Unfolding Meaning in Context:

The Dynamics of Conceptual Similarity

Jelena Mirković[1, 2*] and Gerry T. M. Altmann[3]

[1] York St John University

School of Psychological and Social Sciences

York St John University

Lord Mayor's Walk

York YO31 7EX

UK

[2] University of York

Heslington

York YO10 5DD

UK

[2] University of Connecticut

Department of Psychological Sciences

University of Connecticut

406 Babbidge Road, Unit 1020

Storrs, CT 06269-1020

USA

[*] Corresponding author: jelena.mirkovic@york.ac.uk

FINAL PREPRINT PRIOR TO PUBLICATION
Mirkovic, J., & Altmann, G. T. M. (2019). Unfolding meaning in context: The dynamics of conceptual similarity. Cognition, 183, 19-43. https://doi.org/10.1016/j.cognition.2018.10.018

Abstract

How are relationships between concepts affected by the interplay between short-term contextual constraints and long-term conceptual knowledge? Across two studies we investigate the consequence of changes in visual context for the dynamics of conceptual processing. Participants' eye movements were tracked as they viewed a visual depiction of e.g. a canary in a birdcage (Experiment 1), or a canary and three unrelated objects, each in its own quadrant (Experiment 2). In both studies participants heard either a semantically and contextually similar "robin" (a bird; similar size), an equally semantically similar but not contextually similar "stork" (a bird; bigger than a canary, incompatible with the birdcage), or unrelated "tent". The changing patterns of fixations across time indicated first, that the visual context strongly influenced the eye movements such that, in the context of a birdcage, early on (by word offset) hearing "robin" engendered more looks to the canary than hearing "stork" or "tent" (which engendered the same number of looks), unlike in the context of unrelated objects (in which case "robin" and "stork" engendered equivalent looks to the canary, and more than did "tent"). Second, within the 500 ms post-word-offset eye movements in both experiments converged onto a common pattern (more looks to the canary after "robin" than after "stork", and both more than due to "tent"). We interpret these findings as indicative of the dynamics of activation within semantic memory accessed via pictures and via words, and reflecting the complex interaction between systems representing context-independent and context-dependent conceptual knowledge driven by predictive processing.

Keywords: conceptual processing; semantic cognition; eye-movements; dynamics; predictive processing

Word count: 8143

Introduction

Our knowledge of the concept of a robin comprises all its diverse properties, including its shape and color, the fact that it can fly, and that it is a bird. This intuitive understanding of concepts is shared with many cognitive and neural models of semantic memory which assume that concept knowledge is represented in terms of features constituting them (e.g., Barsalou, 1999; Cree & McRae, 2003; Glenberg, 1997; Rogers & McClelland, 2004; Smith, Shoben, & Rips, 1974; Taylor, Moss, & Tyler, 2007). Although the specific nature of the features varies across these models, they typically represent conceptual knowledge as a static featural space, with concepts from the same category (e.g., robin, canary) closer in the space than concepts from different categories (e.g., robin, tent), and with these relationships relatively fixed.

However, many studies have demonstrated that the activation of individual conceptual properties is highly dynamic. For example, in visual object recognition Yee, Huffstetler, and Thompson-Schill (2011) showed that the activation of an object's conceptual shape (e.g., round for a pizza) precedes activation of that object's conceptual function (e.g., that it is edible). In spoken word recognition, Moss, McCormick, and Tyler (1997) showed that the functional properties of man-made objects are activated before their perceptual properties (e.g., visual form). The difference between the studies showing function before form (e.g., Moss et al., 1997) and the studies showing form before function (e.g., Yee et al., 2011) highlights the additional issue of the dynamics of processing in semantic memory accessed via words and accessed via objects or their visual depictions (cf. Lupyan & Thompson-Schill, 2012; Paivio, 1991). The non-static nature of semantic cognition is further illustrated by classic studies in word recognition which have demonstrated that sentential context provides a strong constraint on the activation of conceptual properties. For example, recalling the word

'piano' after hearing *The man lifted the piano* is more accurate following a contextually appropriate cue '*something heavy'* relative to a plausible, but contextually inappropriate cue, e.g., '*something with a nice sound*' (e.g., Barclay, Bransford, Franks, McCarrel, & Nitsch, 1974; Tabossi, 1988). Furthermore, Kalénine, Mirman, Middleton, and Buxbaum (2012) have demonstrated that sentential context interacts with the time-course of conceptual processing: in their study, the differential activation of functional vs. thematic features was further modified by context, in that the typically later activation of functional features was elicited in an earlier time window with an appropriate sentential context (see also Lee, Middleton, Mirman, Kalénine, & Buxbaum, 2013).

In the current study we explore the consequences of the context and time dependent nature of conceptual processing for relationships *between* concepts. Specifically, we investigate first, how the relationship between concepts in semantic memory changes with changes in *visual context*, and second, how this relationship changes *across time* during which a word associated with one of these concepts is heard. For example, does hearing *The man lifted the piano* change the relationship between pianos and violins, and pianos and boulders, and moreover do these relationships change over time depending on the dynamics of activation of the feature <heavy>? As Tabossi (1988) observed, a critical issue is "*when* (…) contextual information becomes effective" (p. 153), and the notion that, as time unfolds, contextual influences may change.

Our starting point is the observation that on hearing a word, visual attention can be directed towards an object that is semantically related to, but is not, the object referred to by that word (Dahan & Tanenhaus, 2005; Huettig & Altmann, 2005; Mirman & Magnuson, 2009; Yee & Sedivy, 2006). Thus, hearing '*robin*' will engender looks towards a canary, and in proportion to the semantic similarity between robins and canaries (Huettig & Altmann, 2005; Huettig, Quinlan, McDonald, & Altmann, 2006; Mirman & Magnuson, 2009), and

similarly on hearing '*stork*', which also overlaps conceptually with canaries. (In fact, to the same degree as '*robin*' according to semantic similarity norms (Landauer & Dumais, 1997; McRae, Cree, Seidenberg, & McNorgan, 2005).)

However, what would be the consequences of depicting the canary in a visual context affording robins but not storks – for example, in a domestic birdcage (Figure 1)? Specifically, what is the extent to which visual attention to the canary would reflect the *context-independent* conceptual similarity between canaries and robins and canaries and storks versus the *visual-context dependent* constraints that afford robins but not storks?[1] Moreover, how would the relationship between the concepts change over the time-course of the spoken word unfolding within the visual context? How would this relationship change *after* the word has unfolded?

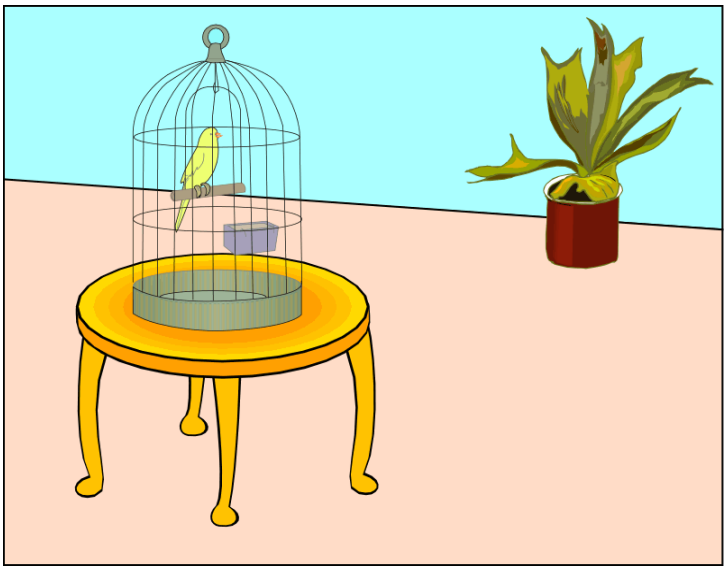


Figure 1. Constrained visual context for the target object: canary (from Experiment 1).

[1] We use the term "contextual dependence" to refer to dependencies between a particular item and the specific, or episodic, context in which it occurs. This usage of the term is different from other uses in the literature that relate to the features' representational status (e.g., Barsalou, 1982).

Standard models of semantic memory predict that the shifts in visual attention to objects in a scene upon hearing a word would reflect context-independent relationships between concepts accessed via visual depictions or via words. For example, looks to the canary would be similar when hearing *'robin'* and when hearing *'stork'* (assuming they overlap with canaries to a similar extent, although even if they overlap to differing extents, the logic we describe below remains the same), and both patterns of looks would be dissimilar to the one when hearing an unrelated word, e.g. '*tent*' (e.g., Huettig & Altmann, 2005; Huettig et al., 2006; Mirman & Magnuson, 2009; Yee & Sedivy, 2006). However, if shifts in visual attention also reflect *context-dependent* relationships between concepts in semantic memory, an alternative prediction would be that in the context of Figure 1 there would be more looks to the canary when hearing *'robin*' than when hearing *'stork*'. Moreover, while there would also be more looks to the canary after hearing '*robin*' than after hearing '*tent*', whether or not there would be more looks to the canary after hearing '*stork*' than after hearing *'tent'* (storks would not fit in the depicted birdcage) would depend on the strength of the contextual constraint afforded by the visual context (the canary *in a cage*). In other words, the presence of the cage in the visual context would change the nature of the relationship between canaries, robins, and storks: Unlike the static models of semantic memory where these three concepts share the same area of the semantic space as defined by e.g. the proportion of shared features, the visual context would shift their relative positions in this space making robins more similar to canaries and both dissimilar to storks.

An additional issue is how the activation of properties over time would influence the relationships between these concepts. Rogers and Patterson (2007) have suggested that, in a distributed model of semantic memory, activation changes over time such that the more information is shared among concepts (for example, [context-independent] category/domain-level information) the earlier it is available in processing, and that over time increasingly

more specific information becomes available (see also Clarke & Tyler, 2015). This type of conceptual dynamics would predict a late-emerging effect of the visual context, because the context influences the fit between the scene and the word regarding *specific* information (e.g., the size of robins and storks). Thus, only early in processing would looks to the canary be similar when hearing *'robin'* and when hearing *'stork'* (and both dissimilar to when hearing *'tent'*). The context-dependent similarity between canaries and robins, relative to canaries and storks, given that the birdcage affords robins but not storks, would emerge later. This pattern of looks would indicate that early conceptual processing is sensitive to context-independent relationships between concepts, and that context-dependent relationships emerge later in time as the activation spreads through the semantic network[2].

A different prediction arises from evidence for early effects of context in studies of visual object recognition (see Bar, 2004; and Oliva & Torralba, 2007, for reviews). For example, in a sentence-picture verification study using sentential contexts describing different states of an object (e.g., flying duck vs. sitting duck, '*The ranger saw a duck in the air, 'The ranger saw a duck in the lake*), participants were not only faster to recognize the object when its depicted state (e.g., flying duck) matched the one described in the sentential context (duck in the air), but crucially this effect was reflected in the early modulation of the M1 component of neural activity in the occipital cortex, with a stronger response to the state-matching picture relative to both the state-mismatching and an unrelated picture within the first 100 ms from picture onset (Hirschfeld, Zwitserlood, & Dobel, 2011). Thus, in the context of a visual scene affording robins but not storks, access to context-relevant specific information, including the relationship between the cage and the canary and the constraints this places on the canary

[2]As pointed out below, the standard featural models of semantic memory (e.g., Cree & McRae, 2003; Rogers & McClelland, 2004; Taylor et al., 2007) are not in principle incompatible with time and context-sensitivity.

(e.g., size), should be available early. Hence early in the processing of '*robin*' or '*stork*', looks to the canary depicted in the context of a birdcage would reflect higher context-dependent similarity between canaries and robins than between canaries and storks. These predictions also fit a broader set of models of predictive processing which describe the brain as a 'proactive' organ continuously anticipating upcoming input (e.g., Bar, 2007; Clark, 2013). In this view, context preactivates likely object representations and in this way enacts top-down constraints (Trapp & Bar, 2015). The predictive processing driven by the context would result in a change in the semantic space such that the greater context-dependent similarity between canaries and robins relative to storks would be evident early in processing. What remains unclear in this account is whether and at what point in time there would be evidence of the context-*independent* similarity between the three (i.e. as defined by all three being members of the bird category).

