

Variability driven animacy effects: evidence of structural, not conceptual differences in processing animates and inanimates

Vanja Ković

Department of Psychology, University of Novi Sad

Kim Plunkett

Department of Experimental Psychology, University of Oxford

Gert Westermann

Department of Psychology, Oxford Brookes University

The present eye-tracking study demonstrates that when animate and inanimate object pictures are presented within a single-study, there are no systematic differences between processing these two categories objects. Although participants were taking less time to initiate their first gaze towards animate than to inanimate objects, a result compatible with findings of Proverbio et al. (2007), it turned out that this quicker initiation of the first look in animates was driven by mammals and reptiles only and did not apply to insects or aquatic animals, most probably due to the structural differences within these subcategories. Fixations in this study do not cluster around certain features or areas of the objects for either animate or inanimate categories. Moreover, detailed analysis of looking behaviour does not reveal a clear animate-inanimate distinction.

Thus, given the failure of finding systematic differences between animates and inanimates when assessed using various looking behaviour measurements, the results do not support the prediction from modality specific conceptual account. In fact, these results are more in agreement with an alternative, distributed account of semantic representation that explains processing differences by structural differences between animate and inanimate objects.

Key words: animate, inanimate objects, eye-tracking, mental representations

The eye-tracking research presented here was largely inspired by the work done using the visual-world paradigm in examining allocation of participants' attention to a set of objects on a visual scene, while hearing a description of the scene (Kamide et. al 2003; Huettig & Altmann, 2005; Dahan & Tanenhaus,

2005; Altman & Kamide, 2007). These studies have demonstrated how unfolding language, but also the unfolding mental world, can guide participants' attention towards certain parts of a visual scene. Following on from this and under the assumption that object recognition typically involves matching mental representations of objects stored in memory to representations extracted from visual images (Mozer, 2002), the eye-tracking methodology was considered as a valuable approach in investigating what visual features people attend to in the early stages of object recognition and if language can modify looking behaviour towards single objects. Labelling of the objects prior to visual presentation was expected to evoke a mental representation of a particular object and to allow closer examination of which features or parts of the objects extracted from the visual image form the basis of mental representation.

One of the two major theoretical accounts regarding processing of animate and inanimate objects proposed by Warrington & McCarthy (1987) and Warrington & Shallice (1984) suggested that animate objects are more easily recognised and described by visual features whereas inanimates rely more on functional features. Thus, as a result of these discrepancies we see a feature-based segregation, or modular organisation of conceptual representations of different types of semantic knowledge at the brain level.

The differences between processing of the animate and inanimate objects described by Ković et al. (2009a) and Ković et al. (2009b) seem to suggest employment of different visual processing strategies for animate and inanimate objects. Based on the results of these two studies, it could be argued that the animate objects were processed in a similar way because they have more salient visual features in comparison to the inanimates, as suggested by feature-based account of memory organisation. From the same perspective, inanimates might rely more on functional features that are not directly present in the picture of an object (i.e., chair – sitting, apple – eating, piano – playing) and this is why we see inconsistencies in visual processing of these objects.

On the other hand, it is plausible that participants, when exposed to those pictures, were in the specific semantic context and thus they exhibit strategic looking behaviour in the context of the task they were given. Also, the majority of the animals in the Ković et al. (2009a) belonged to the same semantic category, namely – mammals, whereas, the inter-group variation within the inanimates was much bigger, that is, inanimates belonged to variety of categories such as fruit, furniture, vehicles etc (Ković et al., 2009b). In order to control for these two factors and find an explanation for the previously observed differences between animates and inanimate the current study containing both animates and inanimates was run.

Furthermore, in order for the labelling effect to be more effective, an inter-stimulus-interval of 500ms was introduced between the offset of the auditory stimuli and the onset of the visual stimuli to allow more time for the mental representation of the objects to be evoked before displaying the visual stimuli.

According to the feature-based theory of semantic organisation (McPherson & Holcomb, 1999; Sitnikova et al., 2006; Warrington & McCarthy, 1987; Warrington & Shallice, 1984; West & Holcomb, 2002), participants in this study are expected to process animate objects in a more consistent manner in comparison to inanimates (as reported in Ković et al. (2009a) and Ković et al. (2009b)). That is to say, in the mixed-design study where their looking behaviour cannot be driven by a strategic response in any of the two categories the same differences should be found. Also, participants are expected to be quicker at initiating eye-movements to animate than to inanimate pictures (Proverbio et al., 2007). Finally, given that the participants were given a longer inter-stimulus-interval, that is, enough time to evoke mental representation prior to the visual presentation of the objects, participants were expected to demonstrate a less diffuse pattern of eye-fixations in the naming in comparison to the non-naming condition.

However, if in the mixed design study we would no longer see these differences in visual processing of animate and inanimate objects we could conclude that the differences reported previously were driven by strategic looking behaviour in the context of the task they were given as well as lack of a good control of the category item variability. Such a result would support an alternative account, namely a distributed, unitary account which suggests that all semantic information is processed within a unitary neural system; Tyler et al. (2000).

Method

Participants. Twenty-four healthy, normal, right-handed participants took part in the study. They were all first year Oxford University undergraduate students, native speakers of English, with normal hearing and normal or corrected to normal vision and they were all given course credits for their participation. None of the participants were excluded from the study.

Stimuli. While the majority of the animate objects in Ković et al. (2009a) were mammals, there was much more within-category variability in both the animate and inanimate categories in the current experiment. In order to have better control over within category variability in this study, four sub-categories were selected for both the animate and inanimate categories and five items were then selected within each of the sub-categories (see Table 1).

Table 1. Animate and inanimate objects – categories and items

Animate objects	Mammals:	dog	elephant	giraffe	mouse	rabbit
	Reptiles:	alligator	crocodile	lizard	snake	turtle
	Insects:	ant	bee	butterfly	fly	mosquito
	Aquatic animals:	clam	crab	fish	octopus	shark
Inanimate objects	Food:	apple	bread	cake	cheese	pizza
	Furniture:	bed	chair	lamp	table	TV
	Vehicles:	bike	boat	bus	car	plane
	Clothes:	coat	hat	shoe	sock	tie

Visual stimuli: All of the visual stimuli were photographs of real animate and inanimate objects. The majority of the pictures were chosen from the CD-ROM Graphic Interchange Format Data (Hemera 2000) and some of them were chosen from commercial internet pages and edited using the Adobe Photoshop CS software. For each of the pictures the background was removed and the 10% of a grey background was introduced to reduce brightness on the screen. Similar to Ković et al. (2009a), for each of the chosen animate/inanimate object label three versions of the corresponding static images were chosen, so that the whole sample consisted of 120 (40x3) images in total. All of the pictures were of the same size, 400x400 pixels, and were presented to the participants in the left profile view using the Presentation software.

