggest that these regions mediate the storage of object-associated actions in long-term memory (e.g., Buxbaum & Saffran, 1998; Pelgrims, Olivier, & Andres, 2010). As noted, these skilled use representations appear to have the characteristics of semantic memory, including relatively sustained activation (e.g., Campanella & Shallice, 2011).

Using "conflict" objects as targets, we explicitly controlled the nature of the feature similarity between the target and competitors so that they overlapped in either function-based or structure-based features, but not both. However, in everyday settings, many objects present no conflict. A drinking glass, wine bottle, pitcher, and soda can, for example, are both used and moved with a clench action. One possibility is that with arrays of objects that overlap in both function- and structure-based attributes we may expect to see slightly offset but potentially additive patterns of interference as the "pick up" followed by the "use" competition is processed. Following from this, we might expect that patients with apraxia should show normal fast structure-based competition but diminished or absent subsequent function-based competition (c.f., Jax & Buxbaum, under revision). Such questions are of interest for future investigation.

Action Verb Context facilitates activation of action features in both targets and distractors—Both experiments demonstrated that verbal context modulates the activation of action features. We had speculated that because of closer ties to semantics, effects on competition might be more robust for the `used' than `picked up' verb. In fact, we observed significant effects for both action verbs. Experiment 1 demonstrated earlier emergence of competition effects and greater overall competition in the `picked up' compared to `saw' context. In Experiment 2, target fixations began earlier and targets were detected more rapidly in the `used' than `saw' context.

It is interesting to note that the `picked up' verb increased competition from action-relevant distractors, whereas the `used' verb effectively reduced the effects of competition (as witnessed by earlier target detection). This was the case even when we took into account the degree to which the target `fit' the action verb context based on normative values. In other words, this finding does not simply reflect varying degrees of ability to `rule out' distractors. Although speculative, one possibility is that these different patterns reflect the fact that the verb phrase "*used the X*" highly constrains a precise gesture, whereas the phrase "*picked up the X*" specifies a broader range of action parameters. For example, one uses a calculator with a very specific action: a forefinger-poke gesture that is vertical in orientation and aimed downward. The phrase "*used the calculator*" thus may result in very specific motor simulation, rendering "similar" only objects that are used in precisely the same way (and, since these are rare, thereby reducing overall competition from distractors). In contrast, one picks up a calculator with different parameters depending on its orientation with respect to oneself; as such, visual information from the array is more dominant than verbal information in specifying grasp parameters, and any objects in the array that are plausibly picked up similarly become contenders for the control of action, resulting in heightened competition.

The present data are broadly consistent with past findings showing that preparing an action may facilitate the processing of targets and nearby distractors that are congruent with that action (e.g., Allport, 1987, 1989; Bekkering & Neggers, 2002; Botvinick et al., 2009; Craighero et al., 1999; Craighero, Fadiga, Umiltà, & Rizzolatti, 1996; Hannus, Cornelissen, Lindemann, & Bekkering, 2005; Pavese & Buxbaum, 2002; van Elk, van Schie, Neggers, &

Bekkering, 2010). Several recent studies (Gutteling, Kenemans, & Neggers, 2011; Moore & Armstrong, 2003; Neggers et al., 2007; Ruff et al., 2006), in support of the premotor theory of attention (e.g., Rizzolatti, Riggio, & Sheliga, 1994; Rizzolatti, Riggio, Dascola, & Umiltá, 1987), suggest that motor preparation results in modulation of attention through common neural mechanisms underlying movement preparation and attentional processing. It is possible that the enhanced competition effects we observed with an action verb context similarly occur through the mediation of activation in motor-related brain areas. Action verbs or sentences describing actions have been associated with activations in primary and/ or pre-motor regions (Buccino et al., 2005; Hauk, Johnsrude, & Pulvermüller, 2004; Pulvermuller, 2005; Raposo et al., 2009). Processing object names has been shown to elicit similar motor and pre-motor area activations (Grafton, Fadiga, Arbib, & Rizzolatti, 1997; Rueschemeyer, Lindemann, van Rooij, van Dam, & Bekkering, 2010) as well as mental simulation of object shape and orientation (Borghi & Riggio, 2009) and associated hand actions (Willems, Toni, Hagoort, & Casasanto, 2009). Our study does not provide direct evidence regarding underlying neural substrates, however, and further research is needed to support this hypothesis.