Below, we explore these alternative views of processing dynamics in semantic memory by examining the probability through time of fixating a visual depiction of a canary in a cage (Experiment 1), and of fixating the same canary among unrelated objects (Experiment 2), in three different conditions: First, when the picture is accompanied by a semantically *and contextually* similar word ('*robin'*); second, when it is accompanied by a semantically *but not contextually* similar word ('*stork'*), and third, by both semantically and contextually unrelated word ('*tent'*). (We use the term '*contextually similar*' relative to the constraining contexts of Experiment 1 – in Experiment 2 there are no contextual constraints beyond the unrelated objects presented concurrently in the display.) We assume the standard linking hypothesis concerning the relationship between conceptual representation, language, and visual attention (Allopenna, Magnuson, & Tanenhaus, 1998; Altmann & Kamide, 2007; Tanenhaus, Magnuson, Dahan, & Chambers, 2000), and interpret the probability of fixation at each time point as an indication of the activation level of those components of the conceptual

representations activated by the target word (e.g., '*stork*') that overlap with those associated with the target object (the canary). At issue is whether the effects of context are early or late, how these effects change across time, and crucially how they impact on the conceptual similarity of the target object (the canary) to objects that are related (robins and storks are also birds) but differ in how they fit the visual context.

Experiment 1

In this experiment, individual target objects (e.g., a canary) were presented in the context of a visual scene (e.g., the canary was depicted in a birdcage; Figure 1). For standard models of semantic memory, the relationship between canaries and robins and storks is determined by their semantic similarity defined by their shared features, with features being fixed dimensions of the semantic space. For measures of semantic similarity derived from these models (Landauer & Dumais, 1997; McRae et al., 2005), canaries are as similar to robins as they are to storks (see Table 1 below). However, in the context of a visual scene affording robins but not storks (due to their size and likelihood of fitting into a cage) the conceptual representation of a canary may be more similar to a robin than to a stork. In fact, in this particular context, storks may be no more similar to canaries than to completely unrelated concepts (e.g., *tent*). If this is the case, participants' eye movements to the canary (in Figure 1) should reflect this *context-dependent* semantic similarity: There should be more looks to the canary in response to the word '*robin*' than in response to either '*stork*' or '*tent*'. At issue here is first, the extent to which visual context changes the relationship between the concepts, and second, *when* such change is observed in the eye movement record.

Method

Participants

Forty-two native speakers of English from the University of York participated in the study for monetary remuneration (£4) or course credit after providing a written informed consent.

Stimuli

The stimuli consisted of 21 quadruplets: a visual target (a canary), a semantically and contextually related word (‘*robin*’), a semantically but not contextually related word (‘*stork*’), and a semantically and contextually unrelated word (‘*tent*’) (see Appendices A and B for the full set of items). Semantically related words (‘*robin*’*,* ‘*stork*’) were from the same semantic category as the target object, and were equally semantically similar to the target, as measured by the semantic similarity norms of McRae and colleagues (McRae et al., 2005), as well as by LSA norms (Landauer & Dumais, 1997) (for the McRae norms: $t(40) = -.067$, $p = .95$; for LSA: $t(40) = -.55$, $p = .59$; Table 1). Unrelated words were semantically unrelated to the target object. The words in the three conditions were equated in length (number of phonemes and spoken duration) and word frequency (Table 1). The words were digitally recorded by a native English speaker in a sound attenuated booth, sampled at 44.1 KHz.

The visual stimuli consisted of 21 images representing the target object (e.g., canary) in the context of two unrelated objects (e.g., table, pot plant) and one object that contained or otherwise constrained the target object (e.g. a birdcage, with the canary inside the cage; see Appendix B for the full set of experimental images). The two unrelated objects were unrelated to the three spoken words. The visual scenes were designed such that they afforded the semantically *and contextually similar* item, but not the semantically equally similar but contextually-dissimilar item. For example, due to size restrictions, only a robin can fit the cage in Figure 1, but not a stork.

Table 1. Properties of the spoken words in the three conditions.

		Semantically and contextually related	Semantically but not contextually related	Unrelated
Semantic similarity with the target	McRae norms	0.347	0.349	0
	LSA	0.282	0.313	0.066
Length	number of phonemes	5.1	5.1	4.9
	spoken duration (ms)	531	539	529
Log frequency (British National Corpus)		6.2	6.4	6.8

The visual scenes were created using commercially available ClipArt. They were displayed as 800 x 600 px images on a computer screen with a resolution of 1024 x 768 px. All target objects were clearly visible in all images.

In addition to 21 experimental visual scenes, there were 43 filler scenes, similar in complexity and content to the experimental items. The spoken word for 32 of the filler items referred to an object presented in the scene. In 11 filler items the spoken word was unrelated to any of the objects presented in the scene. Thus within each list of items 50% were visual scenes where the spoken word referred to an entity in the scene, 28% were scenes where the word was unrelated to any entity in the scene, and 22% were scenes where the word was related to an entity in the scene.

There were three lists of 64 items, such that participants saw each visual scene only once. Within each list, a third of the experimental items was presented with a semantically and contextually related word, a third was presented with a semantically but not contextually

related word, and a third was presented with an unrelated word. Thus the semantic and contextual relationship between target objects and auditory words was manipulated both within participants and within items.

Procedure

Participants were seated in front of a 22-inch display monitor, with their eyes approximately 60 cm away from the monitor. They wore an EyeLink II head-mounted eye-tracker, sampling at 250 Hz. The auditory stimuli were presented via two loudspeakers located at each side of the display monitor.

Participants were instructed to look at the visual scene and listen to the words played over the loudspeakers and try to understand them (the 'look-and-listen' task, Altmann and Kamide (1999)).

A drift-correction dot was presented at the onset of each trial. After the participant looked at the dot, it was replaced by the visual stimulus. After 3s, the auditory stimulus was played over the loudspeakers. The visual stimulus stayed on the screen for an additional 2s post-word offset. A 9-point calibration procedure was performed after every 8 trials. There were 4 practice trials before the main experimental block. The entire session lasted approximately 20 minutes.

Data Analysis

We used mixed effects modeling with empirical logit transformed proportion of fixations to the target object (canary) aggregated over 50 ms bins as the outcome measure, separately by participants and by items (Barr, 2008; [dataset] Mirković & Altmann, 2018)[3]. The analyses were run using the lme4 package in R (Bates, Machler, Bolker, & Walker, 2015). To

[3] The analyses using raw proportions (suggested by a reviewer based on Donnelly & Verkuilen, 2017) yielded the same pattern of findings for both Experiments.

minimize the influence of the variation in the duration of the spoken words, we present the analyses in two 500 ms windows, with the first window synchronized to the word onset, and the second to the word offset, on a trial by trial basis (the average duration of the auditory stimuli was 532 ms). Time and Word (semantically and contextually similar (e.g., *robin*), semantically but not contextually similar (e.g., *stork*), and semantically and contextually unrelated (e.g., *tent*)) were included as fixed factors. Random effects for participants and items were included with maximal inclusion of random slopes that allowed the model to converge (Barr, Levy, Scheepers, & Tily, 2013). In Experiment 1, to model Word effects Helmert coding was used to first compare looks to the canary when hearing the semantically and contextually related word (e.g., *robin*) relative to the semantically but not contextually related (*stork*) and unrelated (*tent*) words, and second to compare the latter two (*stork* and *tent*). Significance was calculated using the Satterthwaite approximation.

Results

The time course of fixations to the target objects is presented in Figure 2, and model parameters in Table 2. As the spoken word unfolded in time (the first 500 ms time window from word onset), there was an overall increase in looks to the target object (e.g., canary) (the effect of time in Table 2). Crucially, the rate of the increase was significantly higher while hearing '*robin'* than while hearing '*stork'* or *'tent'* (time x robin vs. stork+tent), with no difference between *'stork'* and *'tent'* (time x stork vs. tent) over this time period. Thus the visual context had an early effect on the relationship between the concepts, with the looks to the canary reflecting a higher overlap with the semantically and contextually related *'robin'* relative to the semantically but not contextually related *'stork'*. Crucially, for most of the duration of the spoken word the looks to the target did not reflect the category/domain-level

similarity, as they were no different while hearing the semantically but not contextually related *'stork'* relative to the completely unrelated *'tent'* (Figure 2).

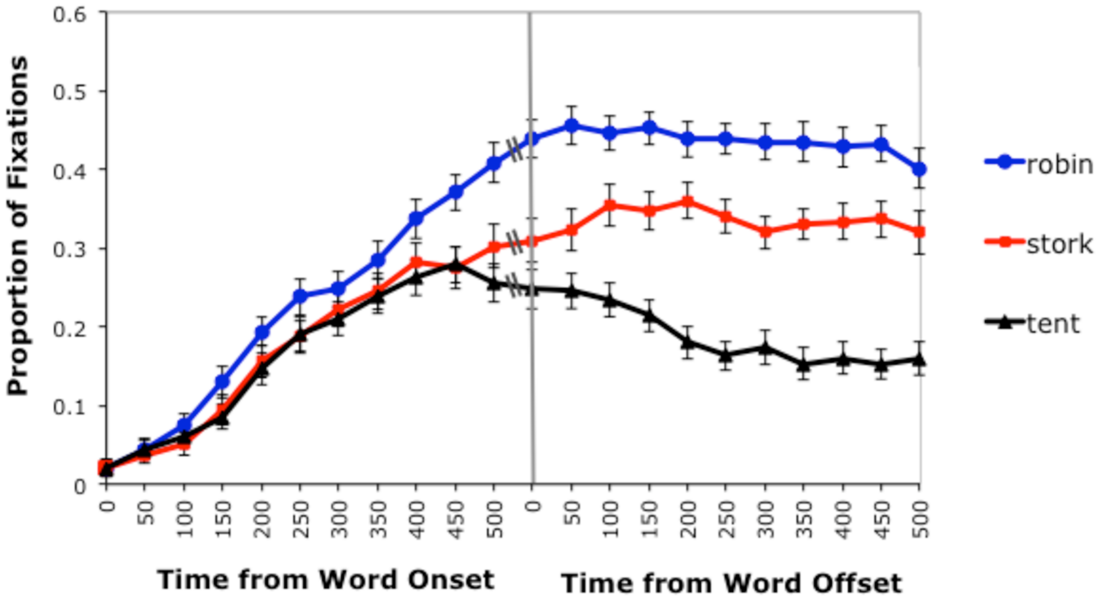


Figure 2. Experiment 1: The time-course of fixations to the target object (e.g., *canary*) in the context of a visual scene. The left half of the figure represents fixations synchronized to the onset of the auditory stimulus (semantically and contextually related: '*robin'*, semantically but not contextually related: '*stork*', unrelated: '*tent*'), and the right half to the offset of the auditory stimulus, on a trial-by-trial basis. Proportions of fixations are presented in all figures for ease of interpretation. Error bars represent standard error.

A difference between *'stork'* and *'tent'* started to emerge only at the end of the spoken word. The analyses of the looks in the 500 ms window synchronized to the word offset showed that at word offset there continued to be more looks to the canary after hearing *'robin'* relative to both *'stork'* and *'tent'* (Table 3: robin vs. stork+tent), and a difference emerged between *'stork'* and *'tent'* (stork vs. tent). Over this time period a significant difference in the rate of change in the looks to the canary after hearing *'stork'* relative to *'tent'* also emerged (time x stork vs. tent). Thus the looks to the canary immediately after word offset continued to show a higher similarity to the semantically and contextually related

'robin', but now the looks also reflected the category/domain-level higher overlap with *'stork'* relative to *'tent'*.