Auditory stimuli: The forty selected labels (see Table 1) were recorded in stereo within the carrier phrase: ‘Look at the <target>’ at 44.11 kHz sampling rate into signed 16-bit files. The other two, non-naming phrases for the non-naming conditions (‘Look at the picture’ and ‘What’s this?’) were recorded on the same session. These two conditions were considered as control conditions – one of which being neutral and the other more exploratory. All of the stimuli were further edited to remove background noise, head and tail clicks and to match for peak-to-peak amplitude by using the GoldWave 5.10 software.

Experimental design. The experiment consisted of six experimental conditions that is, the two animacy conditions (animate and inanimate objects) and three auditory conditions within each of the animacy conditions (‘Look at the <target>!’, ‘Look at the picture!’ and ‘What’s this?’). There were 120 trials in total (20 per condition). A typical trial involved presentation of the fixation cross for 2000ms, during which either a sentence containing the name of the animate or inanimate objects (i.e. ‘Look at the <dog>’) or non-naming sentence (‘Look at the picture!’ or ‘What’s this?’) was uttered. There was an inter-stimulus-interval of 500ms between the offset of the auditory stimuli and the onset of the visual stimuli (see Figure 1).

The ISI was introduced in order to give participants a bit more time to evoke the mental representation of the object in the naming condition before the object was presented on the screen. The visual stimuli were presented at the offset of the fixation cross and remained on the screen for 2000ms promptly.

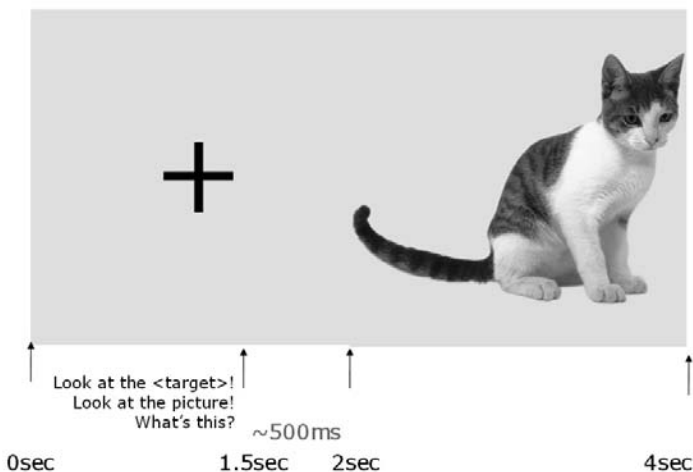


Figure 1. The time course of the stimuli presentation

In this study all of the pictures were presented in the left-profile only and the presentation of the auditory conditions and animacy conditions was counterbalanced across participants using Latin Square order. The presentation order of stimuli was randomised for each subject (see Figure 2).







<p>Cond. I: (Look at the 'target!')</p>	+		+		...
<p>Cond. II: (Look at the picture!)</p>	+		+		...
<p>Cond. III: (What's this?)</p>	+		+		...

Figure 2. The three experimental conditions

Procedure. Participants were seated in a darkened room approximately a metre away from the monitor displaying centrally presented visual stimuli ($\sim 6^\circ$ of visual angle). In the brief instruction at the beginning of the study participants were instructed to focus at the fixation cross when presented on the screen and look freely when the visual stimuli were displayed, as well as to pay attention to the auditory stimuli presented to them through the loudspeakers.

The experiment started once the participant settled down and found the most comfortable position. In order to make the procedure more natural for participants, the chin-rest was not used in the current study. Participants were asked to sit as still as possible and the option for automatically correcting for small head movements was activated as soon as the calibration procedure was completed successfully that is, as soon as the automatic recording of participants' eye-movements started. The rest of the procedure was exactly the same as described in Ković et al., (2009a).

Apparatus. The eye-tracking methodology and procedure were the same as described in Ković et al., (2009a), except that the option for automatic on-line adjusting and correcting for small head movements was activated, given that in this study the chin-rest was not used.

Measurements. All of the eye-tracking measurements for assessing participants' looking behaviour were the same as described in Ković et al., (2009a) eye-tracking study.

Results

Analysis of the first look. A 3x2 ANOVA with factors: Auditory condition ('Look at the picture!', 'Look at the <target>!' and 'What's this?') and Animacy (Animate, Inanimate) revealed a significant effect of Animacy ($F(1,342)=22.76$,

$p < .001$), but not of the Auditory ($F(2,342)=0.78$, $p=0.46$) condition regarding initiation of the first look. The interaction effect was not significant.

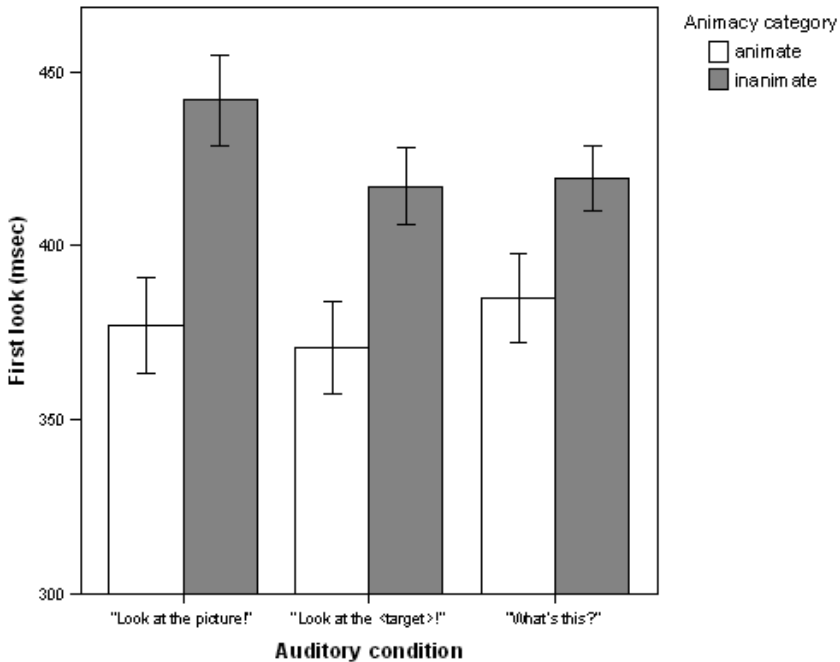


Figure 3. Average initiation of the first look times: animates vs. inanimates

Planned comparisons showed a significant difference between animate ($M=376.98\text{ms}$, $s.e.m.=13.82$) and inanimate ($M=441.89\text{ms}$, $s.e.m.=13.09$) object processing in the 'Look at the picture!' condition ($t(1,118)=3.41$, $p < .001$). Similarly, in the 'Look at the <target>!' condition participants took less time to initiate eye-gaze towards animate ($M=370.58\text{ms}$, $s.e.m.=13.28$) than towards inanimate objects ($M=417.07\text{ms}$, $s.e.m.=11.02$), $t(1,188)=2.69$, $p < .008$. A similar pattern of results was observed in the 'What's this?' condition: $M(\text{animate})=384.86\text{ms}$ ($s.e.m.=12.78$), $M(\text{inanimate})=419.38\text{ms}$ ($s.e.m.=9.53$), $t(1,188)=2.16$, $p < .032$, see Figure 3.

