The Effect of Incremental Context on Conceptual Processing: Evidence from Visual World and Reading Experiments

Diego Frassinelli (d.frassinelli@sms.ed.ac.uk)
Frank Keller (keller@inf.ed.ac.uk)
Institute for Language, Cognition and Computation
School of Informatics, University of Edinburgh
10 Crichton Street, Edinburgh EH8 9AB, UK

Christoph Scheepers (Christoph.Scheepers@glasgow.ac.uk)
Institute of Neuroscience and Psychology, University of Glasgow
58 Hillhead Street, Glasgow, G12 8QB, UK

Abstract

The analysis of the internal structure of concepts reveals the presence of a substantial amount of contextual information. Even though this interaction is easily recognizable, it is not clear how contextual information is processed and included into concept representations. The aim of this paper is to shed light on this question by analyzing the effect that an increasing amount of context exerts on conceptual processing. We report a self-paced reading experiment and a visual world experiment to test two hypotheses about the integration of context information: the *incremental activation hypothesis* suggests that the degree of facilitation in concept processing increases with the amount of context available; and the *immediate activation hypothesis* states that once a sufficient amount of contextual support is reached, no more facilitation occurs. Our data are compatible with the latter account.

Keywords: incremental processing; context effects; eye movements; visual world; self-paced reading.

Introduction

Contextual information plays an essential role in different cognitive domains, including language processing, visual processing, or reasoning. In this paper we investigate the effect that context exerts on conceptual processing (Murphy, 2002). A widely recognized way to analyze the internal structure of concepts is the use of semantic feature norms (Wu & Barsalou, 2009; McRae, Cree, Seidenberg, & McNorgan, 2005). In experiments eliciting feature norms, participants enumerate features associated with a target concept; these lists can then be used to shed light on the internal structure of concepts, including the role of context (Frassinelli & Keller, 2012). However, feature norms are static, and do not allow to study the time-course of concept processing. Real-time data are required to investigate how conceptual representations are constructed, and how they interact with conceptual information. Relevant prior work includes a study by Huettig, Quinlan, McDonald, and Altmann (2006), which analyzed the activation of conceptual information over time using the visual world paradigm (Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995). However, Huettig et al. (2006) were not interested in contextual effects and used contexts that were as neutral as possible. Frassinelli and Keller (2012) replicated this study, but also introduced contextual variability as a factor, comparing the effect of a neutral context and two biasing contexts associated with the target concept in different ways. However, Frassinelli and Keller (2012) focused their analysis on the concept region, providing only indirect evidence regarding effects taking place in the previous part of the sentence, where contextual information is integrated.

The aim of the present paper is to study contextual constraints on the processing of context, focusing in particular on the question of how context is integrated with conceptual information. According to Federmeier and Kutas (1999), contextual facilitation effects can occur in different tasks (and affect, e.g., reading time, lexical decision times, pronunciation time); effects appear both at the level of lexical priming and at the level of the entire sentence. Based on this assumption, we aim to clarify the relation between single contextual words and the entire sentence. We constructed sentence materials that provide a differential number of context words that bias comprehension towards the target concept. This design allows us to determine how such facilitation (or bias) occurs, and to distinguish two possible hypotheses: the *incremental activation hypothesis*, which suggests that the degree of facilitation in concept processing increases with the amount of context available; and the *immediate activation hypothesis*, which states that once a sufficient amount of contextual support is reached, no more facilitation occurs.

We performed two experiments. In a self-paced reading experiment, we tested if the amount of context available has an effect on the ease of conceptual processing, measured as the reading time for the target concept. In a visual world study, we presented the target concept pictorially, allowing us to measure the degree of facilitation (i.e., the number of looks) that occurs while the context words are processed, giving us access to the time course of conceptual integration.