Table 2. Model parameters for empirical logit transformed proportions of looks to the target in Experiment 1 synchronized to the word onset.

Fixed factors	By participants				By items			
	ß	SE	t	p	ß	SE	t	p
time	3.86	.16	23.47	<.001	4.63	.33	13.90	<.001
robin vs. stork+tent	.00	.08	.01	>.1	.06	.09	.70	>.1
stork vs. tent	-.04	.15	-.30	>.1	-.05	.17	-.29	>.1
time x robin vs. stork+tent	.63	.13	4.87	<.001	.49	.17	2.98	.003
time x stork vs. tent	.20	.22	.90	>.1	.15	.29	.52	>.1

Note: Random effects structure: intercept and slopes for time, robin vs. stork + tent, and stork vs. tent (by participants and by items).

Table 3. Model parameters for empirical logit transformed proportions of looks to the target in Experiment 1 synchronized to the word offset.

Fixed factors	By participants				By items			
	ß	SE	t	p	ß	SE	t	p
time	-.47	.21	-2.18	.035	-.45	.29	-1.58	>.1
robin vs. stork+tent	.48	.07	6.84	<.001	.50	.12	4.18	<.001
stork vs. tent	.37	.15	2.43	.019	.44	.22	1.98	.047
time x robin vs. stork+tent	.20	.11	1.78	.076	.27	.16	1.75	.080
time x stork vs. tent	1.35	.20	6.88	<.001	1.21	.27	4.47	<.001

Note: Random effects structure: intercept and slopes for time, robin vs. stork + tent, and stork vs. tent (by participants and by items).

Discussion

Given prior semantic similarity effects using the same paradigm (e.g., Huettig & Altmann, 2005; Huettig et al., 2006; Mirman & Magnuson, 2009; Yee & Sedivy, 2006), these findings

demonstrate that visual context can exert a strong influence on the relationship between concepts, making semantically *but not contextually* related concepts no more similar to each other than are completely unrelated concepts. This runs counter to the views of semantic memory which describe concepts in terms of fixed points in semantic space (e.g., Cree & McRae, 2003; Landauer & Dumais, 1997; McRae et al., 2005). Some of these models are not in principle incompatible with our findings (e.g., Cree & McRae, 2003; Rogers & McClelland, 2004), but what our findings clearly demonstrate is that they must be extended to be able to accommodate the dynamical change of the semantic space resulting from contextual constraints. Moreover, the finding that the *early* looks reflected the effect of the visual context fits with the predictions of models of predictive processing (Bar, 2007; Trapp & Bar, 2015) in that the visual scene pre-activated the context-relevant features allowing for the effect to emerge early as the semantically and contextually related spoken word was unfolding. Importantly, the category/domain-level similarity between concepts (both robin and stork different from tent) only emerged later, in the looks post-word offset. It is important to note that our assumption here is that the dynamics of semantic activation during the 3 s delay between display onset and word onset is constant across conditions (e.g., Chen & Mirman, 2015), and that any changes in that dynamic as a function of word type, as evidenced by the eye movement measure, will reflect the interaction between the prior semantic activation driven by the visual image, and the semantic dynamics associated with recognition of the unfolding word in the context of the visual scene. We return to the implications of these findings in the General Discussion.

Experiment 2

To further test the early context-dependent change in the relationship between concepts in semantic memory, in Experiment 2 we presented the same target objects in the context of three unrelated objects (Figure 3). In this case the available visual context should not afford robins any more than storks (or tents), and thus this context should reveal standard category/domain-level semantic similarity effects (e.g., Huettig & Altmann, 2005; Yee & Sedivy, 2006), i.e. more looks to the canary after hearing both *robin* and *stork* relative to *tent.* Note that the concept of the canary is accessed by its visual depiction, which may influence the dynamics of activation of its visual properties (Moss et al., 1997; Yee et al., 2011). Similar to Experiment 1, a distributed model of conceptual processing (e.g., Rogers & Patterson, 2007) would predict that the early looks to the canary would reflect conceptual differences at the context-independent (category/domain) level. Thus we would expect more early looks to the canary when hearing both *robin* and *stork* relative to *tent* and no difference between *robin* and *stork* (to the extent that they are equally similar to the concept canary), while later looks may reflect specific, visual form-related differences between the concepts. Models of predictive processing (e.g., Bar, 2007) would converge on similar predictions, as the objects in the visual context are unrelated and thus no specific context-relevant property would be expected to preactivate.

Figure 3. Example item from Experiment 2.

Method

Participants

Thirty nine native speakers of English from the University of York participated in the study for monetary remuneration (£4) or course credit.

Stimuli

The visual stimuli consisted of 21 images representing the target object (e.g., canary) in the context of three unrelated objects (e.g., drums, paintbrush, foliage). Each object was located at the center of a 400 x 300 pixel quadrant. The three unrelated objects were also unrelated to the three auditory words. As in Experiment 1, the image dimensions were 800 x 600 pixels, displayed on the screen with a resolution of 1024 x 768 pixels. We used the same visual depiction of the target object (the same Clipart image) as in Experiment 1 (see Appendix B for the images). The same words (and audio files) were used as in Experiment 1.

We included 43 filler items, using the same objects as in the fillers from Experiment 1. Similar to the experimental items, each filler image contained four objects. Thirty-two filler images were presented with a word that named one of the four objects, and 11 filler images were presented with a word unrelated to any of the objects.

The location of the target object was counterbalanced such that it was equally likely to occur in all four quadrants. There were three lists of 64 items, employing the same design as Experiment 1.

Procedure and Data Analysis

The procedure and data analyses ([dataset] Mirković & Altmann, 2018) were the same as in Experiment 1. The only exception was the Helmert coding scheme: given the predictions, the coding scheme compared first, the looks to the canary while hearing *'robin'* and *'stork'* relative to *'tent'*, and second, the looks to the canary while hearing *'robin'* relative to *'stork'*.

Results

The time course of fixations to the target objects in the context of three other unrelated objects is presented in Figure 4.

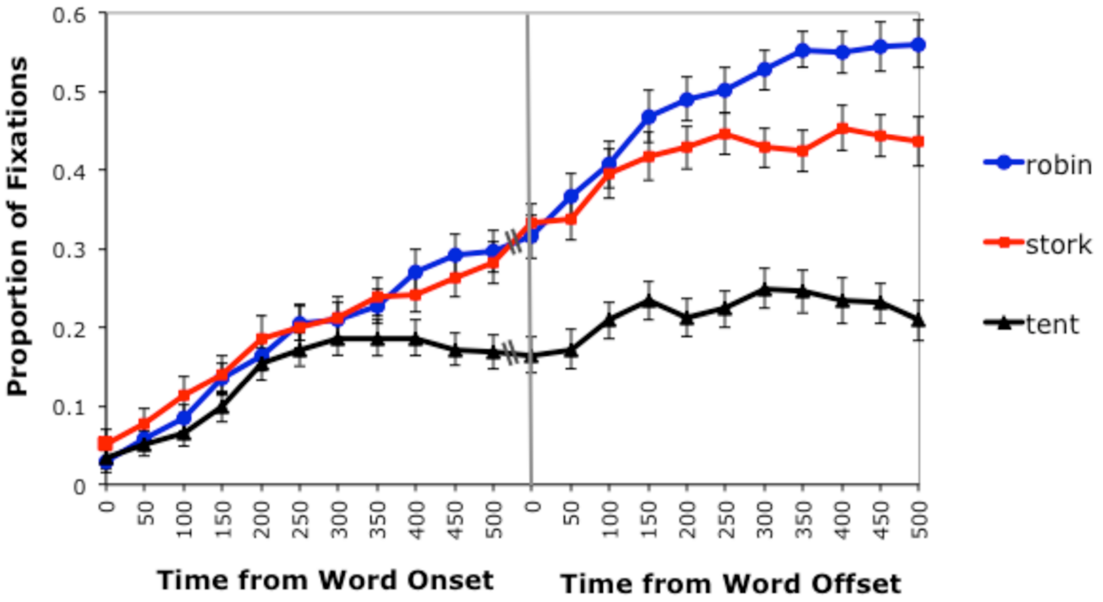


Figure 4. Experiment 2: The time-course of fixations to the target object (*canary*) presented in the context of three unrelated objects. The left half of the figure represents fixations synchronized to the onset of the auditory stimulus ('*robin*', '*stork*', '*tent*'), and the right half to the offset of the auditory stimulus, calculated on a trial-by-trial basis. Error bars represent standard error.

As the word unfolded in time there was an overall increase in the looks to the target object (Table 4: time). As illustrated in Figure 4, the rate of the increase in looks to the canary was higher while hearing *'robin'* and *'stork'* relative to *'tent'* (time x robin+stork vs. tent): looks to the canary plateaued midway through hearing unrelated *'tent*', while they continued to rise while hearing *'robin'* and *'stork'*. There was a faster rate of increase for *'robin'* relative to *'stork'* (time x robin vs. stork), but unlike Experiment 1, at word offset the looks to the canary were exactly the same after having just heard *'robin'* or *'stork'* (Table 5: robin vs. stork), and in both cases there were more looks for both relative to *'tent'* (Table 5: robin+stork vs. tent). Over this 500 ms period post word offset, looks to the canary after hearing *'robin'* or *'stork'* continued to rise faster relative to *'tent'* (Table 5: time x robin+stork vs. tent). There was now also a clear faster increase after *'robin'* relative to

'stork' (Table 5: time x robin vs. stork): looks to the canary after hearing *'stork'* plateaued midway through this period, while they continued to rise after hearing *'robin'*.

Table 4. Model parameters for empirical logit transformed proportions of looks to the target in Experiment 2 synchronized to the word onset.

Fixed factors	By participants				By items			
	ß	SE	t	p	ß	SE	t	p
time	2.69	.25	10.55	<.001	3.20	.28	11.48	<.001
robin+stork vs. tent	.02	.09	.26	>.1	.02	.09	.22	>.1
robin vs. stork	-.16	.17	-.95	>.1	-.26	.20	-1.33	>.1
time x robin+stork vs. tent	.59	.13	4.54	<.001	.67	.17	3.83	<.001
time x robin vs. stork	.56	.23	2.48	.013	.98	.30	3.25	.001

Note: Random effects structure: intercept and slopes for time, robin + stork vs. tent, and robin vs. stork (by participants and by items).

Table 5. Model parameters for empirical logit transformed proportions of looks to the target in Experiment 2 synchronized to the word offset.

Fixed factors	By participants				By items			
	ß	SE	t	p	ß	SE	t	p
time	1.17	.26	4.58	<.001	1.33	.35	3.76	.001
robin+stork vs. tent	.56	.10	5.84	<.001	.54	.14	3.80	<.001
robin vs. stork	.02	.15	.12	>.1	.04	.24	.19	>.1
time x robin+stork vs. tent	.51	.14	3.67	<.001	.75	.17	4.44	<.001
time x robin vs. stork	1.08	.24	4.51	<.001	1.08	.29	3.68	<.001

Note: Random effects structure: intercept and slopes for time, robin + stork vs. tent, and robin vs. stork (by participants and by items).

These findings show that when an object (e.g., a canary) is presented in the context of unrelated objects, the conceptual correlates of the unfolding semantically related words (e.g., robin, stork) are initially perceived as similar to the conceptual correlates of the target object. As illustrated in Figures 4 and 5, this category/domain-level similarity between the concepts

was reflected in the parallel rise of looks to the target object as the category-related words unfolded. The dynamics of looks during this period is strikingly different from the looks to the same target object while hearing the same spoken words in Experiment 1, when that target object was presented in a meaningful visual scene (Figures 2 and 5). In that case, the visual scene modulated the activation of the conceptual representations: whereas in the context of unrelated objects '*robin*' and '*stork*' "travel together" as they unfold reflecting a similar degree of conceptual overlap with the canary, the same is not the case in the context of a meaningful visual scene. Here, the concept associated with '*robin*' is more similar to the conceptual representation of the canary *given the visual context* than is the concept associated with '*stork*', and hence the latter "travels together" with the unrelated *'tent'*.