Concluding remarks—We have provided evidence for task-incidental activation of both structure-based and function-based action features in a picture-word matching task not requiring an object-related motor response. These data provide a fine-grained description of the time course of activation of each of these action feature types, helping to refine models of object-related action and attention, including the 2AS model (e.g., Buxbaum & Kalenine, 2010). We have also provided evidence that an action verb context modulates the activation of these action features, bringing this work into contact with a rich prior literature in the domains of `selection for action' and context-sensitive semantic processing. Finally, this work provides a basis for future investigations in patients with disorders of `use' action knowledge (e.g., ideomotor apraxia). As noted, we might expect that patients with apraxia should show normal early structure-based competition but diminished or absent subsequent function-based competition (c.f., Jax & Buxbaum, under revision). Alternatively, it is possible that relatively normal function-based activations will be evident on this implicit task, despite patients' deficits on tasks explicitly assessing knowledge of skilled object manipulation. Investigations currently underway in our laboratory will assess these possibilities.

Acknowledgments

References

Allopenna PD, Magnuson JS, Tanenhaus MK. Tracking the Time Course of Spoken Word Recognition Using Eye Movements: Evidence for Continuous Mapping Models. Journal of Memory and Language. 1998; 38(4):419–439.

Allport, A. Selection for action: Some behavioral and neurophysiological considerations of attention and action. In: Heuer, H.; Sanders, AF., editors. Perspectives on perception and action. Hillsdale; NJ: Erlbaum: 1987. p. 395-419.

Allport, A. Foundations of cognitive science. The MIT Press; Cambridge, MA, US: 1989. Visual attention; p. 631-682.

Altmann, GT.; Kamide, Y. Now you see it, now you don't: Mediating the mapping between language and the visual world. In: Henderson, JM.; Ferreira., F., editors. The interface of language, vision,

and action: Eye movements and the visual world. Psychology Press; New York, NY: 2004. p. 313-345.

Altmann GTM, Kamide Y. The real-time mediation of visual attention by language and world knowledge: Linking anticipatory (and other) eye movements to linguistic processing. Journal of Memory and Language. 2007; 57(4):502–518.

Bekkering H, Neggers SFW. Visual Search Is Modulated by Action Intentions. Psychological Science. 2002; 13(4):370–374. [PubMed: 12137141]

Borghi AM, Riggio L. Sentence comprehension and simulation of object temporary, canonical and stable affordances. Brain Research. 2009; 1253:117–128. [PubMed: 19073156]

Botvinick MM, Buxbaum LJ, Bylsma LM, Jax SA. Toward an integrated account of object and action selection: A computational analysis and empirical findings from reaching-to-grasp and tool-use. Neuropsychologia. 2009; 47(3):671–683. [PubMed: 19100758]

Bub DN, Masson MEJ. Gestural knowledge evoked by objects as part of conceptual representations. Aphasiology. 2006; 20(9–11):1112–1124.

Bub DN, Masson MEJ. On the nature of hand-action representations evoked during written sentence comprehension. Cognition. 2010; 116(3):394–408. [PubMed: 20579981]

Bub DN, Masson ME, Cree S. Evocation of functional and volumetric gestural knowledge by objects and words. Cognition. 2008; 106:27–58. [PubMed: 17239839]

Buccino G, Riggio L, Melli G, Binkofski F, Gallese V, Rizzolatti G. Listening to action-related sentences modulates the activity of the motor system: A combined TMS and behavioral study. Cognitive Brain Research. 2005; 24(3):355–363. [PubMed: 16099349]

Buxbaum LJ. Ideomotor apraxia: a call to action. Neurocase. 2001; 7:445–458. [PubMed: 11788737]

Buxbaum LJ, Coslett HB. Spatio-motor representations in reaching: evidence for subtypes of optic ataxia. Cognitive Neuropsychology. 1998; 15(3):279–312.