Experiment 1: Self-Paced Reading

The aim of this experiment was to analyze the effect that context information has on participants' reading time for the target concept. As widely discussed in the reading literature (Morris, 1994), higher coherence between the context and the target word is reflected in lower reading time. We therefore predict that reading times for a target word representing a

In Markus Knauff, Natalie Sebanz, Michael Pauen, and Ipke Wachsmuth, eds., *Proceedings of the 35th Annual Conference of the Cognitive Science Society*, 460–465. Berlin, 2013.

concept are reduced in proportion to the number of context words presented that bias the reader towards the concept.

Method

Materials We used the same 24 concepts as Huettig et al. (2006) (but we dropped their semantically related condition, as it is not relevant for the present experiment). We embedded these concepts into a sentential context using the following general structure:

(1) *location* – *actor* – verb – *object* – **target concept** – spill-over region

The target word is in **bold**, the three context words in *italics* (see below for an example). For each target concept, we identified three context words which were highly related to it (high-biasing (HB) words) and three context words that were unrelated to it (low-biasing (LB) words). Of the resulting eight possible combinations of LB and HB context words, we chose four, illustrated by the following examples:

(2) `All LB context` (*None*): On the *path*, the *man* was holding a *box* full of **mushrooms** carefully.

(3) `HB location context` (*Loc*): In the *forest*, the *man* was holding a *box* full of **mushrooms** carefully.

(4) `HB location and actor context` (*LocAct*): In the *forest*, the *picker* was holding a *box* full of **mushrooms** carefully.

(5) `All HB context` (*All*): In the *forest*, the *picker* was holding a *basket* full of **mushrooms** carefully.

This resulted in 96 experimental sentences: four contexts for each of the 24 concepts.

Norming Studies In order to make sure that the context words we chose had the biasing effect we expected them to have, we conducted a series of norming studies on Amazon Mechanical Turk.

First, 20 participants performed a sentence plausibility judgment task: they assessed how plausible the experimental sentence was by rating it on a scale from 1 (completely implausible) to 7 (completely plausible). A sentence was considered plausible when the averaged rating was higher than 4. This process allowed us to identify those sentences that were not completely plausible; they were replaced and re-tested.

A sentence completion study then evaluated the predictability of the target concept from the sentence context. Twenty new participants had to complete each sentence (with the target concept removed) by typing in a noun. An Anova showed a statistically significant difference between the *None* (6.4% of correct answers) and the *All* (43%) condition ($F(1,3) = 1.84, p < .001$). The *Loc* (19.2%) and the *LocAct* (14.4%) were not significantly different ($F(1,3) = 0.64$ and $F(1,3) = 0.40$).

Twenty-four further participants performed a word completion study: the three context words appeared on the screen one after the other and then participants had to type a word related to the context word. The aim of this study was to exclude syntactic effects (word order, but also the effect of the verb) in the completion study. The outcome of this experiment was in line with the sentence completion experiment, despite the fact that the context words were presented in isolation rather than in a sentence.

Finally, sixty participants performed a word association study in order to test the associations between the six context words (three LB and three HB words) and the target word in the sentence. According to an Anova, the three HB context words were equally strongly associated with the target word. Furthermore, all of them were more associated with the target than the LB words.

Procedure The 96 experimental sentences were distributed over four lists of 24 items each according to a Latin square design. Twenty-seven fillers were added and the list randomized for each participant. Twenty yes/no questions about the sentence were also included.

Thirty-four native English speakers from the University of Edinburgh took part in the experiment after giving informed consent and were paid £5. Each saw one of the lists. We excluded one participant based on a low percentage of correct answers ($< 50\%$) and another participant with a reading time averaging 2.5 standard deviations above the grand mean (as suggested in Hofmeister (2011)).

We used a moving-window self-paced reading procedure (Just, Carpenter, & Woolley, 1982), in which participants read a sentence on the screen at their own pace. At the beginning of each trial, all the words in the sentence are masked with dashes and separated by spaces; participants had to press the space bar to uncover the next word and hide the previous one. For this experiment, we used the software package Linger (version 2.94) on Apple computers.

