Seeing what's next: Processing and anticipating language referring to objects

Cover credits: text-image.com and mazwebs at sxc.hu

ISBN: 978-90-76203-52-2

Printed and bound by Ipskamp Drukkers b.v.

Seeing what's next: Processing and anticipating language referring to objects

Proefschrift

ter verkrijging van de graad van doctor

aan de Radboud Universiteit Nijmegen

op gezag van de rector magnificus prof. mr. S.C.J.J. Kortmann,

volgens besluit van het college van decanen

in het openbaar te verdedigen op maandag 25 november 2013

om 13.00 uur precies

door

Joost Rommers

geboren op 10 december 1985

te Rijswijk

Promotor:

Prof. dr. A.S. Meyer

Copromotor:

Dr. F. Huettig (MPI for Psycholinguistics)

Manuscriptcommissie:

Prof. dr. J.M. McQueen

Prof. dr. H. Schriefers

Dr. C. Dobel (University of Münster)

The research reported in this thesis was supported by the Max-Planck-Gesellschaft zur Förderung der Wissenschaften, München, Germany.

Contents

Chapter 1

General Introduction

General introduction

Adult readers and listeners processing language activate, construct, and integrate the meaning of the input they encounter using a wealth of prior knowledge built up over years of experience. While much of this knowledge is likely to be quite abstract in nature (for example, the meaning of the word "meaning"), some of the knowledge represents directly observable physical aspects of the world around us, such as the appearance of objects. For instance, reading the word "apple" might activate knowledge of the round shape of the object. This thesis is concerned with the mechanisms and representations underlying the processing of references to objects in sentence context.

Processing references to objects

Several studies suggest that listeners and readers activate visual representations of objects when they are referred to by words. For example, lexical decision responses to words (e.g., “apple”) have been found to be faster when a prime word with a visually similar referent (e.g., “ball”, also a round object) preceded it than when an unrelated prime preceded it (Schreuder, Flores d'Arcais, and Glazenborg, 1984). Even implied changes of the shape of an object appear to be able to influence performance. For instance, a sentence such as “The ranger saw the eagle in the sky” implies that the eagle had its wings stretched out, whereas the sentence “The ranger saw the eagle in the nest” implies that the eagle was sitting with its wings tucked in. In an experiment where participants read one of these sentences, this visual information influenced subsequent performance on a visually presented object (e.g., the eagle). Specifically, decision latencies on whether or not the object had been mentioned in the preceding sentence were found to be shorter when the object's shape matched the shape implied in the preceding sentence than when these shapes did not match (Zwaan, Stanfield, & Yaxley, 2002). It has been claimed that such visual representations are routinely activated (Wassenburg & Zwaan, 2010). Visual representations play an important role in theories of embodied cognition according to which language, thought, and memory re-activate sensory experiences. For example, Prinz (2002, p.119) has suggested that conceptual knowledge consists of "representational codes that are specific to our perceptual systems". Similarly, Barsalou (1999, p.577) has argued that "perceptual memory can function symbolically, standing for referents in the world, and entering into symbol manipulation. As collections of perceptual symbols develop, they constitute the representations that underlie cognition". Given the importance ascribed to visual representations in such views, it would be of considerable interest to investigate whether visual representations are activated in a stable manner in different task environments. This thesis addressed the generality of the influence of visual representations on performance by investigating whether effects of implied shape and orientation can be obtained across a range of tasks or instead are task-dependent (Chapter 2).

Anticipating references to objects

As readers and listeners interpret sentences, the activation of knowledge does not only include many different kinds of information, it also occurs remarkably rapidly. Several studies attest to a strong form of incremental sentence comprehension: prediction (Altmann & Kamide, 1999; DeLong, Urbach, & Kutas, 2005; Federmeier & Kutas, 1999; van Berkum, Brown, Kooijman, Zwitserlood, & Hagoort, 2005; Wicha, Moreno, & Kutas, 2004). Thus, listeners and readers do not only process words as soon as they come in, they can also think ahead about what might be coming next. Prediction plays an important role in current views on language comprehension and cognition in general, where the brain is sometimes portrayed as a "prediction machine". For instance, minimizing "free energy" or "prediction error" (the difference between what is predicted and what is actually presented; Friston, 2005) has been proposed to be a fundamental principle of brain function (Friston, 2010; Huang, 2008). Now that several studies have established that listeners and readers can predict upcoming information, the next step is to investigate more detailed aspects of the predictive process.

This thesis treats cues, contents, and mechanisms as three basic components of predictive language processing. The cues refer to the parts and kinds of context that are used to generate predictions. The contents of predictions refer to the information that becomes pre-activated. Finally, the term mechanisms refers to the underlying processes that enable readers and listeners to move from the (conscious or unconscious) identification of one or more cues to the generation of one or more predictions, and how these predictions influence performance.

Take as an example the sentence “In 1969 Neil Armstrong was the first man to set foot on the moon”, where the last word is highly predictable. The cues that may lead to prediction of the word “moon” are present in the part of the sentence preceding the final word. For instance, the parts “1969” and “Neil Armstrong” are strong cues enabling the generation of the prediction that the word “moon” might be mentioned at some point later in the sentence. The sentence also contains syntactic structure, and, if spoken, prosodic information, which could both be used to anticipate when the word “moon” might be mentioned. The contents of the prediction may span a range of attributes of “moon”, from its syntactic category and its meaning to its phonological, acoustic, or orthographic form. Finally, the mechanisms involved in predicting that "moon" might be mentioned might be specific to language comprehension, or related to language production, or even to non-linguistic cognitive processes.

Several earlier studies have specified contextual cues that can be used to generate predictions. These cues range from general discourse context (Kaiser & Trueswell, 2004; van Berkum, Brown, Zwitserlood, Kooijman, & Hagoort, 2005), lexical-syntactic information (De Ruiter, Mitterer, & Enfield, 2006), syntactic structure (Arai & Keller, 2013), prosody (Weber, Grice, & Crocker, 2006), the selectional restrictions of verbs (Altmann & Kamide, 1999), combined information from subject and verb (Kamide, Altmann, & Haywood, 2003) and case-marking (Kamide, Scheepers, & Altmann, 2003) to visually presented events (Knoeferle, Crocker, Scheepers, & Pickering, 2005). The present thesis did not add further to that list, but instead focused on contents and mechanisms.

Regarding the contents of predictions, studies have suggested that listeners and readers can predict upcoming syntactic structure (Arai & Keller, 2013; Carminati, van Gompel, Scheepers, & Arai, 2008; Staub & Clifton, 2006), functional semantic aspects of word meaning (Altmann & Kamide, 1999; Federmeier & Kutas, 1999), specific upcoming lexical items (DeLong, Urbach, & Kutas, 2005; van Berkum, Brown, Kooijman, Zwitserlood, & Hagoort, 2005; Wicha, Moreno, & Kutas, 2004), and the orthographic forms of words (Dikker, Rabagliati, Farmer, & Pylkkänen, 2010; Laszlo & Federmeier, 2009). This thesis investigated whether object shape representations might also become pre-activated as listeners anticipate upcoming information (Chapter 3).

Regarding the mechanisms of predictions, some theoretical proposals have linked predictive language processing to the anticipation of events (Altmann & Mirković, 2009), and some proposed predictive mechanisms are so general that they are thought to be able to provide a unified account of perception, cognition, and action (Clark, 2013). Interestingly, not every reader or listener is expected to predict equally strongly or consistently, both in language processing and in other domains. This thesis therefore investigated whether there are relationships between individual differences in predictive behavior in verbal and non-verbal tasks (Chapter 4). Prediction in language comprehension has also been linked to language production. For instance, Pickering and Garrod (2007; in press) proposed that predictions could be generated using the language production system. On this account, listeners covertly produce what they would say as a speaker in the given context. This thesis investigated how far such production-comprehension links go, specifically whether any production-like motor preparation might occur when reading predictable sentences (Chapter 5).

Methodology

The method relied on for addressing the question of context-dependence in the processing of references to objects was the measurement of participants' response latency. Participants read sentences that implied a certain shape or orientation for an object, and the effects of this visual information on button press latencies and object naming latencies in response to visually matching and mismatching objects were examined. These methods have previously been shown to be influenced by visual representations that become activated during language processing (e.g., Stanfield & Zwaan, 2001).

For most research on prediction, methods that sample continuous signals at a high rate are useful. Two such methods were used: Eye-tracking and electroencephalography (EEG). Both of these methods have a temporal resolution in the order of milliseconds and do not require the participant to engage in any meta-linguistic tasks.

The eye-tracking experiments made use of the visual world paradigm, where participants listen to utterances while viewing a display of objects on a computer screen. An eye-tracker combined the location of the center of the pupil, the corneal reflection created by the use of infrared light, and a short calibration procedure to track the eye movements of the participants as they mapped spoken input onto visually presented objects. Eye movements recorded in this task have been shown to be closely linked to language comprehension (Cooper, 1974; Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995). A particularly useful feature for research on prediction is the phenomenon of anticipatory eye movements, which demonstrate that listeners can predict upcoming information. For instance, in an experiment by Altmann and Kamide (1999), participants heard sentences such as "The boy will eat the cake" (with a constraining verb) or "The boy will move the cake" (with a non-constraining verb). While listening, participants viewed a display in which a cake was the only edible object. After the constraining verb but not after the non-constraining verb, participants already looked more at the cake than at other objects before the cake had been mentioned. Chapters 3 and 4 investigated such anticipatory eye movements further.

The EEG experiments made use of paradigms where participants read or listened to predictable sentences. The EEG signal, recorded through electrodes placed on the scalp, represents fluctuations in electrical activity generated by the brain, embedded in a large amount of noise from other sources. In most analyses the EEG was averaged in the time domain across multiple trials to create event-related potentials (ERPs), time-locked to a stimulus of interest. The rationale behind the use of ERPs is that during such averaging, most of the random noise that is unrelated to the processing of the stimulus cancels out, making visible the stable components that are consistently evoked by stimuli. Such components have been labeled according to latency and polarity and have become associated with specific aspects of language processing. For instance, the N400, a negative component that peaks at around 400 ms after onset of a meaningful stimulus, is a sensitive index of semantic processing (for review, see Kutas & Federmeier, 2011). Chapters 3 and 5 examined effects of visual information and predictability on ERPs. Because during ERP averaging some information (called induced activity) is lost, additional analyses of the spectral content of the EEG signal were performed in one of the projects (Chapter 5).

Thesis outline

Chapter 2 addressed the issue whether visual representations influence performance in a routine (automatic) or context-dependent manner. To this end, participants read sentences that implied a particular visual shape or orientation of an object. After each sentence, a picture was presented. In different experiments, participants named the picture, decided whether it had been mentioned in the sentence they had just read, or named the picture after having created a visual image of the situation described in the sentence. If implied visual representations are routinely activated, it would be expected that these representations influence performance across the whole range of tasks.

Chapter 3 investigated the contents of predictions during sentence comprehension. Specifically, it was examined whether object shape representations can form part of the pre-activated information. Participants listened to predictable words (e.g., "moon") in sentences (e.g., the Dutch translation equivalent of "In 1969 Neil Armstrong was the first man to set foot on the moon"). In one experiment, their eye movements were tracked as they looked at four-object displays containing one critical object. Depending on the condition, the critical object corresponded to the predictable word's referent (e.g., the moon), or was an object with a visually similar shape (e.g., a tomato), or was an unrelated object (e.g., rice). If shape representations can form part of the contents of predictions, participants should launch anticipatory eye movements not only toward predictable words' referents, but also to visually similar objects. In another experiment, participants listened to the same sentences while their EEG was recorded. No visual displays were presented. Now the predictable words were replaced by words referring to objects with a referent with a similar shape (e.g., "tomato"), or to words with referents of a different shape (e.g., "rice"). If shape representations can form part of the contents of predictions, including in situations where no objects are visually presented, the two types of word replacement should be processed differently.

Predictive behavior occurs not only in language comprehension, but also in other cognitive tasks. Chapter 4 investigated whether predictive behavior in verbal tasks relates to predictive behavior in non-verbal tasks. Participants took part in a variant of the eye-tracking experiment of Chapter 3. In addition, their performance on several other tasks was assessed. One of the tasks was a spatial cueing task based on Posner, Nissen, and Ogden (1978), where participants pressed a button to indicate the location (left or right) of an X symbol on the screen. Before the X symbol was presented, they saw a briefly presented arrow cue that was predictive of the location of the upcoming stimulus. The degree to which participants' responses are facilitated by such arrow cues can be taken as a measure of anticipatory spatial attention. If language-mediated anticipatory eye movements are related to anticipatory spatial attention in nonverbal tasks, individual differences in performance on the two tasks should correlate. Several other tasks were included to evaluate possible alternative interpretations.

Finally, Chapter 5 addressed the potential role of the language production system in aspects of language comprehension, such as prediction. An EEG study examined whether this link could go as far as motor preparation for predictable items. Participants were presented with objects (e.g., a broom) after predictive lead-in sentences (e.g., “He swept the floor with a…”) or neutral lead-in sentences (e.g., “He saw the drawing of a…”). Participants named the objects or just viewed them. If reading a predictable sentence is similar to preparing to complete the sentence out loud, the effects of predictability on the EEG signal should be similar in both tasks. Chapter 6 summarizes and discusses the findings.

References

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

Altmann, G. T. M., & Mirković, J. (2009). Incrementality and Prediction in Human Sentence Processing. *Cognitive Science, 33*(4), 583-609.

Arai, M., & Keller, F. (2013). The use of verb-specific information for prediction in sentence processing. *Language and Cognitive Processes, 28(4),* 525-560.

Barsalou, L.W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences, 22*(4), 577-609.

Carminati, M.N., Van Gompel, R.P.G., Scheepers, C., & Arai, M. (2008). Syntactic priming in comprehension: The role of argument order and animacy. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34,* 1098-1110.

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences, 36(3),* 181-204.

Cooper, R. M. (1974). The control of eye fixation by the meaning of spoken language: A new methodology for the real-time investigation of speech perception, memory, and language processing. *Cognitive Psychology, 6*, 84-107.

DeLong, K. A., Urbach, T. P., & Kutas, M. (2005). Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. *Nature Neuroscience, 8*(8), 1117-1121.

Dikker, S., Rabagliati, H., Farmer, T. A., & Pylkkänen, L. (2010). Early occipital sensitivity to syntactic category is based on form typicality. *Psychological Science, 21(5),* 629-34.

De Ruiter, J. P., Mitterer, H., & Enfield, N. J. (2006). Projecting the end of a speaker's turn: A cognitive cornerstone of conversation. *Language, 82(3),* 515-535.

Federmeier, K. D., & 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.

Friston, K. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society London B: Biological Sciences*, 360(1456): 815-36.

Friston, K. (2010). The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience,*11(2), 127-38.

Huang, G. (2008). Is this a unified theory of the brain? *New Scientist,* 2658, 30–33.

Kaiser, E. & Trueswell, J.C. (2004). The role of discourse context in the processing of a flexible word-order language. *Cognition, 94(2),* 113-147.

Kamide, Y., Altmann, G.T.M., & Haywood, S.L. (2003). Prediction and thematic information in incremental sentence processing: Evidence from anticipatory eye movements. *Journal of Memory and Language, 49,* 133-156.

Kamide, Y., & Scheepers, C., & Altmann, G.T.M. (2003). Integration of syntactic and semantic information in predictive processing: Cross-linguistic evidence from German and English. *Journal of Psycholinguistic Research, 32,* 37-55.

Knoeferle, P., Crocker, M.W., Scheepers, C., & Pickering, M.J. (2005). The influence of the immediate visual context on incremental thematic role-assignment: evidence from eye-movements in depicted events. *Cognition, 95,* 95-127.

Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: Finding meaning in the N400 component of the event-related brain potential (ERP). *Annual Review of Psychology, 62*, 621-647.

Laszlo, S. and Federmeier, K. D. (2009). A beautiful day in the neighborhood: An event-related potential study of lexical relationships and prediction in context. *Journal of Memory and Language, 61,* 326-338.

Pickering, M. J., & Garrod, S. (2007). Do people use language production to make predictions during comprehension? *Trends in Cognitive Sciences, 11*(3), 105-110.

Pickering, M. J., & Garrod, S. (in press). An integrated theory of language production and comprehension. *Behavioral and Brain Sciences*.

Posner, M. I., Nissen, M. J., & Ogden, W. C. (1978). Attended and unattended processing modes: The role of set for spatial location. In: H.L. Pick, & I.J. Saltzman (Eds.), *Modes of perceiving and processing information*. Hillsdale, N.J.: Lawrence Erlbaum Associates.

Prinz, J. J. (2002). *Furnishing the mind: concepts and their perceptual basis*. Cambridge, MA: MIT Press.

Schreuder, R., Flores d'Arcais, G.B., & Glazenborg, G. (1984). Effects of perceptual and conceptual similarity in semantic priming. *Psychological Research, 45*, 339-354.

Stanfield, R.A., & Zwaan, R.A. (2001). The effect of implied orientation derived from verbal context on picture recognition. *Psychological Science, 12*(2), 153-156.

Staub, A., & Clifton, C., Jr. (2006). Syntactic prediction in language comprehension: Evidence from either...or. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32,* 425-436.

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

van Berkum, J. J. A., Brown, C. M., Zwitserlood, P., Kooijman, V., & Hagoort, P. (2005). Anticipating upcoming words in discourse: Evidence from ERPs and reading times. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*(3), 443-467.

Wicha, N. Y. Y., Moreno, E. M., & Kutas, M. (2004). Anticipating words and their gender: An event-related brain potential study of semantic integration, gender expectancy, and gender agreement in Spanish sentence reading. *Journal of Cognitive Neuroscience, 16*(7), 1272-1288.

Wassenburg, S.I., & Zwaan, R.A. (2010). Readers routinely represent implied object rotation: The role of visual experience. *Quarterly Journal of Experimental Psychology, 63*(9), 1665-1670.

Weber, A., Grice, M., & Crocker, M. W. (2006). The role of prosody in the interpretation of structural ambiguities: A study of anticipatory eye movements. *Cognition, 99,* B63-B72.

Zwaan, R.A., Stanfield, R.A., & Yaxley, R.H. (2002). Language comprehenders mentally represent the shapes of objects. *Psychological Science, 13*(2), 168-171.

Chapter 2

Object shape and orientation do not routinely influence performance during language processing

Rommers, J., Meyer, A. S., & Huettig, F. (in press). Object shape and orientation do not routinely influence performance during language processing. *Psychological Science.*

Abstract

The role of visual representations during language processing remains unclear: They could be activated as a necessary part of the comprehension process, or they could be less crucial and influence performance in a task-dependent manner. In the present study participants read sentences which implied a specific shape or orientation of an object. Subsequently they named a picture of that object (Experiments 1 and 3) or decided whether it had been mentioned in the sentence (Experiment 2). Orientation information did not reliably influence performance in any of the experiments. Shape representations influenced performance to a degree dependent upon specific circumstances, namely most strongly when participants were asked to compare a sentence to a picture or when they were explicitly asked to use mental imagery while reading the sentences. Thus, in contrast to previous claims, implied visual information often does not provide a crucial contribution to the comprehension process during normal reading.

Object shape and orientation do not routinely influence performance during language processing

Much research has suggested that listeners and readers activate visual and motor representations of objects that are referred to in spoken or written utterances (for review, see Zwaan, 2004). Since Barsalou's (1999) seminal paper, many of these findings have been interpreted in terms of *embodied cognition*, which is the view that high-level cognitive processes such as language, memory, and thought involve re-enactment or simulation of perception and action states. This idea contrasts sharply with amodal theories of knowledge (e.g., Kintsch, 2008), where the format of high-level cognitive processes is abstracted away from perceptual processes. However, amodal propositions or features can in principle also be used to capture perceptual representations, and theories of embodied cognition assume that "a simulator produces simulations that are *always* partial and sketchy, *never* complete" (Barsalou, 1999, p.586; italics in the original). It is therefore doubtful whether any experimental evidence could falsify either theory with regard to representational format (cf. Mahon & Caramazza, 2008; for an analogous debate in mental imagery, see Kosslyn, 1994; Pylyshyn, 1981).

Regardless of whether the representational format of conceptual representations is modal, amodal, or both, from a psycholinguistic viewpoint the evidence for the activation of visual representations during language processing remains intriguing. The question arises whether perceptual representations deserve a more explicit role in models of language processing than is currently the case. To answer this question, the conditions under which such representations become activated must be determined.

One hypothesis is that the activation of visual representations is a crucial component of language processing: Whenever we process a word referring to an object (e.g., a tomato), we activate a visual representation specifying its shape (e.g., round) and this representation forms a necessary part of the situation model built up during comprehension. Thus, visual representations may routinely be activated during language processing (e.g., Wassenburg & Zwaan, 2010).

This claim is supported by studies showing that visual representations were activated although they were presumably not needed for the task. For instance, in a sentence-picture verification study, Zwaan, Stanfield, and Yaxley (2002) presented participants with sentences implying a certain shape of an object (e.g., "The ranger saw the eagle in the sky" or "The ranger saw the eagle in the nest", implying that the eagle had its wings stretched out or tucked in, respectively). Participants were faster to decide that a visually presented object (e.g., an eagle) had been mentioned in the sentence when its shape corresponded to the shape implied in the sentence than when these shapes mismatched. This suggests that visual representations of shape were activated. The same pattern was seen when the verification task was replaced by picture naming.

Similarly for spatial orientation, Stanfield and Zwaan (2001) found that participants were faster to indicate that an object (e.g., a vertically oriented nail) had been mentioned in a preceding sentence when its orientation matched the orientation implied in that sentence (e.g., "The carpenter hammered the nail into the floor") than when these orientations mismatched (e.g., "The carpenter hammered the nail into the wall"). Recently, Zwaan and Pecher (2012) replicated the orientation and shape effects in an online study, using the sentence-picture verification task. Pecher, van Dantzig, Zwaan, and Zeelenberg (2009) observed the same effects in a memory task, where old/new judgments on the pictures were influenced by implied orientation and shape from sentences read 45 minutes earlier. This suggests that representations of shape and orientation were activated during sentence reading and retained over time.

An alternative hypothesis is that visual representations are not crucial for language comprehension. Visual representations become activated during language processing, but this may be a by-product of the way information cascades through the cognitive system (Mahon & Caramazza, 2008). Moreover, whether, or to what extent, visual representations are activated may depend on the particular situation in which a concept is instantiated. A prediction from such a view is that the activation of visual representations is limited to certain task situations.

Although perceptual and/or motor representations have often been claimed to become activated automatically or routinely (e.g., Pulvermüller, 2005; Wassenburg & Zwaan, 2010), some proponents of the embodied cognition hypothesis have also suggested that the activation of perceptual representations in linguistic or conceptual tasks may be context-dependent (e.g., van Dam, van Dijk, Bekkering, & Rueschemeyer, 2012). For instance, Pecher, Zeelenberg, and Barsalou (2003) observed that property verification on a given trial (e.g., "apple" - "green") was faster and more accurate after a trial involving the same sensory modality (e.g., "diamond" - "sparkle") than after a trial involving a different modality (e.g., "airplane" - "noisy"). Furthermore, Solomon and Barsalou (2004) found that performance in property verification depended upon perceptual variables, but only when thorough processing was encouraged through the presence of associatively related fillers. Moreover, studies that have investigated priming between shape-related word pairs have yielded inconsistent results. In a lexical decision task, Schreuder, Flores d'Arcais, and Glazenborg (1984) observed perceptual priming (e.g., "ball" - "apple"), but Pecher, Zeelenberg, and Raaijmakers (1998) failed to replicate the priming effect, except when visual representations had been emphasized by shape decision tasks prior to the experiment. Recall that our aim was not to evaluate (dis)embodied theories, but to investigate the orthogonal issue whether visual representations routinely influence performance.

Other support for context-dependence comes from visual world eye-tracking studies, where participants listen to spoken sentences while simultaneously viewing arrays of objects. Here, upon hearing or anticipating a spoken word (e.g., "snake"), participants rapidly shift their eye gaze to visually similar objects (e.g. a cable; Dahan & Tanenhaus, 2005; Huettig & Altmann, 2007; Rommers, Meyer, Praamstra, & Huettig, 2013). However, this does not happen when the objects are replaced by printed words (Huettig & McQueen, 2011). Regarding color representations, while Connell and Lynnott (2009) observed facilitation from congruent colors, Connell (2007) observed interference. For orientation also both facilitation (Zwaan, Stanfield, and Yaxley, 2001) and interference (Richardson, Spivey, Barsalou, & McRae, 2003) have been reported. Finally, Kang, Yap, Tse, and Kurby (2011) failed to replicate an effect of object size reported by Sereno, O'Donnell, and Sereno (2009).

In the present study, we ask whether information about the shape and orientation of objects mentioned in written sentences routinely influences performance. The studies discussed above differed in many ways, including the materials, the timing of the stimuli, and the tasks. We used the same materials in all experiments, but, across experiments, varied the task. Participants read sentences and named pictures (Experiment 1), explicitly compared sentences to pictures (Experiment 2), or used imagery before naming pictures (Experiment 3). If the activation of visual representations is an inherent part of the comprehension process, it should influence performance across a range of processing tasks. If visual representations influence performance in a task-dependent manner, larger effects should be observed in Experiments 2 and/or 3 compared to Experiment 1.

Experiment 1

Participants

Fifty-two native speakers of Dutch (45 female) with an average age of 20 years (range 17-26 years) were paid for their participation. All had normal or corrected-to-normal vision and reported no language impairments.

Stimuli and design

The stimuli consisted of the 92 black-and-white picture-sentence quadruplets used in Pecher, van Dantzig, Zwaan, and Zeelenberg (2009). In 52 quadruplets the *orientation* of an object was manipulated, and in 40 quadruplets the *shape* of an object was manipulated, as in the studies discussed above. Each sentence-picture combination occurred in one of four counterbalanced lists along with 92 fillers. The pictures varied in size (width range: 63-720 pixels, height range: 63-540 pixels). The items occurred in random order with the same condition (match, mismatch, and filler) not occurring more than three times in succession.