Buxbaum LJ, Kalénine S. Action knowledge, visuomotor activation, and embodiment in the two action systems. Annals of the New York Academy of Sciences. 2010; 1191:201–218. [PubMed: 20392282]

Buxbaum LJ, Kyle KM, Grossman M, Coslett HB. Left inferior parietal representations for skilled hand-object interactions: Evidence from stroke and corticobasal degeneration. Cortex. 2007; 43(3): 411–23. [PubMed: 17533764]

Buxbaum LJ, Sirigu A, Schwartz MF, Klatzky R. Cognitive representations of hand posture in ideomotor apraxia. Neuropsychologia. 2003; 41:1091–1113. [PubMed: 12667544]

Campanella F, Shallice T. Manipulability and object recognition: is manipulability a semantic feature? Experimental Brain Research. 2011; 208:369–383.

Cant JS, Westwood DA, Valyear KF, Goodale MA. No evidence for visuomotor priming in a visually guided action task. Neuropsychologia. 2005; 43(2):216–226. [PubMed: 15707906]

Craighero L, Bello A, Fadiga L, Rizzolatti G. Hand action preparation influences the responses to hand pictures. Neuropsychologia. 2002; 40(5):492–502. [PubMed: 11749979]

Craighero L, Fadiga L, Rizzolatti G, Umiltà C. Action for perception: A motor-visual attentional effect. Journal of Experimental Psychology: Human Perception and Performance. 1999; 25(6): 1673–1692. [PubMed: 10641315]

Craighero L, Fadiga L, Umiltà CA, Rizzolatti G. Evidence for visuomotor priming effect. Neuroreport. 1996; 8(1):347–349. [PubMed: 9051808]

Dahan D, Tanenhaus MK. Looking at the rope when looking for the snake: Conceptually mediated eye movements during spoken-word recognition. Psychonomic Bulletin & Review. 2005; 12:453–459. [PubMed: 16235628]

Dahan D, Magnuson JS, Tanenhaus MK, Hogan EM. Subcategorical Mismatches and the Time Course of Lexical Access: Evidence for Lexical Competition. Language and Cognitive Processes. 2001; 16(5):507–34.

Folstein MF, Folstein SE, McHugh PR. "Mini-mental state" A practical method for grading the cognitive state of patients for the clinician. Journal of Psychiatric Research. 1975; 12(3):189–198. [PubMed: 1202204]

Garofeanu C, Kroliczak G, Goodale M, Humphrey GK. Naming and grasping common objects: a priming study. Experimental Brain Research. 2004

Glenberg AM, Kaschak MP. Grounding language in action. Psychonomic Bulletin & Review. 2002; 9(3):558–565. [PubMed: 12412897]

Grafton ST, Fadiga L, Arbib MA, Rizzolatti G. Premotor cortex activation during observation and naming of familiar tools. Neuroimage. 1997; 6:231–236. [PubMed: 9417966]

Gutteling TP, Kenemans JL, Neggers SFW. Grasping Preparation Enhances Orientation Change Detection. PLoS ONE. 2011; 6(3):e17675. [PubMed: 21408131]

Hallett, PE. Eye movements. In: Boff, KR.; Kaufman, L.; Thomas, JP., editors. Handbook of perception and human performance. Wiley; New York, NY: 1986. p. 10.1-10.112.