Results

We analyze the reading times associated with the target word, which can be assumed to index the amount of effort associated with processing the target concept. Table 1 shows the mean reading times for the target concept across the four context conditions. The results indicate that reading time decreases in proportion with the number of HB words in the context. Table 2 reports a linear mixed-effects model (LME) with log-transformed reading time as the dependent variable. In the model, the factors `Loc`, `LocAct`, and `All` were contrast coded against the reference level `None`. `Participant` and `Item` were included as random intercepts, and as random slopes under `Loc`, `LocAct`, and `All` (thus implementing a maximal random effects structure, D. Barr, Scheepers, and Tily (2013)). The LME model shows that *LocAct* and *All* are statistically different from the *None* condition. After a HB context, participants spend less time reading the target concept compared to a LB context. The difference in reading time between *None* and the *Loc* does not reach significance.

Condition	Reading Time
None	390.1 ± 25.2
Loc	358.5 ± 21.4
LocAct	346.0 ± 15.5
All	336.7 ± 10.5

Table 1: Reading time (in ms) with standard errors for the target concept in the four contextual conditions.

Predictor	Coefficient
(Intercept)	12.771^{***}
Loc	-0.067
LocAct	-0.097^{*}
All	-0.094^{*}
$^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$	

Table 2: Coefficients for the mixed effects model for the reading time data in Table 1.

Discussion

In this self-paced reading experiment, we looked at the relation between the amount of HB information in the context and the reading time of the target concept. The hypotheses we started with predict two distinct outcomes: the incremental activation hypothesis predicts that the reading time of the target concept is reduced in proportion to the number of HB context words that are present. On the other hand, based on the immediate activation hypothesis, we expect that there is a threshold on the amount of contextual information that is required before an effect of context on reading time occurs.

Descriptively, the reading times in Table 1 are compatible with incremental activation: each additional contextual concept results in a further reduction in reading time. However, the mixed model analysis in Table 2 shows significant differences only between *LocAct* and *None* and *All* and *None*; *Loc* on its own does not have a significant effect. Furthermore, post-hoc tests failed to show significant differences between the reading times in the three context conditions (*Loc*, *LocAct* and *All*). This is a pattern that would be expected under the immediate activation hypothesis: the contextual threshold has been reached at *LocAct*, and additional context words do not significantly pre-activate the target concept any further.

Taken together, the results of this experiment are inconclusive. We therefore performed a follow-up experiment which directly tests our two hypotheses by measuring the amount of activation the target concept receives during the processing of each context word.

Experiment 2: Visual World

The aim of this experiment was to test whether the activation of concepts happens gradually (more activation with every new context word, as predicted by incremental activation), or at once (the first context word triggers full activation, which then declines, as predicted by immediate activation). In a visual world study, we measure the amount of activation for the target concept at each context word in terms of the proportion of looks received by the object corresponding to the target concept.

Method

Materials We used visual scenes that consisted of four black and white line drawings extracted from the Snodgrass and Vanderwart (1980) collection: one target object and three distractors randomly arranged in four quadrants. (These were the same stimuli as in Frassinelli and Keller (2012), already normed by Huettig et al. (2006)). The sentence materials were the same as in Experiment 1, and the stimuli instantiated the same 24 target concepts. As an example, consider the visual stimulus in Figure 1, which corresponds to the sentences in (2)–(5).

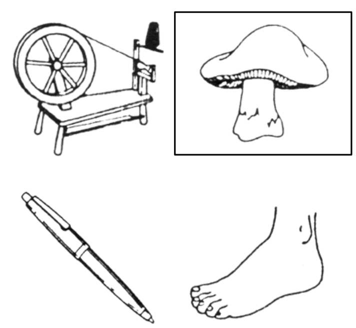


Figure 1: Example of the scene for the target concept **mushroom** (the box is not shown to the participants).