Apparatus and Procedure

Participants were tested individually in a sound-damped booth, seated in front of a 17-inch iiyama HM703UT monitor, a buttonbox (built in-house) and a Sennheiser microphone. The experiment was performed using NBS Presentation® software (version 14.1). Each trial started with a black screen which was presented for 200 ms. Then a sentence appeared in white letters of the font Arial, size 15, in the center of the screen. After participants pressed a button to indicate that they had read the sentence, a white fixation cross appeared which remained on the screen for 250 ms. Then a picture was presented for 3 seconds. Participants named the picture and their speech was recorded. After every eighth trial there was an optional break. Participants were instructed to read each sentence carefully and to name each picture as quickly and accurately as possible.

Analysis

Naming latencies were manually measured using a speech waveform editor. Responses different from the object name mentioned in the preceding sentence were discarded (note that on experimental trials the object was always mentioned in the preceding sentence). One of the orientation items had to be excluded due to an error in the construction of the materials. The data analysis followed Zwaan and Yaxley (2002). For both the orientation and the shape items, the response latencies were aggregated to medians by participants and subjected to a 2 (Condition: match, mismatch) × 2 (Picture version) × 4 (List) repeated-measures ANOVA, with List as a between-subjects variable (because of the counterbalanced design, no item analyses are reported; cf. Raaijmakers, Schrijnemakers, & Gremmen, 1999).

As some of the experiments yielded null effects that are surprising given previous findings, we examined the effects seen in each experiment further in two ways. First, because traditional *p*-values never allow one to accept the null hypothesis, we computed an approximation to Bayesian posterior probabilities (p_{BIC}) from the ANOVA following Wagenmakers (2007) and Masson (2011). In case of a null effect we report the evidence for the null hypothesis, $p_{\mathrm{BIC}}(\mathrm{H}_0|\mathrm{D})$, otherwise we report the evidence for the alternative hypothesis, $p_{\mathrm{BIC}}(\mathrm{H}_1|\mathrm{D})$, but note that these sum to one. Second, we conducted a more powerful analysis using linear mixed-effects regression models (Baayen, Davidson, & Bates, 2008). The responses were log-transformed to reduce skewness and analyzed with a model including the fixed factor Condition (match, mismatch) and random intercepts and slopes by participant, picture, and sentence, using the R (version 10.2.1; R Development Core Team, 2009) libraries lme4 (version 0.999375-34) and languageR (version 1.0). The Match condition was mapped onto the intercept. This model was compared to a model without the fixed effect of Condition but with the same random effects structure, using a likelihood ratio test. Error rates were low and are reported in Table S1 in the supplemental Accuracy Analyses available online and in the Appendix of this chapter.

Results and discussion

Separate analyses were carried out for the orientation and shape items. For the *orientation* items, we discarded 128 incorrect responses and 2 trials due to a technical error (5%). The 9 ms difference between conditions (see Figure 1) was not significant, $F(1,48)=.999$, $p=.323$, $\eta^2_p=.020$.

There was a Condition × List interaction, *F*(3,48)=3.920, *p*=.014, η^2_p=.197, and a three-way Condition × Picture version × List interaction, *F*(3,48)=2.781, *p*=.051, η^2_p=.148. The $p_{BIC}(H_0|D)$ for the main effect of Condition was .808, which according to Raftery's (1995) classification constitutes "positive evidence" for the conclusion that there was no effect of orientation. The mixed-effects model did not indicate an effect of Condition either, β=.0084 (8 ms), $\chi^2(1)$=1.480, *p*=.224.

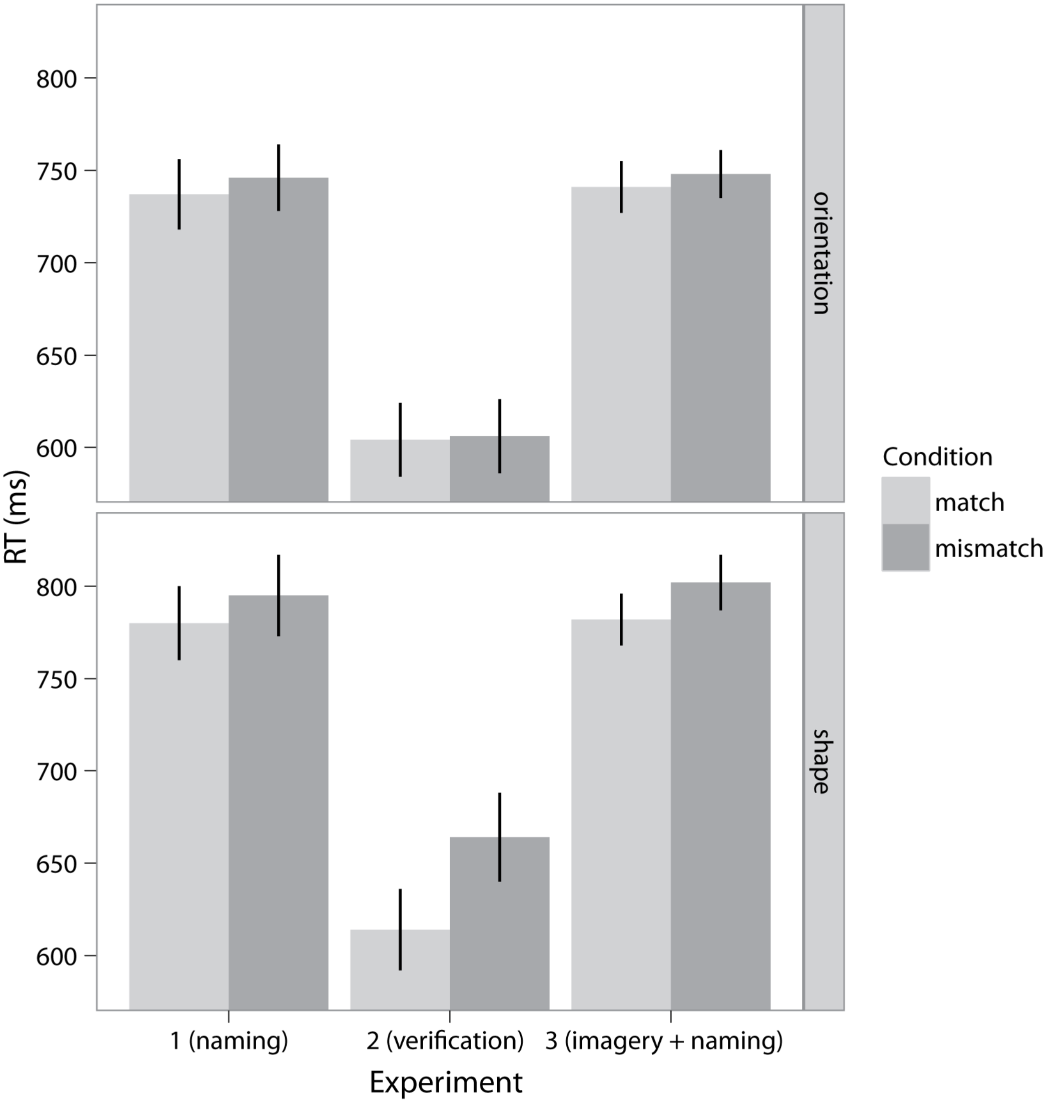


Figure 1. Average reaction times in Experiments 1, 2 and 3, based on participant medians. Upper panel: Orientation. Lower panel: Shape. Error bars indicate standard error.

For the *shape* items, we discarded 228 incorrect responses, and 24 trials due to a technical error (11.6%). The 16 ms difference (Figure 1) was not significant, $F(1,48)=.759$, $p=.388$, $\eta^2_p=.016$, and there was positive evidence for the null hypothesis, $p_{BIC}(H_0|D)=.827$. There was a trend for a Condition × List interaction, $F(3,48)=2.511$, $p=.070$, $\eta^2_p=.136$, and a Condition × Picture version × List interaction, $F(3,48)=4.352$, $p=.009$, $\eta^2_p=.214$. Only the mixed-effects model showed an effect of Condition, $\beta=.02418$ (19 ms), $\chi^2(1)=5.540$, $p=.019$. A joint analysis of the orientation and shape medians in a 2 (Type: orientation, shape) × 2 (Condition: match, mismatch) ANOVA yielded no Type × Condition interaction, $F<1$, suggesting that both effects were comparably small.

In sum, with the same analyses (ANOVAs on medians) and sample size that Zwaan and Yaxley (2002) used, neither the orientation nor the shape manipulation had a significant effect on naming latencies. This is surprising, given the congruency effects found for both manipulations by Pecher et al. (2009), who used the same materials in a memory task, and the findings by Zwaan et al. (2002, Experiment 2), who used materials similar to our shape-items in a naming task and found a 33 ms effect. However, in the mixed-effects model, a small effect appeared for shape. In the next experiment, we examined whether stronger effects would be obtained with the same materials in a sentence-picture verification task, where participants are explicitly required to relate the sentences to the pictures (following Stanfield and Zwaan, 2001, and Zwaan et al., 2002, Experiment 1).

Experiment 2

Participants

Forty-four native speakers of Dutch (38 female) with an average age of 21 years (range 18-26 years) were paid for their participation. All had normal or corrected-to-normal vision and reported no language impairments. None had taken part in Experiment 1.

Stimuli

The same stimuli as in Experiment 1 were used.

Procedure

Experiment 1 was the same as Experiment 2, except for the task. Participants were asked to indicate as quickly and accurately as possible whether the depicted object had been mentioned in the preceding sentence, by means of a button press on a button box. A green button on the left indicated a yes-response and a red button on the right a no-response. When the participant pressed one of the buttons, the picture disappeared from the screen.

Results and discussion

The analysis was the same as for Experiment 1. For the *orientation* items, fifty-four incorrect trials were discarded (2.4%). The ANOVAs on medians showed that the mean difference of 1 ms (see Figure 1) was not significant, $F(1,40)=.627$, $p=.433$, $\eta^2_p=.015$, with positive evidence for the null hypothesis, $p_{BIC}(H_0|D)=.825$. There was a trend towards a Condition × Picture version × List interaction, $F(3,40)=2.624$, $p=.064$, $\eta^2_p=.164$. In the mixed-effects model the effect of Condition also did not reach significance, $\beta=.02263$ (14 ms), $\chi^2(1)=3.224$, $p=.073$.

For the *shape* items, sixty-three incorrect trials (3.5%) were discarded. On average participants were 50 ms slower in the mismatch condition than in the match condition (Figure 1), $F(1,40)=33.455$, $p<.001$, $\eta^2_p=.455$, with very strong evidence for the alternative hypothesis, $p_{BIC}(H_1|D)=.999$. Condition further interacted with List, $F(3,40)=4.300$, $p=.010$, $\eta^2_p=.244$, with Picture version, $F(1,40)=11.443$, $p=.002$, $\eta^2_p=.222$, and there was a Condition × Picture version × List interaction, $F(3,40)=13.420$, $p<.001$, $\eta^2_p=.502$. In the mixed-effects model there was also a significant effect of Condition, with responses in the mismatch condition being slower than in the match condition, $\beta=.06724$ (44 ms), $\chi^2(1)=3.224$, $p<.001$. A joint analysis of the orientation and shape medians yielded a Type × Condition interaction, $F(1,43)=11.964$, $p=.001$, $\eta^2_p=.218$, indicating that the shape effect was larger than the orientation effect.

The shape congruency effect was larger in the verification task (Experiment 2; 50 ms) than in the naming task (Experiment 1; 16 ms). In a joint analysis of the medians from Experiment 1 and 2 this was confirmed by a Condition × Experiment interaction, $F(1,94)=6.977$, $p=.010$, $\eta^2_p=.069$. This supports task-dependence, since the participants apparently activated and used visual representations of shape more systematically in the verification than in the naming task. To carry out the verification task, the participants had to compare the content of the sentences with the depicted objects, which can be done using visual representations. If this is the underlying mechanism, one would expect to see a shape and possibly an orientation congruency effect in a naming task if the readers are explicitly instructed to recruit visual representations.

Experiment 3

Participants

Eighty-eight native speakers of Dutch (74 female) with an average age of 20 years (range 18-28 years) were paid for their participation. All had normal or corrected-to-normal vision and reported no language impairments. None had participated in Experiment 1 or 2.

Stimuli

The same stimuli as in Experiments 1 and 2 were used.

Procedure

The procedure was the same as in Experiment 1, where participants read sentences and named pictures, except that they were now asked to try to create a visual image of the situation described in each sentence before they pressed the button to continue. To remind participants of this instruction, every other "break" screen was replaced by an "imagery rating" screen where the participants were asked to indicate how well they had been able to create visual images of the situations described in the sentences they had just read, on a scale of 1 ("very badly, my images were vague, dark or even absent") to 10 ("very well, my images were as bright and lively as normal visual perception").

Results and discussion

For the *orientation* items, we discarded 193 incorrect responses (4.3%). The 7 ms advantage for the match condition (Figure 1) was not significant, $F(1,84)=1.632$, $p=.205$, $\eta^2_p=.019$, with positive evidence for the null hypothesis, $p_{BIC}(H_0|D)=.825$. There was a Condition × List interaction, $F(3,84)=3.476$, $p=.020$, $\eta^2_p=.110$, and a Condition × Picture version × List interaction, $F(3,84)=5.659$, $p=.001$, $\eta^2_p=.168$. The mixed-effects model did not indicate an effect of Condition either, $\beta=0.006347$ (5 ms), $\chi^2(1)=0.995$, $p=.318$.

For the *shape* items, we discarded 330 incorrect responses (9.4%). On average participants were slower by 20 ms to name the pictures in the mismatch condition than in the match condition (Figure 1), which yielded a main effect of Condition, $F1(1,84)=10.108$, $p=.002$, $\eta^2_p=.107$, with positive evidence for the alternative hypothesis, $p_{BIC}(H_1|D)=.941$. There was a Condition × List interaction, $F(3,84)=3.721$, $p=.014$, $\eta^2_p=.117$, and a Condition × Picture version × List interaction, $F(3,84)=7.261$, $p<.001$, $\eta^2_p=.206$. The mixed-effects model also indicated an effect of Condition, $\beta=.0.029034$ (23 ms), $\chi^2(1)=9.100$, $p=.003$. Thus, in contrast to Experiment 1, we now observed a clear effect of implied shape in an object naming task, though in joint analyses it was not significantly larger than the orientation effect, Type × Condition interaction $F(1,87)=1.460$, $p=.230$, $\eta^2_p=.017$, nor larger than the shape effect in Experiment 1, Condition × Experiment interaction $F(1,138)=.006$, $p=.937$, $\eta^2_p=.000$. We take the result to indicate that recruiting mental imagery during sentence reading facilitated picture recognition, in turn shortening naming latencies.[1]

Experiment 4

While in the previous experiments object shape information influenced performance in a task-dependent manner, orientation representations barely influenced performance at all. Does this reflect an inherently smaller influence of orientation representations relative to shape representations on response latencies? In Experiment 4, a rating study, we examined an alternative, namely whether the two item sets differed in how well the visual representations implied by the sentences corresponded to the pictures on the screen.

Participants

Forty native speakers of Dutch (33 female) with an average age of 20 years (range 17-23 years) were recruited from the same participant pool as the previous participants and also paid for their participation. None of them had participated in Experiment 1, 2, or 3.

Stimuli and design

The stimuli from the previous experiments were used. Of the 40 participants, 10 were randomly assigned to each of the four lists, which included only one possible sentence-picture combination for each item. The shape and orientation items were now presented in separate test blocks so that the participants could carry out the same type of judgment (shape or orientation match) on successive trials. Block order was counterbalanced between participants.

---

[1] We had previously run a version of Experiment 3 with 24 participants (replaced by a larger sample on reviewer request), which yielded similar results: An effect of shape (30 ms), but not of orientation (-2 ms), thus supporting our conclusions.

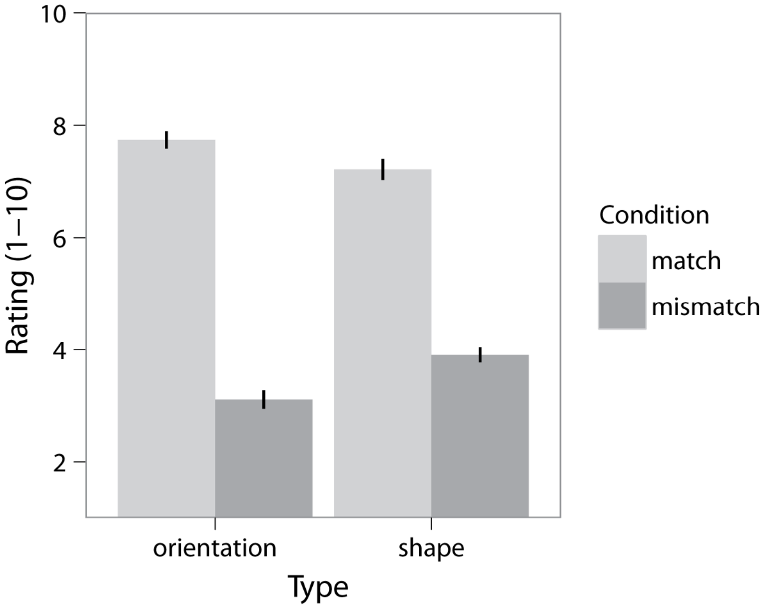


Figure 2. Average ratings for the materials in Experiment 4 in the different conditions. Error bars indicate standard error of the participant mean.

Procedure

The experiment was performed using the online experiment package WebExp (Keller, Gunasekharan, Mayo, & Corley, 2009). On each trial a sentence was presented above a picture. Participants were asked to look carefully at each sentence and picture and to rate the fit between the sentence and the picture in terms of object shape or orientation (depending on the block at hand), using a scale of 1 (poorest fit) to 10 (best fit).

Results and discussion

Mean ratings are shown in Figure 2. For the *orientation* items, ratings were on average 4.6 points higher in the match than the mismatch condition, $F(1,36)=247.044$, $p<.001$, $\eta^2_p=.873$. For the *shape* items, ratings were 3.3 points higher in the match than the mismatch condition, $F(1,36)=173.664$, $p<.001$, $\eta^2_p=.828$. Importantly, a joint analysis showed that the rating difference between match and mismatch was larger for the orientation items than the shape items, Type × Condition interaction $F(1,36)=29.796$, $p<.001$, $\eta^2_p=.453$. This means that the pattern of results - only shape but no reliable orientation congruency effects - cannot be ascribed to poorer quality of the orientation items.

General discussion

We compared the activation of two types of visual representations, shape and orientation, across three different task settings. Our research was motivated by the question whether the striking demonstrations of the activation of visual representations during language tasks (e.g., Stanfield & Zwaan, 2001; Zwaan et al., 2002) reflect routine or task-dependent processes. The results make two main contributions.

First, effects of implied orientation appear to be very difficult to obtain. At present we cannot resolve this discrepancy between earlier studies that have found significant effects and ours, which failed to find orientation effects. In the present study, effects of orientation were absent not only in naming tasks where this had not been tested before, but even in sentence-picture verification where we did observe a clear effect of shape. The ratings obtained in Experiment 4 confirmed that the items in the orientation set were well chosen. The larger effects for shape than orientation can therefore not be attributed to a difference in item quality, but are likely to depend on cognitive factors, such as the importance of shape in object recognition (Biederman & Cooper, 1991). Orientation is more viewpoint-dependent and thus less characteristic of objects. Although our results do not exclude that orientation representations could be routinely activated (Wassenburg & Zwaan, 2010), they do cast doubt on claims that orientation representations routinely influence performance.

Second, the results advance our understanding of the role of visual representations in language processing by showing that the influence of shape representations is task-dependent. Shape influences were weaker during naming (Experiment 1) than during sentence-picture verification (Experiment 2) and when participants had been instructed to use imagery

(Experiment 3). The influence of implied shape representations appears to occur 'on demand', rather than being an inherent consequence of the reading process.

Because the only difference between Experiments 1 and 3 were the imagery instructions, our data suggest that the use of imagery can be a mediating factor between language processing and the activation of visual representations. In conceptual processing research, this idea has previously been rejected based on the absence of correlations between the size of the modality switch effect discussed in the Introduction (Pecher, Zeelenberg, and Barsalou, 2003) and visual imagery measures (Pecher, van Dantzig, & Schifferstein, 2009). The present study provides more direct experimental evidence and supports the involvement of imagery. Interestingly, in a joint analysis of Experiments 1 and 3 we did not observe an Experiment × Condition interaction. One interpretation of this pattern of results is that some participants in Experiment 1 spontaneously used imagery without explicit instruction to do so. Consistent with this idea, Stanfield and Zwaan (2001, p. 156) mentioned that in a related unpublished study (Stanfield, 2000), 25% of the participants reported trying to actively generate images. Thus, the relationship between imagery and the activation of visual representations is clearly worth further investigation.

In sum, our findings paint a different picture of the role of visual representations than previous studies. They suggest that orientation representations only play a minor role during language comprehension, and that the influence of shape representations is mediated by task demands and likely the use of imagery. During everyday reading tasks, implied visual information often does not provide a crucial contribution to the comprehension process.

Acknowledgements

The authors thank Ronald Fisher and John Nagengast for technical support, and the research assistants of our lab for testing part of the participants and measuring speech onsets. We are grateful to Diane Pecher for kindly providing the experimental materials and suggesting the Bayesian analyses, and to her and an anonymous reviewer for comments.

References

Baayen, R.H., Davidson, D.J., & Bates, D.M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*(4), 390-412.

Barsalou, L.W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences, 22*(4), 577-609.

Biederman, I., & Cooper, E.E. (1991). Evidence for complete translational and reflectional invariance in visual object priming. *Perception, 20*(5), 585-593.

Connell, L. (2007). Representing object colour in language comprehension. *Cognition, 102*, 476-485.

Connell, L., & Lynott, D. (2009). Is a bear white in the woods? Parallel representation of implied object color during language comprehension. *Psychonomic Bulletin & Review, 16*(3), 573-577.

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*, 453-459.

Huettig, F., & Altmann, G. T. M. (2007). Visual-shape competition during language-mediated attention is based on lexical input and not modulated by contextual appropriateness. *Visual Cognition, 15*, 985-1018.

Huettig, F., & McQueen, J. M. (2011). The nature of the visual environment induces implicit biases during language-mediated visual search. *Memory & Cognition, 39*, 1068-1084.

Kang, S.H.K., Yap, M.J., Tse, C.S., & Kurby, C.A. (2011). Semantic size does not matter: "Bigger" words are not recognized faster. *Quarterly Journal of Experimental Psychology, 64*(6), 1041-1047.

Keller, F., Gunasekharan, S., Mayo, N., & Corley, M. (2009). Timing accuracy of Web experiments: A case study using the WebExp software package. *Behavior Research Methods, 41*(1), 1-12.

Kintsch, W. (2008). Symbol systems and perceptual representations. In M. De Vega, A. Glenberg & A. Graesser (Eds.), *Symbols and Embodiment.* (pp. 145-164). Oxford: Oxford University Press.

Kosslyn, S.M. (1994). *Image and brain: The resolution of the imagery debate*. MIT Press.

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*(1-3), 59-70.

Masson, M.E.J. (2011). A tutorial on a practical Bayesian alternative to null-hypothesis significance testing. *Behavior Research Methods, 43*, 679-690.

Pecher, D., van Dantzig, S., & Schifferstein, H.N.J. (2009). Concepts are not represented by imagery. *Psychonomic Bulletin & Review, 16*, 914-919.

Pecher, D., van Dantzig, S., Zwaan, R.A., & Zeelenberg, R. (2009). Language comprehenders retain implied shape and orientation of objects. *Quarterly Journal of Experimental Psychology, 62*(6), 1108-1114.

Pecher, D., Zeelenberg, R., & Barsalou, L. W. (2003). Verifying different-modality properties for concepts produces switching costs. *Psychological Science, 14(2),* 119-124.

Pecher, D., Zeelenberg, R., & Raaijmakers, J.G.W. (1998). Does pizza prime coin? Perceptual priming in lexical decision and pronunciation. *Journal of Memory and Language, 38*, 401-418.

Pulvermüller, F. (2005). Brain mechanisms linking language and action. *Nature Reviews Neuroscience, 6,* 576-582.

Pylyshyn, Z.W. (1981) The imagery debate: analogue media versus tacit knowledge. *Psychological Review, 88,* 16-45.

Raaijmakers, J.G.W., Schrijnemakers, J.M.C., & Gremmen, F. (1999). How to deal with "the language-as-fixed-effect fallacy": Common misconceptions and alternative solutions. *Journal of Memory and Language, 41,* 416-426.

Raftery, A.E. (1995). Bayesian model selection in social research. In P.V. Marsden (Ed.), *Sociological methodology 1995* (pp. 111–196). Cambridge: Blackwell.

R Development Core Team. (2009). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.

Richardson, D.C., Spivey, M.J., Barsalou, L.W., & McRae, K. (2003). Spatial representations activated during real-time comprehension of verbs. *Cognitive Science, 27*(5), 767-780.

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.

Schreuder, R., Flores d'Arcais, G.B., & Glazenborg, G. (1984). Effects of perceptual and conceptual similarity in semantic priming. *Psychological Research, 45*, 339-354.

Sereno, S.C., O'Donnell, P.J., & Sereno, M.E. (2009). Size matters: Bigger is faster. *Quarterly Journal of Experimental Psychology, 62*(6), 1115-1122.

Solomon, K. O., & Barsalou, L. W. (2004). Perceptual simulation in property verification. *Memory & Cognition, 32(2),* 244-259.

Stanfield, R.A., & Zwaan, R.A. (2001). The effect of implied orientation derived from verbal context on picture recognition. *Psychological Science, 12*(2), 153-156.

van Dam, W. O., van Dijk, M., Bekkering, H., Rueschemeyer, S-A. (2012). Flexibility in embodied lexical-semantic representations. *Human Brain Mapping, 33,* 2322-2333.