Hannus A, Cornelissen FW, Lindemann O, Bekkering H. Selection-for-action in visual search. Acta Psychologica. 2005; 118(1–2):171–191. [PubMed: 15627415]

Hauk O, Johnsrude I, Pulvermüller F. Somatotopic Representation of Action Words in Human Motor and Premotor Cortex. Neuron. 2004; 41(2):301–307. [PubMed: 14741110]

Huettig, F.; Altmann, GTM. The online processing of ambiguous and unambiguous words in context: Evidence from head-mounted eye-tracking. Psychology Press; New York: 2004.

Huettig F, Altmann GTM. Word meaning and the control of eye fixation: semantic competitor effects and the visual world paradigm. Cognition. 2005; 96(1):B23–B32. [PubMed: 15833303]

Huettig F, Altmann GTM. Visual-shape competition during language-mediated attention is based on lexical input and not modulated by contextual appropriateness. Visual Cognition. 2007; 15:985–1018.

Jax, SA.; Buxbaum, LJ. Response interference between functional and structural object-related actions is increased in patients with ideomotor apraxia. Under Revision

Jax SA, Buxbaum LJ. Response interference between functional and structural actions linked to the same familiar object. Cognition. 2010; 115(2):350–355. [PubMed: 20156619]

Jax SA, Rosenbaum DA. Hand path priming in manual obstacle avoidance: Rapid decay of dorsal stream information. Neuropsychologia. 2009; 47(6):1573–1577. [PubMed: 18597796]

Jeannerod, M. Fundamentals of cognitive neuroscience. Blackwell Publishing; Malden: 1997. The cognitive neuroscience of action.

Johnson-Frey SH. The neural bases of complex tool use in humans. Trends in Cognitive Sciences. 2004; 8(2):71–78. [PubMed: 15588811]

Kalénine, Solène; Mirman, D.; Middleton, EL.; Buxbaum, LJ. Temporal dynamics of activation of thematic and functional action knowledge during auditory comprehension of artifact words. under review

Kamide Y, Altmann GT, Haywood SL. The time-course of prediction in incremental sentence processing: Evidence from anticipatory eye movements. Journal of Memory and Language. 2003; 49(1):133–156.

Kukona A, Fang S-Y, Aicher KA, Chen H, Magnuson JS. The time course of anticipatory constraint integration. Cognition. 2011; 119(1):23–42. [PubMed: 21237450]

Masson MEJ, Bub DN, Newton-Taylor M. Language-based access to gestural components of conceptual knowledge. The Quarterly Journal of Experimental Psychology. 2008; 61(6):869–882. [PubMed: 18470818]

Masson MEJ, Bub DN, Warren CM. Kicking calculators: Contribution of embodied representations to sentence comprehension. Journal of Memory and Language. 2008; 59(3):256–265.

Mirman D, Magnuson JS. Dynamics of activation of semantically similar concepts during spoken word recognition. Memory & Cognition. 2009; 37:1026–1039.

Mirman D, Dixon JA, Magnuson JS. Statistical and computational models of the visual world paradigm: Growth curves and individual differences. Journal of Memory and Language. 2008; 59(4):475–494. [PubMed: 19060958]

Mirman D, Strauss TJ, Dixon JA, Magnuson JS. Effect of Representational Distance Between Meanings on Recognition of Ambiguous Spoken Words. Cognitive Science. 2010; 34(1):161–173. [PubMed: 20354577]

Moore T, Armstrong KM. Selective gating of visual signals by microstimulation of frontal cortex. Nature. 2003; 421(6921):370–373. [PubMed: 12540901]

Müsseler J, Steininger S, Wühr P. Can actions affect perceptual processing? The Quarterly Journal of Experimental Psychology Section A. 2001; 54:137–154.