Procedure The 96 sentences in the experiment were spoken by the speech synthesis system Festival (Clark, Richmond, & King, 2007) using an HMM voice (Roger), so as to reduce possible effects of prosody or speaker variation.

In order to counterbalance order or position effects, we rotated the four objects on the screen. The resulting 384 items were distributed over 32 lists of 24 items each according to a Latin square design. Twenty-five fillers were added and the list randomized for each participant. Nine yes/no questions about the sentence were also presented.

Thirty-four native English speakers from the University of Edinburgh took part in the experiment after giving informed consent and were paid £5. Each saw one of the lists.

Participants were sat in front of a 21” multi-scan monitor with a resolution of 1024 $\times$ 768 pixels and their eye movements were recorded using an EyeLink II head-mounted eye-tracker with a sampling rate of 500 Hz. Only the dominant eye was tracked. At the beginning of the experiment and after every ten trials, the eye-tracker was recalibrated using a nine-point randomized calibration. Before each trial, drift correction was performed. At the beginning of each trial the scene appeared on the screen, and the sentence began to play at the same time; the scene disappeared 1500 ms after the end of the sentence. The experiment lasted approximately 20 minutes.

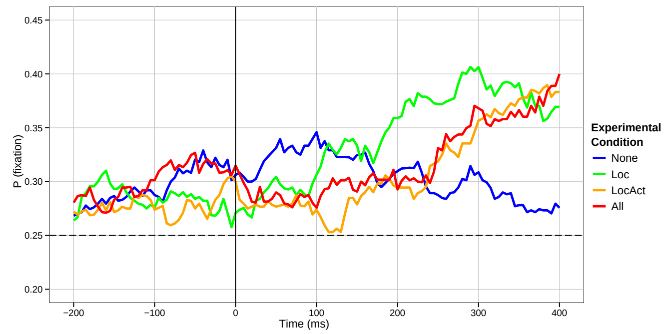
Data Analysis The analysis is based on the proportion of fixations on the target object across experimental conditions. We excluded out-of-screen fixations and blinks from the analysis.

In order to analyze the effects exerted by a context word before and after its acoustic offset, we aligned the fixation probabilities at that point (0 ms). In order to exclude any overlap between two regions of analysis in the sentence we calculated the minimum amount of time between the onset and the offset of the context word (150 ms) and between the offset of the context word and the onset of the following one (400 ms for location, actor, and concept; 150 ms for object). The vertical line shows the offset of the context word, while the horizontal dotted line indicates the probability of randomly fixating on one of the four objects depicted on the screen (25% of total fixations).

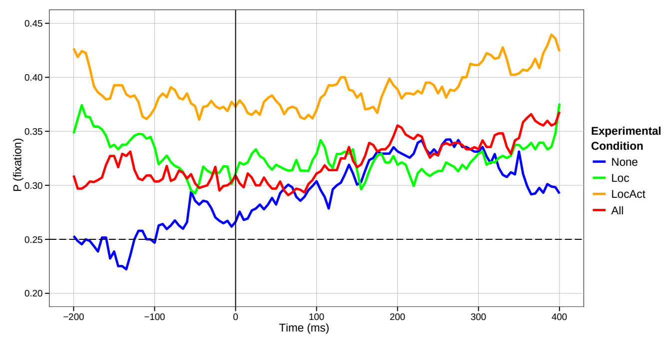
For each context word we report an LME analysis of the results. As suggested by D. J. Barr (2008), the dependent variable of our models is the *empirical logit* of the fixation probability calculated for each bin. We used a bin size of 5 ms. To compare the effects produced by HB and LB contexts, we included three factors in contrast coding: each factor encodes the differences between the reference level `None` (coded as $-.5$) and one of the three other conditions (`Loc`, `LocAct`, `All`; coded as .5). The continuous factor `Time` shows variations over time. In order to identify the minimal model that best fits our data, we used the best-path forward model selection procedure (recommended by D. Barr et al. (2013) if a model with full random effects structure fails to converge). We report only the coefficients and the significance levels for the minimal model, i.e., we show only the main effects and the interactions included during the selection procedure. All models included `Participant` and `Item` as random intercepts, as well as random slopes for `Context` and `Time`.