Wagenmakers, E.-J. (2007). A practical solution to the pervasive problems of p values. *Psychonomic Bulletin & Review, 14,* 779–804.

Wassenburg, S.I., & Zwaan, R.A. (2010). Readers routinely represent implied object rotation: The role of visual experience. *Quarterly Journal of Experimental Psychology, 63*(9), 1665-1670.

Zwaan, R.A. (2004). The immersed experiencer: toward an embodied theory of language comprehension. In B. H. Ross (Ed.), *The Psychology of Learning and Motivation* (pp. 35-62). New York: Academic Press.

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.

Appendix A: Sentence stimuli

Table 1: Orientation items

Object	Vertical	Horizontal
schaar	Rebecca kraste met de schaar haar initialen in het natte cement.	Rebecca kraste met de schaar haar initialen in de boom.
honkbalknuppel	Dirk probeerde met zijn honkbalknuppel het plafond aan te raken.	Dirk probeerde met zijn honkbalknuppel de deurknop aan te raken.
strijkijzer	Frank zette het strijkijzer op het rekje, hopend dat hij niet al te laat was.	Frank drukte het strijkijzer op de broek, hopend dat hij niet al te laat was.
schaal	Jan goot het allerlaatste beetje saus uit de schaal.	Jan deed het allerlaatste beetje saus in de schaal.
bezem	Janneke had de bezem in de kast gezet.	Janneke had de bezem onder het bed gelegd.
fles	Jessica had de fles in de emmer met ijs gedaan, want ze verwachtte een romantisch avondje.	Jessica had de fles in het wijnrek gedaan, want ze verwachtte een romantisch avondje.
schroef	John wilde net de schroef in de vloer van de kast draaien.	John wilde net de schroef in de zijkant van de kast draaien.
wasknijper	Laura deed de losse wasknijper aan de waslijn	Laura deed de losse wasknijper aan het lichtkoordje
gitaar	Mickey zette zijn gitaar in het rek, terwijl hij zich voorbereidde voor het concert.	Mickey tokkelde op zijn gitaar, terwijl hij zich voorbereidde voor het concert.
potlood	Rik had zijn potlood met een touwtje aan de koelkast gehangen.	Rik had zijn potlood bovenop de koelkast gelegd.
pen	De accountant had de pen in zijn borstzakje gedaan.	De accountant had de pen achter zijn oor gedaan.
zoutvaatje	Het bange jochie had het zoutvaatje op tafel gezet.	Het bange jochie had het zoutvaatje op tafel omgegooid.
spijker	De timmerman sloeg de spijker in de vloer.	De timmerman sloeg de spijker in de muur.
rups	De rups kroop langs de lantaarnpaal.	De rups kroop langs de stoep.
kwast	Toen de schilder klaar was liet hij de kwast in een glas achter.	De schilder had de kwast uit de emmer gehaald en begon met schilderen.
sleutel	De laboratoriumassistent had de sleutel van de rattenkooi op het haakje gelaten.	De laboratoriumassistent had de sleutel van de rattenkooi in het slot gelaten.
bijl	Jeske had de bijl in de kast gezet.	Jeske had de bijl onder de kast gelegd.
wijnglas	De ober vulde het wijnglas terwijl hij de jurk van Marieke bewonderde.	De ober gooide het wijnglas om terwijl hij de jurk van Marieke bewonderde.
krukje	Marcus was op het krukje gaan zitten.	Marcus had het krukje omgegooid.
schoen	Daniel plette met zijn schoen een mug op de muur.	Daniel plette met zijn schoen een mug op de grond.
aansteker	Jan gaf Wendy een vuurtje met zijn aansteker.	Jan gaf Wendy een por met zijn aansteker.
baby	De moeder hield de baby vast zodat hij een boertje kon laten.	De moeder legde de baby neer om hem een schone luier te geven.
rits	Wesley deed de rits van zijn regenjas dicht.	Wesley deed de rits van zijn weekendtas dicht.
boor	De klusser maakte met de boor een gat in het plafond.	De klusser maakte met de boor een gat in de muur.
theepot	Susan schonk het allerlaatste beetje water uit de theepot.	Susan deed het allerlaatste beetje water in de theepot.
kegel	De bowler had de laatste kegel niet geraakt.	De bowler had de laatste kegel ook geraakt.
zaklantaarn	Het jochie hield de zaklantaarn tegen zijn kin om een griezelmonster te spelen.	Het jochie hield de zaklantaarn naar voren om een griezelschaduw te maken.
prei	Als alternatief had Michiel een prei in de	Als alternatief had Michiel een prei op het

	vaas gedaan.	aanrecht gelegd.
liniaal	Dirk mat de hoogte van het raam met zijn liniaal.	Dirk mat de breedte van het raam met zijn liniaal.
speer	De krijger had zijn speer in de grond gestoken terwijl hij ging eten.	De krijger had zijn speer in een boom gestoken terwijl hij ging eten.
stempel	Jessica kreeg een stempel op haar hand.	Jessica kreeg een stempel op haar wang.
tarwehalm	De boer had de tarwe nog niet geoogst.	De boer had de tarwe net geoogst.
ton	Jos had de ton bij de herberg neergezet.	Jos rolde de ton naar de herberg toe.
vuist	De actievoerder stak zijn vuist hoog in de lucht.	De actievoerder sloeg met zijn vuist op tafel.
broek	De naaister deed de broek aan de paspop.	De naaister haalde de broek door de naaimachine.
stift	Lars stak zijn stift in het gaatje in de grond.	Lars stak zijn stift in het gaatje in de muur.
punaise	De verkoper prikte het prijskaartje met een punaise op het vloerkleed.	De verkoper prikte het prijskaartje met een punaise op het wandkleed.
vork	Aan tafel prikte Rik met zijn vork in zijn spruitjes.	Aan tafel prikte Rik met zijn vork in zijn broertje.
slak	De slak liet een slijmspoor achter op de tafelpoot.	De slak liet een slijmspoor achter op de balkonvloer.
sok	Robert liep op zijn sokken door het huis.	Robert lag met zijn sokken uitgestrekt op de bank.
slagboom	Toen ze bij de slagboom kwamen konden ze meteen doorrijden.	Toen ze bij de slagboom kwamen moesten ze meteen stoppen.
toiletrol	Er stond een volle rol toiletpapier in de kast.	Er hing een volle rol toiletpapier in de wc.
touw	Ze pakten het touw vast om te gaan rotsklimmen.	Ze pakten het touw vast om te gaan touwtrekken.
bord	Eveline zag haar oude bord in het afdruiprek.	Eveline zag haar oude bord in de gootsteen.
tennisracket	Ella strekte zich uit om met haar tennisracket de lamp aan het plafond te raken.	Ella strekte zich uit om met haar tennisracket de lamp aan de wand te raken.
roeispaan	De roeier probeerde met de roeispaan de bodem te raken.	De roeier probeerde met de roeispaan de wal te raken.
revolver	De cowboy mikte zijn revolver op zijn tegenstander tijdens het duel.	De cowboy deed zijn revolver in zijn holster vooraf aan het duel.
plug	Erik deed een plug in het plafond voordat hij de gordijnen ging ophangen.	Erik deed een plug in de muur voordat hij de gordijnen ging ophangen.
hamer	Marijke klopte heel zachtjes met de hamer tegen de keukendeur.	Marijke klopte heel zachtjes met de hamer op de keukenvloer.
schroevedraaier	Nella gebruikte een schroevendraaier om een lamp aan het plafond vast te maken.	Nella gebruikte een schroevendraaier om een lamp aan de muur vast te maken.
zwaard	De ridder bood met 2 handen zijn zwaard aan de koning aan.	De ridder liet zijn handen rusten op het zwaard dat in de grond stak.
rugzak	De reiziger liet zijn rugzak op zijn rug.	De reiziger liet zijn rugzak in het bagagerek.

Table 2: Shape items

Object	Shape 1	Shape 2
ballon	Dirk pakte een ballon uit het zakje.	Dirk pakte een ballon uit de lucht.
vleermuis	Johan observeerde een vleermuis in de grot.	Johan observeerde een vleermuis in de lucht.
boek	Iemand had een boek op de plank achtergelaten.	Iemand had een boek op het kopieerapparaat achtergelaten.
vliegtuig	Het vliegtuig bevond zich al in de loods.	Het vliegtuig bevond zich al in de lucht.
appel	Er zat ook een appel in de zak.	Er zat ook een appel in de salade.
brood	Elwin deed brood in de oven.	Elwin deed brood in de broodrooster.
kaas	Er lag kaas op de hapjesschaal.	Er lag kaas op de boterham.

champignons	Piet deed de champignons in het doosje.	Piet deed de champignons in de omelet.
ananas	Mark had ook ananas in de fruitsalade gedaan.	Mark had ook ananas bij de fruitmand gedaan.
ui	Janneke deed de uien in het boodschappenmandje.	Janneke deed de uien in de braadpan.
kip	Ella haalde de kip uit de oven.	Ella haalde de kip uit het hok.
sigaret	De schoonmaakster haalde een sigaret uit het pakje.	De schoonmaakster haalde een sigaret uit de asbak.
vis	Erik deed een vis in de vijver.	Erik deed een vis in de oven.
ijshockeyer	De coach riep naar de ijshockeyer op het ijs.	De coach riep naar de ijshockeyer op de bank.
adelaar	De jager zag een adelaar in de lucht.	De jager zag een adelaar in het nest.
ei	Iemand had een ei in de koelkast achtergelaten.	Iemand had een ei in de koekenpan achtergelaten.
maïskolf	Er lag een maïskolf op de akker.	Er lag een maïskolf op het bord.
zeilboot	De zeilboot was op de aanhangwagen.	De zeilboot was op het meer.
overhemd	Linda wilde het overhemd van de plank pakken.	Linda wilde het overhemd van de hanger pakken.
tissue	Evert haalde een tissue uit de doos.	Evert haalde een tissue uit de prullenbak.
tomaat	Jos had ook tomaten in de sla gedaan.	Jos had ook tomaten in de tas gedaan.
handdoek	Bernard wilde de handdoek van het rek pakken.	Bernard wilde de handdoek van de vloer pakken.
watermeloen	Laura presenteerde de watermeloen op de hapjesschaal.	Laura presenteerde de watermeloen op de jaarbeurs.
spaghetti	Bobby haalde de spaghetti uit het doosje.	Bobby haalde de spaghetti uit de pan.
benzinepomp	Tom stond bij de benzinepomp en gooide zijn tank vol.	Tom stond bij de benzinepomp en vroeg de weg.
CD-speler	Carla stopte de nieuwe CD in haar CD-speler.	Carla draaide de nieuwe CD op haar CD-speler.
tandpasta	Angela deed de tandpasta op de tandenborstel.	Angela deed de tandpasta in het winkelmandje.
voordeur	Bob stond bij de voordeur en zwaaide zijn vriendin uit.	Bob stond bij de voordeur en belde bij zijn vriendin aan.
paraplu	Ellen hield de paraplu in haar auto.	Elen hield de paraplu in de lucht.
koffer	Jacob pakte zijn koffer in voor de lange reis.	Jacob pakte zijn koffer op voor de lange reis.
landkaart	De handelsreiziger had de wegenkaart in zijn handen.	De handelsreiziger had de wegenkaart in het dashbordkastje.
laptop	Op de tafel stond zijn nieuwe laptop.	In de kast stond zijn nieuwe laptop.
auto	Lisa ging haar auto stofzuigen.	Lisa ging haar auto inzepen.
lipstick	Voor het feestje gebruikte Marta een nieuwe lipstick.	Voor het feestje kocht Marta een nieuwe lipstick.
rolmaat	De timmerman hield de rolmaat langs de tafel.	De timmerman legde de rolmaat op de tafel.
nietmachine	Stefan deed de nietjes in de nietmachine.	Stefan deed de blaadjes in de nietmachine.
taart	De dikke man keek begerig naar de taart in de vitrine.	De dikke man keek begerig naar de taart op zijn schoteltje.
vuilniszak	Paulien zag de vuilniszak op de stoeprand.	Paulien zag de vuilniszak in de kast.
mobieltje	Naima hield het mobieltje in haar hand tijdens het gesprek.	Naima hield het mobieltje in haar tas tijdens het gesprek.
vel papier	Bruno gooide het vel papier in de prullenbak.	Bruno gooide het vel papier op de stapel.

Appendix B: Accuracy analyses

Accuracy results are shown in Table S1. For the *orientation* items, there were no effects of Condition on accuracy and there was positive evidence for the null hypothesis (Experiment 1: $F(1,48)=0.231$, $p=.633$, $\eta^2_p=.005$, $p_{BIC}(H_0|D)=.851$; Experiment 2: $F(1,40)=0.966$, $p=.332$, $\eta^2_p=.024$, $p_{BIC}(H_0|D)=.817$; Experiment 3: $F(1,84)=0.060$, $p=.807$, $\eta^2_p=.001$, $p_{BIC}(H_0|D)=.901$). For the *shape* items, there were small effects (2-3%), yielding weak to positive evidence for the alternative hypothesis (Experiment 1: $F(1,48)=5.696$, $p=.021$, $\eta^2_p=.106$, $p_{BIC}(H_1|D)=.715$; Experiment 2: $F(1,40)=8.210$, $p=.007$, $\eta^2_p=.170$, $p_{BIC}(H_1|D)=.893$; Experiment 3: $F(1,84)=7.061$, $p=.010$, $\eta^2_p=.077$, $p_{BIC}(H_1|D)=.787$).

Table S1: Accuracy (%)

Type	Experiment	Match	Mismatch
Orientation	1 (naming)	95 (0.7)	95 (0.6)
	2 (verification)	97 (0.5)	98 (0.4)
	3 (imagery + naming)	96 (0.4)	96 (0.4)
Shape	1 (naming)	91 (1.0)	88 (1.0)
	2 (verification)	97 (0.6)	95 (0.7)
	3 (imagery + naming)	92 (0.7)	90 (0.7)

Note. Standard errors of the means are indicated between brackets.

Chapter 3

The contents of predictions in sentence comprehension: Activation of the shape of objects before they are referred to

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.

Abstract

When comprehending concrete words, listeners and readers can activate specific visual information such as the shape of the words' referents. In two experiments we examined whether such information can be activated in an anticipatory fashion. In Experiment 1, listeners' eye movements were tracked while they were listening to sentences that were predictive of a specific critical word (e.g., “moon” in “In 1969 Neil Armstrong was the first man to set foot on the moon”). 500 ms before the acoustic onset of the critical word, participants were shown four-object displays featuring three unrelated distractor objects and a critical object, which was either the target object (e.g., moon), an object with a similar shape (e.g., tomato), or an unrelated control object (e.g., rice). In a time window before shape information from the spoken target word could be retrieved, participants already tended to fixate both the target and the shape competitors more often than they fixated the control objects, indicating that they had anticipatorily activated the shape of the upcoming word's referent. This was confirmed in Experiment 2, which was an ERP experiment without picture displays. Participants listened to the same lead-in sentences as in Experiment 1. The sentence-final words corresponded to the predictable target, the shape competitor, or the unrelated control object (yielding, for instance, “In 1969 Neil Armstrong was the first man to set foot on the moon/tomato/rice"). N400 amplitude in response to the final words was significantly attenuated in the shape-related compared to the unrelated condition. Taken together, these results suggest that listeners can activate perceptual attributes of objects before they are referred to in an utterance.

The contents of predictions in sentence comprehension: Activation of the shape of objects before they are referred to

In sentence comprehension, readers and listeners often anticipate upcoming information (Altmann & Kamide, 1999; DeLong, Urbach, & Kutas, 2005; Federmeier & Kutas, 1999; van Berkum, Brown, Zwitserlood, Kooijman, & Hagoort, 2005; Wicha, Moreno, & Kutas, 2004). For instance, when a word is highly predictable, differential event-related brain potential (ERP) effects can be observed when a preceding adjective matches or mismatches with the expected word in grammatical gender (van Berkum et al., 2005; Wicha et al., 2004) or in phonological form (DeLong et al., 2005). Anticipation plays an important role in current views of sentence comprehension (Altmann & Mirkovic, 2009; Federmeier, 2007; Gibson, 1998; Kamide, 2008; Levy, 2008; Pickering & Garrod, 2007). Current theories focus on the underlying cognitive processes, with a prominent view being that predictions in comprehension are generated by the language production system (Pickering & Garrod, 2007).

An issue that has not received much attention concerns the kinds of information that listeners or readers pre-activate when they anticipate upcoming words. For understanding the mechanisms underlying prediction it is important to determine what listeners and readers do, or do not, predict. So far, studies have focused on functional semantic features or categories. Evidence for the involvement of this kind of information in predictions comes from anticipatory eye movements observed in the visual world paradigm, where participants listen to sentences and look at displays with multiple objects (Cooper, 1974; Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995; for review see Huettig, Rommers, & Meyer, 2011). For instance, in a seminal study by Altmann and Kamide (1999), participants heard sentences such as "The boy will eat the cake" or "The boy will move the cake" while viewing a display featuring several objects of which a cake was the only edible object. The authors found that, before the cake had been referred to, participants were more likely to initiate eye movements towards the cake when the verb was "eat" than when it was "move". This demonstrates that they predicted which object would be referred to on the basis of functional attributes (e.g., edibility) implied by the verb.

Other evidence for the involvement of functional semantic information in predictions comes from ERP studies. The component of interest is the N400, a centroparietally distributed component of negative polarity that peaks around 400 ms after onset of a content word and is considered a sensitive index of semantic processing (Kutas & Federmeier, 2000, 2011; Kutas & Hillyard, 1980). Federmeier and Kutas (1999) presented participants with contexts such as "They wanted to make the hotel look more like a tropical resort. So along the driveway, they planted rows of ...", which were followed by a predictable word (e.g., "palms"), or an unexpected word from the same semantic category (e.g., "pines"), or an unexpected word from a different category (e.g., "tulips"). N400 amplitude to the unexpected words was reduced when the word was related to the expected word (e.g., "pines") compared to when it was unrelated (e.g., "tulips"), even though in both of these conditions the predictable word was never presented. This suggests that semantic category information had been activated based on the context.

In the present study we were specifically interested in another kind of information that can be activated when words are processed, namely perceptual information - that is, physical attributes of objects. We focused on shape information. There is evidence from several sources that information about object shape can be activated *after* the relevant word has been processed. For instance, in sentence-picture verification experiments, Zwaan, Stanfield and Yaxley (2002) observed congruency effects between a shape implied in a sentence and a shape shown in a picture. Furthermore, several studies have observed perceptual priming in lexical decision, where a response to a target word (e.g., "coin") is faster when a preceding prime word has a referent with a shape similar to the target word's referent (e.g., "pizza") than when the prime word has a referent with a different shape (e.g., "table"; Moss, McCormick, & Tyler, 1997; Schreuder, Flores d'Arcais, & Glazenborg, 1984; but see Pecher, Zeelenberg, & Raaijmakers, 1998). Kellenbach, Wijers, and Mulder (2000) showed that at the neurophysiological level this shape priming was reflected in the N400, with N400 amplitude being smaller when prime and target had similarly shaped referents than when these were different. Finally, several visual world experiments showed that upon hearing a word such as “snake”, listeners were more likely to move their eyes to objects with a shape similar to the referent (e.g., a cable) than to objects with a different shape (Dahan & Tanenhaus, 2005; Huettig & Altmann, 2007).

In sum, there is strong evidence that shape information can be activated when words are processed. However, it is unknown whether this type of information already becomes activated when words are anticipated. For instance, do listeners expecting to hear about a moon activate visual representations of the object before they actually hear "moon"? As mentioned above, there is evidence for a high degree of specificity of predicted lexical information, since specific words can be anticipated (DeLong et al., 2005; van Berkum et al., 2005; Wicha et al., 2004). Listeners thus seem to prepare in advance for the input they might receive. Whether this degree of specificity extends to the level of perceptual attributes is unknown. One might, for instance, speculate that in many contexts shape information is not crucial for understanding the meaning of an utterance and would therefore be activated late or not at all. Alternatively, it could be the case that anticipation of a specific word entails some activation of all associated information, regardless of its relevance for understanding the current utterance.

A second issue we explored in the present study concerns the influence of the participants' task on the activation of shape information. Although previous studies have shown that shape representations can be activated in language comprehension, it is not known to what extent the reported shape effects depend on the participants' tasks. For lexical decision (Schreuder et al., 1984), it is conceivable that the activation of shape information stems from task properties, as being able to activate a mental image for a stimulus is a cue that distinguishes words from non-words. For sentence-picture verification (Zwaan et al., 2002), one could argue that the instruction to compare sentences to pictures encourages readers or listeners to create mental images corresponding to the sentence content. In the present study, participants were simply asked to listen attentively to spoken sentences. No additional task was given. Two online techniques were used to measure pre-activation of visual representations of upcoming words' referents: Eye-tracking during a visual world task, and recording of event-related potentials during a passive listening task. We asked whether participants, in anticipation of specific critical words, would activate the shape of the referents.

In Experiment 1, we used the target-absent version of the visual world paradigm in which a fully matching target word referent is excluded from the visual display on some trials (cf. Huettig & Altmann, 2005; Huettig et al., 2011, for detailed discussion). Participants listened to sentences that were predictive of a specific word (e.g., "moon" in "In 1969 Neil Armstrong was the first man to set foot on the moon"). Five-hundred ms before the acoustic onset of the target word (i.e. "moon"), participants were shown displays of four objects. These were three unrelated distractor objects and one critical object. Depending on the condition, the critical object was either the target object (e.g., a moon), or a shape competitor of the target object (e.g. a tomato), or a control object with a different shape than the target (e.g. rice). The participants' eye movements were recorded. We expected to observe anticipatory eye movements to the target object (e.g., the moon) before it was referred to. More importantly, if specific object shape representations form part of the contents of predictions for upcoming words, the expectation of the target concept (e.g., the moon) should also lead to anticipatory eye movements to a shape competitor (e.g., to a tomato).

We indeed observed such anticipatory eye movements. However, the visual world experiment involved the presentation of pictorial stimuli, which may encourage the activation of information concerning the physical properties of the referent objects shape (for discussion, see Huettig et al., 2011; Mitchell, 2004; for experimental evidence, see Huettig & McQueen, 2011). To determine whether shape information is also preactivated in the absence of pictorial information, a second experiment was conducted. This was an ERP experiment similar to Federmeier and Kutas's (1999) study. The participants listened to the same lead-in sentences as in Experiment 1. The final word of the sentence was either a highly predictable word (e.g., "moon"), or a word referring to an object with a similar shape (e.g., "tomato"), or an unrelated word (e.g., "rice"). We examined whether the N400 component in response to these words would reflect the semantic anomaly and, more importantly, would be sensitive to the referents' shape as well.

Experiment 1

Participants

Forty-five participants (34 women, mean age 21 years, range 18-29 years) from the Radboud University Nijmegen and the HAN University of Applied Sciences gave informed consent and were paid to take part in the experiment. All were native speakers of Dutch who had normal hearing, normal or corrected-to-normal vision and no history of language disorders.

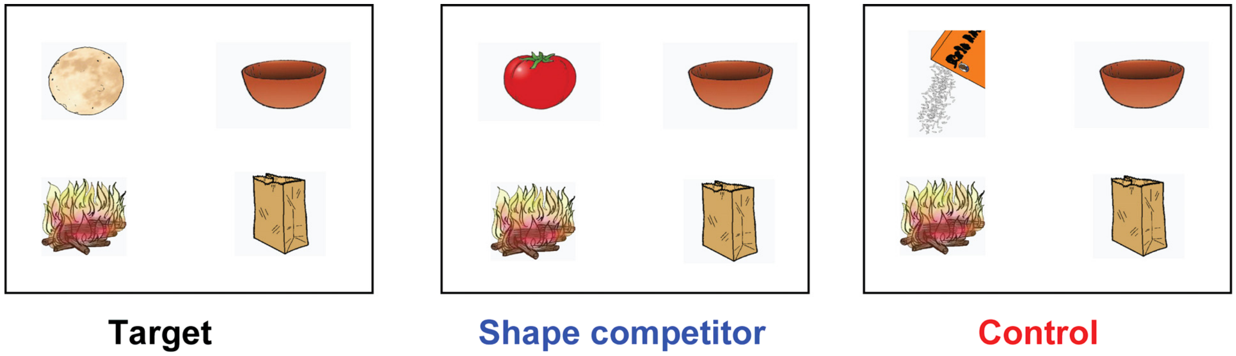
Materials and Design

The experiment consisted of 96 experimental trials and 32 filler trials. On each trial the participants heard a sentence and saw a visual display featuring four objects. On experimental trials one of the four objects was, depending on the condition, the highly predictable target (e.g., a moon), or the shape competitor (e.g., a tomato) or an unrelated control object (e.g., rice). The remaining three objects were unrelated distractor objects. On filler trials four objects were presented of which one was referred to in the spoken sentences. Thus, half of the sentences referred to an object in the accompanying visual display.

The sentences were statements with an average length of 14 words. They were selected from an initial set of 408 sentences. The materials had been pretested on cloze probability to verify that the final words of the experimental sentences were highly predictable. Fifteen participants (13 female, mean age 20 years, range 18-23) were asked to complete each lead-in sentence, that is, the fragment up to the final word, with the first word that came to mind, without trying to be original. The cloze probability of the target words was the proportion of participants who chose to complete the sentence fragment with the word in question (Taylor, 1953).The average cloze probability for the selected sentences was 0.72 (*SD* = 0.30), indicating that the critical words were predictable. The cloze probability for the names of the shape competitors and control objects was zero; that is, none of the participants provided them as continuations of the lead-in sentences. The filler sentences were chosen from the same set and were therefore similarly predictable.

The selected sentences were spoken at a relaxed pace by a female native speaker of Dutch and recorded in a sound-attenuating booth in a single recording session (mono, 44 kHz sampling rate, 16 bit sampling resolution), along with additional sentences used in Experiment 2. The onset times of the critical words were marked using Praat (Boersma & Weenink, 2009).

