



Original Articles

Are visual processes causally involved in “perceptual simulation” effects in the sentence-picture verification task?



Markus Ostarek^{a,b,*}, Dennis Joosen^a, Adil Ishag^c, Monique de Nijs^a, Falk Huettig^a

^a Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands

^b International Max Planck Research School for Language Sciences, The Netherlands

^c International University of Africa, Khartoum, Sudan

ARTICLE INFO

Keywords:

Language comprehension
Conceptual processing
Perceptual simulation
Embodied cognition
Sentence-picture verification

ABSTRACT

Many studies have shown that sentences implying an object to have a certain shape produce a robust reaction time advantage for shape-matching pictures in the sentence-picture verification task. Typically, this finding has been interpreted as evidence for perceptual simulation, i.e., that access to implicit shape information involves the activation of modality-specific visual processes. It follows from this proposal that disrupting visual processing during sentence comprehension should interfere with perceptual simulation and obliterate the match effect. Here we directly test this hypothesis. Participants listened to sentences while seeing either visual noise that was previously shown to strongly interfere with basic visual processing or a blank screen. Experiments 1 and 2 replicated the match effect but crucially visual noise did not modulate it. When an interference technique was used that targeted high-level semantic processing (Experiment 3) however the match effect vanished. Visual noise specifically targeting high-level visual processes (Experiment 4) only had a minimal effect on the match effect. We conclude that the shape match effect in the sentence-picture verification paradigm is unlikely to rely on perceptual simulation.

1. Introduction

In theoretical and empirical efforts to understand conceptual processing during language comprehension recent work has focused on two main problems. The first is concerned with an accurate description of the informational content that is activated as we process language, whereas the second deals with the nature of the neural and cognitive mechanisms that are used to provide this information. Even though both are closely related, it is crucial to address both separately (Barsalou, 1999, 2016; Binder, 2016; Borghesani & Piazza, 2017; Mahon & Caramazza, 2008; Mahon, 2015).

Regarding conceptual content, an overwhelming body of evidence suggests that language processing involves the contextualized retrieval of a multitude of conceptual features that, together, constitute their meanings (Anderson et al., 2016; Binder & Desai, 2011; Binder et al., 2016; Collins & Loftus, 1975; Cree & McRae, 2003; Fernandino et al., 2016; Fernandino, Humphries, Conant, Seidenberg, & Binder, 2016; Huettig & McQueen, 2007; Vigliocco, Meteyard, Andrews, & Kousta, 2009; Vigliocco, Vinson, Lewis, & Garrett, 2004). This view is theoretically appealing because it nicely accounts for the high degree of conceptual flexibility (Barsalou, 1993; Hoenig, Sim, Bochev,

Herrnberger, & Kiefer, 2008; Ostarek & Huettig, 2017a; van Dam, van Dijk, Bekkering, & Rueschemeyer, 2012; Yee & Thompson-Schill, 2016) by conceiving of conceptual processing as a form of ad hoc sampling from a feature space that is constrained by both long-term memory and immediate context.

Recent behavioural and neuroimaging studies have begun to unravel the underlying mechanisms and started painting a multifaceted picture of a widely distributed system that includes modality-specific processes (Fernandino et al., 2016; Hauk, Johnsrude, & Pulvermüller, 2004; Lewis & Poeppel, 2014; Ostarek & Huettig, 2017b, 2017a; Vukovic, Feurra, Shpektor, Myachykov, & Shtyrov, 2017), different stages of convergence possibly culminating in a modality-independent central hub (Bruffaerts et al., 2013; Fernandino et al., 2016; Patterson, Nestor, & Rogers, 2007; Ralph, Jefferies, Patterson, & Rogers, 2017), and flexible retrieval mechanisms (Kan & Thompson-Schill, 2004).

The present study focuses on one particular semantic feature; object shape. Visual world eye-tracking studies indicate that processing nouns referring to concrete objects activates information about their typical shapes (Dahan & Tanenhaus, 2005; Huettig & Altmann, 2007). As many objects can occur in multiple different shapes, listeners often need to incorporate contextual information in order to retrieve the appropriate

* Corresponding author at: MPI Psycholinguistics, Wundtlaan 1, Nijmegen, The Netherlands.

E-mail address: markus.ostarek@mpi.nl (M. Ostarek).

shape representations. Using the sentence-picture verification task, a classic experiment by Zwaan, Stanfield, and Yaxley (2002) provided evidence that contextually appropriate shape information is readily activated during sentence comprehension. In that paradigm, participants read or listen to sentences about objects that are implied to have a certain shape (e.g., *The ranger saw the eagle in the sky*; implying outstretched wings). Shortly after sentence offset, in the critical conditions a picture appears of the mentioned object either in matching (e.g., an eagle with outstretched wings) or mismatching shape (an eagle with closed wings). Participants then have to indicate as quickly and accurately as possible whether the object was mentioned in the sentence or not by pressing one of two buttons. The critical finding (Zwaan & Pecher, 2012; Zwaan et al., 2002) is shorter response latencies in the matching condition, suggesting that the sentences activate information about object shape that is specific enough to produce a priming effect on the verification judgement. Although there has been some debate about the replicability of congruency effects of this type (Papesh, 2015; Rommers, Meyer, & Huettig, 2013) and about reproducibility more generally (Pashler & Wagenmakers, 2012; Wagenmakers et al., 2016), the shape match advantage, at least in the sentence-picture verification paradigm, has proven to be very robust and reproducible (Engelen, Bouwmeester, de Bruin, & Zwaan, 2011; Rommers et al., 2013; Zwaan & Pecher, 2012).

Previous studies have implicitly or explicitly gone further and suggested that the reaction time advantage in the match condition indicates the kind of process that provides shape information, namely the process of perceptual simulation (Engelen et al., 2011; Pecher, van Dantzig, Zwaan, & Zeelenberg, 2009; Yaxley & Zwaan, 2007; Zwaan & Pecher, 2012; Zwaan et al., 2002). According to that account, accessing conceptual shape information (e.g., about a flying eagle) involves the approximate re-instatement of sensory processes that are active during visual perception of relevant objects (e.g., of a flying eagle).

However, one does not need to invoke simulation in order to explain the behavioural pattern, as studies using the sentence-picture verification paradigm can only tell us something about the kind of information that is accessed, but not about the kinds of processes and representations involved. One way to get at the latter question is to study the neural correlates of the shape match effect. Hirschfeld, Zwitserlood, and Dobel (2011) conducted a magnetoencephalography study using the sentence-picture verification paradigm to assess changes in neural activity for shape matching vs. mismatching pictures. They observed a stronger positivity to pictures following shape matching vs. mismatching sentences in occipital cortex at ca. 120 ms after picture onset (M1), suggesting a top-down modulation of early visual processing as a function of shape match vs. mismatch. However, changes in the way the target picture was visually processed do not necessarily imply that visual processes were activated during comprehension. Indeed, that scenario would predict repetition suppression, not enhancement. Therefore, the data are consistent with top-down input from higher-level cortical areas. Thus, this approach still cannot answer whether visual processes were involved in sentence comprehension, as, similar to RT paradigms, what is measured is the effect of the comprehension process on picture verification that happens only after sentence comprehension is accomplished (Mahon & Caramazza, 2008).

One direct way of testing the hypothesis that visual processes are functionally involved in visual information retrieval is to interfere with visual processing during language comprehension and assess whether visual information retrieval is impaired. Recent studies have demonstrated that dynamic low-level visual noise patterns can selectively interfere with the retrieval of visual information during auditory single word processing (Ostarek & Huettig, 2017a) and in a property verification task (Edmiston & Lupyan, 2017), and they can strongly diminish the effectiveness of a word cue on a subsequent picture discrimination task (Edmiston & Lupyan, 2017). Here, we used the visual noise technique to interfere with visual processing while participants

were listening to sentences to directly probe the functional role of perceptual simulation in the sentence-picture verification task.