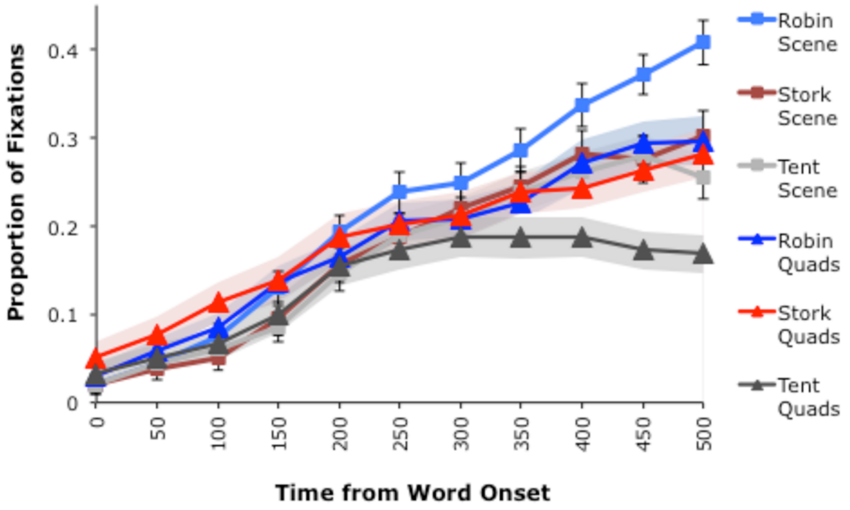


Figure 5. Fixations to the target object (*canary*) in the context of a scene (Experiment 1: squares) or in the context of three unrelated objects presented in different quadrants (Experiment 2: triangles) from word onset. Error bars (for the scene context) and ribbons (for the quadrants) represent standard error.

To further assess the conceptual dynamics in the two visual contexts, we pooled the data from the two experiments together. We compared the looks to the target object at word offset, and the time course for 500 ms after word offset. The analyses at word offset are analogous

to more traditional analyses of eye movements that, rather than assessing eye movements over an entire time-course, assess the looks to the target at a particular moment in time. These analyses reflect the outcome of the early interaction between visual and linguistic processing in the time leading up to word offset. The analyses of the time-course synchronized to the word offset assess the dynamics of the eye movements from the moment when both the visual context and the linguistic stimuli were fully available for processing onward.

The analyses at word offset used the empirical logit-transformed proportions of fixations at the target synchronized to the word offset, while the time-course analyses used the proportions over 50 ms bins starting from the word offset for 500 ms. In both analyses, the fixed factors included visual context (scene vs. quadrants, i.e. Experiment 1 vs. Experiment 2), and word type using Experiment 2 contrasts. The use of Experiment 2 contrasts, comparing looks after hearing *'robin'* and *'stork'* relative to *'tent'*, and then *'robin'* vs. *'stork'*, allowed us to assess more specific hypotheses than an ANOVA-style omnibus interaction, and in particular the key visual context x *'robin'* vs. *'stork'* interaction. This interaction assesses the extent to which the visual context changes the relationship between semantically related concepts, i.e. the extent to which context-dependent similarity influences processing.

As illustrated in Figure 6, at word offset there were overall more looks to the canary immediately upon hearing both '*robin'* and '*stork'* relative to '*tent',* and upon hearing '*robin'* relative to '*stork'* (Table 6, by-participant analysis). Crucially, there were significantly more looks to the canary upon hearing '*robin'* than upon hearing '*stork'* in the context of the visual scene but not in the context of three unrelated objects (context x robin vs. stork interaction in Table 6, by-participant analysis). Given that none of the effects were significant in the by-item analysis (which is likely under-powered, Brysbaert and Stevens (2018)), these findings should be taken with caution. However, overall, these analyses provide evidence that by word

offset the visual scene affording robins but not storks makes the conceptual correlates of the word '*robin'* more similar to the concept of the canary than the conceptual correlates of the word *'stork*', unlike the context of unrelated objects.

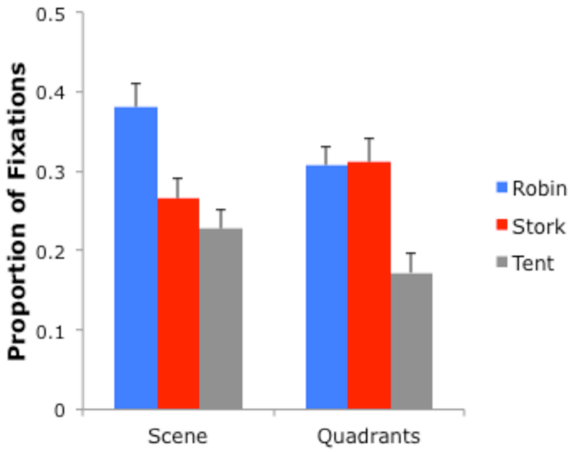


Figure 6. Fixations on the target object (*canary*) in the context of a scene (Experiment 1) and in the context of three unrelated objects presented in different quadrants (Experiment 2) at word offset. Error bars represent standard error.

Table 6. Model parameters for empirical logit transformed proportion of looks at word offset for the pooled data.

Fixed factors	By participants				By items			
	ß	SE	t	p	ß	SE	t	p
context	.17	.12	1.42	>.1	.28	.22	1.28	>.1
robin+stork vs. tent	.38	.07	5.30	<.001	.39	.10	3.97	<.1
robin vs. stork	.25	.12	2.04	.044	.32	.18	1.75	.087
context x robin+stork vs. tent	-.23	.14	-1.60	>.1	-.19	.20	-0.94	>.1
context x robin vs. stork	.67	.25	2.67	.008	.57	.37	1.55	>.1

Note: Random effects structure: intercept and slopes for robin vs. stork (by participants and by items).

The time-course of looks to the target object in two visual contexts synchronized to the word offset (Figure 7) provides further support to the finding that visual context influences the dynamics of conceptual processing. The key influence of the visual context on the

dynamics of eye movements is illustrated in the three-way interaction between visual context, time and the robin vs. stork contrast (Table 7): First, while already at the start of this period (word offset) in the context of a scene (affording robins but not storks) there are more looks to the canary upon hearing '*robin'* than upon hearing '*stork',* and in both conditions more relative to '*tent'*, in the context of three unrelated objects there is initially the same amount of looks to the canary upon hearing '*robin'* and upon hearing '*stork'*, and in both cases more relative to '*tent'*. Second, while this pattern of looks is maintained in the scene context (particularly for '*robin'* and '*stork'*, with a small decrease upon hearing '*tent'*), in the context of unrelated objects the eye movements continue to evolve such that in a period approximately 300 ms after word offset there is a rise in the looks to the canary upon hearing '*robin'*, and not much further change upon hearing '*stork'* or '*tent'*. These analyses again demonstrate that the visual context crucially changes the relationship between concepts as reflected in the eye movements to the target object.

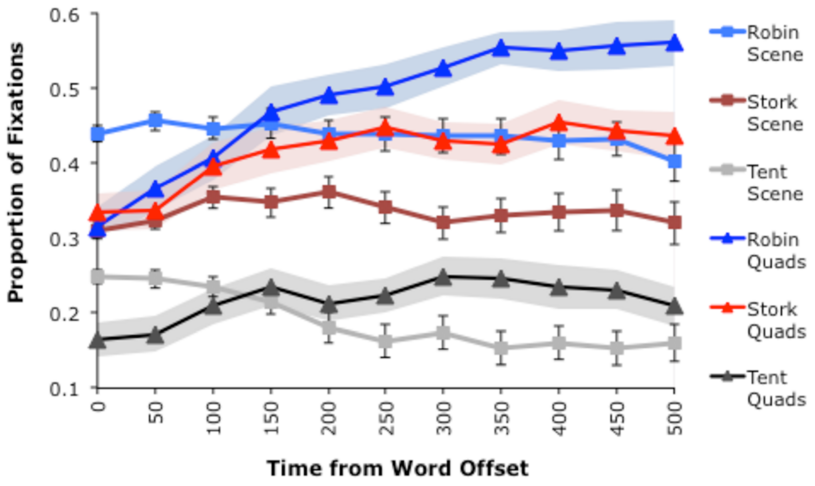


Figure 7. Fixations to the target object (*canary*) in the context of a scene (Experiment 1) or in the context of three unrelated objects presented in different quadrants (Experiment 2). The eye movements were synchronized to the word offset on a trial-by-trial basis. Error bars (for the scene context) and ribbons (for the quadrants) represent standard error.

Table 7. Model parameters for empirical logit transformed fixation proportions for the pooled data over the 500 ms period synchronized to the word offset.

Fixed factors	By participants				By items			
	ß	SE	t	p	ß	SE	t	p
context	.19	.13	1.49	>.1	.16	.24	.69	>.1
time	.35	.17	2.13	.037	.44	.23	1.93	.061
robin+stork vs. tent	.49	.06	7.74	<.001	.51	.10	5.34	<.001
robin vs. stork	.28	.10	2.84	.005	.29	.16	1.79	.079
context x time	-1.63	.33	-4.94	<.001	-1.78	.45	-3.92	<.001
context x robin+stork vs. tent	-.13	.13	-1.00	>.1	-.07	.19	-.38	>.1
context x robin vs. stork	.52	.20	2.66	.009	.48	.32	1.52	>.1
time x robin+stork vs. tent	.64	.09	7.21	<.001	.75	.12	6.48	<.001
time x robin vs. stork	.35	.15	2.30	.021	.44	.20	2.22	.027
context x time x robin+stork vs. tent	.27	.18	1.50	>.1	-.01	.23	-.05	>.1
context x time x robin vs. stork	-1.45	.31	-4.72	<.001	-1.27	.40	-3.19	.001

Note: Random effects structure: intercept and slopes for time, robin + stork vs. tent, and robin vs. stork (by participants and by items).

Discussion

The initial pattern of looks to the target object found in the context of unrelated objects (Experiment 2) is in line with the standard models of semantic memory (e.g., Cree & McRae, 2003; Landauer & Dumais, 1997; Rogers & McClelland, 2004; Taylor et al., 2007), and with behavioral findings showing the sensitivity of eye movements to context-independent category/domain-level semantic similarity (e.g., Huettig & Altmann, 2005). Thus looks to the target object (e.g., canary) increase while hearing semantically related words (robin, stork) relative to unrelated words (tent). Interestingly, the conceptual relationships (reflected in the eye movements) even in this seemingly neutral context change across time, in that later eye movements reveal higher similarity between robins and canaries than between storks and canaries (as evidenced by more fixations on the canary in response to '*robin*' than to '*stork*'). We interpret these findings as reflecting different ways of accessing the conceptual system

via pictures and via words (e.g., Lupyan & Thompson-Schill, 2012): the visual depiction of the canary provides specific information about the visual-form (e.g., size, shape, color) unlike a spoken or a printed word. Thus we hypothesize that the greater similarity in visual properties between canaries and robins relative to storks (for example, with regards to size) led to the late-emerging difference between robins and storks. The reason for its late emergence here is that size was a less relevant featural dimension in the displays used in Experiment 2 than it was in the context of the visual scenes used in Experiment 1, where size was key in respect of the interpretation of, and integration across, the elements of the scene.[4] Thus, from the perspective of accounts of predictive processing (e.g., Bar, 2007), there was no pre-activation of a specific context-related feature given the neutral context of the unrelated objects. This findings is also compatible with the late emerging differences in specific names in the model of Rogers and Patterson (2007), implementing standard models of semantic cognition.

The joint analyses across the two experiments and two visual contexts provide additional evidence to demonstrate the crucial influence of context on the dynamics of conceptual processing, which we discuss further below.