To create the visual displays for the experimental trials, 96 sets of six objects were composed consisting of three critical objects (either the Target, Shape competitor or Control picture, depending on the sentence it was paired with) and three distractor pictures which were unrelated to the other pictures in shape, semantics, or phonology (example see Figure 1). In designing the experiment, we aimed to make sure that differences in the participants' eye



Target **Shape competitor** **Control**

Figure 1. Example displays for each of the three conditions for the sentence "In 1969 Neil Armstrong was the first man to set foot on the moon". Shown along with the three distractors (a bowl, a fire, and a bag): the Target condition (with a moon), the Shape competitor condition (with a tomato), and the Control condition (with rice).

movements could be unambiguously attributed to the relationship of the critical objects to the sentence contexts, rather than properties of the sentence-final words or the corresponding objects themselves. Therefore, each critical object appeared in all three experimental conditions; that is, as a correct target on one list, as a Shape competitor on another list, and as a Control object on another list (see design below). In addition to the experimental sets, 32 four-object sets were composed to be used on filler trials.

The pictures were selected from Rossion and Pourtois (2004) and other sources or created by an artist. Colored pictures were used to facilitate the recognition of the objects (Rossion & Pourtois, 2004) and to ensure that participants would not confuse similarly-shaped pictures with the actual referents of the spoken words. When only a black-and-white picture was available in the picture data base, the artist colored the picture.

The pictures were pretested on three dimensions: familiarity, visual complexity, and name agreement. This was done to make sure that the critical pictures and the distractors did not differ in any obvious way and to establish that the objects were named as we anticipated. Fifteen participants (12 female, average age 20 years, range 18-26 years) from the same participant pool as used in the main experiment were paid for participation. Each picture was presented in the center of the screen until a response was recorded. First, the participants rated each picture for visual complexity on a scale from 1 ("very simple") to 7 ("very complex") by pressing the corresponding key on a keyboard. In the instructions, visual complexity was defined as the amount of details and the complexity of the lines in the pictures. The participants were asked to judge the complexity of the pictures rather than the complexity of the objects that the pictures represented. They were encouraged to use the full 7-point scale. Next, the participants viewed the pictures again and indicated their familiarity with each object on a scale from 1 (“very unfamiliar”) to 7 (“very familiar). Here, they were asked to indicate how often they came in contact with the object or thought about it. Finally, the participants typed in the name of each object as they would call it. The average visual complexity of target pictures (3.09, *SD* = .80) and distractor pictures (3.12, *SD* = .76) was very similar, *t* (382) = .378, *p* = .706. The same held for the average familiarity of the target pictures (4.70, *SD* = .88) and distractor pictures (4.67, *SD* = .85), *t* (382) = .302, *p* = .763. Average name agreement was high (distractor objects: 81%, *SD* = 20%; target objects: 76%, *SD* = 28%; *t* (382) = 1.904, *p* = .058).

All pictures were resized to fit into a square of 210 by 210 pixels, corresponding to 7 degrees of visual angle for the participant. They appeared in the four corners of the 1024 × 768 pixels screen at a distance of 256 or 768 pixels from the left and 192 or 576 pixels from the top.

The experimental and filler items were distributed across three presentation lists. In each list each experimental sentence was played once, accompanied by one of the three corresponding displays. In each list, each third of the 96 experimental sentences was combined with displays showing the target, the shape competitor or the unrelated competitor. Across the three lists, each sentence appeared once in each condition (Target, Shape competitor, Control). The 32 filler items appeared on all lists. Each list was used for 15 participants, who each received different randomizations of trials and picture positions on the screen. The same condition or the same target picture position never occurred on more than three consecutive trials.

Procedure

Participants performed a look-and-listen task as in previous experiments investigating anticipatory eye movements (e.g., Altmann & Kamide, 1999). They were asked to listen to the sentences carefully and told that they were free to look at whatever they wanted to, but that they should not take their eyes off the screen.

Participants were tested individually in a dimly illuminated room. They were seated in front of a screen with their chin on a chin rest. The movements of each participant's right eye were recorded with an Eyelink 1000 Tower Mount eye tracker sampling at 1000 Hz (except for one participant whose right eye could not be calibrated; the left eye was used instead). The spoken sentences were played through headphones.

Each trial started with a central fixation circle that remained on the screen until the participant fixated it. Then a blank screen appeared. A spoken sentence started playing, and 500 ms before the critical word was spoken four pictures appeared on the screen. Thus, in contrast to typical visual world experiments there was no long preview of the visual display before the critical target word was heard. This was done to avoid visual priming of the target concept (e.g. priming of the shape of the concept moon by the visual presence of a tomato in the display; see Huettig et al., 2011, for further discussion). The pictures remained on the screen until 2000 ms after sentence offset. Then a blank screen appeared for 500 ms. After every trial, a central fixation circle appeared, allowing for drift correction.

Data were coded in terms of fixations, saccades, and blinks using the algorithm provided in the Eyelink software. A fixation was coded as a fixation on an object if it occurred within a square interest area of 300 by 300 pixels around the centre of that object.

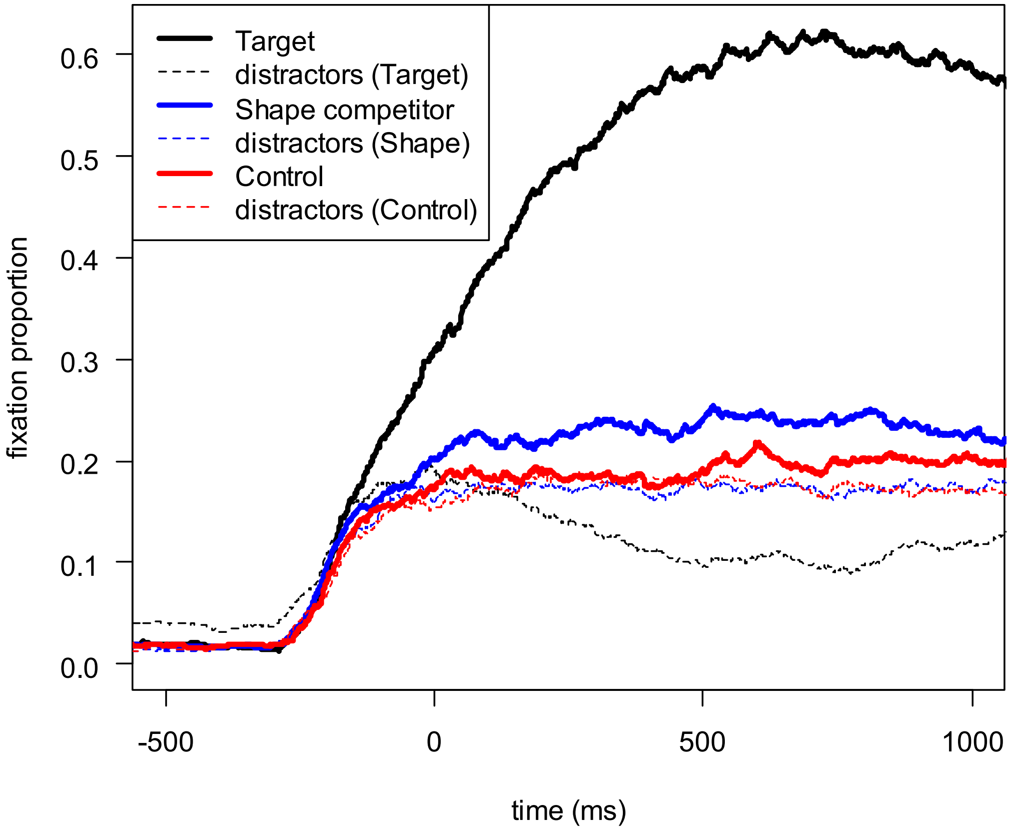


Figure 2. Results of Experiment 1. Time-course graph showing fixation proportions to Targets, Shape competitors, and Control objects (solid lines) along with fixation proportions averaged across the three corresponding unrelated distractors (dashed lines). Display onset was at -500 ms, time zero indicates critical word onset.

Results

Figure 2 shows a time-course graph of the proportions of fixations to the critical pictures (Target, Shape competitor, and Control) and the average proportions of fixations to the three distractor pictures across items and participants. Recall that each display included only one of the three critical objects. The three dotted lines show the average proportions of looks to distractors in displays featuring targets, shape competitors, and control objects, respectively. The proportions indicate the number of trials in which participants were fixating each location on the screen. Time zero represents the acoustic onset of the critical spoken word.

The graph shows that, unsurprisingly, participants were far more likely to fixate upon the targets than on the competitors or distractors. More importantly, they were also more likely to fixate upon the Shape competitors than upon the unrelated Control objects. Finally, from about 500 ms onward the Control objects were fixated somewhat more frequently than the corresponding distractors, perhaps because due to counterbalancing they were referred to in another sentence in the experiment. We had no hypotheses about fixations at this point in time. Since it takes on average about 200 ms to program and initiate a saccadic eye movement (Saslow, 1967), we can safely assume that any fixation bias to either Target object or Shape competitor until 200 ms *after* the acoustic onset of the target word is of anticipatory nature. Thus, for the statistical analyses we chose a 450-ms time window starting at 250 ms before word onset (the point in time when fixations were first directed to any of the objects, see Figure 2) until 200 ms after word onset (the approximate point in time when fixations could first be based on information retrieved from the spoken target word).

Fixation proportions within this time window were transformed to log odds, the appropriate scale for assessing effects on a categorical dependent variable, using the empirical logit function (Barr, 2008). The average log odds of looks to the unrelated distractors was subtracted from the log odds of looks to the Target/Shape/Control object to create the dependent variable, which indicates the strength of any bias toward each experimental picture over the unrelated distractor pictures. This dependent variable was analyzed using linear mixed-effects regression models, which allow for simultaneous inclusion of participants and items as random factors (Baayen, Davidson, & Bates, 2008). *P*-values were calculated assuming that the *t*-values were drawn from a normal distribution, which is justified for datasets the size of the current experiment (Barr, 2008). The full model included a fixed effect of Condition (Target, Shape competitor, Control) and the maximal possible random effects structure consisting of random intercepts and slopes for Condition by participant, sentence, and picture. The Control condition was mapped onto the intercept as a baseline to compare the Shape competitor and Target conditions to. This model was compared to the same model without the fixed effect of Condition using a likelihood ratio test. Including Condition improved model fit, χ^2 (2) = 61.741, $p < .001$. There was no significant bias for the Control pictures (e.g., rice) over the unrelated distractors, $\beta = .04925$, $SE = .14250$, $t = .346$, $p = .730$. There was a higher bias for Target pictures (e.g., moon) than for Control pictures, $\beta = 2.02858$, $SE = .21299$, $t = 9.524$, $p < .001$. Importantly, participants also showed a higher bias for Shape competitors (e.g., tomato) than Control pictures, $\beta = 0.41310$, $SE = .19203$, $t = 2.151$, $p = .031$.[2]

Discussion

Using a variant of the visual world paradigm, we observed that participants anticipated which object a speaker would refer to next. When listening to sentences such as "In 1969 Neil Armstrong was the first man to set foot on the moon", participants looked significantly more to the picture of a moon than to unrelated distractor objects before the word "moon" was heard. This is in line with earlier studies (e.g., Altmann & Kamide, 1999) demonstrating that listeners anticipate upcoming words. More importantly, we also found that participants looked significantly more at objects similar in shape to the target object (e.g. a tomato) than to unrelated objects (e.g., rice) before the word "moon" was heard. To our knowledge, the present research is the first to report such an effect of pre-activation of the visual form of the referents of upcoming words.

---

[2] A repeated-measures ANOVA on by-participant means yielded the same results, including a bias toward Target objects over unrelated distractors that was higher than that for Control objects, F (1,44) = 153.177, $p < .001$, and a bias toward Shape competitors over unrelated distractors that was higher than that for Control objects, F (1,44) = 6.676, $p = .013$.

In this experiment, the visual displays only appeared 500 ms before the onset of the critical word. This preview period for the pictures is shorter than the preview commonly used in the visual world paradigm, which often spans at least the duration of the sentence. Nevertheless, a preview time of 200 ms has been used before, and a bias toward Shape competitors was observed (Huettig & McQueen, 2007). We chose a short preview period to minimize the impact of the visual stimuli on the processing of the spoken sentences, in particular to ensure that any priming of visual representations would be very minimal. However, to assess whether anticipatory activation of shape information would also occur in the total absence of any visual stimuli, a second experiment was conducted.

Experiment 2

In Experiment 2, participants listened to the same sentences as in Experiment 1 (Correct condition) and to sentences where the final highly predictable word was replaced by the name of the shape competitor (Shape condition) or the unrelated control object (Unrelated condition; see Table 1 for an example). The participants were asked to listen to the sentences for comprehension while their electroencephalogram (EEG) was recorded. The EEG component of interest was the N400 in response to the final word of the sentences. Since the sentences in the Shape and Unrelated condition both ended in a contextually anomalous word, we expected a more pronounced N400 in both of these conditions than in the Correct condition (e.g., Federmeier & Kutas, 1999; Kutas & Hillyard, 1980). Importantly, the N400 has also been shown to be sensitive to visual attributes of objects being referred to (Kellenbach et al., 2000). If shape representations are involved in predictions of upcoming meaning in the absence of pictorial information, the N400 amplitude should therefore be attenuated for the Shape condition relative to the Unrelated condition.

Table 1

Example sentences

Condition	Example
Correct	In 1969 zette Neil Armstrong als eerste mens voet op de maan.
	In 1969 Neil Armstrong was the first man to set foot on the moon.
Shape	In 1969 zette Neil Armstrong als eerste mens voet op de tomaat.
	In 1969 Neil Armstrong was the first man to set foot on the tomato.
Unrelated	In 1969 zette Neil Armstrong als eerste mens voet op de rijst.
	In 1969 Neil Armstrong was the first man to set foot on the rice.

Note. English translations are provided in Italics. Critical words are underlined.

Participants

Twenty-four students (22 women, mean age 20 years, range 18-23 years) from the Radboud University Nijmegen and the HAN University of Applied Sciences gave informed consent and were paid to take part in the experiment. All were right-handed, native speakers of Dutch who had normal hearing, normal or corrected-to-normal vision and no history of neurological or language disorders.

Materials and Design

There were three conditions: Correct, Shape, and Unrelated (see Table 1). In the Correct condition, the same 96 sentences were presented as in Experiment 1. For the Shape and Unrelated condition, the final word of the sentences was replaced by the name of the Shape competitor or Control object, respectively. Thus, the Unrelated condition in Experiment 2 roughly corresponds to the Control condition in Experiment 1. In total, the materials consisted of 288 Dutch sentences, formed by recombining critical words and lead-in sentences (for an example, see Appendix A).

Each critical word occurred once in each condition, thus controlling for lexical characteristics. For each sentence context the Shape and Unrelated word came from the same semantic category (e.g., *tomato* and *rice* are both food), thus controlling for previously demonstrated effects of semantic category on N400 amplitude (Federmeier & Kutas, 1999).

The mean spoken sentence durations in the three conditions were similar, Correct: 5334 ms (*SD* = 1291), Shape: 5325 ms (*SD* = 1212), Unrelated: 5413 ms (*SD* = 1232), *F* (2,190) = 1.86, *p* > .1. Written versions of the sentences and words were pretested on two dimensions, the shape similarity of the referents of the words and the plausibility of the sentences. Different participants took part in each pretest, and none of them participated in the main experiment.

The goal of the shape similarity ratings was to verify that the referents of the words in the Shape condition were more highly related in shape to the referents of the Correct word than the referents of the words in the Unrelated condition. The 96 shape-related pairs (e.g. tomato and moon) and the 96 unrelated pairs (e.g. rice and moon) were presented along with 80 other pairs of words. Twenty-four participants (13 women, mean age 25 years, range 21-32 years) rated the shape similarity of the referents of the word pairs on a scale of 1 ("completely different shapes") to 7 ("exactly the same shape"). Each person received a different random order of items, and the left-to-right order of the members of the pairs was counterbalanced across participants. Participants were instructed to use the entire scale from 1 to 7 and to ignore any similarities in meaning between the words. The ratings were averaged across participants for each word pair (pooling over the two left-to-right orders). The ratings confirmed that for the word pairs selected for the ERP experiment, the average ratings were higher in the Shape condition (mean = 4.2, *SD* = 0.9) than in the Unrelated condition (mean = 1.9, *SD* = 0.7), *t* (94) = 14.1, p < .001.

In addition, we wanted to ensure that the words from the Shape and Unrelated conditions would be equally difficult to integrate with the preceding context. In order to test this, plausibility ratings were obtained for the sentences. Forty-two participants (31 women, mean age 21 years, range 18-34) were randomly assigned to one of three lists. Every list consisted of 184 sentences in which no context or critical word was repeated and each condition was represented by 40 or 48 items depending on the list (this was at a stage before selecting the final set of 32 items per list). Forty-eight plausible sentences were used as fillers on every list, and each participant received the sentences in a different random order. They were asked to indicate for every sentence, on a scale from 1 ("very implausible") to 7 ("very plausible"), how plausible the described situation would be in everyday life. The ratings were averaged across participants for each of the 96 selected critical words in each condition. These ratings were submitted to a repeated-measures analysis of variance (ANOVA) with Condition (Correct, Shape, Unrelated) as the within-word factor. This by-items analysis yielded an effect of Condition, F (2, 190) = 3946.25, $p < .001$. Planned comparisons indicated that, as expected, the sentences in the Correct condition (6.4, *SD* = 0.5) were more plausible than those in the Shape condition (1.5, *SD* = 0.4), F (1, 95) = 5697.32, $p < .001$, and also more plausible than those in the Unrelated condition (1.5, *SD* = 0.4), F (1, 95) = 5454.39, $p < .001$. Importantly, the difference between the Shape and the Unrelated conditions was not significant, F (1, 95) = 0.03, $p = .870$. Also, on an item by item basis, shape ratings were not correlated with plausibility ratings, $r = .02$, $p = .749$. Thus, there was no evidence that one of the two types of semantic violation (Shape or Unrelated) was easier to integrate in the sentence context than the other.

The sentences were distributed across three lists. On every list each critical word appeared once, and each condition was represented by 32 items. Sixty-four filler sentences were added to every list, of which 48 were correct and 16 contained a semantic violation. In total, 50% of the sentences on every list contained a violation. Each list was used for eight participants, who received different randomizations of the items.

Procedure

Participants were tested individually in a dimly illuminated room. They were seated in front of a screen and two loudspeakers and were asked to relax, and to move and blink as little as possible while carefully listening to the sentences. The experiment started with two practice blocks of eight trials each to familiarize the participants with the way the trials were structured. After the practice blocks the sentences were presented in twenty short blocks of approximately two minutes. Each trial started with a blank screen with a duration of 50 ms, followed by a short warning tone. 750 ms later, a fixation cross (+) appeared on the center of the screen, which participants were asked to fixate on. After 750 ms, a spoken sentence was played. The fixation cross remained on the screen until 1500 ms after sentence offset. Participants had been asked not to blink or move while the fixation cross was on the screen. Then three asterisks (* * *) were presented for 3000 ms. During this period participants were free to blink and move their eyes. After every block participants could take a break for as long as needed before resuming the experiment by pressing a button on a button box.

EEG recording and analysis

EEG was recorded from 128 active Ag/AgCl electrodes mounted in a cap according to the 10-5 system (Oostenveld & Praamstra, 2001). Recordings were performed relative to common mode sense (CMS) and driven right leg (DRL) electrodes placed just anterior to the Fz electrode. Horizontal eye movements were monitored using electrooculography (EOG) electrodes positioned laterally to the left and right eyes. Two electrodes were placed at the mastoids. The signals were amplified by Biosemi ActiveTwo amplifiers with a lowpass filter at 128 Hz and sampled with a frequency of 512 Hz.

The data were referenced to the average of the left and right mastoids. Bipolar horizontal EOG was computed as the difference between the left and right EOG electrodes. A bandpass filter of 0.05-30 Hz was applied and the continuous EEG was segmented into epochs from 300 ms before until 1000 ms after critical word onset. These epochs were baseline-corrected through subtraction of the mean signal in the 150 ms before critical word onset. Trials containing blinks were removed using participant-specific thresholds for the three frontal electrodes (Fp1, Fpz, Fp2). In a second step, other artifacts were removed using a threshold of + and -100 µV. This was done on a single-channel basis, as in high-density electrode setups the probability increases that some channels are contaminated by artifacts, which can easily lead to an unacceptably low number of trials when removing entire trials (cf. Junghöfer, Elbert, Tucker, & Rockstroh, 2000). Average ERPs were then computed for each participant and condition separately.

Mean voltage measures were taken in two predefined time windows: 300-500 ms (Federmeier & Kutas, 1999; van den Brink et al., 2006) and 500-700 ms after word onset (McCallum, Farmer, & Pocock, 1984). Besides having been used in previous studies, the 500-700 ms time window was based on the knowledge that the N400 can be longer in duration with auditory presentation compared to visual presentation (for discussion, see Kutas & van Petten, 1994). Furthermore, in comparison to most auditory ERP studies, some of our critical words were relatively long in terms of number of syllables (range 1-5, mean 2.1, *SD* 1.1) and duration in ms (range 289-1058, mean 603, *SD* 174). This might delay the point at which the words became recognized, and our effect of interest depended on recognition of the words.

The average ERPs were submitted to nonparametric cluster-based permutation tests (Maris & Oostenveld, 2007) using the Matlab toolbox Fieldtrip (Oostenveld, Fries, Maris, & Schoffelen, 2011). The permutation test determined which channels showed a significant effect by means of a clustering algorithm based on the physiologically plausible assumption that ERP effects are clustered together in space over neighboring electrodes. The test compares two conditions at a time and works as follows (for details see Maris & Oostenveld, 2007). First, a dependent-samples t-test compares the conditions at every data point (in this case, every electrode), and data points that do not exceed a significance level of .05 are zeroed. Adjacent non-zero data points (at neighbouring electrodes) are combined into clusters for each of which the cluster-level *t*-value is the sum of all *t*-values within the cluster. Then a null-distribution is created by randomly assigning subject averages to one of the two conditions a user-specified number of times, and computing the cluster-level statistics for each randomization. Finally, the observed cluster-level test statistics are compared against the null-distribution. When the observed statistic falls in one of the 2.5th percentiles of the null-distribution, the effect is considered significant. Note that this test can only compare two conditions at a time, so we tested each of the three possible contrasts between the three conditions.

In our analysis the distance at which channels were considered neighbors was set such that each channel had an average of 5.1 neighbors. The number of simulations was set to 500. The critical alpha level was set to .05 one-tailed, as the direction of the N400 effect was known and we expected an attenuated N400 effect when the semantically anomalous word had a referent with a similar shape (Shape condition) compared to the Unrelated condition. Reported *t*-values refer to the cluster-level statistics ("sum-*t*").

Results

As can be seen in Figure 3, after word onset the ERPs to the Correct condition remained essentially flat. In contrast, relative to the Correct condition, both types of semantic violation (the Shape and the Unrelated conditions) elicited a slow potential change of negative polarity between ~150 to 800 ms. These negative waves had an early onset, which could be a reflection of mismatches between expected and encountered segmental information similar to the N200 reported by van den Brink, Brown, and Hagoort (2001). These early negativities were not of

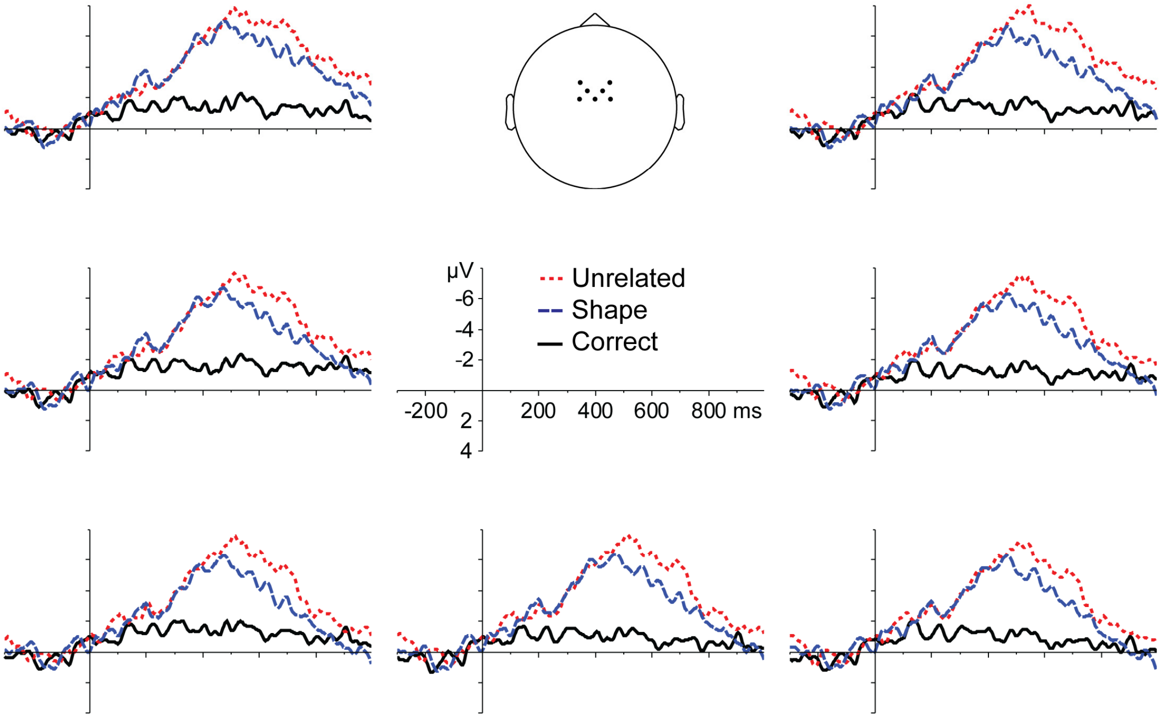


Figure 3. ERP waveforms at seven electrodes of which the locations are indicated on a head map (top), from left to right, top to bottom: FC1, FC2, FCC1h, FCC2h, C1, Cz, C2. Word onset is at zero ms. Negative is plotted up.

interest for the present study, and further discussion will focus on comparisons between conditions in the later N400 time windows.