2. Experiment 1

The basic rationale for this experiment was that interfering with basic visual processing while participants were listening to sentences should significantly reduce the usually observed shape-match effect if it relies on perceptual simulation. Conversely, if the match effect is independent of visual simulation, visual interference should not have an impact on the match advantage. Experiment 1 used the same kind of visual interference that was recently shown to impair access to visual information during semantic processing (Edmiston & Lupyan, 2017; Ostarek & Huettig, 2017a), consisting of dynamically changing Mondrian-type masks that are usually used for continuous flash suppression and are designed to maximally interfere with basic visual processing (Tsuchiya & Koch, 2005). We predicted that visual interference would decrease the match advantage based on four considerations: (1) the match effect pertains to visual shape information, (2) processing of shape information in early visual cortex has been shown to be modulated in the sentence-picture verification task (Hirschfeld et al., 2011), (3) previous studies reported interference effects of visual noise on semantic processing of single words (Edmiston & Lupyan, 2017; Ostarek & Huettig, 2017a), and (4) the intuitive proposal that contextually embedded language tends to engage more specific representations and might thus be more likely to activate modality-specific processes than single words (Kurby & Zacks, 2013; Zwaan, 2014).

2.1. Method

2.1.1. Participants

We recruited 115 healthy participants with normal or corrected-to-normal vision and normal hearing from the local MPI subject database. Four had to be excluded due to technical failure, and one due to excessive error rates (> 20%), resulting in 110 participants that were used for analysis. We opted for a higher number of participants compared to previous studies using this paradigm based on the fact that our design included the additional factor of Visual Condition (visual noise vs. blank screen) and the conviction that high-powered studies are needed in the field of experimental psychology (Pashler & Wagenmakers, 2012). Participants received a payment of 6 euros. The study was covered by ethics approval from Radboud University Nijmegen.

2.1.2. Materials, set-up, and design

We used the materials from the original Zwaan et al. (2002) study that were provided by Rommers, Meyer, Praamstra, and Huettig (2013). They included 40 quadruplets of pairs of sentences implying shape A or shape B and corresponding pairs of pictures of the mentioned objects in shape A or shape B, and there were 40 filler sentences paired with target pictures that are not mentioned in the sentence. In the original design, every participant saw one of four sentence-picture combinations, resulting in four lists. In the present study, the additional factor of Visual Condition (visual noise vs. blank screen) was added such that every sentence-picture pair was still only shown once to each participant, but across participants every pair occurred equally often in the visual noise and blank screen condition, resulting in eight lists.

Participants were seated 60 cm from the screen and placed their head on a chin rest. Presentation (Neurobehavioral Systems) was used to control the display of target pictures and visual noise as well as the sentences that were played back on headphones. Auditory sentences were used instead of written sentences to be able to interfere with visual processing during sentence comprehension. The task was to listen to the sentences and to decide as quickly and accurately as possible by pressing one of two buttons (left/right on a house-built button box,

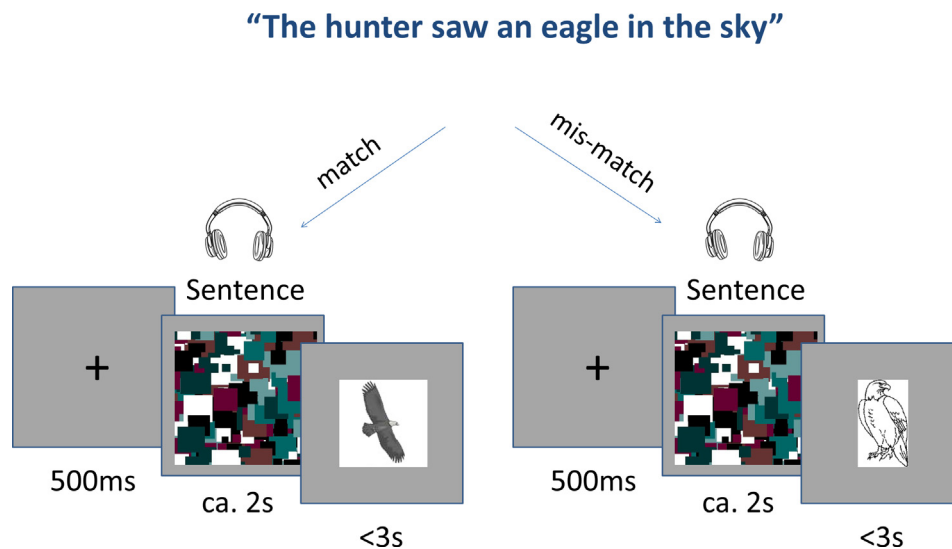


Fig. 1. Illustration of the design and trial structure with one of the sentences in the visual noise condition.

counterbalanced across participants) whether the subsequently displayed picture represented an object that was mentioned in the sentence or not.

Every trial (see Fig. 1) started with a fixation cross at the centre of the screen (500 ms) followed by an auditory sentence (ca. 2 s on average). Sentences were accompanied by visual noise in half of the trials. It consisted of 80 masks that were all generated by randomly superimposing 1000 rectangles of different colours and sizes (similar to Hesselmann, Hebart, & Malach, 2011). For every trial a random order was generated for the 80 masks and they were displayed at a rate of ca. 10 Hz until 250 ms after sentence offset, at which point the target picture was presented. Once a button was pressed or 3 s elapsed, the next trial started. After half of the filler trials, a comprehension question appeared on the screen to encourage participants to listen closely to the sentences.

2.1.3. Analysis

Prior to analysis, fillers and trials with incorrect responses or with RTs faster than 300 ms or slower than 2500 ms were excluded. We then removed trials with RTs 2.5 SDs or higher from the grand condition means. RTs were standardized by subtracting the mean and dividing by the SD for analysis. The resulting dataset was analysed using linear mixed effects modelling as implemented in the R package lme4 (Bates, Mächler, Bolker, & Walker, 2014). The full model included Match Condition (match vs. mismatch) and Visual Condition (visual noise vs. blank screen) and their interaction as fixed effects and by-participant and by-sentence random intercepts and slopes for Match Condition, Visual Condition, as well as the interaction term. The fixed effects predictors were coded as (1, -1). To obtain p-values, we computed type 3 conditional F-tests with Kenward-Roger approximation for degrees of freedom as implemented in the Anova function of the car package (Fox & Weisberg, 2011), which calls the function KRmodcomp of the pbkrtest package (Halekoh & Højsgaard, 2014). Of main interest was whether we would find a reduction of the match effect in the visual noise condition, as reflected in the interaction between Match Condition and Visual Condition. We additionally performed planned follow-up analyses looking at the effect of Match Condition in the visual noise condition and the blank screen condition separately. Specifically, we used dummy coding of the Visual Condition factor to obtain the simple effects of Match Condition for blank screen and visual noise trials, respectively. The analysis scripts and raw data can be found at the website of the Open Science Framework (<https://doi.org/10.17605/OSF.IO/HNDG2>).

2.2. Results and discussion

The results showed a significant main effect of Visual Condition (estimate = -0.034, SE = 0.014, $t = -2.48$, $p = 0.016$) with slower responses in visual noise trials ($M = 758$ ms, $SD = 252$ ms) compared to blank screen trials ($M = 741$ ms, $SD = 251$ ms). We also obtained a main effect of Match Condition (estimate = -0.071, SE = 0.030, $t = -2.35$, $p = 0.021$) with shorter RTs in the shape-matching condition ($M = 735$ ms, $SD = 241$ ms) compared to the mismatching condition ($M = 764$ ms, $SD = 261$ ms), thus replicating the match effect. As can be seen in Fig. 2, there was, however, no evidence for an interaction between the two factors ($t < 1$). The match effect was present both in blank screen trials (estimate = -0.076, SE = 0.036, $t = -2.13$, $p = 0.037$) and in visual noise trials (estimate = -0.067, SE = 0.03, $t = -2.26$, $p = 0.027$).