Myung, J.-yoon; Blumstein, SE.; Sedivy, JC. Playing on the typewriter, typing on the piano: manipulation knowledge of objects. Cognition. 2006; 98(3):223–243. [PubMed: 16399263]

Myung, J.-yoon; Blumstein, SE.; Yee, E.; Sedivy, JC.; Thompson-Schill, SL.; Buxbaum, LJ. Impaired access to manipulation features in Apraxia: Evidence from eyetracking and semantic judgment tasks. Brain and Language. 2010; 112(2):101–112. [PubMed: 20064657]

Neggers SFW, Huijbers W, Vrijlandt CM, Vlaskamp BNS, Schutter DJLG, Kenemans JL. TMS Pulses on the Frontal Eye Fields Break Coupling Between Visuospatial Attention and Eye Movements. Journal of Neurophysiology. 2007; 98(5):2765–2778. [PubMed: 17699696]

Pavese A, Buxbaum LJ. Action matters: The role of action plans and object affordances in selection for action. Visual Cognition. 2002; 9(4/5):559–590.

Pelgrims B, Olivier E, Andres M. Dissociation between manipulation and conceptual knowledge of object use in the supramarginalis gyrus. Human Brain Mapping. 2010; 32(11):1802–1810. [PubMed: 21140435]

Pisella L, Binkofski F, Lasek K, Toni I, Rossetti Y. No double-dissociation between optic ataxia and visual agnosia: Multiple sub-streams for multiple visuo-manual integrations. Neuropsychologia. 2006; 44:2734–2748. [PubMed: 16753188]

Pulvermuller F. Brain mechanisms linking language and action. Nat Rev Neurosci. 2005; 6(7):576–582. [PubMed: 15959465]

Raposo A, Moss HE, Stamatakis EA, Tyler LK. Modulation of motor and premotor cortices by actions, action words and action sentences. Neuropsychologia. 2009; 47(2):388–396. [PubMed: 18930749]

Rizzolatti, G.; Riggio, L.; Sheliga, BM. Attention and performance series. The MIT Press; Cambridge, MA, US: 1994. Space and selective attention. Attention and performance 15: Conscious and nonconscious information processing; p. 232-265.

Rizzolatti G, Riggio L, Dascola I, Umiltá C. Reorienting attention across the horizontal and vertical meridians: Evidence in favor of a premotor theory of attention. Neuropsychologia. 1987; 25(1, Part 1):31–40. [PubMed: 3574648]

Rizzolatti G, Matelli M. Two different streams form the dorsal visual system: anatomy and functions. Experimental Brain Research. Experimentelle Hirnforschung. Expérimentation Cérébrale. 2003; 153(2):146–157.

Rueschemeyer S-A, Lindemann O, van Rooij D, van Dam W, Bekkering H. Effects of Intentional Motor Actions on Embodied Language Processing. Experimental Psychology. 2010; 57(4):260–266. [PubMed: 20178948]

Ruff CC, Blankenburg F, Bjoertomt O, Bestmann S, Freeman E, Haynes J-D, Rees G, et al. Concurrent TMS-fMRI and Psychophysics Reveal Frontal Influences on Human Retinotopic Visual Cortex. Current Biology. 2006; 16(15):1479–1488. [PubMed: 16890523]

Salverda AP, Altmann GTM. Attentional capture of objects referred to by spoken language. Journal of Experimental Psychology: Human Perception and Performance. 2011; 37:1122–1133. [PubMed: 21517215]

Sirigu, Angela; Duhamel, J-R.; Cohen, L.; Pillon, B.; Dubois, B.; Agid, Y. The Mental Representation of Hand Movements After Parietal Cortex Damage. Science. 1996; 273(5281):1564–1568. [PubMed: 8703221]

Symes E, Tucker M, Ellis R, Vainio L, Ottoboni G. Grasp preparation improves change detection for congruent objects. Journal of Experimental Psychology: Human Perception and Performance. 2008; 34(4):854–871. [PubMed: 18665731]

Tanenhaus MK, Spivey-Knowlton MJ. Eye-Tracking. Language and Cognitive Processes. 1996; 11:583–588.