Results

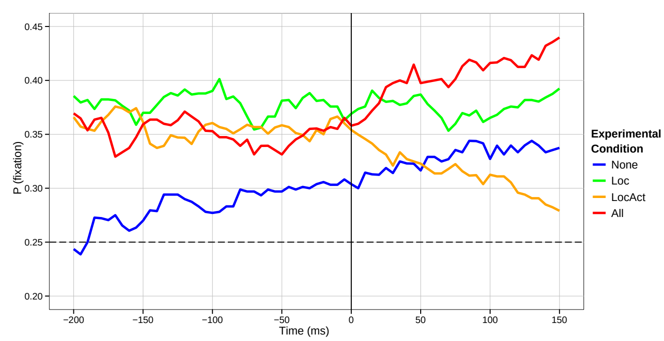
Location Word The first context word we analyze is location. The plot in Figure 2(a) shows the probability of fixating the target object at this context word. Before its offset, it is already possible to identify a general effect produced by the presence of a location (both HB and LB), as the fixation probabilities are higher than random. However, specific effects appear only 100 ms after the offset of the context word. The *None* (low biasing) condition shows a decrease over time, while an increase in fixation probability is observed in the *Loc* and *LocAct* conditions (compared to *None*, the reference level), which corresponds to the significant interactions `Time:Loc` and `Time:LocAct` (see Table 3, column 1). A similar effect is visible for *All*, but fails to reach significance.



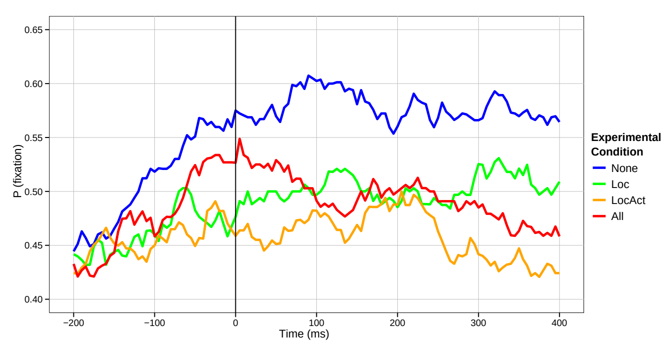
(a) *Location*: Target Fixation probability



(b) *Actor*: Target Fixation probability



(c) *Object*: Target Fixation probability



(d) *Concept*: Target Fixation probability

Figure 2: Fixation probabilities aligned at the offset (0 ms) of the *context words*.

Predictor	Coefficient Location	Coefficient Actor	Coefficient Object	Coefficient Concept
(Intercept)	-0.7548^{***}	-0.6764^{***}	-0.6215^{***}	-0.0138
Time	0.0005	0.0002^{***}	0.0003	0.0003^{***}
Loc	0.1019	-0.7303	0.1955	-0.1110
LocAct	-0.0388	0.4809	-0.0744	-0.2919
All	-	-0.1136	0.2204	-0.0849
Time:Loc	0.0008^{***}	-0.0011^{***}	-	0.0003^{**}
Time:LocAct	0.0003^{**}	-	-0.0024^{***}	-0.0007^{***}
Time:All	-	-	0.0011^{***}	-0.0004^{***}

$^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$

Table 3: Coefficients for the mixed effects model for the data in Figures 2(a), 2(b), 2(c), 2(d). Empty cells indicate that the factor in question was not included during model selection.