[4] It could be argued that our effects in Experiment 1 may have been exacerbated by size becoming a particularly salient feature across trials (albeit implicitly – we observed no differences across word conditions during the 3 second preview period). Similarly, the scenes in Experiment 1 may have led inadvertently to expectations regarding the spoken word (e.g., that it will be related somehow to the more salient object(s) in the scene) that were absent in Experiment 2. At issue is why certain relations would be more expected than others (e.g. making '*robin*' more expected than '*stork*') – indeed, this is exactly the contrast that we intended to study: Contextual relevance, contextual salience, and contextual expectation are closely related; salience and expectation are themselves contextually determined. The conceptual activation dynamics we observed in both studies necessarily reflect contextual dependencies, as intended.

General Discussion

Across two experiments exploring spoken word processing in the context of a visual scene we found that the content of the visual scene crucially determines the relationship between the concepts represented by the words and the objects in the scene. In Experiment 1, in the context of a meaningful visual scene depicting a canary and affording robins but not storks (two concepts otherwise equally similar to the canary), participants' eye movements to the canary reflected the greater context-dependent similarity between canaries and robins relative to canaries and storks. The context dependence was evident early on in the processing of the word, with a clear advantage for *'robin'* relative to *'stork*' at word offset. Moreover, the early looks indicated that in the context of the visual scene storks were no more similar to canaries than were tents – the specifics of the visual context gave greater prominence to featural dimensions that were contextually relevant, thereby influencing early conceptual activation along those dimensions. Over time, the dynamics of activation in semantic memory allowed for context-independent category/domain-level similarities between canaries and storks (e.g., that they are both birds) to emerge. Crucially, the dynamics we observed reflected the relationship between the unfolding conceptual correlates of the auditory word and the (potentially still unfolding) conceptual correlates of the target object and of the context in which it occurred; that is, the observed dynamics do not just reflect the changing relationship between the auditory word and the target object, but between the auditory word and the distinct (but dynamically interacting) elements of the scene more generally. In contrast, in the absence of a meaningful visual scene, i.e. in the context of three objects unrelated to birds or tents in Experiment 2, featural dimensions on which storks and canaries differ were not contextually relevant, and early looks to the canary reflected the context-independent category/domain-level similarity between canaries and robins, and canaries and storks,

relative to tents. Over time, the dynamics of activation in semantic memory accessed by the words and by a specific visual instantiation of the depicted concept allowed for the differences between storks and canaries, and robins and canaries, to emerge. Thus participants' eye movements to objects in a display were modulated by the spoken words in a way that revealed an interaction between the visual context in which the object was situated and both context-dependent and context-independent conceptual properties of heard words and the depicted object. Critically, this interaction was reflected in the differences in the time course with which the activation in semantic memory manifested behaviorally, specifically showing the impact of the shifting semantic space on selective (visual) attention.

We appealed to the predictive processing framework in visual cognition (e.g., Bar, 2007) to account for the early effect of the visual context in Experiment 1, where the context of the visual scene pre-activates context-relevant features which then facilitate the activation of the semantically *and contextually* related conceptual correlates of the unfolding word *'robin'* relative to the semantically but not contextually related correlates of the word *'stork'* (and relative to completely unrelated word *'tent'*). According to this framework, the neural processes supporting cognition allow for continuous anticipation of future outcomes based on current input. Specifically, Bar (2007) suggests that the current input is mapped onto memory representations by a process of similarity-based analogy, and the associative nature of memory allows for the activation to spread through memory networks, constituting the process of forecasting predictions (see Altmann & Mirković, 2009; and Kuperberg & Jaeger, 2016, for related proposals in language comprehension). Crucially, the context in which the input is encountered constrains the activation in memory such that it biases the activation of context-relevant features. This process is accomplished by reciprocal connections between neural networks that encompass what are traditionally considered semantic and episodic memory areas, and pre-frontal regions involved in prediction processes (see Clarke, Taylor,

& Tyler, 2011; Clarke & Tyler, 2015; Dikker & Pylkkanen, 2013, for related evidence). This view is also compatible with recent versions of distributed models of semantic cognition that include a separate semantic control network that allows for greater time and context sensitivity (e.g., Lambon Ralph, Jefferies, Patterson, & Rogers, 2017).

Context effects and prediction are not in principle incompatible with standard cognitive models of semantic memory (e.g., Cree & McRae, 2003; Rogers & McClelland, 2004; Taylor et al., 2007). They can be captured as dynamic changes in the featural semantic space as a function of context. In distributed featural models, the consequence of these changes is a warping of the semantic space, crucially changing the *relationships* between concepts, and making semantically *and contextually* related concepts more similar to each other than semantically but not contextually related concepts (see Çukur, Nishimoto, Huth, & Gallant, 2013 for neural evidence of the warping of semantic networks as a result of task contexts). Alternatively, the same effects may be captured by modeling the fit between the context and the currently processed concepts. Importantly, and as shown in our studies, the context-related change in the semantic space does not obliterate the long-term conceptual knowledge: in Experiment 1, the context-independent category/domain level similarity emerges in later processing, and in the neutral context of Experiment 2 it emerges early in processing, as predicted by standard models of semantic memory (e.g., Rogers & Patterson, 2007).

Regardless of the visual context, and notably even when relative size is irrelevant (Experiment 2), the patterns of eye movements eventually settle to reveal more subtle differences in conceptual overlap than are revealed through either semantic feature norms (McRae et al., 2005) or LSA (Landauer & Dumais, 1997) – namely, that robins are more similar to canaries than are storks, precisely because of their size. The late-emerging difference between robins and storks as they relate to canaries in the neutral context of Experiment 2 provides further evidence in support of the idea that there are differences in

conceptual representations that are activated by words and by the concept's visual depictions (e.g., Lupyan & Thompson-Schill, 2012). In particular, the visual depiction of the concept of a canary is far richer in perceptual detail than the categorical abstraction of the spoken word (see also Edmiston & Lupyan, 2015, 2017). The specific visual properties instantiated in the picture clearly activate different aspects of the conceptual representation than an isolated word, and have further consequences for the dynamics of activation in semantic memory and the relationship between concepts.

The current findings are compatible with models of semantic cognition emphasizing conceptual flexibility (e.g., Kiefer & Pulvermuller, 2012), and broader dynamic views of cognition (e.g., Elman, 2004; Tabor, Juliano, & Tanenhaus, 1997). Our findings shed light on the dynamics of interaction between short-term contextual constrains (i.e. what's often termed episodic knowledge) and long-term conceptual knowledge (i.e. semantic memory), suggesting that the notion of "semantic space" should not be interpreted as applying to a single memory system (e.g. semantic memory as distinct from other memory systems) but should be interpreted as a space defined *across* memory systems and the neural substrates that support them. Such interactions are one of the key issues in semantic cognition (Clarke & Tyler, 2015). What we have demonstrated here is that one crucial consequence of this interaction is an immediate change in the relationship between concepts, such that when seeing a removal man lifting a piano, we do indeed think of pianos as much more similar to boulders than to violins.

Acknowledgments

We would like to thank Kerry Bell for assistance with stimulus preparation and data collection, and Silvia Gennari, Eiling Yee, Gareth Gaskell, and two anonymous reviewers for helpful comments on the article. The work was supported by grants from the Wellcome Trust (076702/Z/05/Z) and ESRC (RES-063-27-0138) awarded to GTMA.

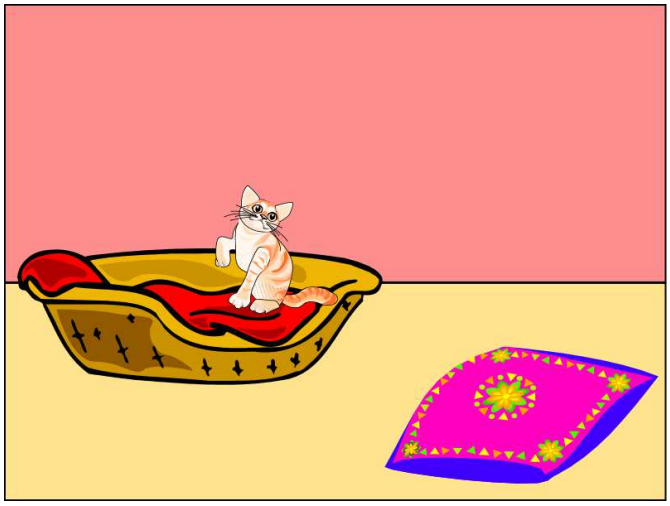
Appendix A

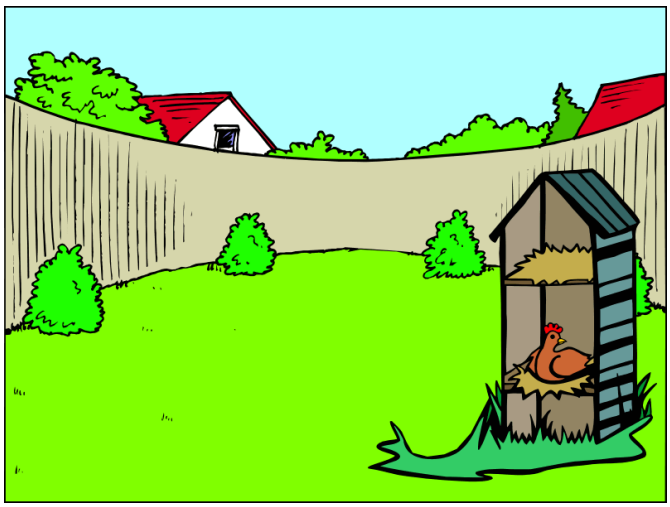
Experimental word stimuli

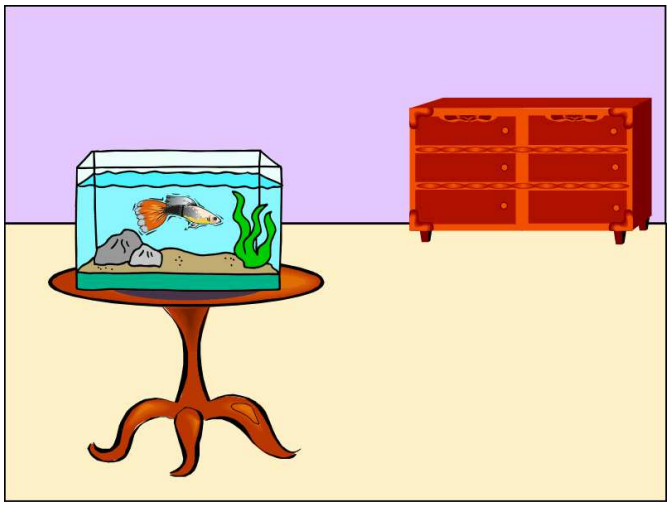
Visual target	*Semantically and contextually related word*	*Semantically but not contextually related word*	*Unrelated word*
cat	porcupine	buffalo	broccoli
chicken	pigeon	ostrich	hatchet
dog	squirrel	ox	blouse
guppy	terrapin	dolphin	saxophone
canary	robin	stork	tent
ant	caterpillar	butterfly	typewriter
car	bike	train	squid
fox	raccoon	bison	cheese
fork	scissors	rake	camel
gloves	scarf	boots	flute
horse	cow	elephant	jacket
knife	gun	rifle	eagle
mixer	toaster	dishwasher	cathedral
motorcycle	skateboard	lorry	biscuit
olive	walnut	aubergine	pillow
pickle	grape	apple	door
rabbit	hamster	pony	anchor
raspberry	peas	pumpkin	yacht
shirt	sweater	coat	goose
snake	worm	crocodile	hammer
trumpet	violin	piano	envelope