Detailed analysis of the initiation of the first saccade across the sub-categories revealed that the difference in processing between animate and inanimate objects may be driven by the Mammals and Reptiles categories which were processed faster than the other two animate and four inanimate categories. The mean initiation of the first look for the Mammal category was $M=328.77\text{ms}$ ($s.e.m.=8.36$) and for the Reptiles: $M=349.63\text{ms}$ ($s.e.m.=10.54$).

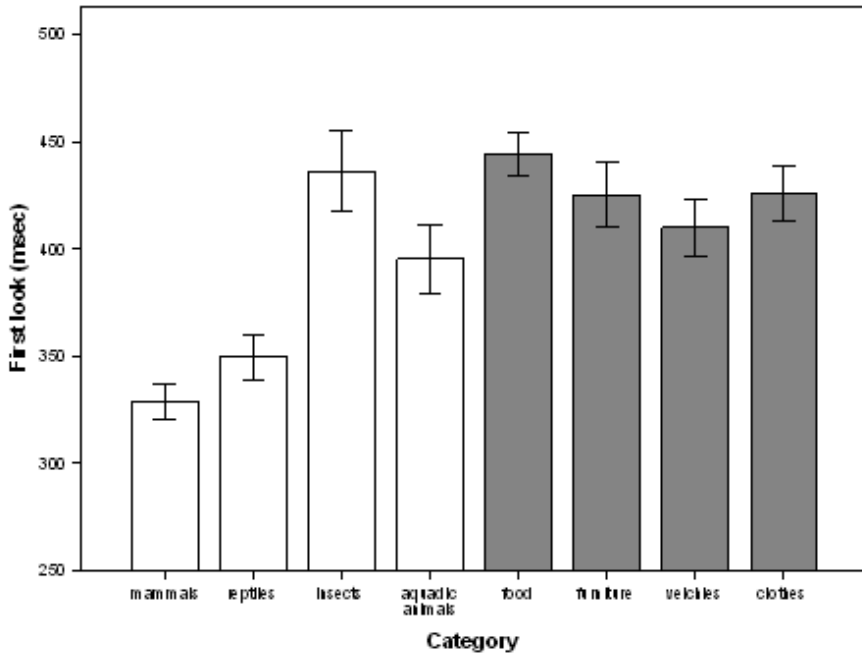


Figure 4. Average initiation of the first look across the categories

For the other six categories the initiation of the first look lasted: Insects $M=436.24\text{ms}$ ($s.e.m.=18.92$), Aquatic animals $M=395.23\text{ms}$ ($s.e.m.=15.99$), Food $M=444.11\text{ms}$ ($s.e.m.=10.39$), Furniture $M=425.17\text{ms}$ ($s.e.m.=15.13$), Vehicles $M=409.77\text{ms}$ ($s.e.m.=12.87$) and Clothes $M=425.42\text{ms}$ ($s.e.m.=5.18$). The initiation of the first look did not differ significantly between Mammals and Reptiles, but Mammals differed significantly from the other six categories (Insects, Aquatic animals, Food, Furniture, Vehicles and Clothes). Reptiles also differed significantly from Insects, Food, Furniture and Clothes sub-categories. All of the differences were significant at the $p<.05$ level and Bonferroni corrections were applied to account for multiple comparisons. There were no other significant differences (see Figure 4).

Analysis of the longest look. The results of 3x2 ANOVA with factors Auditory condition and Animacy showed only a significant effect of Animacy ($F(1,342)=10.31$, $p<.001$), but not of the Auditory ($F(2,342)=0.342$, $p=0.71$) condition. Interaction effect was not significant.

Planned comparisons revealed that the longest look measurement between animate ($M=1070.16\text{ms}$, $s.e.m.=30.43$) and inanimate ($M=937.17\text{ms}$, $s.e.m.=34.79$) objects was significantly different only in the naming ('Look at the <target>') condition ($t(1,118)=2.87$, $p<.005$), but there were no systematic

differences across the other two, non-naming conditions ('Look at the picture!': $M(\text{animate})=1033.88\text{ms}$, $s.e.m.=29.82$, $M(\text{inanimate})=944.19\text{ms}$, $s.e.m.=26.77$, $t(1,118)=0.99$, $p=0.324$; 'What's this?': $M(\text{animate})=1034.47\text{ms}$, $s.e.m.=32.04$, $M(\text{inanimate})=968.59\text{ms}$, $s.e.m.=32.34$, $t(1,118)=1.44$, $p=0.15$), see Figure 5. The post-hoc comparisons across the 8 groups of animate and inanimate objects regarding the longest look measurement revealed a lot of variation between sub-categories. Overall, Mammals, Insects, Aquatic animals and Clothes tended to attract longer looks than the other four categories Reptiles, Food, Furniture and Vehicles. Vehicles tended to attract the shortest looks among all sub-categories.