Thus, our results suggest that shape information was activated (as reflected by the match effect), but low-level visual processes were not necessary for it (as reflected by the irrelevance of visual interference for the match effect). In the context of two recent studies that reported disrupted access to visual information due to the same type of visual noise in paradigms using single words (Edmiston & Lupyan, 2017; Ostarek & Huettig, 2017a), it seems implausible that the visual noise technique did not sufficiently interfere with basic visual processing. By extension, higher level processes can be assumed to have provided the implicit shape information. Regarding the study by Hirschfeld et al. (2011), this result speaks against the possibility that the occipital M1 modulation they observed reflected low-level visual simulations activated during comprehension. This is consistent with the increased positivity observed by Hirschfeld et al. (2011), as opposed to a decrease that would be expected in a priming-via-re-activation explanation (due to repetition suppression).

3. Experiment 2

Experiment 1 did not provide evidence for the hypothesis that the shape match effect in the Sentence-Picture Verification task relies on low-level perceptual simulation. It is important to note, however, that we used a type of visual noise that selectively interferes with the most basic computations related to local colour, edge, and orientation detection involving only horizontal and vertical components. Thus, it remains possible that visual processes higher up in the hierarchy are causally involved in providing conceptual shape information. To test this possibility, we replaced the low-level visual noise with what we might call *mid-level* visual noise. 80 mid-level noise masks were

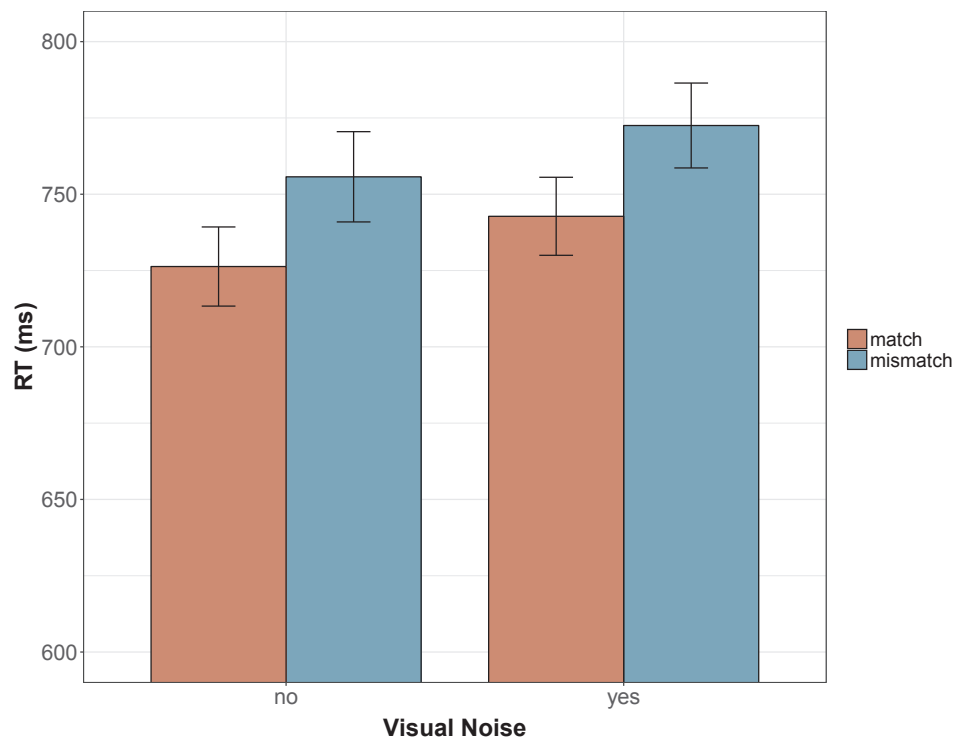


Fig. 2. Experiment 1. Mean RTs to shape-matching vs. mismatching pictures (colour-coded) with and without visual noise during auditory sentence presentation. Error bars indicate 95% confidence intervals.



Fig. 3. Examples of the mid-level visual noise masks used in Experiment 2.

generated by superimposing 30–40 images of random objects (from De Groot, Koelewijn, Huettig, & Olivers, 2016) and distorting them with the Adobe Photoshop functions “shear”, “ripple”, and “crystalize” such that they could no longer be recognised (see Fig. 3). The resulting images were at least as visually complex as real objects but were

unlikely to evoke consistent semantic associations. Our rationale for using these masks as visual noise was that they should interfere with neuron populations with larger receptive fields that are tuned to complex conjunctions of multiple visual features (Peirce, 2015; Vernon, Gouws, Lawrence, Wade, & Morland, 2016). As such, the mid-level

visual noise allowed us to test whether mid-level visual processes are involved in the shape match effect.

3.1. Method

3.1.1. Participants

We recruited 114 participants from the MPI subject database, one of which was excluded due to an error rate higher than 20%.

3.1.2. Materials, set-up, design, and analysis

Everything was identical to Experiment 1 except for the use of mid-level visual noise that was presented at ca. 10 Hz in visual noise trials.

3.2. Results and discussion

The data were trimmed for uncorrected responses and outliers in the same way as in Experiment 1. There was again a significant main effect of Visual Condition (estimate = -0.065 , SE = 0.014 , $t = -4.70$, $p < 0.001$) with slower RTs in the visual noise condition (M = 777 ms, SD = 239 ms) compared to the blank screen condition (M = 744 ms, SD = 231 ms), and a main effect of Match Condition (estimate = -0.069 , SE = 0.033 , $t = -2.10$, $p = 0.039$) with shorter RTs in the match (M = 746 ms, SD = 229 ms) compared to the mismatch condition (M = 774 ms, SD = 241 ms). Again, as Fig. 4 indicates, there was no evidence for an interaction ($t < 1$). The size of the match effect was similar in the blank screen condition (estimate = -0.079 , SE = 0.035 , $t = -2.23$, $p = 0.029$) and in the visual noise condition (estimate = -0.06 , SE = 0.035 , $t = -1.70$, $p = 0.094$).

Thus, we again replicated the match effect but found no evidence that mid-level visual processes were functionally involved.

4. Experiment 3

The results of the first two experiments point to a striking independence of the shape match effect from visual processes, given its



Fig. 5. Examples of the objects used as semantic noise.

prior interpretation (Zwaan et al., 2002; Zwaan, 2003). Given the recent demonstrations of the involvement of low-level visual processes in semantic processing using the same interference technique (Edmiston & Lupyan, 2017; Ostarek & Huettig, 2017a), it is implausible that the visual noise we used was not capable of impeding simulation. Nevertheless, Experiment 3 was designed to ascertain that dynamic visual noise can in principle reduce the match effect in the sentence-picture verification paradigm. To that end, the meaningless visual noise masks were replaced with pictures of intact objects that were again displayed at ca. 10 Hz (henceforth *semantic noise*; see Fig. 5). The critical difference to the previous two experiments was that pictures activate semantic representations, even when they are presented for very short durations and in rapid succession (Potter, 1976; Potter, Wyble, Haggmann, & McCourt, 2014; Thorpe, Fize, & Marlot, 1996). Recent studies have also shown that visual object recognition even involves the rapid activation of object names in adults (McQueen & Huettig, 2014), and even in toddlers (Mani & Plunkett, 2010). As such, semantic noise