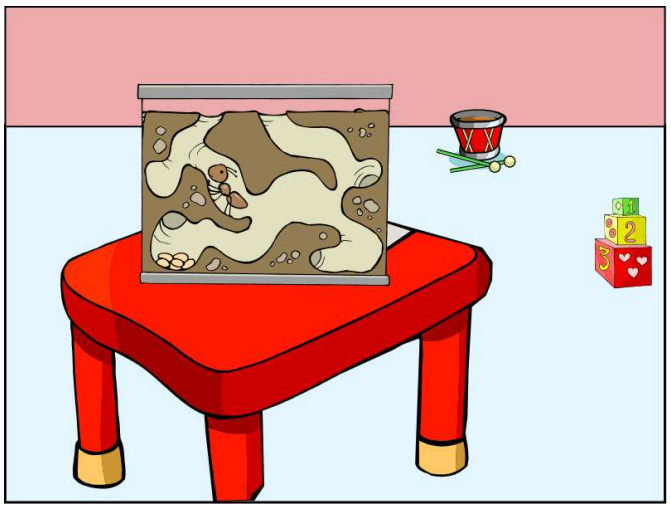
Appendix B

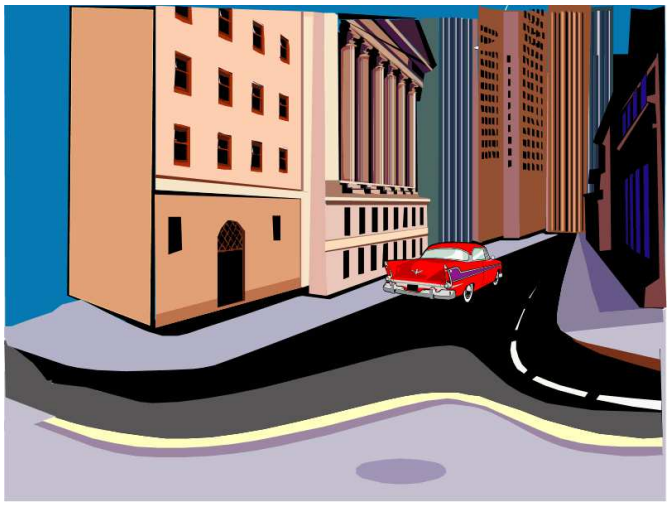
1. Experimental images used in Experiment 1