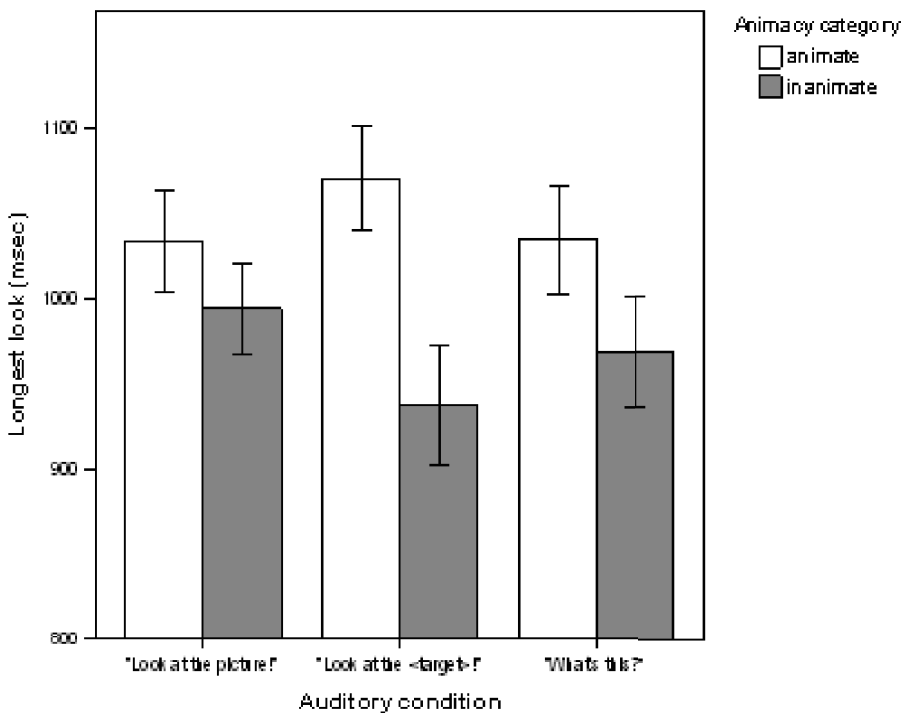


Figure 5. Average duration of longest look: animates vs. inanimates

The mean longest look for the inanimates was: Food: $M=959.64\text{ms}$ ($s.e.m.=33.94$), Furniture: $M=956.47\text{ms}$ ($s.e.m.=34.35$), Vehicles: $M=849.51\text{ms}$ ($s.e.m.=37.03$) and Clothes: $M=1100.99\text{ms}$ ($s.e.m.=30.61$) and for the animate groups it was: Mammals: $M=1049.54\text{ms}$ ($s.e.m.=39.63$), Reptiles: $M=958.49\text{ms}$ ($s.e.m.=34.72$) Insects: $M=1084.59\text{ms}$ ($s.e.m.=33.35$) and Aquatic animals: $M=1047.06\text{ms}$ ($s.e.m.=30.95$).

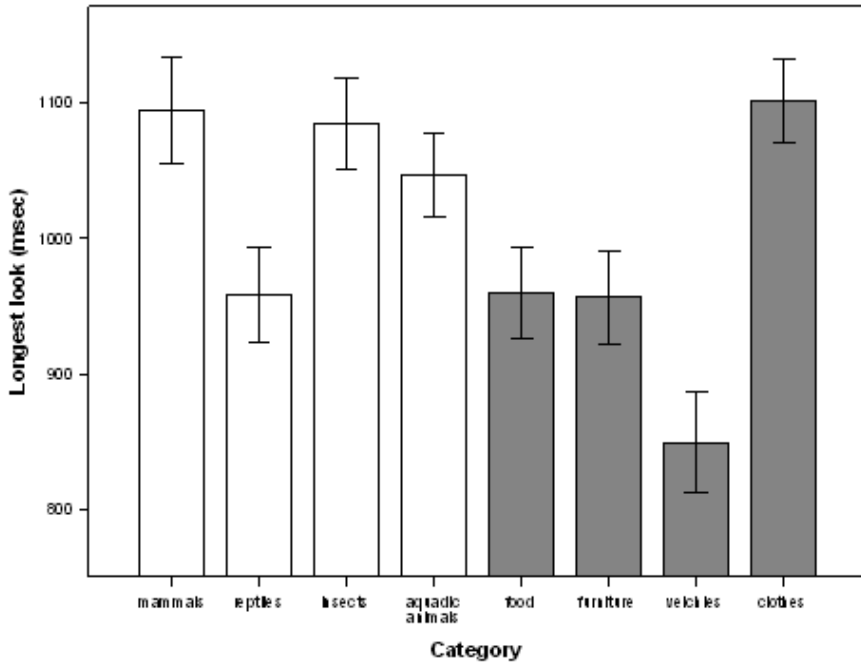


Figure 6. Average duration of longest looks across conditions

Taking into account Bonferroni corrections for multiple comparisons the analyses showed that only Vehicles differed significantly from Mammals, Insects Aquatic animals and Clothes (see Figure 6). All of these differences were significant at the $p < .05$ level.

Analysis of the total looking time. Regarding the total looking time measure, a 3x2 ANOVA with factors Auditory condition and Animacy revealed only a marginally significant effect of Animacy ($F(1,342)=3,81$, $p=0.052$), but no main effect of the Auditory ($F(2,342)=0.07$, $p=0.99$) condition. Interaction effect was not significant.

A detailed analysis across the three Auditory conditions revealed no significant differences in total looking time between animate and inanimate objects: ‘Look at the <target>!’: $M(\text{animate})=1574.99\text{ms}$, $s.e.m.=24.36$, $M(\text{inanimate})=1609.66\text{ms}$, $s.e.m.=20.06$, $t(1,118)=1.09$, $p=0.274$; ‘Look at the picture!’: $M(\text{animate})=1568.61\text{ms}$, $s.e.m.=24.23$, $M(\text{inanimate})=1614.04\text{ms}$, $s.e.m.=21.73$, $t(1,118)=1.39$, $p=0.165$; ‘What’s this?’: $M(\text{animate})=1576.60\text{ms}$, $s.e.m.=25.56$, $M(\text{inanimate})=1602.68\text{ms}$, $s.e.m.=19.68$, $t(1,118)=0.81$, $p=0.42$ (see Figure 7).