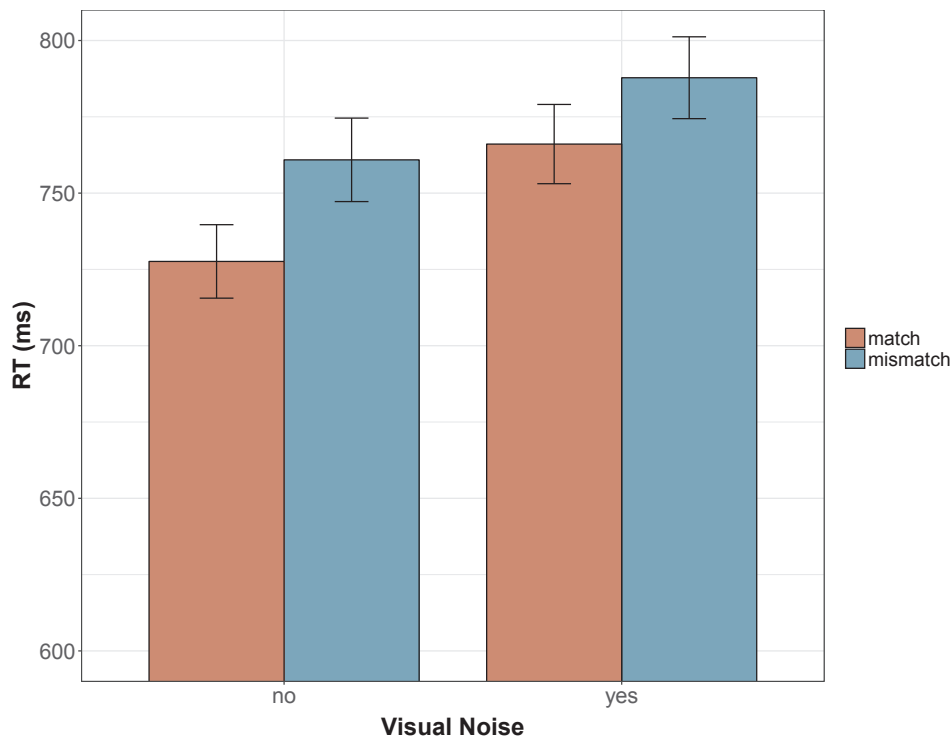


Fig. 4. Experiment 2. Mean RTs to shape-matching vs. mismatching pictures (colour-coded) with and without visual noise during auditory sentence presentation. Error bars indicate 95% confidence intervals.

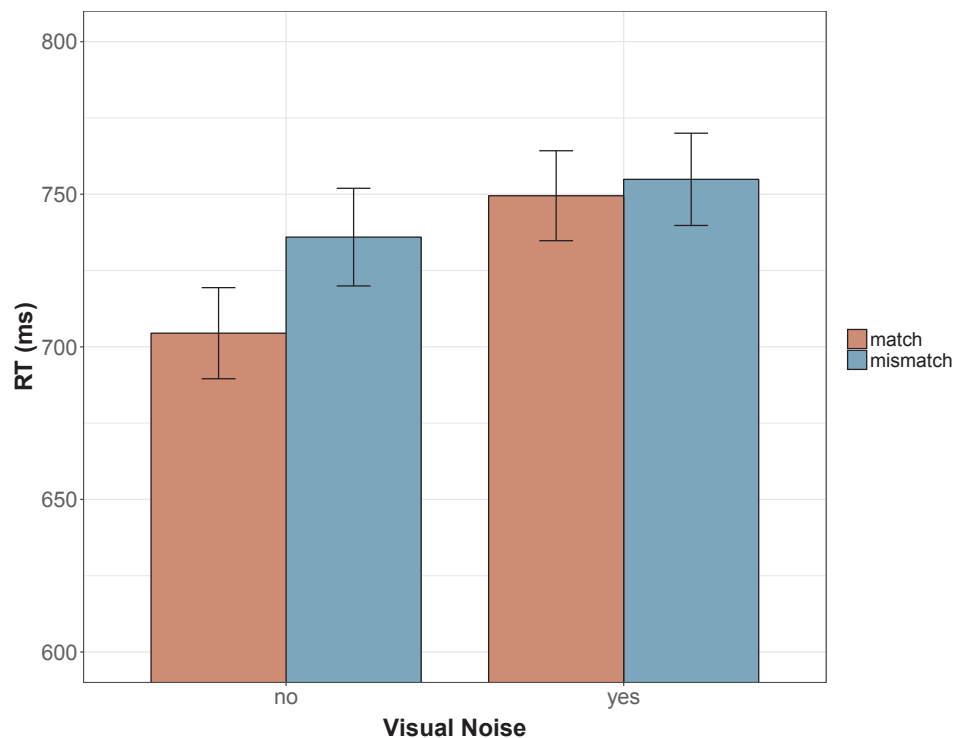


Fig. 6. Experiment 3. Mean RTs of to shape-matching vs. mismatching pictures (colour-coded) with and without visual noise during auditory sentence presentation. Error bars indicate 95% confidence intervals.

can be expected to interfere with the access to conceptual shape information implicit in sentences.

4.1. Method

4.1.1. Participants

We recruited 112 participants with normal or corrected-to-normal vision and normal hearing from the MPI subject database. One participant was excluded due to an error rate > 20%.

4.1.2. Materials, set-up, design, and analysis

Everything was identical to Experiment 1 and 2 except for the use of object pictures as dynamic visual noise (again ca. 10 Hz). To that end, 80 pictures were randomly selected from the de Groot et al. (2016) database with the constraint that they did not represent items mentioned in any of the sentences.

4.1.3. Results and discussion

Due to a large number of participants with very high mean RTs and large SDs, the outlier removal procedure used in the previous experiments resulted in very few observations per condition in some participants and the mixed effects model did not converge. We therefore excluded participants who had mean RTs larger than 1000 ms and SDs higher than 400 ms in at least one condition to reduce noise.¹ The data from the remaining 73 participants (that were trimmed for outliers as in

¹ We also conducted a generalized linear mixed effects analysis with a gamma distribution using all of the data (no outlier removal except for RTs < 300 ms). This analysis is well suited for positively skewed data, as kindly pointed out by one of the reviewers. In short, it similarly showed a decrease of the match effect in the semantic noise condition (estimate = -7.814, SE = 3.842, $t = -2.03$, $p = 0.042$), thus confirming the results above without outlier removal. See supplementary materials for more information and equivalent analyses for the other experiments.

the previous experiments) are summarised in Fig. 6. As in the previous experiments, there was a main effect of Visual Condition (estimate = -0.081, SE = 0.018, $t = -4.50$, $p < 0.001$) with shorter responses in the blank screen condition ($M = 720$ ms, $SD = 209$ ms) compared to the visual noise condition ($M = 752$ ms, $SD = 206$ ms). There was no significant main effect of Match Condition (matching: $M = 727$ ms, $SD = 206$ ms; mismatching: $M = 745$ ms, $SD = 210$ ms; estimate = -0.047, SE = 0.034, $t = -1.38$, $p = 0.172$), but crucially, the interaction was significant (estimate = -0.035, SE = 0.017, $t = -2.11$, $p = 0.042$), reflecting the match effect (of 31 ms) in the blank screen condition (estimate = -0.081, SE = 0.039, $t = -2.11$, $p = 0.038$) compared to the absent match effect (5 ms) in the visual noise condition (estimate = -0.012, SE = 0.036, $t = -0.34$, $p > 0.7$).

Thus, semantic noise strongly interfered with the access to shape information implicit in sentences, demonstrating that dynamic visually presented stimuli can be effective at interfering with the retrieval of conceptual shape information if the relevant system is targeted by the noise. This confirms that the absent impact of visual noise on the match effect in Experiments 1 and 2 was unlikely due to an inability of the present interference technique to diminish it, but rather due to its independence of modality-specific visual processes.

5. Experiment 4

The experiments presented so far indicate that low-level and mid-level processes do not contribute functionally to the shape match effect (at least not to a theoretically interesting extent), whereas semantic noise reduced it to near-zero. One interpretation of these results is that the match effect does not rely on modality-specific visual processes. Alternatively, it is conceivable that shape information is provided by modality-specific visual processes, but at a higher level than tested in Experiments 1 and 2. Specifically, a difference between the visual and



Fig. 7. Examples of the high-level visual noise stimuli.

semantic noise manipulations was that the pictures used as noise in Experiment 3 are processed holistically and require figure-ground discrimination. Thus, it is possible that semantic noise interfered with access to shape information because of interference on a high visual - rather than semantic - level. To tease these two possibilities apart, we ran a final experiment with visual noise consisting of abstract shapes that are perceived as objects but do not evoke any particular semantic associations (see Fig. 7). The rationale was that this type of visual noise taps high-level visual processing related to holistic object perception and figure-ground separation, but has no or very limited effects on (non-visual) semantic processing.