Tanenhaus MK, Spivey-Knowlton M, Eberhard K, Sedivy J. Integration of visual and linguistic information in spoken language comprehension. Science. 1995; 268(5217):1632–1634. [PubMed: 7777863]

Taylor LJ, Zwaan RA. Motor resonance and linguistic focus. The Quarterly Journal of Experimental Psychology. 2008; 61(6):896. [PubMed: 18470820]

Tucker M, Ellis R. On the relations between seen objects and components of potential actions. Journal of Experimental Psychology. Human Perception and Performance. 1998; 24(3):830–846. [PubMed: 9627419]

Tucker M, Ellis R. The potentiation of grasp types during visual object categorization. Visual Cognition. 2001; 8:769–800.

Tucker M, Ellis R. Action priming by briefly presented objects. Acta Psychologica. 2004; 116(2):185–203. [PubMed: 15158182]

Vainio L, Symes E, Ellis R, Tucker M, Ottoboni G. On the relations between action planning, object identification, and motor representations of observed actions and objects. Cognition. 2008; 108(2):444–465. [PubMed: 18452910]

van Elk M, van Schie HT, Neggers SFW, Bekkering H. Neural and Temporal Dynamics Underlying Visual Selection for Action. Journal of Neurophysiology. 2010; 104(2):972–983. [PubMed: 20538783]

Vingerhoets G, Acke F, Vandemaele P, Achten E. Tool responsive regions in the posterior parietal cortex: effect of differences in motor goal and target object during imagined transitive movements. NeuroImage. 2009; 47(4):1832–1843. [PubMed: 19523524]

Willems RM, Toni I, Hagoort P, Casasanto D. Body-specific motor imagery of hand actions: neural evidence from right- and left-handers. Frontiers in Human Neuroscience. 2009; 3:39. [PubMed: 19949484]

Yee E, Sedivy JC. Eye movements to pictures reveal transient semantic activation during spoken word recognition. Journal of Experimental Psychology. Learning, Memory, and Cognition. 2006; 32(1):1–14.