The results of the statistical analyses are shown in Figure 4. The Unrelated condition elicited a significantly more negative voltage than the Correct condition in a large cluster of electrodes covering the fronto-central scalp. This was the case in the first time window, between 300 and 500 ms after word onset, sum-$t = -418.14$, $p < .001$, as well as the second window, between 500 and 700 ms after word onset, sum-$t = -388.18$, $p < .001$. The ERPs for the Shape condition were also significantly more negative-going than those for the Correct condition in both the 300-500 ms and the 500-700 ms time windows, sum-$t = -318.88$, $p < .001$, and sum-$t = -220.32$, $p < .01$, respectively. All of these clusters were widely distributed across the scalp,

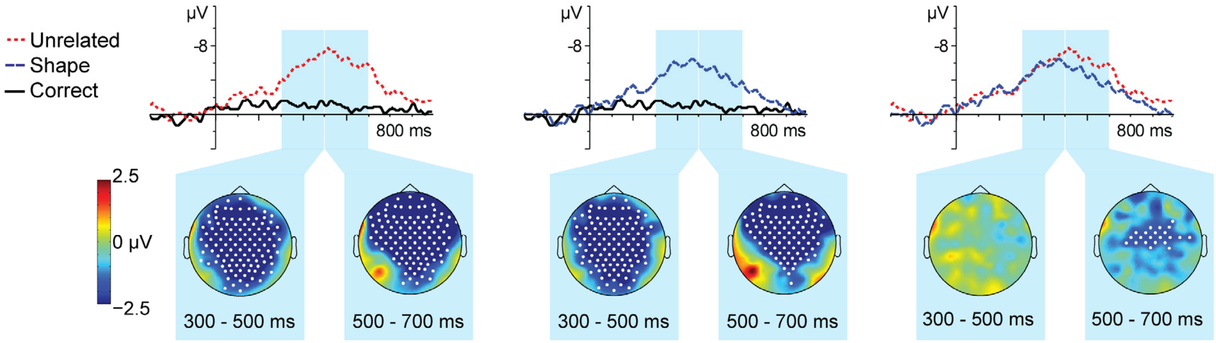


Figure 4. Scalp topographies of the mean difference between each of the conditions (indicated by the two relevant waves taken from the Cz electrode): Unrelated - Correct, Shape - Correct, and Unrelated - Shape. Two time windows are shown. White dots indicate electrodes included in a significant cluster.

consistent with an N400. However, there was a frontal emphasis which is slightly atypical, though the N400 tends to be more frontally distributed during listening than during reading (McCallum, Farmer, & Pocock, 1984) as well as for concrete words compared to abstract words (Kounios & Holcomb, 1994). The more frontal distribution might also be due to offsets of the ERPs of the preceding word or to overlap with the contingent negative variation. Critically, the slow negative wave signaling a semantic mismatch was of lower amplitude in the Shape condition than in the Unrelated condition. The attenuation for the Shape condition had a discrete onset, as reflected in the waveforms of Figures 3 and 4 and in the relevant statistics. Thus, the difference between the Shape condition and the Unrelated condition was not significant in the early time window from 300-500 ms (no clusters were detected), but was significant in the later window from 500-700 ms, sum-t = -46.70, p = .036.[3] The cluster for the difference between the

---

[3] A repeated-measures ANOVA on ERPs averaged over this cluster of electrodes yielded the same pattern of results, including a difference between the Unrelated and Shape conditions in the 500-700 ms time window, F(1,23) = 5.53, p = .028.

Shape and Unrelated condition was smaller than those for the other comparisons.[4]

Discussion

Semantic violations elicited N400 effects relative to correct words. Importantly, the N400 amplitude was significantly attenuated for the Shape condition relative to the Unrelated condition between 500 and 700 ms after word onset. The fact that the N400 amplitude in response to semantic violations was modulated by shape similarity between referents of presented words (e.g., "tomato" vs. "rice") and words that were predictable but not presented (e.g., "moon") is consistent with the hypothesis that predicted information can be very specific and perceptual in nature, involving the expected referent's shape. The results fit with previous studies indicating high specificity of anticipation at the lexical level (DeLong et al., 2005; van Berkum et al., 2005; Wicha et al., 2004) and extend these findings to the semantic/conceptual level.

Before turning to the general implications of our study we will discuss four alternative explanations of the ERP results. First, one might argue that the shape effect (i.e., the difference between the Shape and the Unrelated condition) occurred relatively late (between 500 and 700 ms after critical word onset) and therefore does not reflect the 'typical' N400 component. The effect observed in the present study indeed occurred somewhat later than previously reported semantic category effects in the auditory modality (Federmeier, McLennan, De Ochoa, & Kutas, 2002), though time windows after 500 ms are very common in N400 studies using auditory stimuli (e.g., Holcomb & Andersson, 1993; Holcomb & Neville, 1991; McCallum, Farmer, & Pocock, 1984; Perrin & García-Larrea, 2003). In our study, the shape effect may have arisen

---

[4] Following suggestions made by reviewers of the paper, we further examined whether a more graded effect of shape-relatedness was detectable by correlating the amplitude of the N400 with the ratings of shape relatedness obtained for the individual items. These analyses, which did not yield any significant results, are reported in Appendix B.

relatively late because many of the critical words were polysyllabic, or because shape information might not receive the same degree of priority as semantic category information in facilitating the processing of unexpected words. Determining the origin of the latency difference between semantic and shape information is difficult however, given that the semantic and shape effects were seen in separate studies, differing in materials, languages, and participants.

Related to this, the difference between the expected and unexpected words was observed well before the shape effect. We interpret the early difference as an acoustic rather than a semantic effect, reflecting the differences between the expected and actual speech signal, similar to the N200 reported by van den Brink, Brown, and Hagoort (2001). The ERPs elicited by the Shape and Unrelated conditions overlap initially because in both cases the expected and presented word mismatch acoustically. The later timing of the shape effect follows from the fact that, in contrast to the earlier difference, the shape effect requires access to the meaning of the presented word (e.g., "tomato"). We regard the clear electrophysiological difference between the Shape and Unrelated conditions as the main finding; whether or not this difference should be called an N400 effect is, in our view, of secondary importance.

Second, as we measured the consequences of expectations in terms of N400 amplitude in response to semantic violations, we cannot rule out that the effect reported here might reflect shape activation arising from the process of dealing with disconfirmed predictions. However, two observations argue against this interpretation. First, our plausibility ratings did not differ between the Shape and Unrelated conditions; both conditions were outright semantic violations that should lead to similar reanalysis processes. Second, reanalysis processes, particularly those involved in revising a strong prediction when encountering unexpected input, are usually associated with late positivities instead of the negativity observed here (cf. Federmeier, 2007).

A third alternative might be that our results are due to integration rather than semantic expectancy. We consider this to be very unlikely given that the plausibility ratings for our materials showed that the shape-related violations (e.g., "tomato" instead of "moon") were not more plausible substitutes for the correct words than the unrelated violations (e.g., "rice").

Fourth, it might be the case that the shape-related words were also associatively related to the targets. We tested this possibility, as far as possible, post-hoc using Dutch association norms (de Groot & de Bil, 1987; Lauteslager, Schaap, & Schievels, 1986; van Loon-Vervoorn & van Bekkum, 1991) which had been collected from 100 participants with each participant stating the first association that came to mind, such that association strength ranges from 0 to 100. These revealed no associative relationships between the critical words in either direction for any of the 50 word pairs for which norms were available. Thus, we found no support for this explanation. A related possibility is that the shape effect could have been driven by associations between the shape-related words and words in the context sentence. We examined how strongly the content words in the context sentences (context words hereafter) were associated to each of the three critical words. Thus, for the sentence "In 1969 Neil Armstrong was the first man to set foot on the moon/tomato/rice", we examined whether any of the content words that were not names (e.g., "first", "man", "set", "foot") were associated to the critical words (e.g., "moon", "tomato", or "rice"). Of the 545 context words, 327 appeared in the norms. Of these, 60 words were associated to Correct critical words, with the average association strength being 8.583 (average across all 327 words: 1.490), one word was associated to a Shape-related critical word with an association strength of 11 (average across all words: .039) , and three words were related to Unrelated critical words, with an average strength of 7 (average across all words: .047). These analyses show that there was some degree of associative relatedness between the words in the context and the Correct critical words, but not between the words in the other two conditions.

In sum, we found no evidence to support alternative explanations of our result. Instead, the ERP difference between the Shape and Unrelated conditions most likely arose because the referents of the critical words in the Shape condition were similar in physical shape to the referents of the predicted words, which was not the case for the Unrelated condition.

General Discussion

The present study investigated the type of information activated during predictive sentence processing. More specifically, we asked whether listeners activate information about the shape of objects before the objects are referred to. A secondary question was whether activation of object shape information during language processing, previously mainly observed in meta-linguistic tasks (such as lexical decision or sentence picture verification) or with an extended preview of visual stimuli (as in most previous visual world studies) could be observed when visual stimuli were only presented for a brief period (Experiment 1) or not at all (Experiment 2).

Experiment 1 used the visual world paradigm. The visual stimuli were shown only 500 ms before the onset of the critical spoken word. On critical trials, the final word of the spoken sentences could readily be predicted on the basis of the preceding sentence context. We found that participants were more likely to look at objects that were related in shape to the predictable target than at unrelated objects. Importantly, these fixations were initiated before the predicted word itself was processed. Thus, the shape of the referent object was anticipatorily activated. Experiment 2 was an EEG study. Participants listened to sentences ending in a highly predictable word, or in a semantically anomalous word with a referent that had a shape similar to the referent of the expected word, or in an unrelated word. Both anomalous conditions elicited N400 effects relative to the correct condition. Importantly, the N400 effect was attenuated by shape similarity between the referents of the expected and the encountered words. Our findings that shape similarity influenced both anticipatory eye movements and N400 amplitude provide strong converging evidence for the notion that object shape representations can form part of the contents of listeners' predictions for upcoming meaning.

The experimental results also have implications for the generality of shape effects. The fact that we saw shape effects in both experiments provides evidence that activation of perceptual attributes in language comprehension is neither limited to meta-linguistic tasks nor to a particular type of online method (visual world eye-tracking or ERPs). The present results extend previous research by showing that a relevant visual context is not required for such effects to occur. Of course, our findings do not imply that shape representations, or other representations of physical properties of the objects mentioned, are always activated during language processing. Determining when listeners and readers are likely, or unlikely to activate such information is an important issue for further research.

In addition to clarifying the types of information that can be predicted, the observed shape effect may also be useful in distinguishing between concept-specific prediction versus graded prediction of a range of concepts. In the case of functional semantic features such as edibility, anticipatory eye movements (Altmann & Kamide, 1999) could arise from specific prediction (e.g., in their example, from prediction of cake, especially when a cake is present in the display), or it could arise from prediction of a range of candidates (e.g., edible things) that could plausibly be mentioned next. This is because it is exactly this functional feature (being edible) that allows these objects to play a given role in the context. Either or both of these processes – prediction of one specific concept or of a range of candidates – could also have driven the reduction of N400 amplitude to words semantically related to an expected word in the Federmeier and Kutas (1999) study. In contrast, object shape representations are less likely to be shared across a class of expected objects, because having a similar shape does not necessarily allow different objects (e.g., tomato and moon) to play the same role in a given sentence. Our experiments show that when a specific object (e.g. the moon) is predictable, then this can affect the processing of objects with a similar shape. This evidence for shape pre-activation could potentially be a useful index of the degree of activation of a specific candidate concept (e.g., the moon).

Why does pre-activation of object shape information occur? Further research is required to understand the functionality of the prediction of object shapes. One can speculate that the pre-activation of shape information might represent a link between language and the visual environment, where activating a shape representation would be a useful preparation for object recognition. In visual environments, where speakers sometimes refer to objects that are present, predicting what the shape of an object is could help in directing visual attention to the object before it is actually referred to. Just as prediction of words is thought to speed up language comprehension (e.g., Kutas & Federmeier, 2000), the prediction of shape attributes of objects might, in some contexts, speed up language-vision interactions.

The processing mechanisms underlying the pre-activation of shape information also need to be investigated in further research. Shape information could be activated because the pre-activation of the expected word or concept causes activation to spread to knowledge connected to the lexical and conceptual representation corresponding to the word, which would include shape attributes of the referent (cf. Mahon & Caramazza, 2008). In addition, some form of mediated priming might contribute to the visual anticipation effect we observed (cf. Bar, 2007). For instance, the concept moon might prime the concept round which in turn primes the concept tomato. As Mani and Huettig (in press) have argued, predictive language processing may draw on several mechanisms. An important task for further research is to identify these mechanisms and their impact in language processing.

Conclusions

In sum, two experiments showed, in quite different ways, that listeners anticipated the shape of objects that were about to be mentioned. The results converge with previous findings from other paradigms and show that meta-linguistic tasks or a visual environment displaying the referents is not necessary for obtaining effects of anticipated shape. Predictive language processing is not limited to functional semantic information but can involve the activation of visual representations.

Acknowledgements

We thank Alma Veenstra for lending her voice to our materials, Tilman Harpe for drawing pictures for Experiment 1, Marloes Wensink and Sophie Kirkels for help with association norms, and three anonymous reviewers for comments.

References

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

Altmann, G. T. M., & Mirkovic, J. (2009). Incrementality and Prediction in Human Sentence Processing. *Cognitive Science, 33*(4), 583-609.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*, 390–412.

Bar, M. (2007). The proactive brain: Using analogies and associations to generate predictions. *Trends in Cognitive Sciences, 11(7),* 280-289.

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

Boersma, P., & Weenink, D. (2009). Praat: Doing phonetics by computer (Version 5.1).

Cooper, R. M. (1974). The control of eye fixation by the meaning of spoken language: A new methodology for the real-time investigation of speech perception, memory, and language processing. *Cognitive Psychology, 6,* 84−107.

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.

de Groot, A. M. B., & de Bil, J. M. (1987). *Nederlandse woordassociatienormen met reactietijden*. Lisse: Swets & Zeitlinger.

DeLong, K. A., Urbach, T. P., & Kutas, M. (2005). Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. *Nature Neuroscience, 8*(8), 1117-1121.

Federmeier, K. D. (2007). Thinking ahead: The role and roots of prediction in language comprehension. *Psychophysiology, 44*(4), 491-505.

Federmeier, K. D., & 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.

Federmeier, K. D., McLennan, D. B., De Ochoa, E., & Kutas, M. (2002). The impact of semantic memory organization and sentence context information on spoken language processing by younger and older adults: An ERP study. *Psychophysiology, 39*(2), 133-146.

Gibson, E. (1998). Linguistic complexity: locality of syntactic dependencies. *Cognition, 68*(1), 1-76.

Holcomb, P. J., & Anderson, J. E. (1993). Cross-modal semantic priming: A time-course analysis using event-related brain potentials. *Language and Cognitive Processes, 8(4),* 379-411.

Holcomb, P. J., & Neville, H. J. (1991). Natural speech processing: An analysis using event-related brain potentials. *Psychobiology, 19*(4), 286-300.

Huettig, F., & Altmann, G. T. M. (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.

Huettig, F., & McQueen, J. M. (2011). The nature of the visual environment induces implicit biases during language-mediated visual search. *Memory & Cognition, 39*, 1068-1084.

Huettig, F., Rommers, J., & Meyer, A. S. (2011). Using the visual world paradigm to study language processing: A review and critical evaluation. *Acta Psychologica, 137,* 151-171.

Junghöfer, M., Elbert, T., Tucker, D. M., & Rockstroh, B. (2000). Statistical control of artifacts in dense array EEG/MEG studies. *Psychophysiology, 37*(4), 523-532.

Kamide, Y. (2008). Anticipatory processes in sentence processing. *Language and Linguistics Compass, 2*, 647-670.

Kellenbach, M. L., Wijers, A. A., & Mulder, G. (2000). Visual semantic features are activated during the processing of concrete words: event-related potential evidence for perceptual semantic priming. *Cognitive Brain Research, 10*(1-2), 67-75.

Kounios, J. & Holcomb, P. J. (1994). Concreteness effects in semantic processing: ERP evidence supporting dual-coding theory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20,* 804-823.

Kutas, M., DeLong, K. A., Smith, N. J. (2011). A look around at what lies ahead: Prediction and predictability in language processing. In: M. Bar (Ed.), *Predictions in the Brain: Using Our Past to Generate a Future*. Oxford University Press.

Kutas, M., & Federmeier, K. D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends in Cognitive Sciences, 4*(12), 463-470.

Kutas, M., & Federmeier, K. D. (2011). Thirty Years and Counting: Finding Meaning in the N400 Component of the Event-Related Brain Potential (ERP). *Annual Review of Psychology, 62*, 621-647.

Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science, 207*(4427), 203-205.

Kutas, M., & van Petten, C. (1994). Psycholinguistics electrified: Event-related brain potential investigations. In M.A. Gernsbacher (ed.), *Handbook of Psycholinguistics* (pp. 83–143). San Diego, CA: Academic.

Laszlo, S. and Plaut, D. C. (2011). Simulating event-related potential reading data in a neurally plausible parallel distributed processing model. *Proceedings of the 33th Annual Conference of the Cognitive Science Society*. Mahwah, NJ: Lawrence Erlbaum Associates.

Lauteslager, M., Schaap, T., & Schievels, D. (1986). *Schriftelijke woordassociatienormen voor 549 zelfstandige naamwoorden*. Lisse: Swets & Zeitlinger.

Levy, R. (2008). Expectation-based syntactic comprehension. *Cognition, 106*(3), 1126-1177.

Luck, S. J. (2005). Ten simple rules for designing ERP experiments. In T. C. Handy (Ed.), *Event-Related Potentials: A Methods Handbook* (pp. 17-32). Cambridge, MA: MIT Press.

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*(1-3), 59-70.

Mani, N., & Huettig, F. (in press). Towards a complete multiple-mechanism account of predictive language processing: Commentary on Pickering & Garrod. *Behavioral and Brain Sciences*.

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods, 164*(1), 177-190.

McCallum, W. C., Farmer, S. F., & Pocock, P. V. (1984). The effects of physical and semantic incongruities on auditory event-related potentials. *Electroencephalography and Clinical Neurophysiology, 59*(6), 477-488.

Mitchell, D. C. (2004). On-line methods in language processing: Introduction and historical review. In M. Carreiras, & C. Clifton (Eds.), *The on-line study of sentence comprehension: Eyetracking, ERP and beyond* (pp. 15−32). New York, NY:Psychology Press.

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

Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J.-M. (2011). FieldTrip: Open Source Software for Advanced Analysis of MEG, EEG, and Invasive Electrophysiological Data. *Computational Intelligence and Neuroscience, 2011*, 9.

Oostenveld, R., & Praamstra, P. (2001). The five percent electrode system for high-resolution EEG and ERP measurements. *Clinical Neurophysiology, 112*(4), 713-719.

Otten, M., & van Berkum, J. J. A. (2008). Discourse-Based Word Anticipation During Language Processing: Prediction or Priming? *Discourse Processes, 45*(6), 464-496.

Pecher, D., Zeelenberg, R., & Raaijmakers, J. G. W. (1998). Does pizza prime coin? Perceptual priming in lexical decision and pronunciation. *Journal of Memory and Language, 38*, 401-418.

Perrin, F., & García-Larrea, L. (2003). Modulation of the N400 potential during auditory phonological/semantic interaction. *Cognitive Brain Research, 17,* 36-47.

Pickering, M. J., & Garrod, S. (2007). Do people use language production to make predictions during comprehension? *Trends in Cognitive Sciences, 11*(3), 105-110.

Rossion, B., Pourtois, G. (2004). Revisiting Snodgrass and Vanderwart's object pictorial set: The role of surface detail in basic-level object recognition. *Perception, 33,* 217-236.

Saslow, M. G. (1967). Latency for saccadic eye movement. *Journal of the Optical Society of America, 57,* 1030–1033.

Schreuder, R., Flores d'Arcais, G. B., & Glazenborg, G. (1984). Effects of perceptual and conceptual similarity in semantic priming. *Psychological Research, 45*, 339-354.

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

Taylor, W. L. (1953). Cloze procedure: a new tool for measuring readability. *Journalism Quarterly, 30*, 415-433.

van Berkum, J. J. A., Brown, C. M., Zwitserlood, P., Kooijman, V., & Hagoort, P. (2005). Anticipating upcoming words in discourse: Evidence from ERPs and reading times. *Journal of Experimental Psychology: Learning Memory and Cognition, 31*(3), 443-467.

van den Brink, D., Brown, C. M., & Hagoort, P. (2001). Electrophysiological evidence for early contextual influences during spoken-word recognition: N200 versus N400 effects. *Journal of Cognitive Neuroscience, 13*(7), 967-985.

van den Brink, D., Brown, C. M., & Hagoort, P. (2006). The cascaded nature of lexical selection and integration in auditory sentence processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32*(2), 364-372.

van Loon-Vervoorn, W. A., & van Bekkum, I. J. (1991). *Woordassociatielexicon*. Amsterdam/Lisse: Swets & Zeitlinger.

van Petten, C., & Luka, B. J. (2011). Prediction during language comprehension: Benefits, costs, and ERP components. *International Journal of Psychophsiology, 83,* 176-190.

Wicha, N. Y. Y., Moreno, E. M., & Kutas, M. (2004). Anticipating words and their gender: An event-related brain potential study of semantic integration, gender expectancy, and gender agreement in Spanish sentence reading. *Journal of Cognitive Neuroscience, 16*(7), 1272-1288.

Zwaan, R. A., Stanfield, R. A., & Yaxley, R. H. (2002). Language comprehenders mentally represent the shapes of objects. *Psychological Science, 13*(2), 168-171.

Appendix A

Table A1: Example of one of the twelve sets of 24 sentences in the materials of Experiment 1

Condition	Examples
Correct	Nadat de chocoladeliefhebber haar boterham had besmeerd met boter bestrooide ze hem met flink veel hagelslag. *After the chocolate lover had spread the bread with butter she covered it with a large amount of chocolate sprinkles.*
	In 1969 zette Neil Armstrong als eerste mens voet op de maan. *In 1969 Neil Armstrong was the first man to set foot on the moon.*
	Het echtpaar had pech met het weer en trouwde in de stromende regen. *The couple was unlucky weather-wise and got married in the pouring rain.*
	De Chinees gebruikte altijd eetstokjes bij het eten van zijn witte rijst. *The Chinese person always used chopsticks when eating his white rice.*
	De keukenhulp verkreeg jus d'orange door het persen van een stuk fruit genaamd sinaasappel. *The kitchen assistant obtained orange juice by squeezing a piece of fruit called orange.*
	Het was winter en er vielen uit de lucht dikke witte vlokken sneeuw. *It was winter and from the sky fell big white flakes of snow.*
	Pastasaus krijgt zijn rode kleur door een groente genaamd tomaat. *Pasta sauce gains its red color from a vegetable called tomato.*
	Het was warm want gedurende de hele dag scheen de zon. *It was hot because all day the sun shone.*
Shape	Nadat de chocoladeliefhebber haar boterham had besmeerd met boter bestrooide ze hem met flink veel regen. *After the chocolate lover had spread the bread with butter she covered it with a large amount of rain.*
	In 1969 zette Neil Armstrong als eerste mens voet op de tomaat. *In 1969 Neil Armstrong was the first man to set foot on the tomato.*
	Het echtpaar had pech met het weer en trouwde in de stromende hagelslag. *The couple was unlucky weather-wise and got married in the pouring chocolate sprinkles.*
	De Chinees gebruikte altijd eetstokjes bij het eten van zijn witte sneeuw. *The Chinese person always used chopsticks when eating his white snow.*
	De keukenhulp verkreeg jus d'orange door het persen van een stuk fruit genaamd zon. *The kitchen assistant obtained orange juice by squeezing a piece of fruit called sun.*
	Het was winter en er vielen uit de lucht dikke witte vlokken rijst. *It was winter and from the sky fell big white flakes of rice.*
	Pastasaus krijgt zijn rode kleur door een groente genaamd maan. *Pasta sauce gains its red color from a vegetable called moon.*
	Het was warm want gedurende de hele dag scheen de sinaasappel. *It was hot because all day the orange shone.*
Unrelated	Nadat de chocoladeliefhebber haar boterham had besmeerd met boter

bestrooide ze hem met flink veel zon.
After the chocolate lover had spread the bread with butter she covered it with a large amount of sun.
In 1969 zette Neil Armstrong als eerste mens voet op de rijst.
In 1969 Neil Armstrong was the first man to set foot on the rice.
Het echtpaar had pech met het weer en trouwde in de stromende sinaasappel.
The couple was unlucky weather-wise and got married in the pouring orange.
De Chinees gebruikte altijd eetstokjes bij het eten van zijn witte maan.
The Chinese person always used chopsticks when eating his white moon.
De keukenhulp verkreeg jus d'orange door het persen van een stuk fruit genaamd regen.
The kitchen assistant obtained orange juice by squeezing a piece of fruit called rain.
Het was winter en er vielen uit de lucht dikke witte vlokken tomaat.
It was winter and from the sky fell big white flakes of tomato.
Pastasaus krijgt zijn rode kleur door een groente genaamd sneeuw.
Pasta sauce gains its red color from a vegetable called snow.
Het was warm want gedurende de hele dag scheen de hagelslag.
It was hot because all day the chocolate sprinkles shone.

Note. The same sentence contexts and critical words were recombined to create the different conditions. English translations of the Dutch sentences are provided in Italics. In one case in the English translation a critical word consists of two words (*chocolate sprinkles*), but the actual Dutch critical words were all single words. Critical words are underlined.