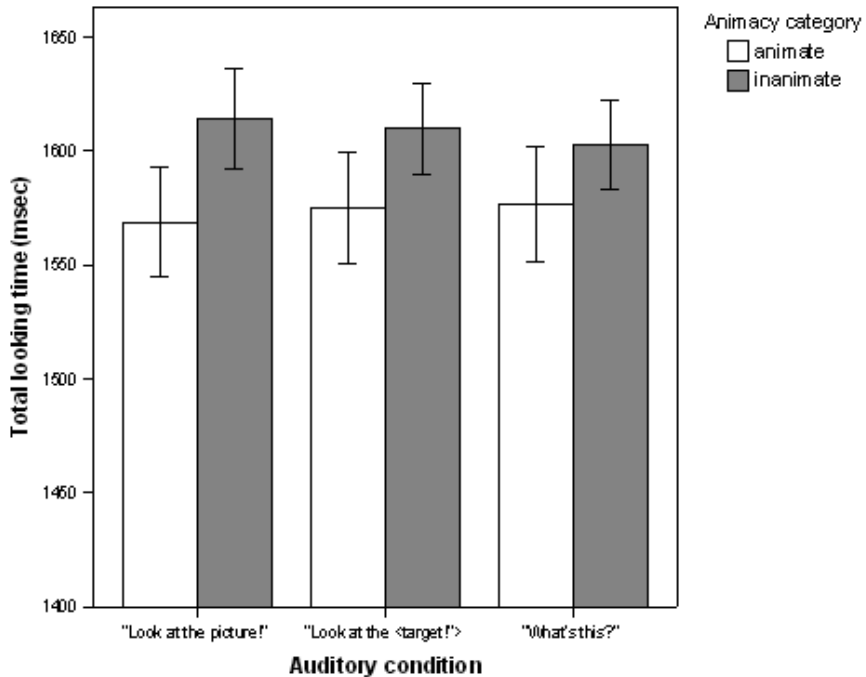


Figure 7. Average duration of total looking time: animates vs. inanimates

Regarding the total amount of time participants spent looking at the objects (TLT), post-hoc comparisons revealed variable patterns across the animate and inanimate sub-categories. Within the animate categories participants looked longer at the Insects than at the other three categories and spent the least amount of time looking at the Vehicles within the inanimates. The mean total looking times for the animate groups were: Mammals: $M=1049.54\text{ms}$ ($s.e.m.=39.63$), Reptiles: $M=958.49\text{ms}$ ($s.e.m.=34.72$), Insects: $M=1084.59\text{ms}$ ($s.e.m.=33.35$) and Aquatic animals: $M=1047.06\text{ms}$ ($s.e.m.=30.95$) and for the inanimates it was Food: $M=959.64\text{ms}$ ($s.e.m.=33.94$), Furniture: $M=956.47\text{ms}$ ($s.e.m.=34.35$), Vehicles: $M=849.51\text{ms}$ ($s.e.m.=37.03$) and Clothes: $M=1100.99\text{ms}$ ($s.e.m.=30.61$).

The detailed analysis showed statistically significant differences both within and between sub-categories. Considering within animate category variability there were significant differences between Insects and Mammals, Reptiles and Aquatic animals as well as between Aquatic animals and Reptiles regarding total looking time. Furthermore, TLT at Food differed from both Furniture and Vehicles and TLT at Furniture differed from Clothes within inanimate categories. Moreover, regarding between sub-categories variability, the analysis showed that

Mammals and Reptiles did not differ from Vehicles; Clothes did not differ from Insects and Aquatic animals, Furniture from Mammals and Aquatic animals and Food from Insects on average, see Figure 8. All the other between sub-categories comparisons were significant at the $p < .05$ level and Bonferroni corrections were applied for this analysis.

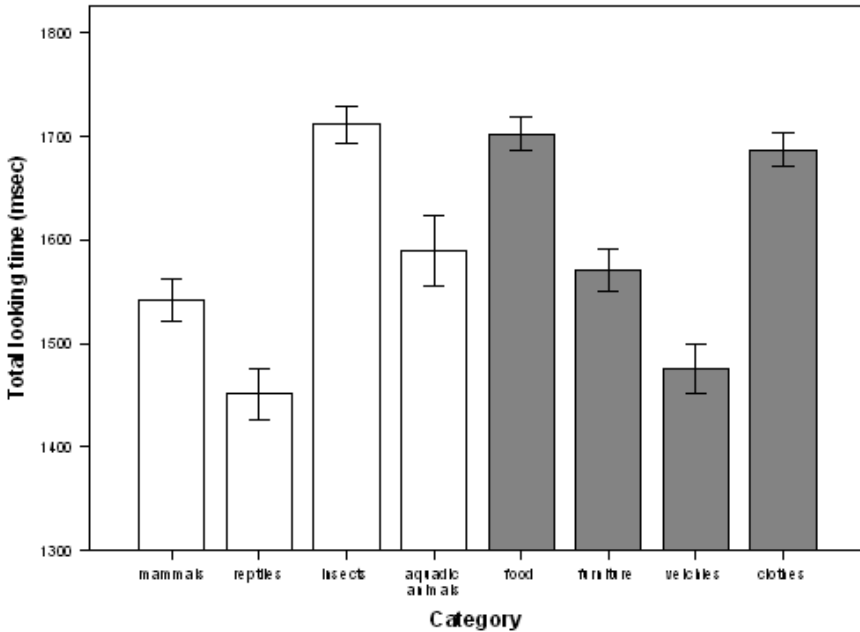


Figure 8. Average duration of total looking time across the conditions

Analysis of number of fixations. A 3x2 ANOVA with factors Auditory condition and Animacy revealed no significant effects of Auditory condition ($F(2,342)=1.35$, $p=0.26$) or Animacy ($F(1,342)=2.08$, $p=0.151$) regarding the number of fixations participants made. The interaction effect was found not to be significant.

Planned comparisons across the three auditory conditions revealed that participants made between six and seven fixations on average and there were no systematic differences between animate and inanimate object processing. The mean number of fixations in ‘Look at the <target>!’ condition was: $M(\text{animate})=6.67$, $s.e.m.=0.12$, $M(\text{inanimate})=6.61$, $s.e.m.=0.10$, $t(1,118)=0.39$, $p=0.693$; and in ‘Look at the picture!’ condition: $M(\text{animate})=6.68$, $s.e.m.=0.12$, $M(\text{inanimate})=6.47$, $s.e.m.=0.10$, $t(1,118)=1.31$, $p=0.186$; and in ‘What’s this?’ condition: $M(\text{animate})=6.81$, $s.e.m.=0.12$, $M(\text{inanimate})=6.81$, $s.e.m.=0.11$, $t(1,118)=0.71$, $p=0.48$), see Figure 9.