Actor Word The plot in Figure 2(b) shows the fixations at the word encoding an actor. In the *Loc* condition, participants tend to fixate less on target object (compared to *None*), an effect that is more evident before the offset of the word. This corresponds to a significant negative interaction `Time:Loc` in the LME (see Table 3, column 2). The plot also seems to indicate an overall higher level of fixations in the *LocAct* condition (compared to *All*, which is identical at this point in the sentence). However, this difference is not significant (no main effect or interaction involving *LocAct* in the LME).

Object Word The next context word analyzed is the object of the sentence. Figure 2(c) shows that before the offset of the object, the HB conditions all show a higher overall fixation proportion compared to *None*. After the offset, the curves diverge: *All* shows a steeper increase than *None* (significant positive interaction `All:Time`, see Table 3, column 3). This is explained by the fact that *All* is the only condition with a HB object. The condition *LocAct*, shows a steep decrease, i.e., the significant negative interaction `LocAct:Time`, while *Loc* remains constant (no significant effects involving `Loc`).

Concept Word Figure 2(d) shows the number of fixations at the point when participants hear the concept associated with target object on the screen (note the different y-axis). At this point, global effects of different amounts of HB information across conditions should be visible. After the offset of the context word, there is an inverse relation between the amount of HB information and the slope of the curves in the *Loc*, *LocAct*, and *All* conditions. The more HB information is available, the sooner fixation proportions decrease. This is consistent with the pattern observed in the reading times of the concept word in Experiment 1.

On the other hand, fixation probability in the *None* condition increases, in particular after word offset, and remains high. The negative interactions `Time:LocAct` and `Time:All` (see Table 3, column 4) are consistent with this observation, indicating a significant decrease in fixations in *LocAct* and *All* compared to *None*. Furthermore, there is a significant positive interaction `Time:Loc`, indicating an increase in fixation probability in this condition compared to *None*.

Discussion

The aim of this experiment was to analyze the effect of incremental context information over time. The analysis of actor and object (see Figure 2(b) and Figure 2(c)) showed few interesting effects. The regions at which it was possible to identify a clear effect of contextual variability were location and concept. Location is the first context word participants are exposed to and it had a strong effect on driving their fixations towards the target concept. This is in line with previous results of visual world studies on language comprehension (Altmann & Kamide, 1999), showing anticipatory eye movements towards a target as a result of predictive spoken language input. Less expected are the outcomes related to the concept area. We found that high-biasing contexts allows participants to identify and process the target object at an early stage: this effect is visible even at the first context word (location in our case). At the concept word itself, we then fail to observe a sustained increase in fixations to the target object. The opposite pattern was observed in the low-biasing context: were we see no increase in fixations at the context words, but a sustained increase once the concept word has been processed. In an HB context, the target word is contextually expected, and thus fixated less, while in the LB context, it is unexpected and thus fixated more.

One possible explanation for this pattern of results (i.e., a decrease in target fixations at the target word after a biasing context) is inhibition of return. This is a well-know effect in eye-movements, which manifests itself in a low probability of returning to a region once it has been fixated (Posner, Rafal, Choate, & Vaughan, 1985). Our failure to find an increase in fixations at the concept region in the HB conditions could be due to inhibition of return, as the target had already been fixated at an earlier point in these conditions (i.e., during anticipatory processing while hearing biasing context words).

General Discussion

The aim of this study was to test the effect that an incremental amount of contextual information exerts on conceptual processing. We ran two experiments: a self-paced reading and a visual world experiment. The former identified a reduction of the time required to read a target concept with increasing amount of biasing context. The visual world study showed increased looks to the target during the processing of the first context word (location) and at the concept word itself.

The results obtained in our visual world study corroborate the findings of Frassinelli and Keller (2012) who found that a biasing context leads to an early recognition and processing of the target concept. Moreover, the results show an expectation effect: over time, the biasing contexts produced an expectation of the target concept and this led to a reduced number of fixations to the target object when the corresponding concept word occurred (potentially involving inhibition of return as the driving mechanism behind this decrease in fixation probability).