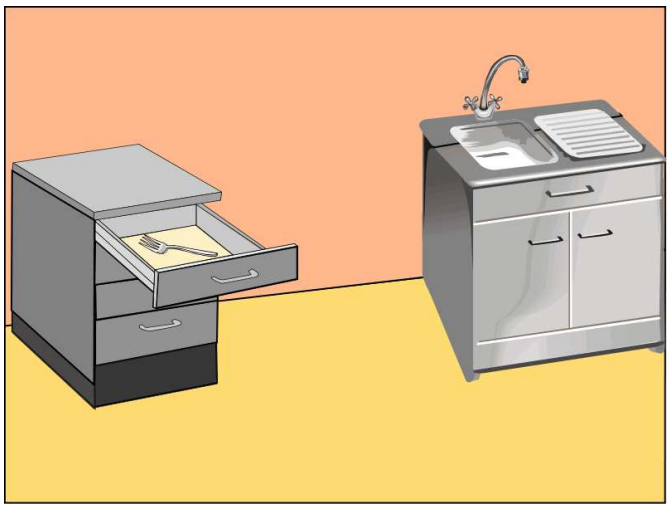


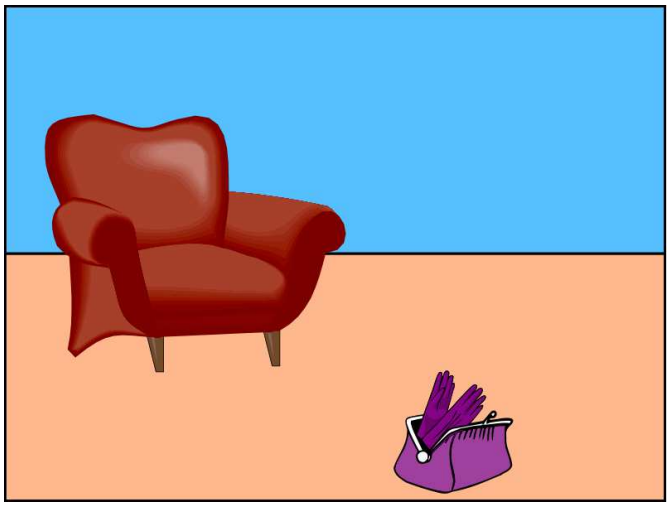


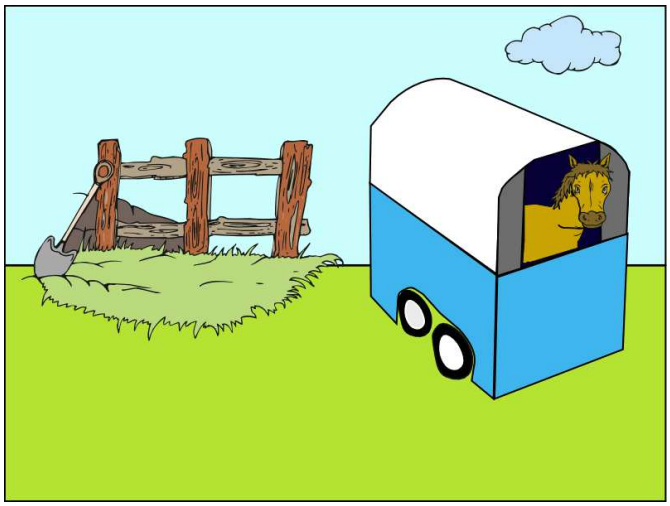


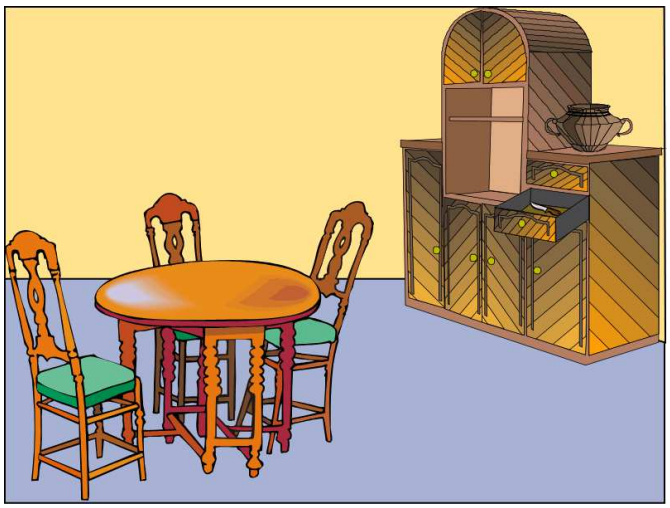


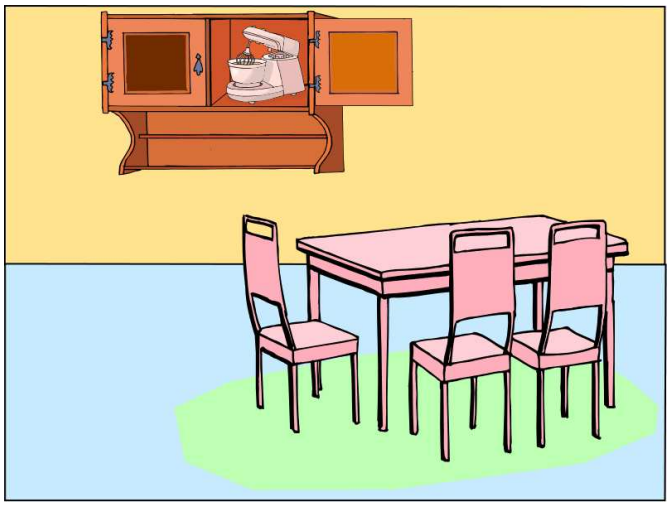


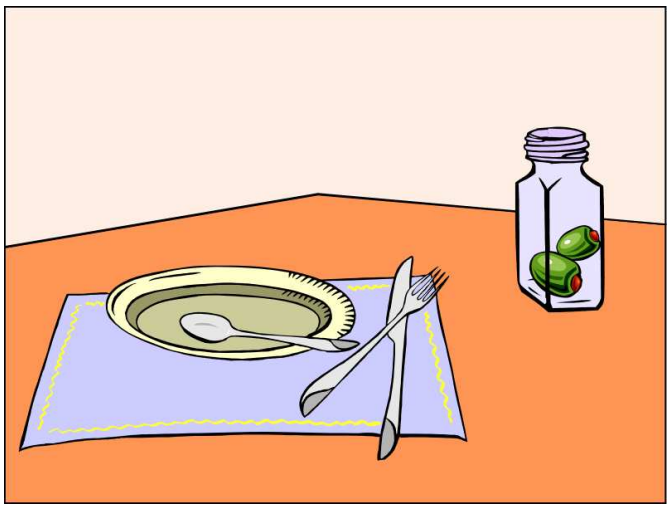


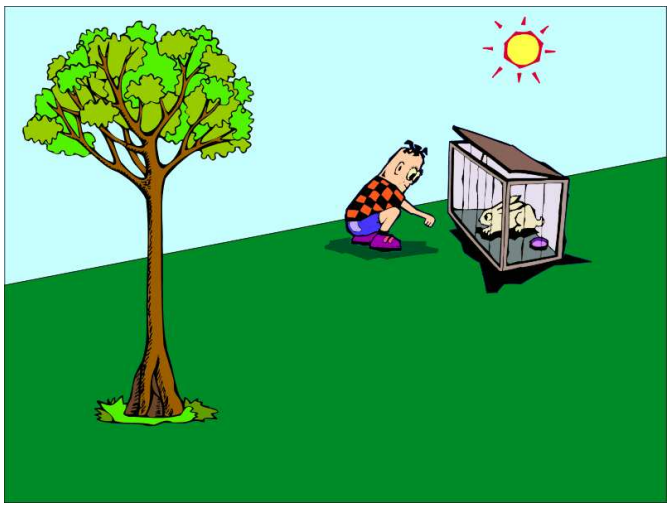


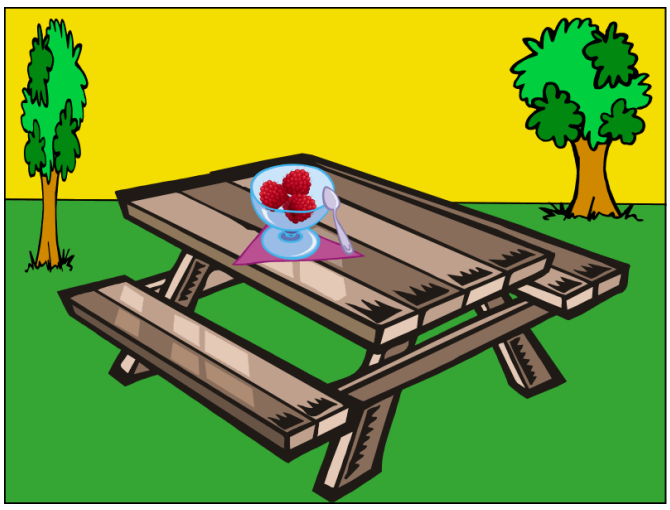


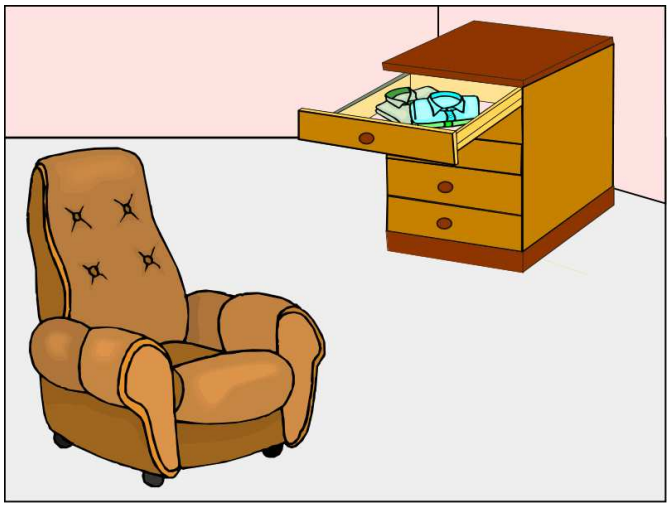


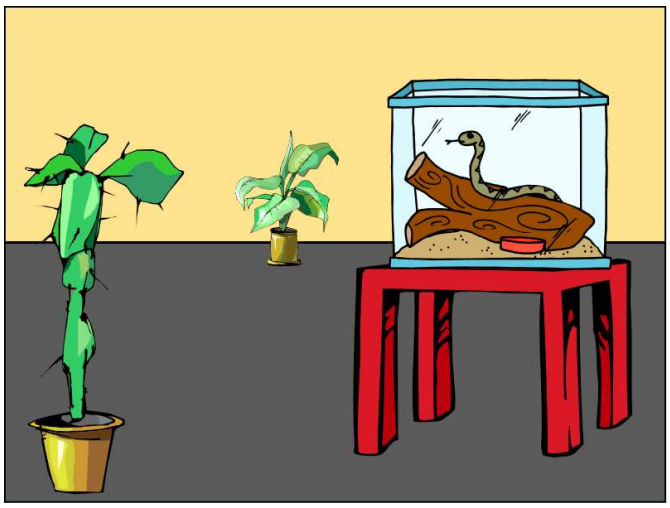






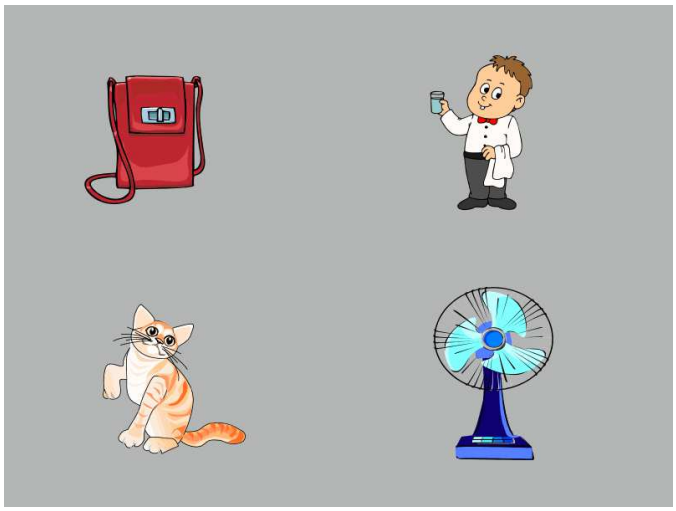


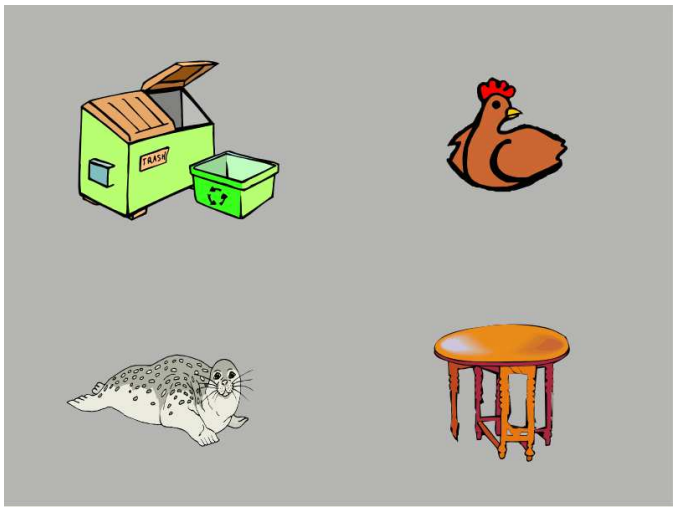


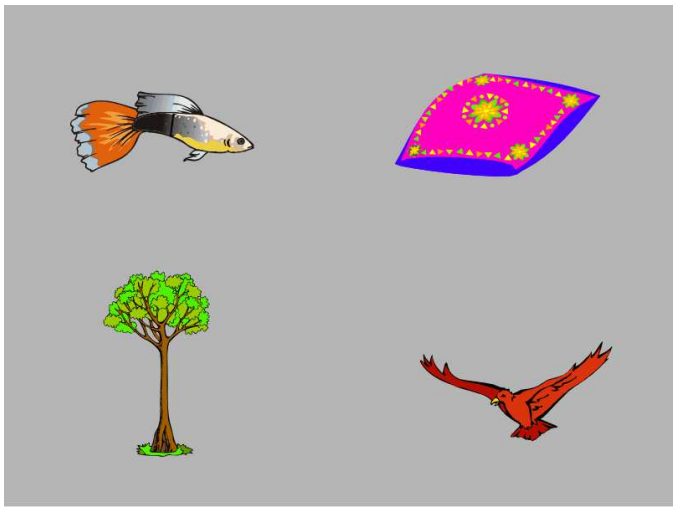


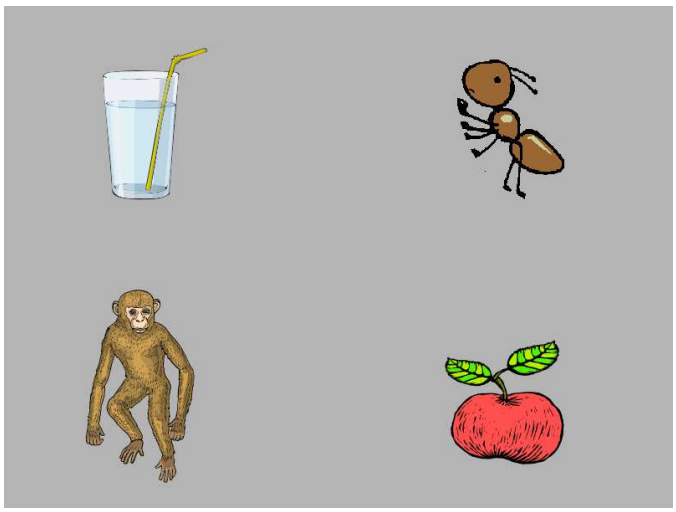

STOP

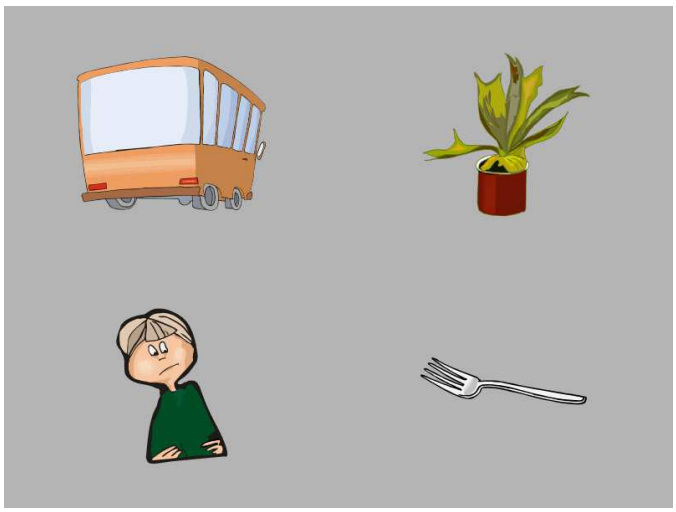
2. Experimental images used in Experiment 2