Table A2: Full set of critical triplets

Correct	Shape	Unrelated
aardappel	hoofd	duim
(potatoe)	(head)	(thumb)
appel	handgranaat	sabel
(apple)	(hand grenade)	(sabre)
arm	honkbalknuppel	tennisbal
(arm)	(baseball bat)	(tennis ball)
augurk	sabel	handgranaat
(pickle)	(sabre)	(hand grenade)
baksteen	mondharmonica	triangel
(brick)	(mouth organ)	(triangle instrument)
banaan	boemerang	bom
(banana)	(boomerang)	(bomb)
banjo	tennisracket	peddel
(banjo)	(tennis racket)	(paddle)
basketbal	hersenen	been

(basketball)	(brain)	(leg)
been	speer	basketbal
(leg)	(javelin)	(basketball)
boemerang	banaan	citroen
(boomerang)	(banana)	(lemon)
bom	citroen	banaan
(bomb)	(lemon)	(banana)
bord	dvd	videoband
(plate)	(DVD)	(video cassette)
casettebandje	snijplank	taartvorm
(cassette tape)	(chopping board)	(cake tin)
chocoladereep	prikbord	klok
(chocolate bar)	(notice board)	(clock)
citroen	bom	boemerang
(lemon)	(bomb)	(boomerang)
dak	plectrum	keyboard
(roof)	(guitar pick)	(keyboard instrument)
deur	parkeerplaats	slagboom
(door)	(parking place)	(gate)
didgeridoo	hockeystick	pingpongbatje
(didgeridoo)	(hockey stick)	(ping pong racket)
duim	komkommer	aardappel
(thumb)	(cucumber)	(potatoe)
dvd	bord	tafel
(DVD)	(plate)	(table)
flat	keyboard	plectrum
(block of flats)	(keyboard instrument)	(guitar pick)
fluit	peddel	tennisracket
(flute)	(paddle)	(tennis racket)
gitaar	pingpongbatje	hockeystick
(guitar)	(ping pong racket)	(hockey stick)
hagelslag	regen	zon
(chocolate sprinkles)	(rain)	(sun)
hamer	lantaarnpaal	put
(hammer)	(lamppost)	(drain)
handgranaat	appel	augurk
(handgrenade)	(apple)	(pickle)
hek	spoor	nummerbord
(fence)	(railway track)	(number plate)
hersenen	basketbal	speer
(brain)	(basketball)	(javelin)
hockeystick	didgeridoo	gitaar
(hockey stick)	(didgeridoo)	(guitar)
hoed	schildpad	slang
(hat)	(turtle)	(snake)
honkbalknuppel	arm	oog
(baseball bat)	(arm)	(eye)

hoofd
(head)
horloge
(watch)
kaas
(cheese)
keyboard
(keyboard instrument)
klok
(clock)
koekje
(biscuit)
komkommer
(cucumber)
kompas
(compass)
ladder
(ladder)
lantaarnpaal
(lamppost)
lp
(LP record)
maan
(moon)
moer
(hardware nut)
mondharmonica
(mouth organ)
muts
(hat/cap)
nummerbord
(number plate)
oog
(eye)
pannenkoek
(pancake)
parkeerplaats
(parking place)
peddel
(paddle)
pingpongbatje
(ping pong racket)
pizza
(pizza)
plectrum
(guitar pick)
prei

aardappel
(potatoe)
worm
(wurm)
schilderij
(painting)
flat
(block of flats)
pannenkoek
(pancake)
spiegel
(mirror)
duim
(thumb)
put
(drain)
slagboom
(gate)
hamer
(hammer)
taartvorm
(cake tin)
tomaat
(tomato)
rotonde
(roundabout)
baksteen
(brick)
slak
(snail)
raam
(window)
tennisbal
(tennis ball)
klok
(clock)
deur
(door)
fluit
(flute)
gitaar
(guitar)
wiel
(wheel)
dak
(roof)
ruitenwisser

komkommer
(cucumber)
slak
(snail)
spiegel
(mirror)
dak
(roof)
chocoladereep
(chocolate bar)
schilderij
(painting)
hoofd
(head)
lantaarnpaal
(lamppost)
parkeerplaats
(parking place)
kompas
(compass)
snijplank
(chopping board)
rijst
(rice)
stoplicht
(traffic light)
tent
(tent)
worm
(wurm)
hek
(fence)
honkbalknuppel
(baseball bat)
prikbord
(notice board)
ladder
(ladder)
banjo
(banjo)
didgeridoo
(didgeridoo)
versnellingspook
(gear shift)
flat
(block of flats)
stuur

(leek)
prikbord
(notice board)
put
(drain)
raam
(window)
regen
(rain)
riem
(belt)
rijst
(rice)
rotonde
(roundabout)
ruitenwisser
(windshield wiper)
sabel
(sabre)
schilderij
(painting)
schildpad
(turtle)
sinaasappel
(orange)
slagboom
(gate)
slak
(snail)
slang
(snake)
sneeuw
(snow)
snijplank
(chopping board)
speer
(javelin)
spiegel
(mirror)
spijker
(nail)
spoor
(railway track)
stokbrood
(baguette)
stoplicht
(traffic light)

(windshield wiper)
chocoladereep
(chocolate bar)
kompas
(compass)
nummerbord
(number plate)
hagelslag
(chocolate sprinkles)
slang
(snake)
sneeuw
(snow)
moer
(hardware nut)
prei
(leek)
augurk
(pickle)
kaas
(cheese)
hoed
(hat)
zon
(sun)
ladder
(ladder)
muts
(hat/cap)
riem
(belt)
rijst
(rice)
casettebandje
(cassette tape)
been
(leg)
koekje
(biscuit)
stoplicht
(traffic light)
hek
(fence)
versnellingspook
(gear shift)
spijker
(nail)

(steering wheel)
pannenkoek
(pancake)
hamer
(hammer)
spoor
(railway track)
sinaasappel
(orange)
schildpad
(turtle)
maan
(moon)
spijker
(nail)
taart
(pie)
appel
(apple)
koekje
(biscuit)
riem
(belt)
regen
(rain)
deur
(door)
horloge
(watch)
hoed
(hat)
tomaat
(tomato)
lp
(LP record)
hersenen
(brain)
kaas
(cheese)
rotonde
(roundabout)
raam
(window)
wiel
(wheel)
moer
(hardware nut)

stuur	taart	prei
(steering wheel)	(pie)	(leek)
taart	stuur	ruitenwisser
(pie)	(steering wheel)	(windshield wiper)
taartvorm	lp	casettebandje
(cake tin)	(LP record)	(cassette tape)
tafel	videoband	dvd
(table)	(video cassette)	(DVD)
tennisbal	oog	arm
(tennis ball)	(eye)	(arm)
tennisracket	banjo	fluit
(tennis racket)	(banjo)	(flute)
tent	triangel	mondharmonica
(tent)	(triangle instrument)	(mouth organ)
tomaat	maan	sneeuw
(tomato)	(moon)	(snow)
triangel	tent	baksteen
(triangle instrument)	(tent)	(brick)
ui	vuist	vinger
(onion)	(fist)	(finger)
versnellingspook	stokbrood	pizza
(gear shift)	(baguette)	(pizza)
videoband	tafel	bord
(video cassette)	(table)	(plate)
vinger	wortel	ui
(finger)	(carrot)	(onion)
vuist	ui	wortel
(fist)	(onion)	(carrot)
wiel	pizza	stokbrood
(wheel)	(pizza)	(baguette)
worm	horloge	muts
(wurm)	(watch)	(hat/cap)
wortel	vinger	vuist
(carrot)	(finger)	(fist)
zon	sinaasappel	hagelslag
(sun)	(orange)	(chocolate sprinkles)

Appendix B

In supplementary analyses of Experiment 2 suggested by reviewers of the paper, we treated the shape-relatedness of the words' referents to the Correct words' referents in the Shape and in the Unrelated condition as a continuous variable and examined whether there was a correlation between this variable and the amplitude of the N400. If the N400 difference between the Shape and Unrelated condition is due to a difference in the relatedness of the shapes of the referents, such a correlation might be seen. The measure for the degree of shape relatedness between a competitor and a target was the average rating given by the participants in the norming study. All analyses were performed on the EEG signal from the shape cluster reported in the main text.

The first analysis was a within-subjects correlation. For each participant we calculated the correlation coefficient between the shape similarity ratings of the items (stemming from other participants) and the corresponding EEG signal. This was done for all items in the Shape and Unrelated conditions and involved no averaging across trials. The average correlation was $r = -0.007$ (r range -0.328 to 0.269) and did not differ significantly from zero, $t\,(23) = -0.237$, $p = 0.815$).

The second analysis was an item analysis. We computed the N400 amplitude for each item across participants and correlated it with the average similarity rating. This correlation was also close to 0 ($r = .055$, $p = .642$). Note that each item was heard by only eight participants in the main experiment and rated by different participants in the rating study.

Finally, we sorted the items, according to the average shape similarity ratings, into four 16-item bins of increasing relatedness and assessed whether the amplitude of the N400 increased across the bins. However, we found that the linear trend was not significant, $F\,(1,23) = 0.002$, $p = .966$.

In our view, these results are not very informative. They do not support our assumption that the N400 difference between the Shape and Unrelated condition indeed arose because the items differed in their shape similarity to the targets. However, being null-results, they do not undermine our argument either. Power was low given that in the analyses above each data point stemmed from maximally 1, 8, or 16 trials, respectively. Future research could obtain ERPs with a higher signal-to-noise ratio using many more items (e.g., 30-60; Luck, 2005) and/or participants (e.g., over 100; Laszlo & Plaut, 2011).

Chapter 4

Looking ahead using arrows or words: A shared tendency to predict in spatially-cued and language-mediated attention

Rommers, J., Meyer, A. S., & Huettig, F. (in preparation). Looking ahead using arrows or words: A shared tendency to predict in spatially-cued and language-mediated attention.

Abstract

During language comprehension listeners often anticipate upcoming information, as shown by anticipatory overt attention being drawn to visually presented objects before they are referred to. We investigated whether the anticipatory mechanisms involved in such linguistic prediction are shared with other domains of cognition. Participants listened to sentences ending in a highly predictable word (e.g., “In 1969 Neil Armstrong was the first man to set foot on the moon”) while viewing displays containing three unrelated distractor objects and a critical object, which was either the target object (e.g., a moon), or an object with a similar shape (e.g., a tomato), or an unrelated control object (e.g., rice). As in an earlier study, language-mediated anticipatory eye movements to targets and shape competitors were observed. Importantly, the effects were systematically related to individual differences in anticipatory attention, as indexed by a spatial cueing task: Participants whose responses were most strongly facilitated by predictive arrow cues also showed the strongest effects of predictable language input on their eye movements. These results suggest that some mechanisms of anticipatory attention are shared across verbal and nonverbal tasks. Possible mechanisms are discussed.

Looking ahead using arrows or words: A shared tendency to predict in spatially-cued and language-mediated attention

During language comprehension readers and listeners often anticipate upcoming information (Altmann & Kamide, 1999; DeLong, Urbach, & Kutas, 2005; Federmeier & Kutas, 1999; van Berkum, Brown, Kooijman, Zwitserlood, & Hagoort, 2005; Wicha, Moreno, & Kutas, 2004). This has, for instance, been established in eye-tracking studies where participants listen to utterances while viewing objects. For example, when participants hear sentences such as "The boy will eat the cake" while viewing a display in which a cake is the only edible object, they look more at the cake than at other objects before the cake is mentioned (Altmann & Kamide, 1999). Such studies demonstrate that listeners can predict upcoming information.

Prediction is also a characteristic of many tasks that do not involve language. For instance, in the Posner spatial cueing paradigm (Posner, Nissen & Ogden, 1978), participants are asked to press a button (left or right) to indicate on which side of the centre a symbol (e.g., an X) occurs. Before stimulus onset, a central endogenous cue is presented, which can be valid (an arrow pointing to the side where the stimulus will occur), neutral (e.g., "+"), or invalid (an arrow pointing to the other side). Response times are shortest after a valid cue, longest after an invalid cue, and intermediate after a neutral cue. The response time difference between the neutral and valid cue conditions is a measure of the benefit from knowing where a stimulus will occur, termed attentional orienting (Posner, 1980). EEG studies have provided evidence for perceptual facilitation and response preparation during this task (for reviews, see Coles, 1989; Mangun, 1995), consistent with the idea that participants use the cues to anticipate the location of the upcoming stimulus (Klein, 1994; Nobre, Rohenkohl, & Stokes, 2012; Wang, Fan, & Johnson, 2004). Another example of non-linguistic prediction comes from tasks where participants observe everyday actions (e.g., a woman lifting a cup to her mouth). Here participants typically fixate the goal region (e.g., the mouth) before the object reaches this location (Hunnius & Bekkering, 2010). Such predictions are thought to involve the motor system (Stapel, Hunnius, van Elk, & Bekkering, 2010).

The question arises to what extent the mechanisms of prediction are specific to language tasks or are shared with other domains of cognition. In a recent study we asked participants to listen to sentences ending in highly predictable words (e.g. "Neil Armstrong was the first man to set foot on the moon"). The visual displays featured three distractor objects and either the target object (e.g., a moon), or an object with a similar shape (e.g., a tomato), or an unrelated control object (e.g., rice). We found that prior to the onset of the target word, both the target object and the shape-related object (but not the control object) were fixated more often than the unrelated distractors, suggesting that visual representations can be pre-activated (Rommers, Meyer, Praamstra, & Huettig, 2013). Here we followed up on this study to investigate the mechanisms of predictive language processing, more specifically the possible relation to prediction in tasks that do not involve linguistic processing.

One way to approach this issue is to investigate individual differences (Cohen, 1994; Kosslyn et al., 2002; Vogel & Awh, 2008). Underwood (1975), for instance, proposed that theories should be formulated in such a way as to allow for individual-differences tests, and moreover that such tests should be the first step in the assessment of a theory. If the performance in two tasks is substantially correlated, this can be regarded as a "go-ahead signal" for a theory proposing that the cognitive components or abilities assessed in the tasks are related.

A possible mechanism specific to language processing is that predictions in language comprehension are generated by the language production system (Chang, Dell, & Bock, 2006; Federmeier, 2007; Pickering & Garrod, 2007, in press; van Berkum et al., 2005). Pickering and Garrod (in press) propose that predictions can be generated via simulations that build on production experience and involve the listeners covertly producing what they would say as a speaker in the given context (a forward model). This hypothesis has gained support from studies showing that listeners who consistently predict upcoming information perform relatively well in production tasks. For instance, performance in the category fluency task, where participants produce as many members of a semantic category (e.g., animals) as they can in one minute, has been found to correlate with the amplitude of electrophysiological components associated with predictive language comprehension (Federmeier, Kutas, & Schul, 2010; Federmeier, McLennan, De Ochoa, and Kutas, 2002; but see Federmeier & Kutas, 2005). Furthermore, anticipatory eye movements to referents of predictable words have been observed in infants with a large productive vocabulary, but not in infants with a small productive vocabulary (Mani & Huettig, 2012).

Prediction in language might also draw on domain-general mechanisms. Many authors introduce the topic of prediction in language processing by drawing analogies to other domains, such as motion extrapolation in catching a Frisbee (van Berkum et al., 2005; see also DeLong et al., 2005; Kamide, 2008; van Petten & Luka, 2012). Pickering and Garrod’s (2007, in press) forward models are inspired directly by action-perception research. Altmann and Mirković (2009) argue that common representations allow listeners to predict both how a sentence will unfold and how the (non-linguistic) events that the sentence refers to will unfold. Furthermore, associations have been put forward as a basis for prediction both in language comprehension (Pickering & Garrod, in press; in addition to the production-based route) and in visual perception (Bar, 2007). Finally, the theory of predictive coding has been suggested as a general neural mechanism for perception and action. According to this theory, predictions are used to "explain away" the bottom-up sensory input as much as possible. Only residual “prediction errors” (the difference between what was predicted and what was actually presented) propagate forward through further processing stages (e.g., Clark, 2013; Friston, 2005). Recently, Clark (2013) proposed that such prediction models hold promise to ground a unified science of mind, brain, and action. Because such a science would include language processing, similar mechanisms should be important for language. Despite these suggestions, to our knowledge individual differences studies of language processing have not considered non-linguistic factors in prediction.

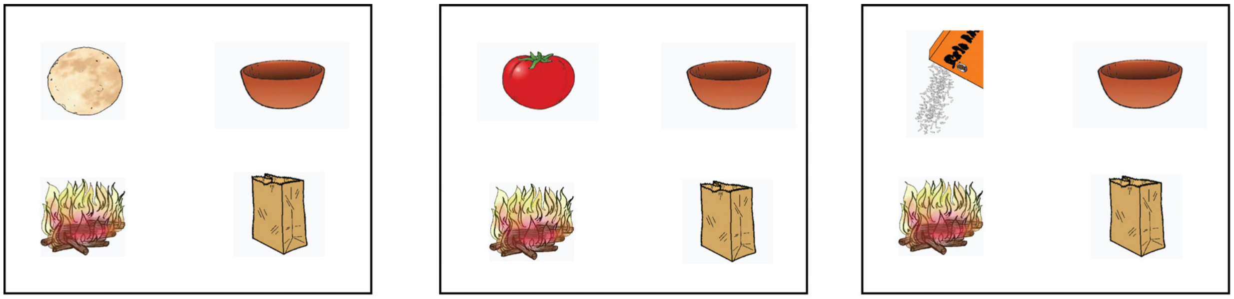
In the current study we investigated whether anticipatory attention in non-linguistic tasks is related to language-mediated anticipatory attention. We replicated the visual world experiment of Rommers et al. (2013) which was described above, except that we now used a longer display preview time: While in the earlier study, the objects appeared only 500 ms before the onset of the critical word, they now appeared before sentence onset. Thus, the participants of the present study could view the objects over the course of the entire spoken sentence, as is the case in most published visual word studies (cf. Huettig, Rommers, & Meyer, 2011). We reasoned that the longer preview time of the displays would allow for more variation in anticipatory eye movements. We examined the relationships between performance in this task and two nonverbal anticipation tasks completed by the same participants. In one task, anticipatory spatial attention was quantified using a classic Posner spatial cueing paradigm (described above). Such tasks have been used widely in the literature, and measure individual differences in orienting with good reliability (Fan, McCandliss, Sommer, Raz, & Posner, 2002). If language-mediated anticipatory eye movements are related to anticipatory spatial attention in non-language tasks, we should see a correlation between performance in these tasks. In addition, we used an action observation task (described above) where action goals could be anticipated. This task has been used with infants and adults, and both groups showed anticipatory effects (Hunnius & Bekkering, 2010). If language-mediated anticipation and action anticipation are based on related mechanisms, we should see a correlation between performance in language-mediated eye movements and the action observation task.

In addition, we collected data from three control tasks. First, we used Raven's advanced progressive matrices test (Raven, Raven, & Court, 1993) to assess and control for differences between participants in terms of fluid intelligence. Second, we used the category fluency task described above as a measure of the speed of accessing words during language production (see Delis, Kaplan, & Kramer, 2001). This task is commonly used in neuropsychological investigations but also in research on bilingualism. Performance is thought to rely primarily on lexical retrieval (in contrast to letter fluency, which has increased demands for executive control; Luo, Luk, & Bialystok, 2010). Third, the participants' vocabulary size was assessed with the Peabody picture vocabulary test (Dunn & Dunn, 1997; Dutch translation by Schlichting, 2005). We reasoned that, even if prediction in language processing makes use of domain-general mechanisms, linguistic knowledge is required as well. In addition, participants with larger vocabularies might be more able to use the sentence context for prediction (Federmeier et al., 2002).

Methods

Participants

Forty-five adult participants (14 men; mean age 21 years, range 18-32 years) were paid for participation in the study. Thirty-six were current or former students of Radboud University Nijmegen or another university, eight were current or former students from the HAN University of Applied Sciences, one had a technical and vocational training background. All were native speakers of Dutch and had normal hearing and normal or corrected-to normal vision.



Target Shape competitor Control

Figure 1. Example displays for each of the three conditions for the sentence "In 1969 Neil Armstrong was the first man to set foot on the moon". Shown along with the three distractors (a bowl, a fire, and a bag): the Target condition (with a moon), the Shape competitor condition (with a tomato), and the Control condition (with rice).

Stimuli, design, and procedure

Participants performed the tasks in the order given below. The fluid intelligence test was administered in a second session within three weeks of the first session.

Category fluency. Participants' speech was recorded with a Sennheiser microphone as they produced as many members of a category (indicated by a word on a screen) as they could in one minute. The categories were animals and professions. The time course of lexical retrieval was characterized by three measures (Rohrer, Wixted, Salmon, & Butters, 1995), averaging across the two categories: The total number of correct responses (excluding mistakes and repetitions), first-response latency (onset time of the first response), and subsequent-response latency (the average interval between the first response and each subsequent response). The latter measure indicates the mean retrieval speed once retrieval has started, that is, excluding task initiation processes.

Visual world experiment. Participants listened to predictable words in sentence contexts (e.g., the Dutch translation equivalent of "In 1969 Neil Armstrong was the first man to set foot on the moon") while looking at visual displays of four objects that differed depending on the condition (see Figure 1). The 96 experimental items and 32 fillers from Rommers et al. (2013) were distributed across three counterbalanced presentation lists, individually randomized with respect to trial order and picture position. Each sentence and picture was presented once.

Participants were asked to listen to the sentences carefully and were told that they were free to look at whatever they wanted to, but that they should not take their eyes off the screen. They were tested individually in a dimly illuminated room, seated in front of a screen with their chin on a chin rest. The movements of the right eye were recorded with an Eyelink 1000 Tower Mount eye tracker sampling at 1000 Hz.

Each trial started with a central fixation circle that remained on the screen until the participant fixated it, allowing for drift correction. Then four objects appeared on the screen. After 1 second, a spoken sentence started playing. The objects remained on the screen until 2000 ms after sentence offset. Then a blank screen appeared for 500 ms. Data were coded in terms of fixations, saccades, and blinks using the algorithm provided in the Eyelink software.

Peabody picture vocabulary test. On each trial, participants heard a word and saw four numbered pictures on the screen. They indicated which of the pictures corresponded to the spoken word by typing in the number (1, 2, 3, or 4) on a keyboard. A percentile score was calculated based on Dutch norms.

Action observation. The materials, interest areas, and time windows from Hunnius and Bekkering (2010) and Stapel et al. (2010) were used. Short videos were presented in which a female actor brought a cup to her mouth, a phone to her ear, or a brush to her hair (the anomalous conditions from the original experiments were excluded). Participants were told that they could look where they wanted without taking their eyes off the screen. Their eye movements were tracked and the stimuli displayed using a Tobii T120 monitor with integrated eye-tracker, sampling at 120 Hz.

The experiment consisted of 87 trials. Following Hunnius and Bekkering (2010), in the first 27 trials each of their three videos was shown nine times (order: cup, phone, brush). Then the twelve stimuli from Stapel et al. (2010; different phones and cups) were presented in random order with each video being shown five times. The frequency of anticipatory looks was computed as the percentage of trials on which the goal region (e.g., the mouth in a video with a cup) was fixated minus the number of trials with a fixation to a control region (e.g., the mouth in a video with a phone) during the lifting phase, when the actor had grasped the object but it had not yet reached its goal.

Posner spatial cueing task. The design followed Posner, Nissen, and Ogden (1978) except that there was a single session rather than three sessions. Participants pressed one of two buttons (left/right) to indicate the location of an X symbol on the screen. Each trial started with a central fixation dot that participants were asked to keep fixating. Then a cue appeared in the position of the dot for 100 ms. The participants were told that on half the trials, this cue was neutral (+), and on the other half, an arrow cue pointing left (<) or right (>) indicated the location of the upcoming X with 80% validity (i.e., the arrow cue was correct 80% of the time). After cue offset, the fixation dot was presented for another 400 ms. Then the target appeared on the left or the right of the centre of the screen for 1700 ms or until a response was made. After every trial, feedback was displayed for 500 ms ("correct", "incorrect", or "too late").

There were seven blocks of 32 trials each; the first block served as practice. Each block consisted of 16 neutral trials, 13 valid trials, and 3 invalid trials in random order, with counterbalanced target position (left, right) and initial fixation dot duration (400, 800, 1200, or 1600 ms). Because we noticed that on a very small portion of trials a response occurred before target onset, which was not registered and usually elicited a second but late response, we excluded trials with response times 2.5 standard deviations from each participant's mean.

Anticipatory attention or “orienting” was quantified as the difference between the mean response latency of the neutral and valid condition.

Raven's advanced progressive matrices. Participants indicated which of eight possible shapes completed a matrix of geometric patterns by clicking on it with a mouse. Items could be skipped and were then shown again at the end of the test with the option to click an "I don't know" button. Participants had 40 minutes to complete 36 items. The time was indicated in the right top corner of the screen. A participant's score was the total number of correct responses.

Results

Figure 2 shows the time course of language-mediated eye movements. Participants tended to fixate upon the Targets much more than on the competitors or distractors, several seconds before critical word onset. They were also more likely to fixate upon the Shape competitors than upon the unrelated Control objects and distractors, especially during the last second before critical word onset.

Because it takes on average about 200 ms to program and initiate a saccadic eye movement (Saslow, 1967), we chose a time window from 200 ms after sentence onset until critical word onset for the statistical analysis. Fixation proportions within this time window were transformed to log odds, the appropriate scale for assessing effects on a categorical dependent variable, using the empirical logit function (Barr, 2008). The average log odds of looks to the unrelated distractors was subtracted from the log odds of looks to the Target/Shape/Control object to create the dependent variable, which indicates the strength of any bias toward each experimental object over the unrelated distractors. The results will be reported below.