Yee E, Huffstetler S, Thompson-Schill SL. Function follows form: Activation of shape and function features during object identification. Journal of Experimental Psychology. General. 2011
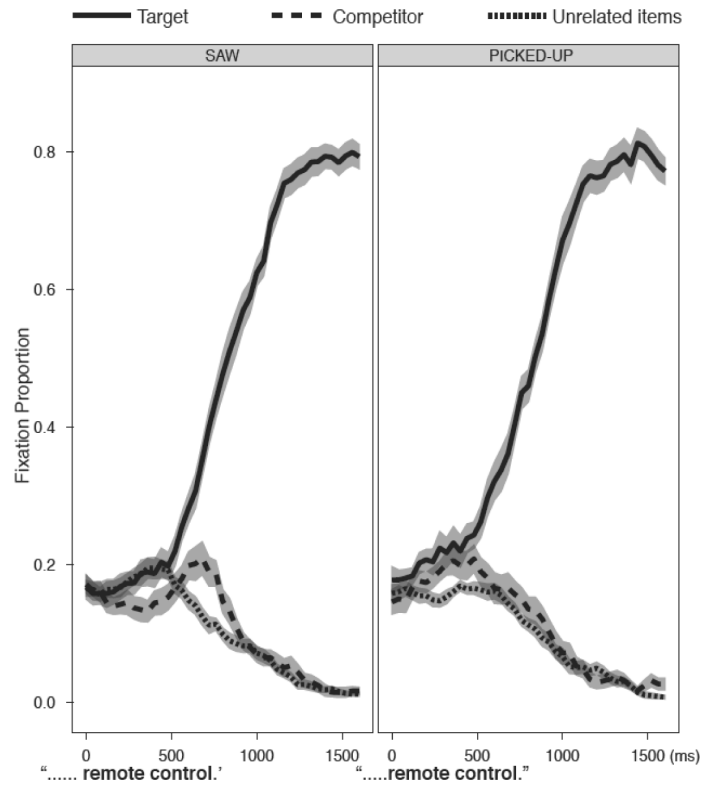
Figure 1.
Fixation proportions for Target (solid line), Competitor (dashed line) and Unrelated items (dotted line) in the neutral `**saw**' context condition (left) and the action verb `**picked up'** context (right). Shades around the data lines show the standard errors.
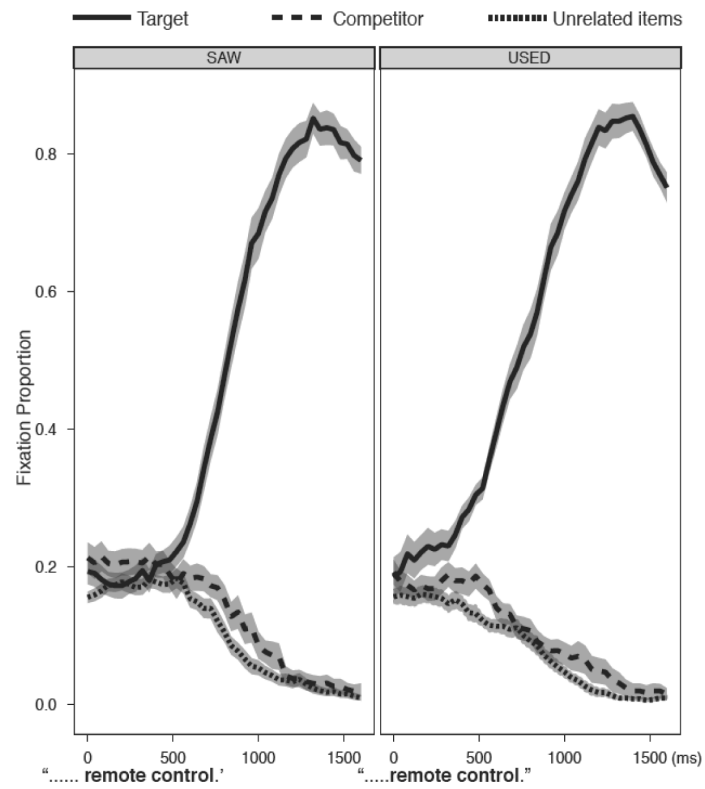
Figure 2.
Fixation proportions for Target (solid line), Competitor (dashed line) and unrelated items (dotted line) in the neutral `**saw**' context condition (left) and the action verb `**used**' context (right). Shades around the data lines show the standard errors.
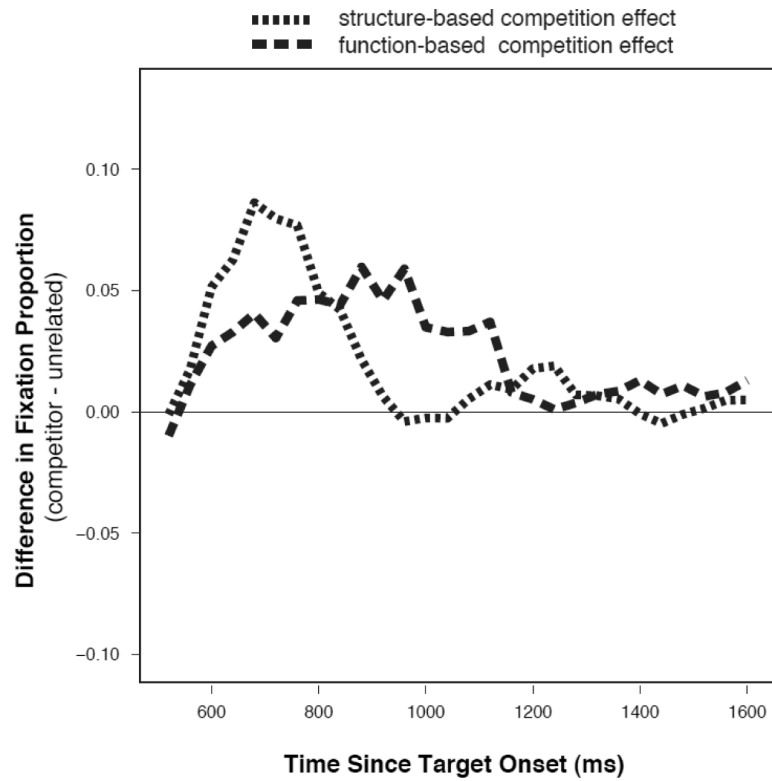
Figure 3.
Point-by-point differences of the competition effects (competitors – unrelated items) from the structure-based competitors (dotted line) and function-based competitors (dashed line) from a more focused time window.