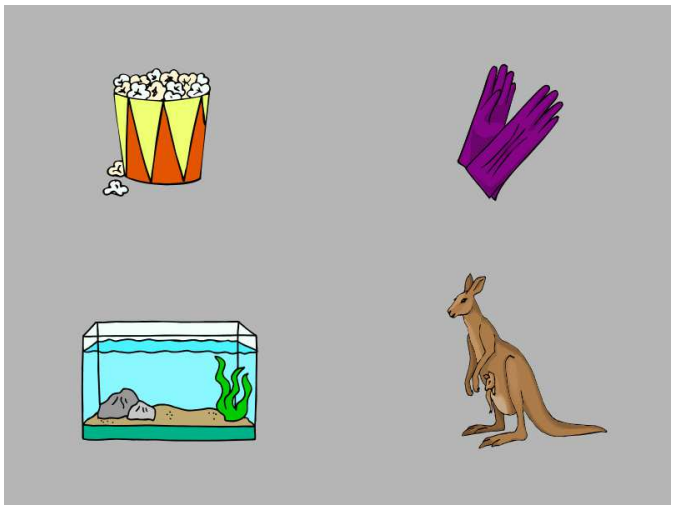


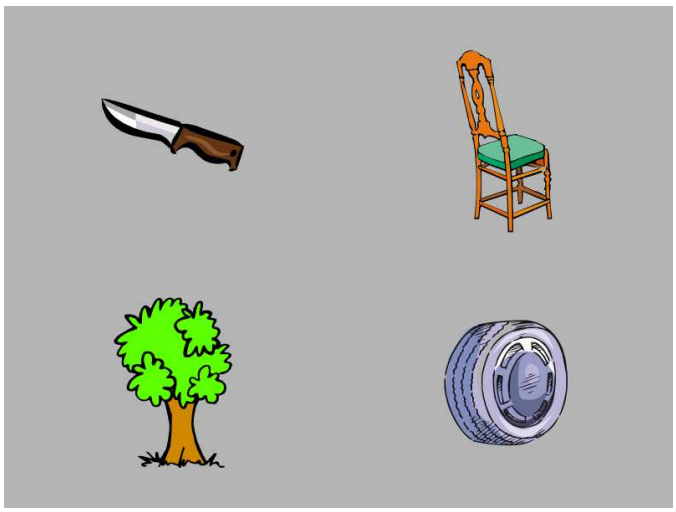


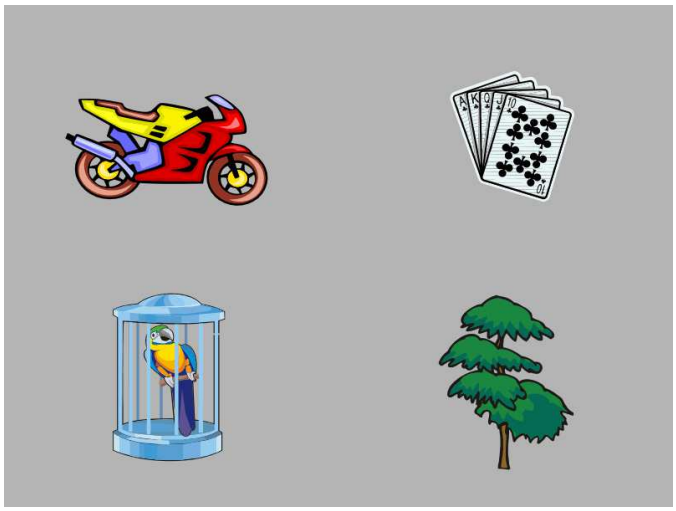


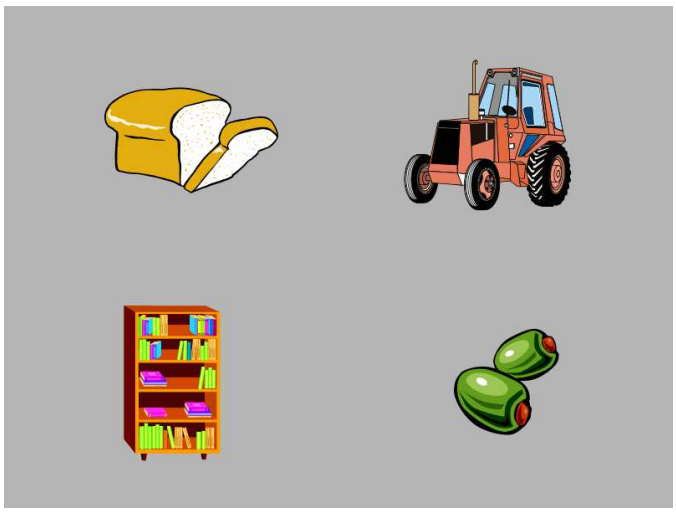


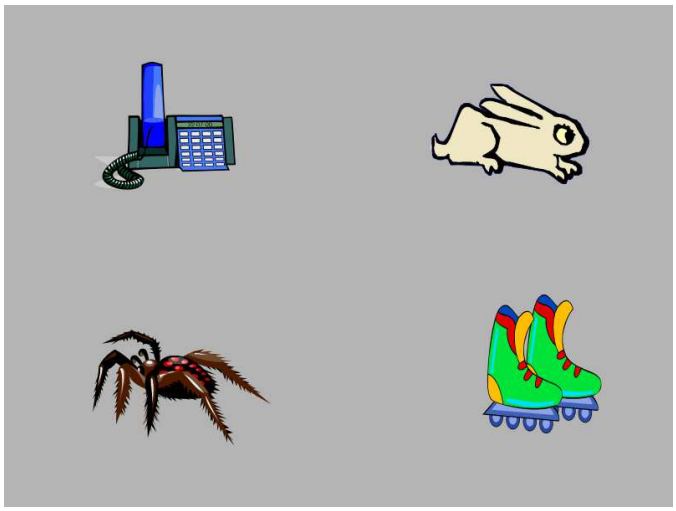


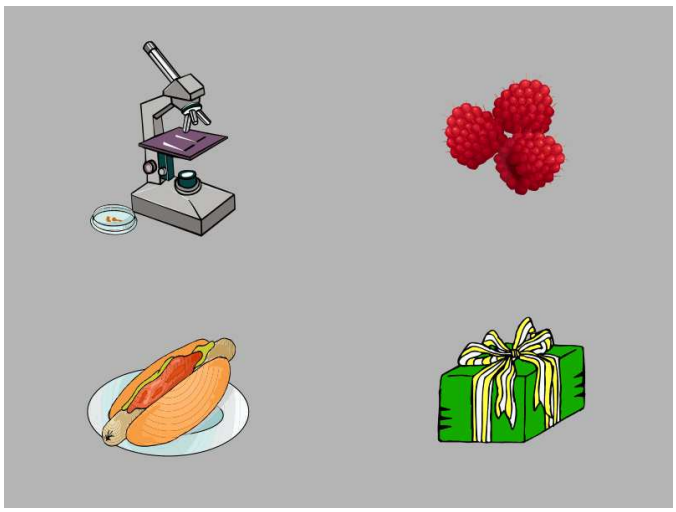


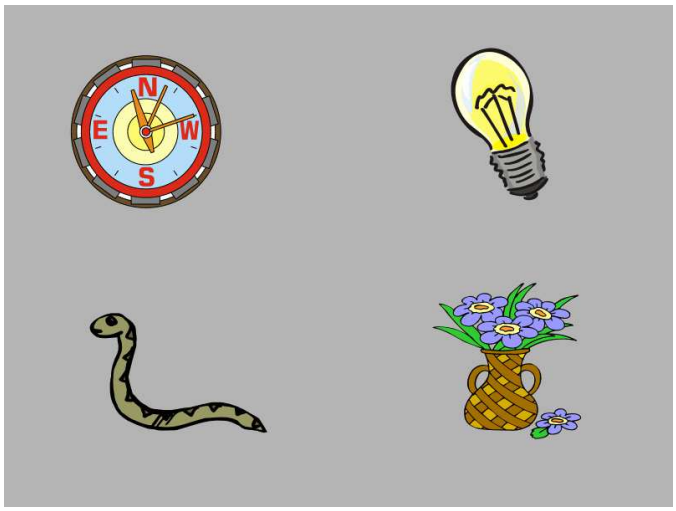











N
E
W
S

Appendix C.

Supplementary Material

Research data are provided on the Open Science Framework (Mirković & Altmann, 2018): https://osf.io/3hjsz/

References

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

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

Altmann, G.T.M., & Kamide, Y. (2007). 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, 57*, 502-518.

Altmann, G.T.M., & Mirković, J. (2009). Incrementality and prediction in human sentence processing. *Cognitive Science, 33*, 583-609.

Bar, M. (2004). Visual objects in context. *Nature Reviews Neuroscience, 5*, 617-629.

Bar, M. (2007). The proactive brain: using analogies and associations to generate predictions. *Trends in Cognitive Sciences, 11*, 280-289.

Barclay, J. R., Bransford, J. D., Franks, J. J., McCarrel, N. S., & Nitsch, K. (1974). Comprehension and semantic flexibility. *Journal of Verbal Learning and Verbal Behavior, 13*, 471-481.

Barr, D. J. (2008). Analyzing 'visual world' eyetracking data using multilevel logistic regression. *Journal of Memory and Language, 59*, 457-474.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68*, 255-278.

Barsalou, L. W. (1982). Context-independent and context-dependent information in concepts. *Memory & Cognition, 10*, 82-93.

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences, 22*, 577-660.

Bates, D., Machler, M., Bolker, B. M., & Walker, S. C. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software, 67*, 1-48.

Brysbaert, M., & Stevens, M. (2018). Power analysis and effect size in mixed effects models: a tutorial. *Journal of Cognition, 1*, 1-20.

Chen, Q., & Mirman, D. (2015). Interaction between phonological and semantic representations: Time matters. *Cognitive Science, 39*, 538-558.

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences, 36*, 181-204.

Clarke, A., Taylor, K. I., & Tyler, L. K. (2011). The evolution of meaning: Spatio-temporal dynamics of visual object recognition. *Journal of Cognitive Neuroscience, 23*, 1887-1899.

Clarke, A., & Tyler, L. K. (2015). Understanding what we see: How we derive meaning from vision. *Trends in Cognitive Sciences, 19*, 677-687.

Cree, G. S., & McRae, K. (2003). Analyzing the factors underlying the structure and computation of the meaning of chipmunk, cherry, chisel, cheese, and cello (and many other such concrete nouns). *Journal of Experimental Psychology: General, 132*, 163-201.

Çukur, T., Nishimoto, S., Huth, A. G., & Gallant, J. L. (2013). Attention during natural vision warps semantic representation across the human brain. *Nature Neuroscience, 16*, 763-770.

Dahan, D., & Tanenhaus, M. K. (2005). Looking at the rope when looking for the snake: conceptually mediated eye movements during spoken-word recognition. *Psychonomic Bulletin and Review, 12*, 453-459.

Dikker, S., & Pylkkanen, L. (2013). Predicting language: MEG evidence for lexical preactivation. *Brain and Language, 127*, 55-64.

Donnelly, S., & Verkuilen, J. (2017). Empirical logit analysis is not logistic regression. *Journal of Memory and Language, 94*, 28-42.

Edmiston, P., & Lupyan, G. (2015). What makes words special? Words as unmotivated cues. *Cognition, 143*, 93-100.

Edmiston, P., & Lupyan, G. (2017). Visual interference disrupts visual knowledge. *Journal of Memory and Language, 92*, 281-292.

Elman, J. L. (2004). An alternative view of the mental lexicon. *Trends in Cognitive Sciences, 8*, 301-306.

Glenberg, A.M. (1997). What memory is for. *Behavioral and Brain Sciences, 20*, 1-19.

Hirschfeld, G., Zwitserlood, P., & Dobel, C. (2011). Effects of language comprehension on visual processing - MEG dissociates early perceptual and late N400 effects. *Brain Lang, 116*, 91-96.

Huettig, F., & Altmann, G. T. (2005). Word meaning and the control of eye fixation: semantic competitor effects and the visual world paradigm. *Cognition, 96*, B23-32.

Huettig, F., Quinlan, P. T., McDonald, S. A., & Altmann, G. T. (2006). Models of high-dimensional semantic space predict language-mediated eye movements in the visual world. *Acta Psycholgica, 121*, 65-80.

Kalénine, S., Mirman, D., Middleton, E. L., & Buxbaum, L. J. (2012). Temporal dynamics of activation of thematic and functional knowledge during conceptual processing of manipulable artifacts. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 38*, 1274-1295.

Kiefer, M., & Pulvermuller, F. (2012). Conceptual representations in mind and brain: theoretical developments, current evidence and future directions. *Cortex, 48*, 805-825.

Kuperberg, G. R., & Jaeger, T. F. (2016). What do we mean by prediction in language comprehension? *Language, Cognition, and Neuroscience, 31*, 32-59.

Lambon Ralph, M. L. A., Jefferies, E., Patterson, K., & Rogers, T. T. (2017). The neural and computational bases of semantic cognition. *Nature Reviews Neuroscience*.

Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review, 104*, 211-240.

Lee, C. , Middleton, E. L., Mirman, D., Kalénine, S., & Buxbaum, L. J. (2013). Incidental and context-responsive activation of structure-and function-based action features during object identification. *Journal of Experimental Psychology: Human Perception and Performance, 39*, 257.

Lupyan, G., & Thompson-Schill, S. L. (2012). The evocative power of words: activation of concepts by verbal and nonverbal means. *Journal of Experimental Psychology: General, 141*, 170-186.

McRae, K., Cree, G. S., Seidenberg, M. S., & McNorgan, C. (2005). Semantic feature production norms for a large set of living and nonliving things. *Behavioral Research Methods, 37*, 547-559.

Mirković, J., & Altmann, G. T. M. (2018). Unfolding meaning in context: The dynamics of conceptual similarity. *Retrieved from osf.io/3hjsz*.

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

Moss, H. E., McCormick, S. F., & Tyler, L. K. (1997). The time course of activation of semantic information during spoken word recognition. *Language and Cognitive Processes, 12*, 695-731.

Oliva, A., & Torralba, A. (2007). The role of context in object recognition. *Trends in Cognitive Sciences, 11*, 520-527.

Paivio, A. (1991). Dual coding theory: Retrospect and current status. *Canadian Journal of Psychology-Revue Canadienne De Psychologie, 45*, 255-287.

Rogers, T. T., & McClelland, J. L. (2004). *Semantic cognition: A parallel distributed processing approach*. Cambridge, MA: MIT Press.

Rogers, T. T., & Patterson, K. (2007). Object categorization: Reversals and explanations of the basic-level advantage. *Journal of Experimental Psychology: General, 136*, 451-469.

Smith, E. E., Shoben, E. J., & Rips, L. J. (1974). Structure and process in semantic memory: A featural model for semantic decisions. *Psychological Review, 81*, 214-241.

Tabor, W., Juliano, C., & Tanenhaus, M. K. (1997). Parsing in a dynamical system: An attractor-based account of the interaction of lexical and structural constraints in sentence processing. *Language and Cognitive Processes, 12*, 211-271.

Tabossi, P. (1988). Effects of context on the immediate interpretation of unambiguous nouns. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14*, 153-162.

Tanenhaus, M. K., Magnuson, J. S., Dahan, D., & Chambers, C. (2000). Eve movements and lexical access in spoken-language comprehension: Evaluating a linking hypothesis between fixations and linguistic processing. *Journal of Psycholinguistic Research, 29*, 557-580.

Taylor, K. I., Moss, H. E., & Tyler, L. K. (2007). The conceptual structure account: A cognitive model of semantic memory and its neural instantiation. In J. Hart & M. Kraut (Eds.), *Neural basis of semantic memory* (pp. 265-301). Cambridge, UK: Cambridge University Press.

Trapp, S., & Bar, M. (2015). Prediction, context, and competition in visual recognition. *Annals of the New York Acadedmy of Sciences, 1339*, 190-198.

Yee, E., Huffstetler, S., & Thompson-Schill, S. L. (2011). Function follows form: Activation of shape and function features during object identification. *Journal of Experimental Psychology: General, 140*, 348-363.

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