5.1. Method

5.1.1. Participants

115 participants with normal hearing and normal or corrected-to-normal vision were recruited from the MPI database and were paid six euros for their time. Two had to be excluded due to error rates higher than 20%.

5.1.2. Materials, set-up, design, and analysis

The experiment was identical to the previous ones except for the 80 visual noise objects which were nonsense-objects designed in Adobe Photoshop. They varied in size, colour, and shape, and were designed to look like possible objects without resembling any particular existing object.

5.1.3. Results and discussion

The data were trimmed for outliers as in the previous experiments. As in Experiment 3, there was an unexpected number of participants with very long RTs that resulted in non-convergence of the mixed effects model. As before, to reduce noise we therefore excluded

participants with RTs slower than 1 s and SD higher than 400 ms (on average) in at least one condition². The remaining 93 participants were used for the analyses reported below. There was a main effect of Visual Condition (estimate = -0.047 , SE = 0.015, $t = -3.14$, $p = 0.003$), with slower responses in the visual noise compared to the blank screen condition. There was also a main effect of Match Condition (estimate = -0.076 , SE = 0.036, $t = -2.09$, $p = 0.04$), with faster responses in the match vs. mismatch condition. However, there was little evidence for an interaction between the two (estimate = -0.016 , SE = 0.014, $t = -1.17$, $p = 0.25$), suggesting that high-level visual noise did not have a robust impact on the size of the match effect (see Fig. 8). There was a descriptive trend towards a reduction of the match effect reflected in the pattern that the match effect was significant in the blank screen trials (estimate = -0.092 , SE = 0.038, $t = -2.42$, $p = 0.018$), but not in the visual noise trials (estimate = -0.06 , SE = 0.040, $t = -1.51$, $p = 0.136$).

6. Exploratory follow-up analyses

In Experiments 1, 2, and 4, based on the high statistical power of the experiments and the small t-values we observed, we interpreted the null-results for the interaction between Visual Condition and Match Condition as speaking against a functional role of visual simulations for the shape match effect. However, instead of a categorical yes/no characterization of this result, it would be preferable to quantify the amount of evidence in favour of an effect of visual noise on the match

² We again conducted a generalized linear mixed effects analysis with a gamma distribution using all of the data (no outlier removal except for RTs < 300 ms). There was strong evidence for a main effect of Visual Condition ($t = -3.35$, $p < .001$) and a main effect of Match Condition ($t = -5.15$, $p < 0.001$), but no evidence for an interaction ($t < 1$, $p > .7$). See the supplementary materials for additional information.

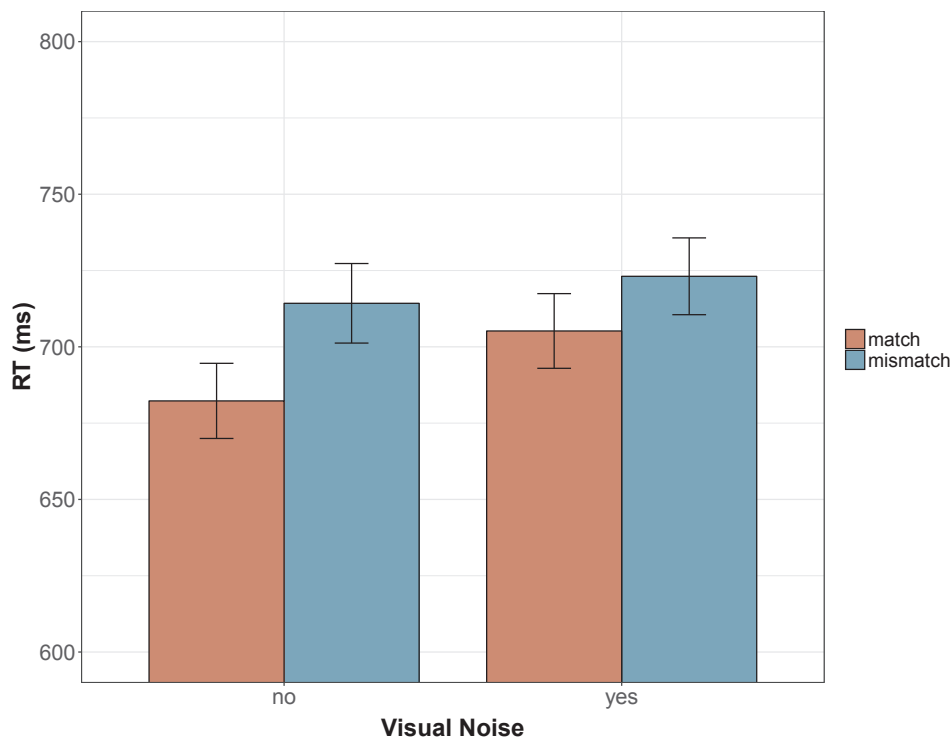


Fig. 8. Experiment 4. Mean RTs of to shape-matching vs. mismatching pictures (colour-coded) with and without visual noise during auditory sentence presentation. Error bars indicate 95% confidence intervals.

effect. To achieve this, we conducted Bayesian follow-up analyses using the *brms* package in R (Bürkner, 2016). The strength of this package is that it allowed us to use the same fixed and random effects structure as in the linear mixed effects models reported above and calculate Bayesian mixed effects estimates and 95% credible intervals (CrIs). The *hypothesis* function allowed us to compute Bayes Factors based on the Savage-Dickey method (Wagenmakers, Lodewyckx, Kuriyal, & Grasman, 2010) which computes an evidence ratio between the fit of a model that assumes the regression weight of the interaction to be zero (null hypothesis) and a model that assumes it to be non-zero taking into account the priors and the new data (BF_{01} ; numbers larger than one indicate evidence in favor of the null). We also calculated directional Bayes Factors as a more sensitive measure of a reduction of the shape match effect (BF_{dir} ; numbers larger than one indicate evidence in favor of the interaction effect in the predicted direction). They are based on an evidence ratio between the fit of a model that assumes the regression weight of the interaction to be negative (i.e. that assumes that noise reduces the shape match effect) to a model that assumes the regression weight to be positive (i.e. that assumes that noise increases the shape match effect). Priors for Experiment 1 were an estimate based on previous experiments using the sentence-picture verification paradigm (Rommers et al., 2013; Zwaan & Pecher, 2012; Zwaan et al., 2002). For the remaining three experiments, the estimates and 95% credible intervals of the previous experiment (including all fixed effects and the intercept) were used as priors for the following experiment, respectively. 100 000 iterations were run per model and the thinning rate was set to 100.

In Experiment 1 (low-level visual noise) and Experiment 2 (mid-level visual noise), the 95% CrIs included zero, the non-directional BFs supported the null hypothesis, and the directional BF provided minimal evidence for an interaction in the predicted direction (Experiment 1: estimate = -0.005 , SE = 0.013, 95% CI: -0.03 , 0.02, $BF_{01} = 7.82$, $BF_{dir} = 1.84$; Experiment 2: estimate = -0.008 , SE = 0.012, 95% CI: -0.031 , 0.016, $BF_{01} = 3.35$, $BF_{dir} = 2.76$). In Experiment 4 (high-level visual noise), although the 95% CrIs included zero and the non-

directional BF favoured the null hypothesis, the directional BF suggested that there is some evidence for an interaction in the predicted direction (estimate = -0.017 , SE = 0.014, 95% CI = -0.044 , 0.012, $BF_{01} = 2.46$, $BF_{dir} = 7.62$). The picture was different in Experiment 3 (semantic noise) where zero was not within the 95% CrI, the non-directional BF provided evidence against the null, and the directional BF provided very strong evidence for an interaction in the predicted direction (estimate = -0.032 , SE = 0.016, 95% CI: -0.063 , -0.002 , $BF_{01} = 0.34$, $BF_{dir} = 63.52$).