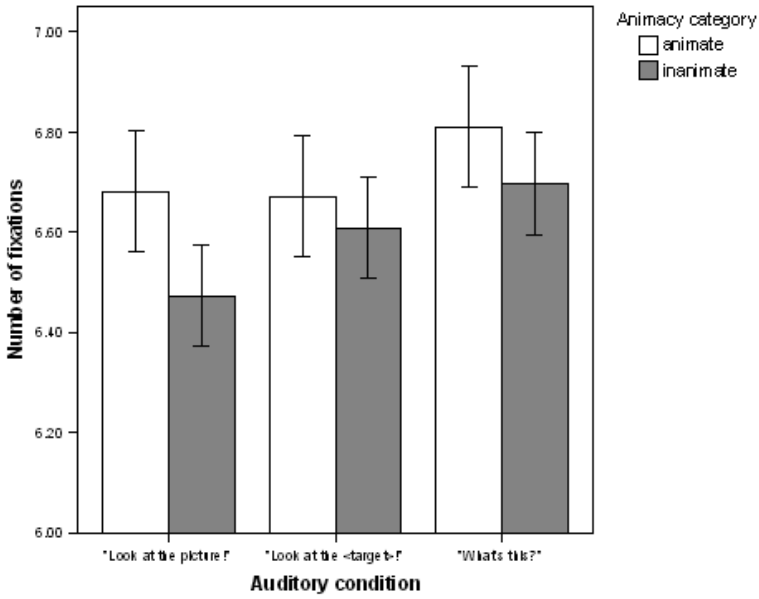


Figure 9. Average number of fixations: animates vs. inanimates

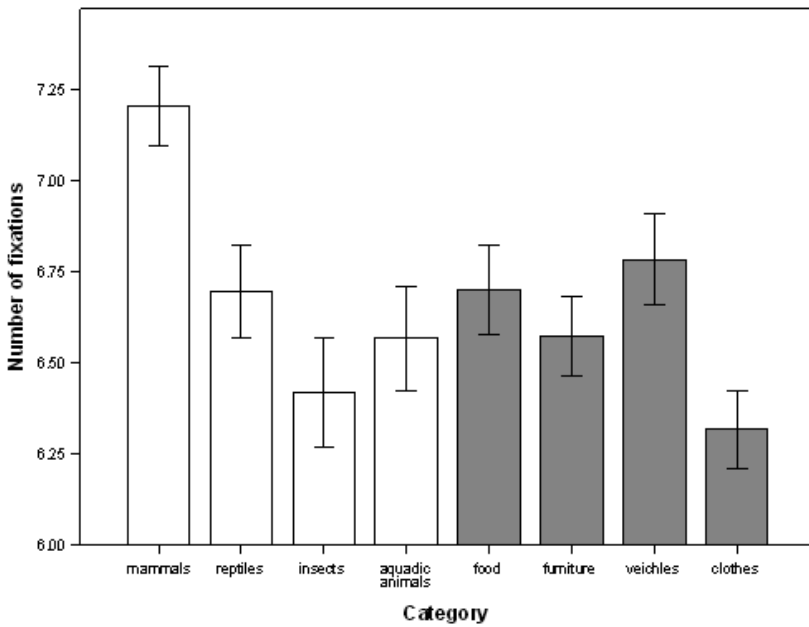


Figure 10. Average number of fixations across the categories

The detailed analysis of the number of fixations participants made revealed that Mammals received significantly more fixations than Insects, Aquatic animals, Furniture and Clothes. These differences, after applying Bonferroni corrections, were all significant at the $p < .05$ level.

The mean number of fixations for the Mammals was: $M=7.20$ (s.e.m.=0.11) and for the other three animate categories it was: Reptiles: $M=6.69$ (s.e.m.=0.13) Insects: $M=6.42$ (s.e.m.=0.15) and Aquatic animals: $M=6.57$ (s.e.m.=0.14), respectively. Inanimate subcategories on average received Food: $M=6.70$ (s.e.m.=0.12), Furniture: $M=6.57$ (s.e.m.=0.11), Vehicles: $M=6.78$ (s.e.m.=0.12) and Clothes: $M=6.31$ (s.e.m.=0.11) fixations (see Figure 10).

Cluster analysis. After extracting all fixations for all of the participants across the 3 experimental conditions, using Ward's method (Ward, 1963) and Clustan software (Wishart, 2004) as in the Ković et al. (2009a), a cluster analysis was performed in order to identify regions of interest where participants tended to focus their attention. Subsequent to cluster analysis, the clusters of fixations were plotted on top of the pictures and presented in different colour for easier interpretation.

Examples of the clusters across the three auditory conditions for one animate and one inanimate object are given in the Figure 11.

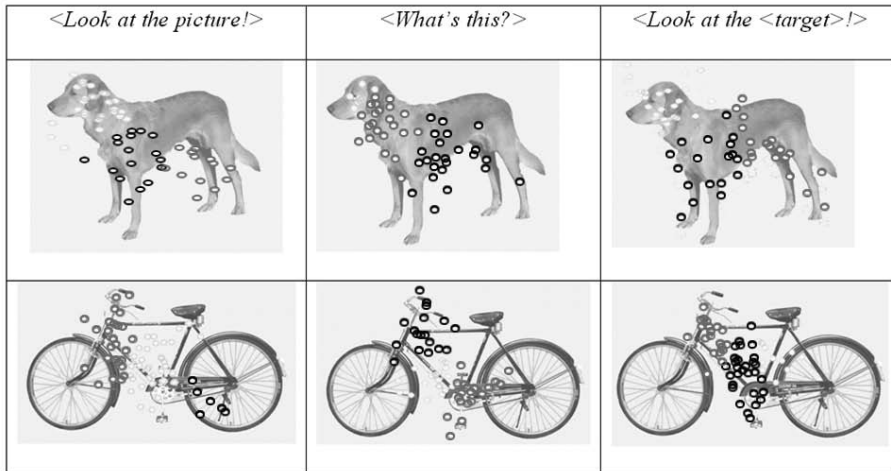


Figure 11. Plotting clusters of fixations on top of the images

The chosen examples demonstrate pictures where all of the fixations clustered in three clusters across all of the naming and non-naming conditions (“Look at the <target>!”, “Look at the picture!” and “What’s this?”, respectively). All of the fixations for a picture of the dog in the “Look at the picture!” condition, clustered in three groups ($F(2,57)= 141.104$, $p < .001$), one of which was around

the head (fixations in brown) and the other two clusters (fixations in green and blue) were around central parts of the body. A similar pattern of fixation distribution was found in the other two conditions, whereby cluster analysis demonstrated three clusters in “What’s this?” ($F(2,55)= 106.816, p<.001$) and in “Look at the dog!” conditions ($F(2,56)= 100.899, p<.001$). Similarly, cluster analysis for a picture of the bike revealed three clusters of fixations in each of the three auditory conditions: $F(2,53)= 44.947, p<.001$ for “Look at the picture!”, $F(2,64)= 123.488, p<.001$ for “What’s this?” and $F(2,60)= 132.38, p<.001$ “Look at the bike!” condition.

The clustering revealed rather dispersed eye-movements in the present mixed-objects design. This time, looking was much more evenly distributed even for the animate objects. Participants still focused on the head of the animal, but there were much fewer fixations in that region and for both animates and inanimates there were more fixations within the centre of the pictures. Generally, participants’ eye-movements for the animate and inanimate objects were much more alike in comparison to Ković et al. (2009a) and Ković et al. (2009b) studies where animates and inanimates were presented independently.