The key novel contribution of the paper is to elucidate the time course of contextual integration. In the introduction, we proposed two possible hypothesis on how context interacts with concept processing: the *incremental activation hypothesis* (the degree of facilitation increases with the amount of context) and the *immediate activation hypothesis* (once sufficient contextual support has accrued, no additional facilitation occurs). The results of Experiment 2 allow us to evaluate these two hypotheses. We found that when participants are exposed to a low-biasing context, we see an increase of fixations to the target concept only when that concept is mentioned. In the high-biasing conditions, on the other hand, this increase occurs already at the first context word, with no further increases at the second or third context word. Also when at the concept word, only a small increase in fixations is observed. It seems that a single context word is sufficient to identify the concept on the screen. Additional context exerts only a confirmatory effect. This pattern of results is compatible with the immediate activation hypothesis: a certain amount of contextual information is sufficient to trigger conceptual processing; additional contextual information does not trigger an incremental increase in concept activation.

In more theoretical terms, these results enhance our understanding of conceptual representation. They indicate that the activation of a specific concept takes place when the overlap between its internal structure and the information extracted from the context reach a certain critical level.

Acknowledgments

This work was funded by the European Research Council under award 203427 "Synchronous Linguistic and Visual Processing". We are grateful to Moreno I. Coco for comments.

References

Altmann, G., & Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition*, *73*(3), 247–264.

Barr, D., Scheepers, C., & Tily, H. (2013). Random-effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255-278.

Barr, D. J. (2008). Analyzing 'visual world' eyetracking data using multilevel logistic regression. *Journal of Memory and Language*, *59*(4), 457–474.

Clark, R., Richmond, K., & King, S. (2007). Multisyn: Open-domain unit selection for the festival speech synthesis system. *Speech Communication*, *49*(4), 317–330.

Federmeier, K., & Kutas, M. (1999). A rose by any other name: Long-term memory structure and sentence processing. *Journal of memory and Language*, *41*(4), 469–495.

Frassinelli, D., & Keller, F. (2012). The plausibility of semantic properties generated by a distributional model: Evidence from a visual world experiment. In N. Miyake, D. Peebles, & R. P. Cooper (Eds.), *Proceedings of the 34th annual conference of the cognitive science society* (pp. 1560–1565). Sapporo.

Hofmeister, P. (2011). Representational complexity and memory retrieval in language comprehension. *Language and cognitive processes*, *26*(3), 376–405.

Huettig, F., Quinlan, P. T., McDonald, S. A., & Altmann, G. T. M. (2006). Models of high-dimensional semantic space predict language-mediated eye movements in the visual world. *Acta Psychologica*, *121*(1), 65–80.

Just, M., Carpenter, P., & Woolley, J. (1982). Paradigms and processes in reading comprehension. *Journal of Experimental Psychology: General*, *111*(2).

McRae, K., Cree, G., Seidenberg, M., & McNorgan, C. (2005). Semantic feature production norms for a large set of living and nonliving things. *Behavior Research Methods*, *37*(4), 547–559.

Morris, R. (1994). Lexical and message-level sentence context effects on fixation times in reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*(1), 92.

Murphy, G. L. (2002). *The Big Book of Concepts*. Cambridge, MA: MIT Press.

Posner, M., Rafal, R., Choate, L., & Vaughan, J. (1985). Inhibition of return: Neural basis and function. *Cognitive Neuropsychology*, *2*(3), 211–228.

Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology. Human Learning and Memory*, *6*(2), 174–215.

Tanenhaus, M., Spivey-Knowlton, M., Eberhard, K., & Sedivy, J. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science*, *268*(5217), 1632–1634.

Wu, L., & Barsalou, L. W. (2009). Perceptual simulation in conceptual combination: Evidence from property generation. *Acta psychologica*, *132*(2), 173–89.