Group-level results for the nonverbal anticipation tasks are shown in Figure 3. In the spatial cueing task, response latencies were longest after invalid cues, shortest after valid cues, and intermediate after neutral cues, all t's > 4.212, ps < .001. In the action observation task,

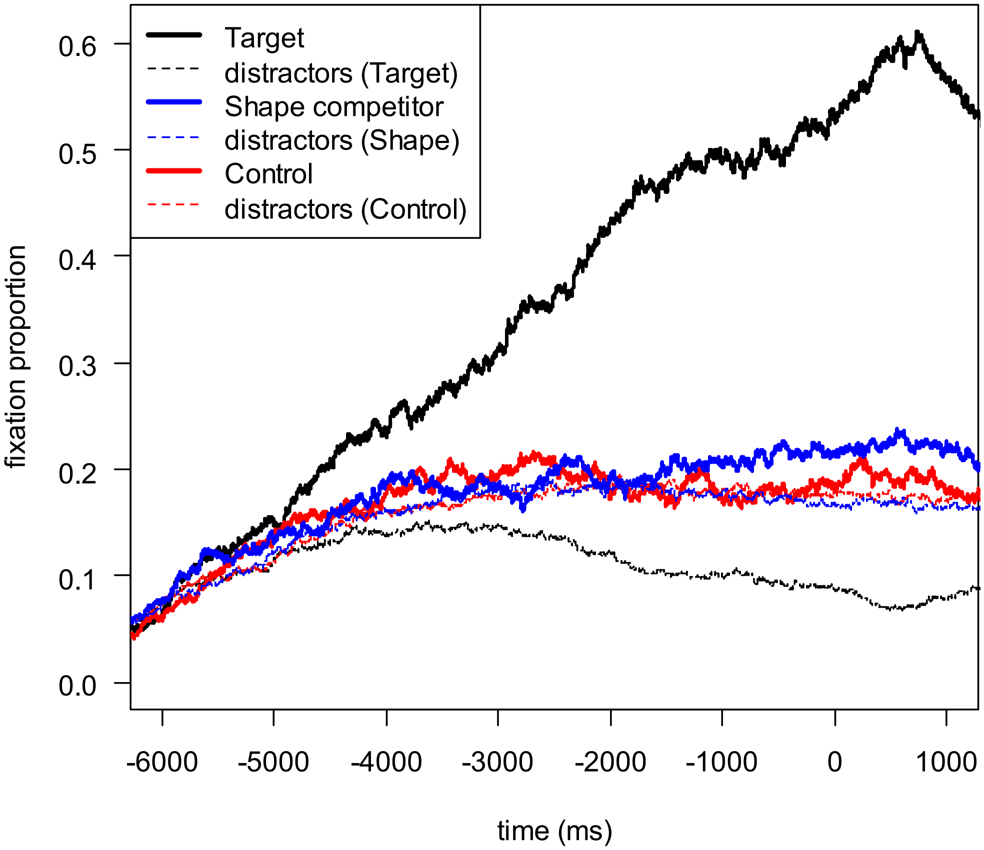


Figure 2. Language-mediated eye movements. Time-course graph showing fixation proportions to Targets, Shape competitors, and Control objects (solid lines) along with fixation proportions averaged across the three corresponding unrelated distractors (dashed lines). The y-axis reflects the proportion of trials in which participants were fixating each type of object. Time zero indicates critical word onset. Anticipatory effects are visible as lines of the same color diverging before time zero.

participants anticipatorily fixated the action goal region more often than the control region, t (44) = 11.834, $p < .001$.

From these and the other background tasks seven measures were derived (see Table 1; The correlations between all measures after averaging by participants are shown in an Appendix). Regarding the background measures, participants with higher intelligence ("ravens") also had higher vocabulary ("pbody"), and participants with more correct responses in the category fluency task ("count") also had higher subsequent-response latencies ("subs") in the same task. In

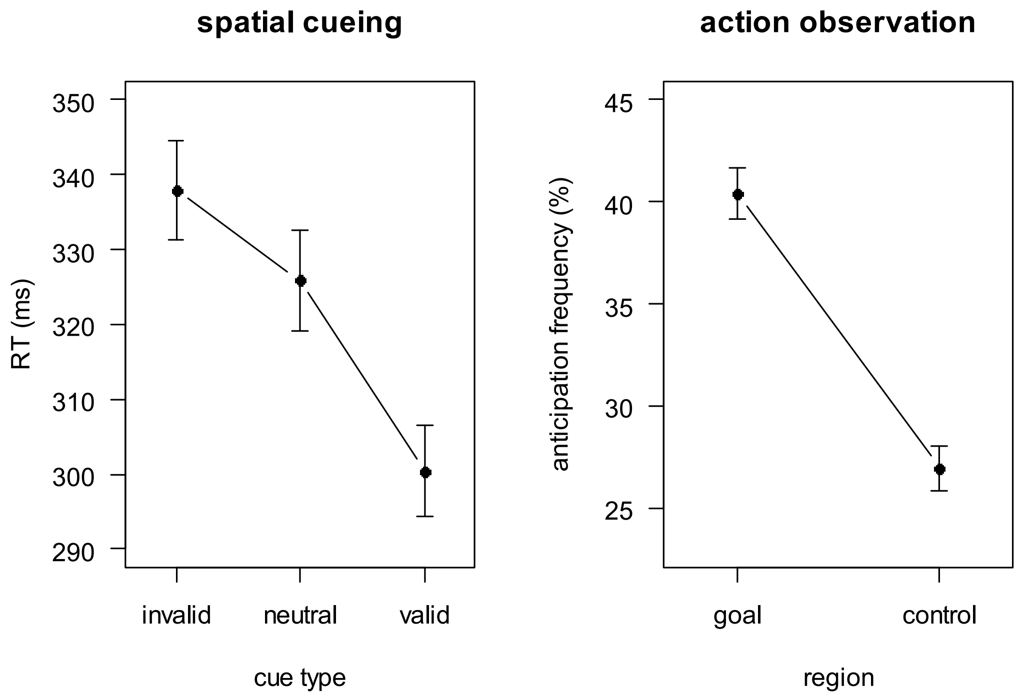


Figure 3. Left panel: Average response latencies in the spatial cueing task. The orienting measure was obtained by subtracting RTs after valid cues from RTs after neutral cues. Right panel: Average anticipation frequency (percentage of trials) in the action observation task. The action goal anticipation measure was obtained by subtracting fixation frequency to the control region from fixation frequency to the goal region. Error bars indicate standard error of the mean.

addition, participants with higher intelligence and/or larger vocabularies tended to start speaking earlier ("onset") in the verbal fluency task. The measures of special interest to this study, orienting and action anticipation, were not correlated with any of the other background measures.

The data were analyzed using linear mixed-effects regression models, which allow for simultaneous inclusion of participants and items as random factors (Baayen, Davidson, & Bates, 2008). Three measures were residualized for being correlated with another measure: Peabody vocabulary size (controlling for Raven's intelligence), category fluency first-response latency (controlling for Raven's intelligence), and category fluency subsequent-response latency (controlling for total responses). Subsequently all background measures were mean-centered,

Table 1

Group performance on the background measures

Task	Measure	Mean	SD
Raven's advanced progressive matrices	Answers correct	23.4	5.3
Posner spatial cueing	Orienting (neutral - valid; ms)	25	21
Action observation	Anticipation frequency (goal - control; %)	13.4	7.6
Peabody picture vocabulary test	Percentile	55	22
Category fluency	Correct responses	20.1	3.9
	First-response latency (s)	2.52	0.89
	Subsequent-response latency (s)	21.93	2.59

reducing collinearity and making the intercept interpretable as a participant with average performance on all measures.

Models were fit separately for the Target, Shape, and Control displays. A backwards elimination procedure was used, starting from an initial model where all background measures were entered as fixed effects, as well as by-participants and by-sentence random intercepts. Non-significant fixed effects were then removed (based on *p*-values calculated assuming that the *t*-values were drawn from a normal distribution; Barr, 2008). The resulting model was compared to the initial full model using a likelihood ratio test. For the remaining fixed effects, by-sentence random slopes were fitted, resulting in the maximal random effects structure possible. This final model was compared to a baseline model with the same random effects structure but without any fixed effects.

Table 2

Coefficients of linear mixed-effects regression models of the language-mediated anticipatory eye movements

Condition	Term	*β*	*SE*	*t*	*p*
Target	Intercept	2.247700	0.174449	12.885	< .001
	Orienting	0.013788	0.007802	1.767	.077
Shape competitor	Intercept	0.180362	0.077023	2.342	.019
	Orienting	0.007452	0.003095	2.408	.016
Control	Intercept	-0.038560	0.079420	-0.486	.627

Results from the final models are shown in Table 2. For anticipatory Target bias, while removal of other fixed effects did not affect model fit, χ^2 (6) = 6.1, *p* = .410, orienting showed a trend towards improving model fit, χ^2 (1) = 3.1, *p* = .078. Participants anticipatorily looked more at the Target object (e.g., the moon) than at the unrelated distractors in the same display, as indicated by the intercept (Table 2). A trend towards an effect of orienting suggested that participants whose responses in the Posner cueing task were facilitated to a higher degree by valid cues also tended to show an increased anticipatory bias to the target (e.g., moon).

For anticipatory Shape competitor bias, while removal of other predictors did not affect model fit, χ^2 (6) = 3.5, *p* = .750, inclusion of the fixed effect of orienting was warranted, χ^2 (1) = 5.6, *p* = .018. Participants anticipatorily fixated Shape competitors more often than unrelated distractors, as indicated by the intercept, and this effect was stronger for participants whose responses were more strongly facilitated by valid cues in the Posner cueing task, as indicated by the effect of orienting (Table 2).

The initial model for the Control condition indicated no significant contribution of any fixed effect, all *p*s > .336, including the intercept. All background measures could be removed without loss of fit, χ^2 (7) = 0.0, *p* = 1. Thus, there was no evidence for a preference for the Control objects over the unrelated distractors, and the background measures did not contribute to explaining variance in these data.

Discussion

The participants preferentially fixated the target objects several seconds before the target name was mentioned, demonstrating that they predicted the objects to be referred to. There was also a smaller but reliable bias toward objects with a shape similar to the upcoming referent (e.g., a tomato), visible at about a second before critical word onset. This is consistent with the notion that predictions can involve the referent's shape. These results replicate our previous findings (Rommers et al., 2013) with a different display preview time.

Our main interest was in comparing the participants' tendency to predict across different tasks. We observed a positive relationship between anticipatory attention in the Posner cueing task and the language-mediated eye movements, despite differences between the tasks in terms of the cues that enabled prediction (visually presented arrows vs. various aspects of linguistic context) and the contents of the predictions themselves (the location of an X symbol vs. upcoming referents). The results are consistent with the notion that some anticipatory mechanisms are shared between verbal and nonverbal tasks.

What could these shared mechanisms be? One possibility is that in both tasks predictions were based on associations. In our visual world task, the sentences included associated words which could have led to pre-activation of the target. Similarly, in the spatial cueing task, the direction of the arrow may have become associated with the location of relevant information on the screen (see Clohessy, Posner, & Rothbart, 2001). Associations are typically thought of as reflecting automatic processes (Shanks, 2007), and there is some evidence that attention shifts induced by spatial cues are automatic (Hommel, Pratt, Colzato, & Godijn, 2001). Bar (2007) has suggested that by taking context into account, the most relevant associations can be selected to generate predictions, thus going beyond simple "passive" spreading activation (cf. Lau, Holcomb, & Kuperberg, in press; Otten & van Berkum, 2007, 2008). In the visual world experiment, such a process may have enabled participants to anticipate that the target (e.g., the moon) rather than other less relevant objects (for example, a rocket) would be mentioned. Predictions could be further verified and updated through a domain-general mechanism such as predictive coding, making use of prediction error. Recent studies suggest that processing can be based on prediction error in language tasks (Dikker & Pylkkänen, in press; Gagnepain, Henson, & Davis, 2012), and future studies could investigate whether this forms part of a wide panoply of predictive mechanisms.

With respect to the other measures, there was no robust relationship between individual differences in the anticipation of action goals and language-mediated eye movements, thus not supporting motor predictions as a common basis. Furthermore, there was no reliable relationship between the tendency to anticipate action goals and anticipatory attention in the spatial cueing task. This rules out an alternative account of the correlation between the spatial cueing task and the language-mediated anticipatory eye movements, namely that both eye movements and spatial cueing effects reflected participants' thresholds for shifting attention. If all participants predicted to the same degree but differed in their thresholds for shifting attention, the participants with low thresholds for shifting attention would show the largest effects in both tasks, thus accounting for the correlation between the two measures without appealing to individual differences in the tendency to predict. However, eye movements in the action observation task also represent shifts of attention. Thus, on this account action observation performance should have correlated with the language-mediated eye movements as well as with the spatial cueing task, neither of which was the case.

Fluid intelligence and vocabulary size did not contribute to explaining individual differences in language-mediated anticipatory eye movements. Thus, it was not the case that only the participants with the best linguistic knowledge or highest intelligence predicted upcoming referents. Of course, these variables may well play a role under more challenging conditions, such as processing of low frequency words or predictions based on complex inferences. Importantly, intelligence and vocabulary were not correlated with attentional orienting, ruling out that the correlation of language-mediated eye movements with attentional orienting was due to one of these factors.

We found no strong relationships between language-mediated anticipatory eye movements and the three language production measures from the category fluency task. Pickering and Garrod (in press) proposed two routes for predictive language processing, the simulation route (based on production experience) and the association route (based on comprehension experience). One possible reason for the absence of a correlation with the category fluency task is that, as discussed above, in the current study participants' predictions may have been based primarily on associations. This might have reduced the involvement of language production mechanisms (corresponding to the simulation route). Another possibility is that production measures are more sensitive when language production develops or declines. This could explain why previous correlations with production measures were found with infant (Mani & Huettig, 2012) or older adult populations (Federmeier, Kutas, & Schul, 2010; Federmeier, McLennan, De Ochoa, & Kutas, 2002). Finally, we cannot rule out that the critical components of production were not tapped by the category fluency task. Further research could investigate the relative involvement of the simulation and association routes across different tasks and materials.

To conclude, the present study provides correlational support for a relationship between participants' anticipatory attention in a verbal and a nonverbal task. In Underwood's (1975) terms, the hypothesis that anticipatory mechanisms are shared across different cognitive domains receives a go-ahead signal.

Acknowledgements

We thank Sabine Hunnius for kindly providing us with the action video materials. We are also grateful to Esther Kroese and Ceci Verbaarschot for collecting the Raven's advanced progressive matrices data; to Juul Coumans, Samantha van den Hoogen, Tim Korbmacher, and Ceci Verbaarschot for help with annotating the category fluency task recordings; and to Mira Cladder-Micus for extracting the Peabody vocabulary percentile scores.

References

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

Altmann, G. T. M., & Mirković, J. (2009). Incrementality and Prediction in Human Sentence Processing. *Cognitive Science, 33*(4), 583-609.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*, 390–412.

Bar, M. (2007). The proactive brain: Using analogies and associations to generate predictions. *Trends in Cognitive Sciences, 11(7),* 280-289.

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

Chang, F., Dell, G. S., & Bock, K. (2006). Becoming syntactic. *Psychological Review, 113(2),* 234-272.

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences, 36(3),* 181-204.

Clohessy, A. B., Posner, M. I., & Rothbart, M. K. (2001). Development of the functional visual field. *Acta Psychologica, 106,* 51-68.

Cohen, R. L. (1994). Some thoughts on individual differences and theory construction. *Intelligence, 18,* 3-13.

Coles, M. G. H. (1989). Modern mind-brain reading: Psychophysiology, physiology, and cognition. *Psychophysiology, 26,* 251-269.

DeGutis, J., Wilmer, J., Mercado, R. J., & Cohan, S. (2013). Using regression to measure holistic face processing reveals a strong link with face recognition ability. *Cognition, 126,* 87-100.

Delis, D. C., Kaplan, E., & Kramer, J. H. (2001). *Verbal fluency subtest of the Delis–Kaplan Executive Function System.* San Antonio, TX: The Psychological Corporation.

DeLong, K. A., Urbach, T. P., & Kutas, M. (2005). Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. *Nature Neuroscience, 8*(8), 1117-1121.

Dikker, S., & Pylkkänen, L. (in press). Predicting language: MEG evidence for lexical preactivation. *Brain & Language.*

Dunn, L.M. & Dunn, L.M. (1997). Peabody Picture Vocabulary Test - Third Edition.

Fan, J., McCandliss, B.D., Sommer, T., Raz, M. & Posner, M.I. (2002). Testing the efficiency and independence of attentional networks. *Journal of Cognitive Neuroscience, 14*, 340-347.

Federmeier, K. D. (2007). Thinking ahead: The role and roots of prediction in language comprehension. *Psychophysiology, 44*(4), 491-505.

Federmeier, K. D., & 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.

Federmeier, K. D., & Kutas, M. (2005). Aging in context: Age-related changes in context use during language comprehension. *Psychophysiology, 42,* 133-141.

Federmeier, K. D., Kutas, M., & Schul, R. (2010). Age-related and individual differences in the use of prediction during language comprehension. *Brain and Language, 115,* 149-161.

Federmeier, K. D., McLennan, D. B., De Ochoa, E., & Kutas, M. (2002). The impact of semantic memory organization and sentence context information on spoken language processing by younger and older adults: An ERP study. *Psychophysiology, 39,* 133-146.

Friston, K. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society London B: Biological Sciences*, 360(1456): 815-36.

Gagnepain, P., Henson, R. N., & Davis, M. H. (2012). Temporal predictive codes for spoken words in human auditory cortex. *Current Biology, 22(7),* 615-622.

Hommel, B., Pratt, J., Colzato, L., & Godijn, R. (2001). Symbolic control of visual attention. *Psychological Science, 12,* 360-365.

Huettig, F., Rommers, J., & Meyer, A. S. (2011). Using the visual world paradigm to study language processing: A review and critical evaluation. *Acta Psychologica, 137,* 151-171.

Hunnius, S. & Bekkering, H. (2010). The early development of object knowledge: A study on infants' visual anticipations during action observation. *Developmental Psychology, 46,* 446-454.

Kamide, Y. (2008). Anticipatory processes in sentence processing. *Language and Linguistics Compass, 2*, 647-670.

Klein, R. M. (1994). Perceptual-motor expectancies interact with covert visual orienting under conditions of endogeneous but not exogeneous control. *Canadian Journal of Experimental Psychology, 48(2),* 167-81.

Kosslyn, S.M., Cacioppo, J.T., Davidson, R.J., Hugdahl, K., Lovallo, W. R., Spiegel, D., Rose, R. (2002). Bridging psychology and biology - the analysis of individuals in groups. *American Psychologist, 57(5),* 341–51.

Lau, E., Holcomb, P. J., & Kuperberg, G. R. (in press). Dissociating N400 effects of prediction from association in single word contexts. *Journal of Cognitive Neuroscience*.

Luo, L., Luk, G., & Bialystok, E. (2010). Effect of language proficiency and executive control on verbal fluency performance in bilinguals. *Cognition, 114,* 29-41.

Mangun, G.R. (1995). Neural mechanisms of visual selective attention. *Psychophysiology, 32,* 4-18.

Mani, N., & Huettig, F. (2012). Prediction during language processing is a piece of cake - but only for skilled producers. *Journal of Experimental Psychology: Human Perception and Performance, 38(4),* 843-847.

Nobre, A. C., Rohenkohl, G., & Stokes, M. G. (2012). Nervous anticipation: Top-down biasing across space and time. In Michael I. Posner (Ed.), *Cognitive Neuroscience of Attention*. New York/London: The Guilford Press.

Otten, M., & Van Berkum, J. J. A. (2007). What makes a discourse constraining? Comparing the effects of discourse message and scenario fit on the discourse-dependent N400 effect. *Brain Research, 1153,* 166-177.

Otten, M., & Van Berkum, J. J. A. (2008). Discourse-based word anticipation during language processing: Prediction or priming? *Discourse Processes, 45,* 464-496.

Pickering, M. J., & Garrod, S. (2007). Do people use language production to make predictions during comprehension? *Trends in Cognitive Sciences, 11*(3), 105-110.

Pickering, M. J., & Garrod, S. (in press). An integrated theory of language production and comprehension. *Behavioral and Brain Sciences*.

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology,* 32(1), 3-25.

Posner, M. I., Nissen, M. J., & Ogden, W. C. (1978). Attended and unattended processing modes: The role of set for spatial location. In: H.L. Pick, & I.J. Saltzman (Eds.), *Modes of perceiving and processing information*. Hillsdale, N.J.: Lawrence Erlbaum Associates.

Raven, J., Raven, J. C., & Court, J. H. (1998). *Raven manual section 4: Advanced progressive matrices*. Oxford, England: Oxford Psychologists Press.

Rohrer, D., Wixted, J. T., Salmon, D. P., & Butters, N. (1995). Retrieval from semantic memory and its implications for Alzheimer's disease. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 1127–1139.

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.

Saslow, M. G. (1967). Latency for saccadic eye movement. *Journal of the Optical Society of America, 57,* 1030–1033.

Schlichting, L. (2005). *Peabody picture vocabulary test-III-NL.* Amsterdam, NL: Harcourt Assessments BV.

Shanks, D. R. (2007). Associationism and cognition: Human contingency learning at 25. *Quarterly Journal of Experimental Psychology, 60(3),* 291-309.

Stapel, J.C., Hunnius, S., van Elk, M. & Bekkering, H. (2010). Motor activation during observation of unusual vs. ordinary actions in infancy. *Social Neuroscience 5,* 451-460.

Underwood, B. J. (1975). Individual differences as a crucible in theory construction. *American Psychologist, 30,* 128-134.

van Berkum, J. J. A., Brown, C. M., Zwitserlood, P., Kooijman, V., & Hagoort, P. (2005). Anticipating upcoming words in discourse: Evidence from ERPs and reading times. *Journal of Experimental Psychology: Learning Memory and Cognition, 31*(3), 443-467.

van Petten, C., & Luka, B. J. (2011). Prediction during language comprehension: Benefits, costs, and ERP components. *International Journal of Psychophsiology, 83,* 176-190.

Vogel, E. K., & Awh, E. (2008). How to exploit diversity for scientific gain: Using individual differences to constrain cognitive theory. *Current Directions in Psychological Science, 17(2),* 171-176.

Wang, H., Fan, J., & Johnson, T. R. (2004). A symbolic model of human attentional networks. *Cognitive Systems Research, 5,* 119-134.

Wicha, N. Y. Y., Moreno, E. M., & Kutas, M. (2004). Anticipating words and their gender: An event-related brain potential study of semantic integration, gender expectancy, and gender agreement in Spanish sentence reading. *Journal of Cognitive Neuroscience, 16*(7), 1272-1288.

Appendix: Correlations

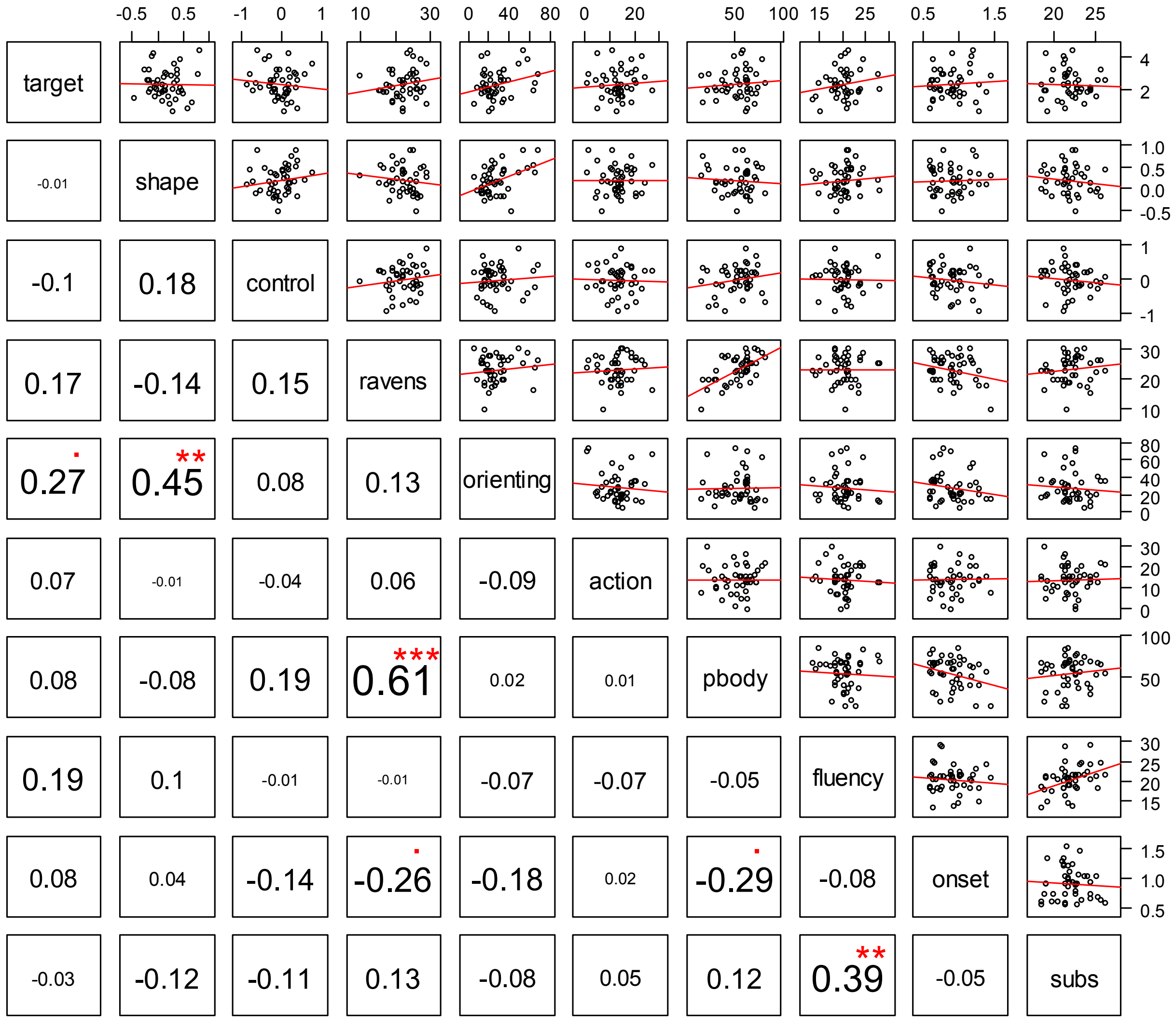


Figure 4. Correlations between inter-individual differences in language-mediated anticipatory eye movements and the background measures. Lower left panels: Pearson correlation coefficient for each pair of measures, with the font size scaled to the absolute r value. Upper right panels: Scatter plots of each measure plotted against each other measure along with a linear regression line. "target", Target bias (Target minus distractors); "shape", Shape competitor bias; "control", Control object bias; "ravens", Raven's advanced progressive matrices score; "orienting", facilitation from valid cues in the Posner cueing task; "action", action goal anticipation frequency; "pbody", Peabody picture vocabulary size; "fluency", category fluency number of correct responses; "onset", category fluency first-response latency (logtransformed ms); "subs", category fluency subsequent-response latency; *** $p < .001$, ** $p < .01$, * $p < .05$, · $p < .1$.