Table 1

Averaged rating scores for visual similarity between targets and distracter images (7: highly similar; 1: not similar at all). Standard deviations are shown in parentheses.

	Target & competitor	Target & unrelated items
Experiment I	3.7 (1.0)	2.8 (1.2)
Experiment II	3.1 (1.1)	3.0 (1.1)

Table 2

Averaged rating scores for context fitness (7: perfect fit; 1: no fit). Standard deviations are shown in parentheses.

	Neutral context			Action verb context		
	Target	Competitor	Unrelated	Target	Competitor	Unrelated
Experiment I	5.2 (1.5)	5.4 (1.4)	5.6 (1.3)	5.7 (1.1)	5.9 (1.1)	4.3 (1.2)
Experiment II	5.1 (1.6)	5.1 (1.7)	5.4 (1.5)	6.1 (1.2)	5.8 (1.3)	4.6 (1.4)

Table 3

Examples for the scheme used for generating filler trials.

Critical/filler	Context	Target	Non-target	Non-target	Non-target
Critical	Neutral	Stapler	Hammer (related)	Chess-piece (unrelated)	Couch (unrelated)
Critical	Action	Stapler	Hammer (related)	Chess-piece (unrelated)	Couch (unrelated)
Filler 1	Neutral	Chess-piece	Hammer	Stapler	Couch
Filler 2	Action	Hammer	Corkscrew	Bathtub	Q-tip
Filler 3	Neutral	Chess-piece	Corkscrew	Bathtub	Q-tip
Filler 4	Action or neutral	Couch	Clipboard	Bathtub	grape

Table 4

Results of estimates, χ^2 and p values for the *Action relatedness* by *Context* interaction effect on each time term (Intercept, Linear, and Quadratic) in Experiment 1. Standard errors and t values for the parameter estimates are shown in the parentheses next to each estimate. Unrelated object and neutral context were treated as the reference level for Action relatedness and Context factors respectively.

Term	**Estimate**	χ^2	*p*
Intercept	0.030 (se=0.015; t=2.031)	3.99	0.046
Linear	−0.181 (se=0.068; t=−2.653)	6.71	0.010
Quadratic	0.111 (se=0.068; t=1.631)	2.60	0.107

Table 5

Results of parameter estimates, χ^2 and p values for time terms (Intercept, Linear, and Quadratic) for the Context effect on target fixations in Experiment 2. Standard errors and t values for the parameter estimates are shown in the parentheses next to each estimate.

Term	**Estimate**	χ^2	*p*
Intercept	0.046 (se=0.021; t=2.209)	4.48	0.034
Linear	−0.24 (se=0.092; t=−2.603)	6.15	0.013
Quadratic	−0.044 (se=0.029; t=−1.533)	2.33	0.127

Table 6

Results of parameter estimates, χ^2 and p values for time terms (Intercept, Linear, and Quadratic) for the comparisons between competition effects in neutral contexts in Experiment 1 and 2. Standard errors and t values for the parameter estimates are shown in parentheses.

Term	**Estimate**	χ^2	*p*
Intercept	0.021 (se=0.004;t=5.280)	27.66	<0.001
Linear	−0.067 (se=0.023;t=−2.871)	8.21	0.004
Quadratic	0.080 (se=0.023;t=3.439)	11.77	0.001