The Bayesian follow-up analyses thus confirmed that there is very little evidence that visual noise robustly diminished the match effect. This was further supported by a linear mixed-effects analysis that pooled the data from all three visual noise experiments (Experiments 1, 2, and 4) to obtain maximal power to detect a small interference effect: The results indicate that Visual Condition ($t = -6.85$) and Match Condition ($t = -2.25$) have robust effects, whereas the interaction does not ($t = -1.22$).

Note that our studies were not designed to investigate whether semantic noise produces more interference than visual noise (as this was not our research question). Experiment 3 (semantic noise) was conducted to provide another demonstration (in addition to the ones in the literature) that the noise manipulation can in principle interfere with semantic processing. Nevertheless, on reviewer request, we assessed statistically whether semantic noise reduced the match effect more strongly than visual noise. Ideally, one would perform an analysis that adds Experiment (i.e. type of noise) as additional factor. This would however yield a design with many more factors/levels for which our sample would be unlikely to have adequate power. To maintain a 2x2 design, for which our sample was intended, we performed a subset analysis without the blank screen trials and coded all types of visual noise (low, mid, and high-level) as “visual”. Thus, we ran a linear mixed effects model with Type of Noise (visual vs. semantic) and Match Condition (matching vs. mismatching) as fixed effects, per-participant random intercepts and slopes for the effect of Match Condition, and per-sentence random intercepts and slopes for both factors and their

interaction (estimate = -0.025 , SE = 0.012 , $t = -1.99$, $p = 0.05$). This result is consistent with the view that semantic noise reduced the match effect more strongly than visual noise. Future studies could compare the relative effect of different types of noise by manipulating this factor within-subjects.

7. General discussion

The sentence-picture verification task has been of great value for the field of language processing, as it is a versatile tool to reveal the contents of conceptual representations that are activated as listeners/readers comprehend sentences. The key insight was that a match vs. mismatch in a feature of interest between a sentence and a following picture modulates response latencies in the verification task to the extent that the feature was activated during sentence processing. Previous studies have shown that listeners activate object shape information implicit in sentences by demonstrating a reaction time advantage in subsequent sentence-picture verification for target pictures that matched the implied object shape (Hirschfeld et al., 2011; Rommers et al., 2013; Zwaan & Pecher, 2012; Zwaan et al., 2002).

Here, we asked which processes enable the retrieval of shape information during online sentence comprehension. The match effect is typically interpreted as indexing perceptual simulation in sentence comprehension (Engelen et al., 2011; Zwaan & Pecher, 2012; Zwaan et al., 2002; Zwaan, 2003). In particular, the idea is that modality-specific visual processes relevant for shape perception are recruited during language comprehension to provide conceptual shape information. Consequently, when a subsequent shape matching picture appears it is processed more efficiently due to the pre-activation of relevant visual processes. However, while the sentence-picture verification paradigm in its basic form is well-suited to uncover the contents of representations activated during sentence comprehension, it does not allow inferences about how this content is represented. As Hirschfeld et al. (2011) pointed out; besides perceptual simulation the match effect is consistent with a top-down effect based on amodal semantic representations (Mahon, 2015) and with task-based expectations/predictions about the target (Rabagliati, Doumas, & Bemis, 2017; Rommers et al., 2013), even though shape congruency was also found to be reflected in the N400 (Coppens, Gootjes, & Zwaan, 2012) and in recognition memory performance (Pecher et al., 2009) when sentences and pictures were temporally decoupled.

To directly probe the functional role of perceptual simulation for the shape match effect, in two high-powered experiments we employed visual noise to interfere with basic visual processing during sentence comprehension and observed very little evidence for a decrease in the match effect. This was despite the use of a visual noise technique that was developed to maximally interfere with visual processing (Tsuchiya & Koch, 2005) and that has recently been shown to selectively hinder access to visual information in single word processing (Edmiston & Lupyan, 2017; Ostarek & Huettig, 2017a). A third experiment using pictures of irrelevant objects as noise (semantic noise) obliterated the match effect, suggesting that relatively high-level cognitive processes drive the match effect. To probe whether these processes are best described as high-level visual or modality-independent semantic processes, we conducted a final experiment in which we used nonsense objects to interfere with high-level visual processes related to figure-ground separation and holistic object perception whilst severely limiting the likelihood of (non-visual) semantic processing. Linear mixed effects analyses did not provide evidence for the view that high-level visual noise robustly reduced the match effect, which was confirmed by exploratory Bayesian analyses that provided only weak evidence for that hypothesis. Our results thus best fit with the view that the shape match effect relies mostly on non-visual semantic processes which only minimally interface with low-level visual processes. The exploratory Bayesian analyses suggest that the likelihood of interference of high-level visual noise was small but non-zero.

This raises interesting questions for future research: One possible account is that interference increases as the activation of semantic information increases. The more high-level a mask, the more likely it presumably is to activate semantic information which then leads to interference with the retrieval of semantic shape information during sentence processing. Related to that, it would be interesting to find out what exactly accounts for the interference effect induced by semantic noise. It is plausible that the semantic system is simply overloaded due to the high rate of objects that are being processed. It is also possible that the irrelevant objects are implicitly labelled. This could be tested by using objects that are easy (e.g., *apple*) vs. difficult (e.g., *dragon fruit*) to label. Alternatively, the presence of certain visual characteristics in the masks might determine the amount of interference. One limitation of our study is that visual and semantic noise masks differed in visual features. Thus, it is possible that the types of visual noise we used lacked visual characteristics that are present in simulations, such as 3D structure, or the presence of different parts or textures. Whereas it seems unlikely that access to these kinds of details plays a bigger role for the shape match effect than access to approximate shape information (see Hirschfeld et al., 2011), more work is needed to fully map out the factors that determine the amount of interference with the retrieval of semantic shape information.

Overall, our findings are theoretically important as they constitute evidence against a strong perceptual simulation-based explanation of the shape match effect. Our findings are compatible with accounts of amodal representation that assume conceptual priming effects to arise in high-level systems with a non-modality-specific representational system (e.g., Fodor, 1975). They are also consistent with the grounding-by-interaction model (Mahon & Caramazza, 2008; Mahon, 2015), which hypothesises amodal representations that are connected to the sensory systems, but only to the extent that sensory states do not measurably affect processing in the amodal system. This model can account for our behavioural results and the MEG data reported by Hirschfeld et al. (2011) by postulating a congruency effect in an amodal conceptual system that optionally interacts with visual processing. One prediction following from this account that could be tested in future studies is that visual interference diminishes the modulation of early visual processing (as observed by Hirschfeld et al. (2011)) without affecting the size of the match effect.

It is important to stress that we are not denying that perceptual simulation contributes to language comprehension, given the large body of evidence for this view (e.g., Barsalou, 2008; Correia et al., 2014; Fernandino et al., 2016; Hauk et al., 2004; Lewis & Poeppel, 2014; Meteyard, Cuadrado, Bahrami, & Vigliocco, 2012; Ostarek & Huettig, 2017b, 2017a; Pulvermüller, 2005; Vukovic et al., 2017). Other paradigms have recently provided compelling evidence that low-level visual processes (likely related to shape) are engaged in the comprehension of concrete object words (Edmiston & Lupyan, 2017; Lewis & Poeppel, 2014; Ostarek & Huettig, 2017b, 2017a). However, the match effect in the sentence-picture verification task seems to depend on higher-level processes.