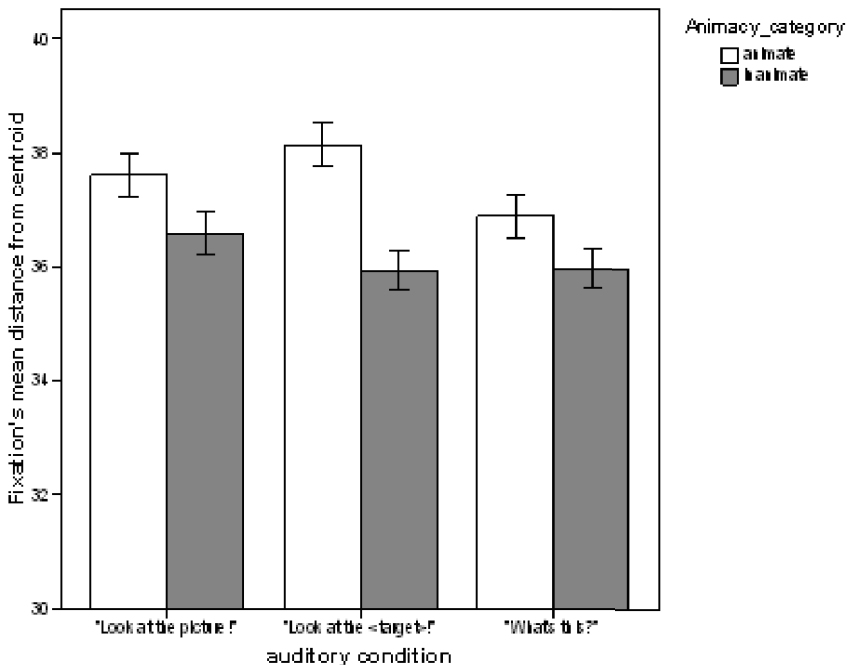


Figure 12. Fixations’ mean distance from clusters’ centroids

To further quantify the clustering results, a mean distance for each fixation from its cluster centroid was calculated and averaged across the naming and

non-naming conditions. A 2x3 ANOVA with factors Auditory condition and Animacy revealed only significant effects of Animacy ($F(1,342)=21.07, p<.001$). Auditory condition ($F(1,342)=2.02, p=0.128$) and Auditory condition x Animacy interaction ($F(2,342)=1.91, p=0.148$) were not significant.

Planned comparisons revealed that the fixations were more dispersed in the “Look at the <target>!” conditions for animates in comparison to inanimates, whereas the distribution of fixations in the other two conditions across animates and inanimates was not statistically different (‘Look at the <target>!’ condition: $M(\text{animate})=38.44, s.e.m.=0.389, M(\text{inanimate})=35.93, s.e.m.=0.342, t(1,118)=4.262, p<.001$; ‘Look at the picture!’ condition: $M(\text{animate})=37.60, s.e.m.=0.372, M(\text{inanimate})=36.59, s.e.m.=0.368, t(1,118)=1.935, p=0.053$; and in ‘What’s this?’ condition: $M(\text{animate})=36.88, s.e.m.=0.389, M(\text{inanimate})=35.97, s.e.m.=0.345, t(1,118)=1.764, p=0.078$), (see Figure 12).

However, the same problem with additional quantification reported in Ković et al. (2009a) and Ković et al. (2009b) was apparent in this study as well. Namely, the difficulty was in defining a clear cut between the areas of interest. Some of the fixations which belonged to certain clusters were very close or overlapping with the neighbouring clusters and thus it was difficult to pursue the analysis which would examine the amount of time spent in a certain area of interest, or the number of fixations participants made within those areas.

Regarding the order in which participants processed animate and inanimate objects, Spearman’s correlation between the order of fixations and cluster membership revealed a significant correlation for animates ($r=.107, p<.005$) and for the inanimates ($r=.055, p<.005$). Furthermore, Spearman’s correlation across the auditory conditions was significant in both animates and inanimates (animates: ‘Look at the <target>!’: $r=.055, p<.005$; ‘Look at the picture!’: $r=.107, p<.005$; and ‘What’s this?’: $r=.078, p<.005$ and inanimates: ‘Look at the <target>!’: $r=.074, p<.005$; ‘Look at the picture!’: $r=.103, p<.005$; and ‘What’s this?’: $r=.055, p<.005$, respectively). These correlations although somewhat weak, suggest that the participants demonstrated consistency regarding the order in which they processed the pictures of animate and inanimate objects. Nevertheless, notice that the correlations reported here were weaker than the correlation observed in Ković et al. (2009a) and could have been driven by the starting and finishing fixations which were mainly located in the central region of pictures irrespective of which image was presented on the screen.

DISCUSSION

The goal of the present study was to systematically compare looking behaviour to animate and inanimate pictures and test if animate objects are processed differently to inanimate objects (as demonstrated in Ković et al. (2009a) and Ković et al. (2009b)), or if the different looking behaviours in those

studies were exhibited due to strategic looking specific to the context of the task participants were given.

The initiation of the first look demonstrated that irrespective of the auditory condition, that is, regardless of whether pictures were named or not, participants were taking less time to initiate their first gaze towards animate than to inanimate objects. This result is compatible with findings of Proverbio et al. (2007). However, a detailed analysis across animate/inanimate object categories revealed that this result was driven by quick initiation of the first eye gaze for mammals and reptiles. Regarding the longest look measurement, participants looked longer at the animate objects, but only in the naming (“Look at the <target>!”) condition. However, there was no systematic difference between animate and inanimate categories: Mammals, Insects, Aquatic animals and Clothes tended to receive longer looks than Reptiles, Vehicles, Food and Furniture. The total looking time was less for animates than for inanimates across all of the auditory conditions, but this difference was not significant. Similar to the longest look, the comparisons across the animate-inanimate categories revealed substantial variation: Reptiles and Vehicles received the least amount of TTL, followed by Mammals, Aquatic animals and Furniture whereas Insects, Food and Clothes received the longest TLT. The number of fixations towards animates and inanimates showed no systematic difference across the auditory conditions, but revealed within-category variation. Mammals received more fixations than any other category and systematically differed from Insects, Aquatic animals, Furniture and Clothes. In summary, when the animate and inanimate object pictures were presented within a single-study, there was no systematic difference in the way participants processed them assessed through initiation of the first look, longest look, total looking time and number of fixations. However, there was a lot of within category variation and no clear-cut difference between processing the animate and inanimate categories.

Furthermore, cluster analysis in this paradigm revealed a more dispersed pattern of fixations to animates in comparison to Ković et al. (2009a) which involved presentation of animate objects only. Fixations did not cluster around certain features or areas of the objects for the inanimate categories either, similar to Ković et al. (2009b). In fact, looking behaviour to animate and inanimate categories in this paradigm was much more similar than when the two categories were presented on their own, suggesting that participants may have demonstrated strategic looking behaviour when presented with animate objects (Ković et al. (2009a)) only. The only significant difference regarding auditory conditions was found for the “Look at the <target>!” condition, with more dispersed fixations for animates than for inanimates, contrary to previous findings which showed the opposite result (Ković et al. (2009a) vs. Ković et al. (2009b)) and contrary to the prediction that animates should exhibit a less dispersed pattern of fixations. The fixations’ mean distance from cluster centroids as a measure of fixations dispersion for both animates and inanimates was very similar to the one reported

in Ković et al. (2009b) and twice as high in comparison to Ković et al. (2009a) where fixations clustered much more closely around cluster centroids.