This is a striking result not only because the shape match effect is considered a hallmark finding for theories of embodied cognition, but also because the paradigm seems such a good candidate for perceptual simulation (which may partly explain why the match effect has usually been interpreted the way it has). Edmiston and Lupyan (2017) observed a clear effect of visual noise in single word-picture-verification where word cues are followed by a matching or mismatching picture in correct and inverted orientation and participants have to indicate which the correctly oriented picture is. These results in combination with ours suggest that single words activate low-level visual processes (likely reflecting typical object features), whereas implicit shape information derived from event-level representations involves abstraction away from the sensory systems. This is consistent with neuroimaging studies that implicated anterior temporal regions with high-level semantic processing and conceptual combination, whereas visual regions are

linked to individual object features, such as size, colour, and shape (Borghesani et al., 2016; Coutanche & Thompson-Schill, 2014). It will be crucial for future studies to precisely delineate what determines the involvement and role of sensory processes in conceptual processing.

The arguments and method presented in this paper can readily be applied to other paradigms relying on congruency between sensory-motor content evoked by linguistic input and a task involving sensory-motor processing: Congruency effects do not provide evidence that sensory-motor systems produce them unless the task which language is found to have an effect on *only* involves sensory-motor processes (Lupyan & Ward, 2013; Ostarek & Huettig, 2017b). Therefore, match effects are a useful first step after which further investigations are required to reveal the underlying mechanisms. Interference techniques are appealing because they go beyond correlational approaches by directly testing causality.

Acknowledgements

We would like to thank Christoph Scheepers, Diane Pecher, and Tom Heyman for their helpful suggestions on previous versions of the manuscript, and Phillip Alday for his advice on Bayesian statistics.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at <https://doi.org/10.1016/j.cognition.2018.08.017>.

References

- Anderson, A. J., Binder, J. R., Fernandino, L., Humphries, C. J., Conant, L. L., Aguilar, M., ... Raizada, R. D. (2016). Predicting neural activity patterns associated with sentences using a neurobiologically motivated model of semantic representation. *Cerebral Cortex*. <http://cercor.oxfordjournals.org/content/early/2016/08/12/cercor.bhw240.abstract>.
- Barsalou, L. W. (1993). Flexibility, structure, and linguistic vagary in concepts: Manifestations of a compositional system of perceptual symbols. *Theories of Memory*, 1, 29–31.
- Barsalou, L. W. (1999). Perceptions of perceptual symbols. *Behavioral and Brain Sciences*, 22(04), 637–660.
- Barsalou, L. W. (2008). Grounded cognition. *Annual Review of Psychology*, 59(1), 617–645. <https://doi.org/10.1146/annurev.psych.59.103006.093639>.
- Barsalou, L. W. (2016). On staying grounded and avoiding quixotic dead ends. *Psychonomic Bulletin & Review*, 23(4), 1122–1142.
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). Fitting linear mixed-effects models using lme4. ArXiv Preprint ArXiv:1406.5823. Retrieved from <https://arxiv.org/abs/1406.5823>.
- Binder, J. R. (2016). In defense of abstract conceptual representations. *Psychonomic Bulletin & Review*, 23(4), 1096–1108.
- Binder, J. R., Conant, L. L., Humphries, C. J., Fernandino, L., Simons, S. B., Aguilar, M., & Desai, R. H. (2016). Toward a brain-based componential semantic representation. *Cognitive Neuropsychology*, 33(3–4), 130–174. <https://doi.org/10.1080/02643294.2016.1147426>.
- Binder, J. R., & Desai, R. H. (2011). The neurobiology of semantic memory. *Trends in Cognitive Sciences*, 15(11), 527–536.
- Borghesani, V., & Piazza, M. (2017). The neuro-cognitive representations of symbols: The case of concrete words. *Neuropsychologia*. Retrieved from <http://www.sciencedirect.com/science/article/pii/S0028393217302397>.
- Borghesani, V., Pedregosa, F., Buiatti, M., Amadon, A., Eger, E., & Piazza, M. (2016). Word meaning in the ventral visual path: A perceptual to conceptual gradient of semantic coding. *NeuroImage*, 143, 128–140.
- Bruffaerts, R., Dupont, P., Peeters, R., De Deyne, S., Storms, G., & Vandenberghe, R. (2013). Similarity of fMRI activity patterns in left perirhinal cortex reflects semantic similarity between words. *Journal of Neuroscience*, 33(47), 18597–18607.
- Bürkner, P.-C. (2016). brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software*, 80(1), 1–28.
- Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, 82(6), 407.
- Coppens, L. C., Gootjes, L., & Zwaan, R. A. (2012). Incidental picture exposure affects later reading: Evidence from the N400. *Brain and Language*, 122(1), 64–69.
- Correia, J., Formisano, E., Valente, G., Hausfeld, L., Jansma, B., & Bonte, M. (2014). Brain-based translation: fMRI decoding of spoken words in bilinguals reveals language-independent semantic representations in anterior temporal lobe. *Journal of Neuroscience*, 34(1), 332–338.
- Coutanche, M. N., & Thompson-Schill, S. L. (2014). Creating concepts from converging features in human cortex. *Cerebral Cortex*, 25(9), 2584–2593.
- 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(2), 163.
- 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 & Review*, 12(3), 453–459.
- De Groot, F., Koelwijn, T., Huettig, F., & Olivers, C. N. (2016). A stimulus set of words and pictures matched for visual and semantic similarity. *Journal of Cognitive Psychology*, 28(1), 1–15.
- Edmiston, P., & Lupyan, G. (2017). Visual interference disrupts visual knowledge. *Journal of Memory and Language*, 92, 281–292.
- Engelen, J. A., Bouwmeester, S., de Bruin, A. B., & Zwaan, R. A. (2011). Perceptual simulation in developing language comprehension. *Journal of Experimental Child Psychology*, 110(4), 659–675.
- Fernandino, Leonardo, Binder, J. R., Desai, R. H., Pendl, S. L., Humphries, C. J., Gross, W. L., ... Seidenberg, M. S. (2016). Concept Representation reflects multimodal abstraction: A framework for embodied semantics. *Cerebral Cortex*, 26(5), 2018–2034. <https://doi.org/10.1093/cercor/bhw020>.
- Fernandino, L., Humphries, C. J., Conant, L. L., Seidenberg, M. S., & Binder, J. R. (2016). Heteromodal cortical areas encode sensory-motor features of word meaning. *Journal of Neuroscience*, 36(38), 9763–9769. <https://doi.org/10.1523/JNEUROSCI.4095-15.2016>.
- Fodor, J. A. (1975). *The language of thought* (Vol. 5). Harvard University Press.
- Fox, J., & Weisberg, S. (2011). *An R companion to applied regression* (2nd ed.). Thousand Oaks, CA: Sage URL: <http://socserv.socsci.mcmaster.ca/jfox/Books/Companion>.
- Halekoh, U., & Hejsgaard, S. (2014). A Kenward-Roger approximation and parametric bootstrap methods for tests in linear mixed models – The R package pbrkrtest. *Journal of Statistical Software*, 59(9), 1–30.
- Hauk, O., Johnsrude, I., & Pulvermüller, F. (2004). Somatotopic representation of action words in human motor and premotor cortex. *Neuron*, 41(2), 301–307.
- Hesselmann, G., Hebart, M., & Malach, R. (2011). Differential BOLD activity associated with subjective and objective reports during “blindsight” in normal observers. *Journal of Neuroscience*, 31(36), 12936–12944.
- Hirschfeld, G., Zwitserlood, P., & Döbel, C. (2011). Effects of language comprehension on visual processing—MEG dissociates early perceptual and late N400 effects. *Brain and Language*, 116(2), 91–96.
- Hoenig, K., Sim, E.-J., Bochev, V., Herrnberger, B., & Kiefer, M. (2008). Conceptual flexibility in the human brain: Dynamic recruitment of semantic maps from visual, motor, and motion-related areas. *Journal of Cognitive Neuroscience*, 20(10), 1799–1814.
- Huettig, F., & Altmann, G. T. (2007). Visual-shape competition during language-mediated attention is based on lexical input and not modulated by contextual appropriateness. *Visual Cognition*, 15(8), 985–1018.
- Huettig, F., & McQueen, J. M. (2007). The tug of war between phonological, semantic and shape information in language-mediated visual search. *Journal of Memory and Language*, 57(4), 460–482.
- Kan, I. P., & Thompson-Schill, S. L. (2004). Selection from perceptual and conceptual representations. *Cognitive, Affective, & Behavioral Neuroscience*, 4(4), 466–482.
- Kurby, C. A., & Zacks, J. M. (2013). The activation of modality-specific representations during discourse processing. *Brain and Language*, 126(3), 338–349.
- Lewis, G., & Poeppel, D. (2014). The role of visual representations during the lexical access of spoken words. *Brain and Language*, 134, 1–10.
- Lupyan, G., & Ward, E. J. (2013). Language can boost otherwise unseen objects into visual awareness. *Proceedings of the National Academy of Sciences*, 110(35), 14196–14201.
- Mahon, B. Z. (2015). What is embodied about cognition? *Language, Cognition and Neuroscience*, 30(4), 420–429.
- Mahon, B. Z., & Caramazza, A. (2008). A critical look at the embodied cognition hypothesis and a new proposal for grounding conceptual content. *Journal of Physiology-Paris*, 102(1), 59–70.
- Mani, N., & Plunkett, K. (2010). In the infant’s mind’s ear: Evidence for implicit naming in 18-month-olds. *Psychological Science*, 21(7), 908–913.
- McQueen, J. M., & Huettig, F. (2014). Interference of spoken word recognition through phonological priming from visual objects and printed words. *Attention, Perception, & Psychophysics*, 76(1), 190–200.
- Meteyard, L., Cuadrado, S. R., Bahrami, B., & Vigliocco, G. (2012). Coming of age: A review of embodiment and the neuroscience of semantics. *Cortex*, 48(7), 788–804.
- Ostarek, M., & Huettig, F. (2017b). Spoken words can make the invisible visible—Testing the involvement of low-level visual representations in spoken word processing. *Journal of Experimental Psychology: Human Perception and Performance*, 43(3), 499.
- Ostarek, M., & Huettig, F. (2017a). A task-dependent causal role for low-level visual processes in spoken word comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 43(8), 1215. <https://doi.org/10.1037/xlm0000375>.
- Papesh, M. H. (2015). Just out of reach: On the reliability of the action-sentence compatibility effect. *Journal of Experimental Psychology: General*, 144(6), e116.
- Pashler, H., & Wagenmakers, E.-J. (2012). Editors’ introduction to the special section on replicability in psychological science: A crisis of confidence? *Perspectives on Psychological Science*, 7(6), 528–530.
- Patterson, K., Nestor, P. J., & Rogers, T. T. (2007). Where do you know what you know? The representation of semantic knowledge in the human brain. *Nature Reviews Neuroscience*, 8(12), 976–987. <https://doi.org/10.1038/nrn2277>.
- Pecher, D., van Dantzig, S., Zwaan, R. A., & Zeelenberg, R. (2009). Language comprehenders retain implied shape and orientation of objects. *The Quarterly Journal of Experimental Psychology*, 62(6), 1108–1114.
- Peirce, J. W. (2015). Understanding mid-level representations in visual processing. *Journal of Vision*, 15(7) 5–5.
- Potter, M. C. (1976). Short-term conceptual memory for pictures. *Journal of Experimental Psychology: Human Learning and Memory*, 2(5), 509.
- Potter, M. C., Wyble, B., Haggmann, C. E., & McCourt, E. S. (2014). Detecting meaning in