Regarding the order in which participants looked at the objects in the current paradigm a weak, but significant correlation was observed for both animates and inanimates, suggesting that the participants demonstrated consistency regarding the order in which they processed the pictures of animate and inanimate objects. The overall correlations reported in the current study were similar to the correlation reported for the pictures of inanimate objects in Ković et al. (2009b) and weaker than the correlation reported for the animate objects in Ković et al. (2009a). Like in these studies, the correlations reported in the current experiment could have been driven by the starting and finishing fixations which were mainly located in the central region of pictures, irrespective of which image was presented on the screen, due to the presentation of the fixation cross.

To conclude, in the mixed design study where animate and inanimate objects were presented together and where within-category variation was better controlled by having equal numbers of animate and inanimate categories with 5 items within each category, there were no systematic differences in processing animate and inanimate objects. Moreover, even when differences between the categories were observed, there were no clear-cut effects between processing animate and inanimate objects. Sometimes, these effects were driven by only one or two sub-categories, like the quicker initiation of the first look in animates which was driven by mammals and reptiles.

Thus, given the failure of finding systematic differences between animates and inanimates when assessed using various looking behaviour measurements, the results of the present study do not support the prediction from modality specific conceptual account. In fact, these results are more in agreement with an alternative, distributed account of semantic representation that explains processing differences by structural differences between animate and inanimate objects. This approach is based on the assumption that animates have more shared and semantically correlated features and less distinct features than inanimate objects (Devlin et al., 1998; Tyler et al., 2000, 2003). Given that such featural structure for animates was weakened to some extent in this study by increasing intra-group variability, a systematic processing for animates that was observed in Ković et al. (2009a) study, disappeared in the present study. This result suggests that the systematicity in processing animate objects found in Ković et al. (2009a) was due to strategic looking specific to the context of the task rather than reflecting a visual-feature-based underlying mental representation in animates.

Finally, the naming and non-naming conditions did not produce systematic differences regarding looking behaviour even in the current paradigm. The longer inter-stimulus-interval between auditory label and presentation of the picture was expected to give participants enough time to evoke a mental representation of objects in the naming condition which would affect their subsequent

looking patterns. However, given that no systematic differences were observed, one possible explanation would be that processing of the familiar objects in this paradigm happens so rapidly even without labels that the naming of the object does not make a difference. In fact, some studies claim that the object recognition occurs in less than 150 ms (Grill-Spector & Kanwisher, 2005), suggesting that object recognition happens before the initiation of the first eye-movement which is estimated to take around 200 ms (Huettig & Altman, 2005; Dahan et al., 2001).

ACKNOWLEDGEMENTS

This work was funded by ESRC (RES-000-23-1322) and Leverhulme Trust grants awarded to Prof. Kim Plunkett and a Clarendon graduate scholarship awarded to Vanja Ković by University of Oxford.

REFERENCES

- 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(4), 502–518.
- Dahan, D., Magnuson, J. S., Tanenhaus, M. K., & Hogan, E. M. (2001). Subcategorical mismatches and the time course of lexical access: Evidence for lexical competition. *Language and Cognitive Processes*, 16, 507–534.
- Dahan, D. & Tanenhaus, M. K. (2005). Looking at the rope when looking for the snake: Conceptually mediated eye movements during spoken-word recognition, *Psychological Bulletin & Review*, 12, 455–459.
- Devlin, J. T., Gonnerman, L. M., Anderson, E. S. & Seidenberg, M. S. (1998). Category-specific semantic deficits in focal and widespread brain damage: a computational account. *Journal of Cognitive Neuroscience*, 10, 77–94.
- Grill-Spector, K. & Kanwisher, N. (2005). Visual recognition: As soon as you know it is there, you know what it is. *Psychological Science*, 16, 152–160.
- Huettig, F. & Altmann, G. T. M. (2005). Word meaning and the control of eye fixation: semantic competitor effects and the visual world paradigm. *Cognition*, 96, 23–32.
- Kamide, Y., Altmann, G. T. M. & Haywood, S. (2003). The time-course of prediction in incremental sentence processing: Evidence from anticipatory eye-movements. *Journal of Memory and Language*. 49, 133–159.
- Ković, V., Plunkett, K. & Westermann, G. (2009a). Eye-tracking study of animate objects, *Psihologija*, 42, 3.
- Ković, V., Plunkett, K. & Westermann, G. (2009b). Eye-tracking study of inanimate objects, *Psihologija*, 42, 4.
- McPherson, W. B. & Holcomb, P. J. (1999). An electrophysiological investigation of semantic priming with pictures of real objects. *Psychophysiology*, 36, 53–65.
- Mozer, M. C. (2002). Frames of reference in unilateral neglect and visual perception: A computational perspective, *Psychological Review*, 109, 156–185.

- Proverbio, A. M., Del Zotto, M. & Zani, A. (2007). The emergence of semantic categorization in early visual processing: ERP indexes of animal vs. artifact recognition. *BMC Neuroscience*, 8, 1, 24.
- Sitnikova, T., West, W. C., Kuperberg, G. R. & Holcomb, P. J. (2006). The neural organization of semantic memory: electrophysiological activity suggests feature-based anatomical segregation. *Biological Psychology*, 71, 326–340.
- Tyler, L. K., Moss, H. E., Durrant-Peatfield, M. R. & Levy, J. P. (2000). Conceptual structure and the structure of concepts: a distributed account of category-specific deficits. *Brain and Language*, 75, 195–231.
- Tyler, L. K., Stamatakis, E. A., Dick, E., Bright, P., Fletcher, P. & Moss, H. (2003). Objects and their actions: evidence for a neurally distributed semantic system. *Neuroimage*, 18, 542–557.
- Ward, J. H. (1963). Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association*, 58, 236–244.
- Warrington, E. K. & McCarthy, R. A. (1987). Categories of knowledge. Further fractionations and an attempted integration. *Brain*, 110, 1273–1296.
- Warrington, E. K., Shallice, T. (1984). Category specific semantic impairments. *Brain*, 107, 829–854.
- Wishart, D. (2004). *ClustanGraphics Primer: A guide to Cluster Analysis, third edition*, Clustan Ltd, Edinburgh.
- West, W. C. & Holcomb, P. J. (2002). Event-related potentials during discourselevel semantic integration of complex pictures. *Cognitive Brain Research*, 13, 363–375.