- RSVP at 13 ms per picture. *Attention, Perception, & Psychophysics*, 76(2), 270–279.
- Pulvermüller, F. (2005). Brain mechanisms linking language and action. *Nature Reviews Neuroscience*, 6(7), 576–582.
- Rabagliati, H., Doumas, L. A., & Bemis, D. K. (2017). Representing composed meanings through temporal binding. *Cognition*, 162, 61–72.
- Ralph, M. A. L., Jefferies, E., Patterson, K., & Rogers, T. T. (2017). The neural and computational bases of semantic cognition. *Nature Reviews Neuroscience*, 18(1), 42–55. <https://doi.org/10.1038/nrn.2016.150>.
- Rommers, J., Meyer, A. S., & Huettig, F. (2013). Object shape and orientation do not routinely influence performance during language processing. *Psychological Science*, 24(11), 2218–2225.
- Rommers, J., Meyer, A. S., Praamstra, P., & Huettig, F. (2013). The contents of predictions in sentence comprehension: Activation of the shape of objects before they are referred to. *Neuropsychologia*, 51(3), 437–447.
- Thorpe, S., Fize, D., & Marlot, C. (1996). Speed of processing in the human visual system. *Nature*, 381(6582), 520.
- Tsuchiya, N., & Koch, C. (2005). Continuous flash suppression reduces negative after-images. *Nature Neuroscience*, 8(8), 1096.
- van Dam, W. O., van Dijk, M., Bekkering, H., & Rueschemeyer, S.-A. (2012). Flexibility in embodied lexical-semantic representations. *Human Brain Mapping*, 33(10), 2322–2333.
- Vernon, R. J., Gouws, A. D., Lawrence, S. J., Wade, A. R., & Morland, A. B. (2016). Multivariate patterns in the human object-processing pathway reveal a shift from retinotopic to shape curvature representations in lateral occipital areas, LO-1 and LO-2. *Journal of Neuroscience*, 36(21), 5763–5774.
- Vigliocco, G., Meteyard, L., Andrews, M., & Kousta, S. (2009). Toward a theory of semantic representation. *Language and Cognition*, 1(2), 219–247.
- Vigliocco, G., Vinson, D. P., Lewis, W., & Garrett, M. F. (2004). Representing the meanings of object and action words: The featural and unitary semantic space hypothesis. *Cognitive Psychology*, 48(4), 422–488.
- Vukovic, N., Feurra, M., Shpektor, A., Myachykov, A., & Shtyrov, Y. (2017). Primary motor cortex functionally contributes to language comprehension: An online rTMS study. *Neuropsychologia*, 96, 222–229.
- Wagenmakers, E.-J., Beek, T., Dijkhoff, L., Gronau, Q. F., Acosta, A., Adams, R. B., ... Zwaan, R. A. (2016). Registered Replication Report: Strack, Martin, & Stepper (1988). *Perspectives on Psychological Science*, 11(6), 917–928.
- Wagenmakers, E.-J., Lodewyckx, T., Kuriyal, H., & Grasman, R. (2010). Bayesian hypothesis testing for psychologists: A tutorial on the Savage-Dickey method. *Cognitive Psychology*, 60(3), 158–189.
- Yaxley, R. H., & Zwaan, R. A. (2007). Simulating visibility during language comprehension. *Cognition*, 105(1), 229–236.
- Yee, E., & Thompson-Schill, S. L. (2016). Putting concepts into context. *Psychonomic Bulletin & Review*, 23(4), 1015–1027.
- Zwaan, R. A. (2003). The immersed experimenter: Toward an embodied theory of language comprehension. *Psychology of Learning and Motivation*, 44, 35–62.
- Zwaan, R. A. (2014). Embodiment and language comprehension: Reframing the discussion. *Trends in Cognitive Sciences*, 18(5), 229–234.
- Zwaan, R. A., & Pecher, D. (2012). Revisiting mental simulation in language comprehension: Six replication attempts. *PloS One*, 7(12), e51382.
- Zwaan, R. A., Stanfield, R. A., & Yaxley, R. H. (2002). Language comprehenders mentally represent the shapes of objects. *Psychological Science*, 13(2), 168–171.