Figure 4 shows the correlations between all measures after averaging by participants. Consistent with the mixed-effects regression models, orienting correlated with anticipatory Shape competitor bias, *p* = .002, marginally with Target bias, *p* = .077, and not with Control object bias, *p* = .598. These correlations were not due to individual differences in the neutral cueing condition, as residual response times in the valid condition after controlling for the neutral condition (cf. DeGutis, Wilmer, Mercado, & Cohan, 2013) yielded essentially the same absolute correlations as in Figure 4 (Shape: *r* = .42; Target: *r* = .25; Control: *r* = .09). Other correlations between the eye-tracking data and the background measures did not reach significance, all *p* > .2.

Chapter 5

Constraining the involvement of language production in comprehension: A comparison of object naming and object viewing in sentence context

Rommers, J., Meyer, A. S., Piai, V., & Huettig, F. (under revision). Constraining the involvement of language production in comprehension: A comparison of object naming and object viewing in sentence context.

Abstract

It has been proposed that readers and listeners use their speech production system to predict upcoming words or concepts. We examined whether this production-based prediction includes anticipatory activation of motor plans. Participants named or passively viewed objects (e.g., a broom) after predictive or neutral lead-in sentences (e.g., “He swept the floor with a…” or “He saw the drawing of a…”). Their EEG was recorded. In the object naming task, a sustained positivity and alpha and beta suppression were observed prior to object onset. This suggests that participants assembled a motor plan in anticipation of the object. In contrast, in the object viewing task, these effects were not observed, though predictability did affect N300 and N400 amplitude after object onset. The results constrain the possible involvement of language production in comprehension to processes preceding motor preparation.

Constraining the involvement of language production in comprehension: A comparison of object naming and object viewing in sentence context

Reading and listening to sentences are fast and efficient processes. A possible explanation for the speed and efficiency of sentence comprehension is that readers and listeners continuously generate hypotheses regarding what might be mentioned next. Anticipating upcoming information plays an important role in current views on language comprehension as well as other cognitive functions (for reviews, see Altmann & Mirković, 2009; Bar, 2007; A. Clark, 2013; Federmeier, 2007; Kamide, 2008; Kutas, DeLong, & Smith, 2011; Pickering & Garrod, in press; van Berkum, 2010). However, the processing mechanisms underlying prediction are not yet clear.

Recently, it has been proposed that predictions during language comprehension are generated by the language production system. Van Berkum, Brown, Zwitserlood, Kooijman, and Hagoort (2005) suggested that readers and listeners recruit their language production system by a process not unlike asking themselves what they would say next if they were the speaker (see also Garrett, 1980, p. 213, for an early suggestion of top-down "analysis-by-synthesis"). In Chang, Dell, and Bock's (2006) model of learning of syntactic structure to support production, generating a word as a prediction of what will come next and generating a word for production are modeled as similar processes. Federmeier (2007) also related prediction to "covert generation processes" that may be integrated with language production. Pickering and Garrod (2007, in press) proposed a full-fledged forward modeling mechanism that simulates what the listener would say if they were the speaker. Finally, McCauley and Christiansen (2011) used the same linguistic units and distributional statistical information to model children's sentence comprehension and production.

The link between prediction and production is supported by studies of individual differences. The amplitude of event-related brain potential components associated with predictive language comprehension has been shown to correlate with performance in the category fluency task, in which participants produce as many members of a semantic category (e.g., animals) as they can in one minute (Federmeier, Kutas, & Schul, 2010; Federmeier, McLennan, De Ochoa, and Kutas, 2002; but see Federmeier & Kutas, 2005). Furthermore, infants with a large productive vocabulary use sentence context to launch anticipatory eye movements to referents before words unambiguously refer to them, while infants with a small productive vocabulary do not do so (Mani & Huettig, 2012).

The abovementioned studies are consistent with the involvement of the production system in predictive language comprehension, but do not allow for a specification of which production processes play a role in comprehension. Predictions can be made at different levels: there is evidence for the pre-activation of semantic/conceptual and visual information (e.g., Altmann & Kamide, 1999; Federmeier & Kutas, 1999; Rommers, Meyer, Praamstra, & Huettig, 2013; Szewczyk & Schriefers, 2013), the grammatical gender of words (van Berkum, Brown, Zwitserlood, Kooijman, and Hagoort, 2005; Wicha, Moreno, & Kutas, 2004), the phonological form of words (DeLong, Urbach, & Kutas, 2005; Gagnepain, Henson, & Davis, 2012), and the pronunciation of words (Brunnellière & Soto-Faraco, 2013). However, it is difficult to test whether the pre-activation of these representations is based on knowledge and processes that are intrinsic to comprehension, or on knowledge and processes also involved in production. This is because conceptual/semantic, syntactic, and phonological representations are necessary both during comprehension and during production (though they need not be the same in production and comprehension; E. Clark & Hecht, 1984; H. Clark & Malt, 1983).

This difficulty does not arise with respect to the final production stage, articulation, since language comprehension obviously does not require listeners and readers to pronounce the predicted words. In the present study, we focused on a closely related process that has a clear role only in language production: motor preparation for articulation. We examined whether motor preparation would also be involved in language comprehension. Given that motor preparation is by definition preparation for production, this would be a strong argument for the involvement of production in comprehension.

We asked whether reading predictable sentences is similar to preparing to complete sentences out loud. On each trial, we presented participants with an object after a lead-in sentence that was predictive of the upcoming object (Predictive condition) or not predictive of the upcoming object (Neutral condition). We recorded their electroencephalogram (EEG), as this is a technique through which both comprehension and production can be tracked. Two tasks were compared: an object naming task, which involves all stages of word production from conceptual preparation to articulation, and an object viewing task with the same predictable sentences.

Previous research on object viewing and object naming in sentence context

Electrophysiological studies of language processing have most commonly averaged the EEG signal across multiple trials in the time domain, timelocked to stimulus onset, to form event-related potentials (ERPs). An ERP component relevant to language comprehension is the N400, a negative deflection that peaks at around 400 ms after onset of a meaningful stimulus, such as a word (Kutas & Hillyard, 1980). Many studies attest to the sensitivity of this component as an index of semantic processing (for a recent review, see Kutas & Federmeier, 2011). For instance, its amplitude in response to a word is inversely related to how likely that word was to occur given the context (Kutas & Hillyard, 1984). So the N400 amplitude in response to the word "broom" in predictive sentences such as "He swept the floor with a broom" is typically attennuated compared to the same word in more neutral sentences such as "He saw the drawing of a broom".

The N400 has also been observed in tasks where instead of words, objects were presented. Manipulations of semantic constraint yielded N400 effects in response to objects similar to those observed in studies that used words, namely smaller N400 amplitude for predictable objects than for objects in a neutral sentence (Federmeier, & Kutas, 2001). N400 effects elicited by anomalous pictures have been found to be similar to those elicited by anomalous words (Nigam, Hoffman, & Simons, 1992) though they tended to have a more frontal scalp distribution (Ganis, Kutas, & Sereno, 1996; Willems, Özyürek, & Hagoort, 2008), indicating partially nonoverlapping neuronal systems for word and object processing. In addition to the N400 effect, Federmeier and Kutas (2001) observed earlier effects of semantic constraint including an N300 effect in the same direction as the N400, which might be specific to object processing (see also McPherson & Holcomb, 1999). Taken together, these findings suggest that the comprehension of pictures in sentences is similar, though not identical, to the comprehension of words in sentences.

The second task used in our study is object naming after reading lead-in sentences. Griffin and Bock (1998) used this task with lead-in sentences in a reaction time study and manipulated both the predictability of the objects based on the sentence context (low, medium, and high) and the lexical frequency of the object names (low, high). The word frequency effect (shorter naming latencies for high frequency object names than for low frequency object names; e.g., Jescheniak & Levelt, 1994) diminished with increasing predictability and disappeared in high-constraint sentences. This shows that sentence contexts can influence the lexical selection of object names.

Object naming studies increasingly use ERPs (for a review, see Ganushchak, Christoffels, & Schiller, 2011). Some studies also employed time-frequency analyses to examine the spectral content of the signal, which provides information complementary to the ERPs. This is because during ERP averaging, oscillations that are time-locked but not phase-locked to the experimental event of interest (induced activity) largely cancel each other out, whereas time-frequency analyses preserve induced activity. The oscillatory activity revealed by such analyses is thought to reflect the dynamic coupling and uncoupling of networks in the brain (for a recent review, see Bastiaansen, Mazaheri, & Jensen, 2012). Different frequency bands have been related to different cognitive functions. Most relevant to the present study, the electrophysiological signal preceding an object naming response is known to show a suppression of alpha band (8-12 Hz) power over motor cortex (Laaksonen, Kujala, Hultén, Liljeström, & Salmelin, 2012). Such alpha (and beta, 13-30 Hz) suppression is also associated with motor preparation in non-verbal tasks (e.g., Chatrian, Peterson, & Lazarte, 1959; Pfurtscheller, Brunner, Schlögl, & Lopes da Silva, 2006). Interestingly, in movement tasks alpha and beta suppression have been found both during preparation to move and during motor imagery (McFarland, Miner, Vaughan, & Wolpaw, 2000; Neuper & Pfurtscheller, 1999; Neuper, Wörtz, & Pfurtscheller, 2006). If readers understanding predictable sentences imagine what they would say next, production-like responses might arise in the alpha and beta bands.

The present study

We examined which aspects of the electrophysiological signal were affected by predictability in the object naming and object viewing tasks, to investigate whether reading predictable sentences is similar to preparing to complete sentences aloud. Our hypotheses were as follows.

In the object naming task we should see shorter naming latencies for objects that were preceded by a Predictive sentence context compared to objects that were preceded by a Neutral sentence context. This would indicate that the predictive sentence context influenced participants' speech production processes, as shown by Griffin and Bock's (1998) study. These differences might be accompanied by differences in the EEG, namely alpha and beta power decreases before speech onset, in line with the results reported by Laaksonen et al. (2012). Moreover, if response preparation during the presentation of Predictive sentences occurs in an anticipatory fashion, these alpha and beta power decreases might be seen before object onset. Because the semantically uninformative Neutral sentences did not enable response preparation, this would result in a difference between the Predictive and Neutral condition, that is, alpha and beta power decreases for the Predictive sentences relative to the Neutral sentences.

In the object viewing task, the participants made no overt responses. However, if they used the sentence context to predict the objects up to the level of motor preparation as they do when they prepare overt responses, we should see the same effects of predictability in the EEG as in the naming task, namely decreases in alpha and beta power. In addition, based on Federmeier and Kutas (2001), we expected to see effects of predictability on the N300 and N400 after object onset. In the production task, these components are difficult to analyze because the EEG signal after object onset is contaminated by motor artifacts caused by speech production (Goncharova, McFarland, Vaughan, & Wolpaw, 2003).

A possible concern about comparisons between language production and comprehension tasks is related to potential differences in attentional demands. Therefore, while acknowledging that one of the main advantages of electrophysiological methods is that participants can read or listen to sentences without any additional task (e.g., van Berkum, 2004), we added occasional comprehension questions to ensure that participants paid attention to the sentences in both tasks.

Methods

Participants

Twenty-two students (nine men) with an average age of 23 years (range 18-33 years) gave informed consent and were paid to participate in the experiment. All were right-handed college-educated native speakers of Dutch. They had normal or corrected-to-normal vision and no history of neurological or speech disorders. Two additional participants were excluded because too many trials were contaminated by EEG artifacts (fewer than 20 trials remained in a condition). The experiment was approved by the local ethics committee.

Materials and design

The materials for the object naming and object viewing tasks were based on those used by Piai, Roelofs, and Maris (submitted), who used them in the object naming task only. They consisted of 126 pictures of objects from the picture database of the International Picture-Naming Project (Szekely et al., 2004), each paired with two sentences. One of the two sentences was predictive of the object's name (Predictive condition) and the other was not predictive of the object's name (Neutral condition). An example is displayed in Figure 1. Twenty-four items were associated with a yes/no comprehension question. The predictability of the object names based on the sentences had been established by Piai et al. based on a separate group of participants who provided the first three words that came to mind as completions for each sentence. The results are repeated here for completeness. Two types of cloze probability were computed, the first-response probability and the overall probability. The first-response probability refers to the proportion of participants who completed the sentence with the object name as their first completion. The overall probability refers to the proportion of participants who completed the sentence with the object name as one of their three completions. Object names had a higher cloze probability in the Predictive sentence context than in the Neutral sentence context, both according to first-response probability (.82 vs. .02, *SD* = .26 and .13, respectively), $t\ (117) = 29.151$, $p < .001$, and overall response probability (.83 vs .02, *SD* = .19 and .06), $t\ (117) = 44.050$, $p < .001$.

One half of the items was presented in the object viewing task and the other half in the object naming task, counterbalanced between participants. Within each task, participants saw each item in both the Predictive and the Neutral condition. The halves were matched on cloze probability, visual complexity of the pictures (based on norms from Szekely et al., 2004), and average naming latency in Piai et al.'s experiment, all *p*s > .690. The order of the items was pseudorandomized for each participant individually. No more than three consecutive trials were

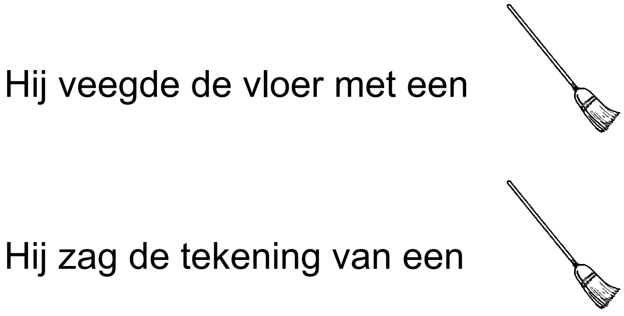


Figure 1. Example of the materials in the object naming and object viewing task. The top sentence (English: "He swept the floor with a") is from the Predictive condition and the bottom sentence ("He saw the drawing of a") is from the Neutral condition.

of the same condition, the comprehension questions were spaced minimally three trials apart, and minimally 21 trials intervened between the two presentations of the same picture (i.e., these were spaced at least one block apart).

Procedure

Participants were tested individually in a dimly illuminated room. They were seated in front of a Sennheiser microphone at a distance of around 80 cm from a screen and asked to relax, and to move and blink as little as possible while relevant stimuli were displayed on the screen. Each object subtended a maximum of 5˚ of visual angle. First, participants were familiarized with the names of the objects. Each object was presented on the screen for 2100 ms. After the object had been on the screen for 600 ms, its name was presented as a printed word superimposed onto the object, after which participants produced the name. Afterwards participants were reminded of any incorrect responses and informed about the next task.

Then participants performed the object viewing task. They were instructed to read the sentences carefully for comprehension. Each trial started with a fixation cross presented in the center of the screen for 500 ms to orient participants towards the stimuli. Then a sentence was presented visually word by word in the center of the screen. Each word was presented for 300 ms, followed by a blank screen for 300 ms. The lead-in sentence was followed by a picture that was presented for 2000 ms. Then three asterisks (* * *) were presented for 2000 ms, indicating that participants were free to blink and carefully move. If the trial was associated with a comprehension question, a question appeared on the screen which participants answered by saying "yes" or "no". Their answers were recorded through the microphone. The sentences were presented in six blocks of 22 trials. After every block participants could take a break for as long as needed before asking the experimenter to resume the experiment.

Subsequently participants performed the object naming task. The block and trial structure were identical to the object viewing task, but now participants were instructed to name the objects. They were instructed to keep their lips apart when not speaking to avoid clicking noises and artifacts. Their speech was recorded through the microphone.

EEG recording

EEG was recorded from 128 active Ag/AgCl electrodes mounted in a cap according to the 10-5 system (Oostenveld & Praamstra, 2001). Recordings were performed relative to common mode sense (CMS) and driven right leg (DRL) electrodes placed just anterior to the Fz electrode. Two additional electrodes were positioned laterally to the left and right eye, two at the mastoids, and one above and one below the left corner of the mouth. The signals were amplified by Biosemi ActiveTwo amplifiers with a lowpass filter at 128 Hz and sampled with a frequency of 512 Hz.

Analysis

Naming responses. For the object naming task, trials were scored as incorrect when an incorrect object name was uttered, when there was a repair or filled pause (e.g., "err"), when a response started before picture onset, or when no response was given. Eight items that had low naming accuracy were excluded. Naming latencies were determined manually using the speech waveform editor Praat (Boersma & Weenink, 2009). Participants' response latencies on correct trials were analyzed using linear mixed-effects regression models, which allow for the simultaneous inclusion of items and participants as random variables (e.g., Baayen, Davidson, & Bates, 2008). The model included a fixed effect of Predictability (Predictive, Neutral) and the maximal possible random effects structure consisting of random intercepts and slopes for Predictability by participant and item (see Barr, Levy, Scheepers, & Tily, 2013). To evaluate the effect of Predictability, this model was compared to a model without the effect of Predictability but with the same random effects structure, using a likelihood ratio test. Naming accuracy was analyzed in the same way, but using logistic regression.

Event-related potentials. The EEG analyses were performed using the open source Matlab toolbox Fieldtrip (Oostenveld, Fries, Maris, & Schoffelen, 2011). The data were referenced to the average of the left and right mastoid channels. The bipolar horizontal electrooculogram was computed as the difference between the channels from the electrodes placed laterally to the eyes. The continuous EEG was segmented into epochs from 800 ms before until 800 ms after picture onset, i.e. including the last word of the sentence which had been matched on length and frequency of occurrence across conditions, and focusing on the time where response preparation and execution would take place. Trials that had been removed from the behavioral analysis were also removed from the ERP analysis. In addition, trials with naming latencies shorter than 300 ms were removed to minimize contamination of the signal by muscle artifacts. Given that in high-density electrode setups the probability increases that some channels are contaminated by artifacts, which can easily lead to an unacceptably low number of trials when removing entire trials (cf. Junghöfer, Elbert, Tucker, & Rockstroh, 2000), channels which consistently exhibited artifacts were interpolated. Trials containing blinks, eye movements, and muscle artifacts that were not part of a task-related response were removed after visual inspection. A Butterworth zero phase-shift lowpass filter at 30 Hz was applied. Exploratory analyses showed that a baseline correction was necessary due to a dominance of low frequency activity in the signal. As the sentences in the Predictive condition varied in the point at which the upcoming object became predictable, there was no clear predictive cue to timelock the signal to. The last words in the sentences had been matched on length and frequency of occurrence across conditions, ensuring that any effects prior to the presentation of the picture could not be attributed to differences between the words in the sentences. Therefore, the data were baseline-corrected through subtraction of the mean signal in the 150 ms before the onset of the last word (i.e., 750-600 ms before picture onset). ERPs were then computed by averaging across trials for each participant and condition separately, from the beginning of the epoch (800 ms before picture onset) up to the point of the earliest remaining overt responses in the Predictive condition of the object naming task (300 ms after picture onset; recall that responses before this point had been removed). A conventional post-object-onset analysis was performed on the object viewing EEG data after picture onset. Here we specifically tested for the presence of N300 and N400 effects found in other studies using the same task (e.g., Federmeier & Kutas, 2001).

For the statistical analysis, the ERPs were submitted to nonparametric cluster-based permutation tests (Maris & Oostenveld, 2007). The permutation test determined which time points and channels showed a significant effect by means of a clustering algorithm based on the physiologically plausible assumption that ERP effects are clustered together in space and time over neighboring electrodes and neighboring samples. This test compares two conditions at a time and proceeded as follows with our settings (for details see Maris & Oostenveld, 2007). First, a dependent-samples *t*-test quantified the difference between the conditions at every data point, and temporally or spatially adjacent data points with a pre-determined significance level of $p < .05$ (two-tailed) were combined into clusters, for each of which the cluster-level t-value was the sum of all t-values within the cluster. Then a null-distribution was created by randomly assigning participant averages to one of the two conditions 1,000 times and computing the cluster-level statistics for each randomization. Finally, the observed cluster-level test statistic was compared against the null-distribution. If the observed statistic fell in one of the 2.5th percentiles of the null-distribution (i.e., $p < .05$), the effect was considered significant. In our analysis the Predictable and Neutral conditions were compared within the object naming and object viewing tasks. For the analysis from the last word until object-naming response, the data from all time points and electrodes over which ERPs had been computed were submitted to the analysis. For the post-object-onset analysis, ERPs were averaged (across time samples, but not across electrodes) within the two time windows used by Federmeier and Kutas (2001) to quantify the N300 (250-350 ms) and the N400 (350-500 ms). The latter tests were one-tailed because we expected larger (more negative) amplitudes in the Neutral compared to the Predictive condition.

Time-frequency representations of power. For the TFR analysis, epochs from 2 seconds before until 2 seconds after object onset were extracted from the data (to allow for a reliable estimate of power in the lower frequencies). The same trials were removed as for the ERP analysis. TFRs of power were computed at frequencies ranging from 4 to 30 Hz, using a sliding time window of three cycles' length advanced in steps of 10 ms and of 1 Hz. Because of high-frequency motor noise preceding naming responses, higher frequencies were not analyzed. The data in each time window were multiplied with a Hanning taper, followed by the fast Fourier transform. Subsequently the TFRs were averaged across trials for each participant and condition in the same time window as the ERPs (i.e., 800 ms pre-object onset to 300 ms post-object onset) In contrast to the ERP analysis, no baseline correction was necessary; instead, power was

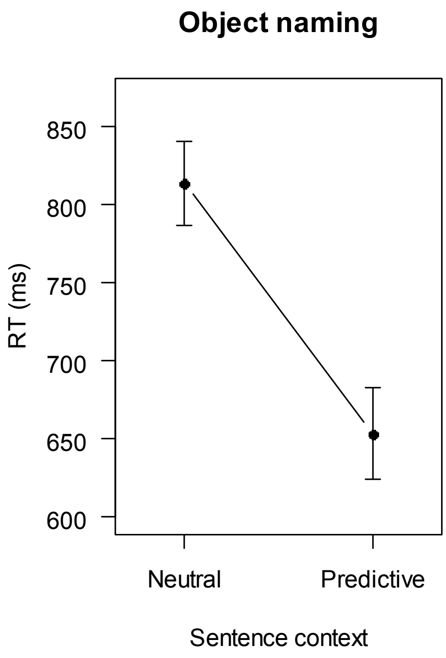


Figure 2. Average naming latencies in the picture naming task. Error bars indicate standard error of participant means.

expressed as the difference between the Predictive and Neutral conditions, normalized (divided) by power in the Neutral condition. The statistical method was the same as for the ERP analysis, except that the test now also clustered in the frequency domain, in addition to the space and time domains.

Results

Behavioral data

Participants answered most comprehension questions correctly, both during the object naming task (mean accuracy = 97.5%, *SD* = 1.2%) and during the object viewing task (mean = 97.5%, *SD* = 1.1%). This demonstrates that they attended to the sentences in both tasks. In the object naming task, naming accuracy was high, both for objects presented in Predictive sentences (mean = 96.5%, *SD* = 2.1%) and for objects presented in Neutral sentences (mean = 97.1%, *SD* =

2.1%). There was no evidence for a difference in naming accuracy between the conditions, as including the fixed effect of Predictability did not improve model fit, β = -0.033, *SE* = 0.258, *z* = 0.126, χ^2 (1) = 0.0052, *p* = .943. Average naming latencies for correct trials (see Figure 2) showed that objects were named faster by on average 160 ms after Predictive than after Neutral lead-in sentences, β = -160, *SE* = 14, χ^2 (1) = 47.513, *p* < .001.

EEG data

On average 2.6 (range 0-5) channels were interpolated per participant. They usually concerned lateral sites (e.g., T7, T8), which are not crucial to the effects of interest. The total percentage of trials removed due to incorrect responses, the 300 ms cutoff, and artifacts was 15.3%, with on average 50 trials remaining per participant per condition. ERPs from both tasks are shown in Figure 3. Early components in all conditions showed the visual N1-P2 complex in response to the final word before picture onset, followed by a relatively slow negative wave. In the object naming task, there was an effect of predictability in the form of a widespread sustained positive deflection for the Predictive relative to the Neutral sentence context, maximal at central electrodes between 300 and 200 ms before picture onset, which was detected as a cluster from -334 ms until the picture-naming response, *p* = .002.

TFRs from both tasks are shown in Figure 4. In the object naming task, a centrally distributed power decrease in the alpha frequency band (8-12 Hz), also spreading into the beta frequency band (12-30 Hz), occurred from around 400 ms before picture onset until 100 ms after picture onset, *p* = .004. This oscillatory suppression for the Predictive relative to the Neutral condition had a widespread distribution, maximal at centroparietal channels. In addition, power increases in the theta (4-7 Hz) and upper-beta band (20-30 Hz) occurred after picture onset, which was reflected in a cluster from around 0 ms (i.e., from picture onset) until the picture-naming response, *p* = .034. Given their timing close to the onset of some of the first responses

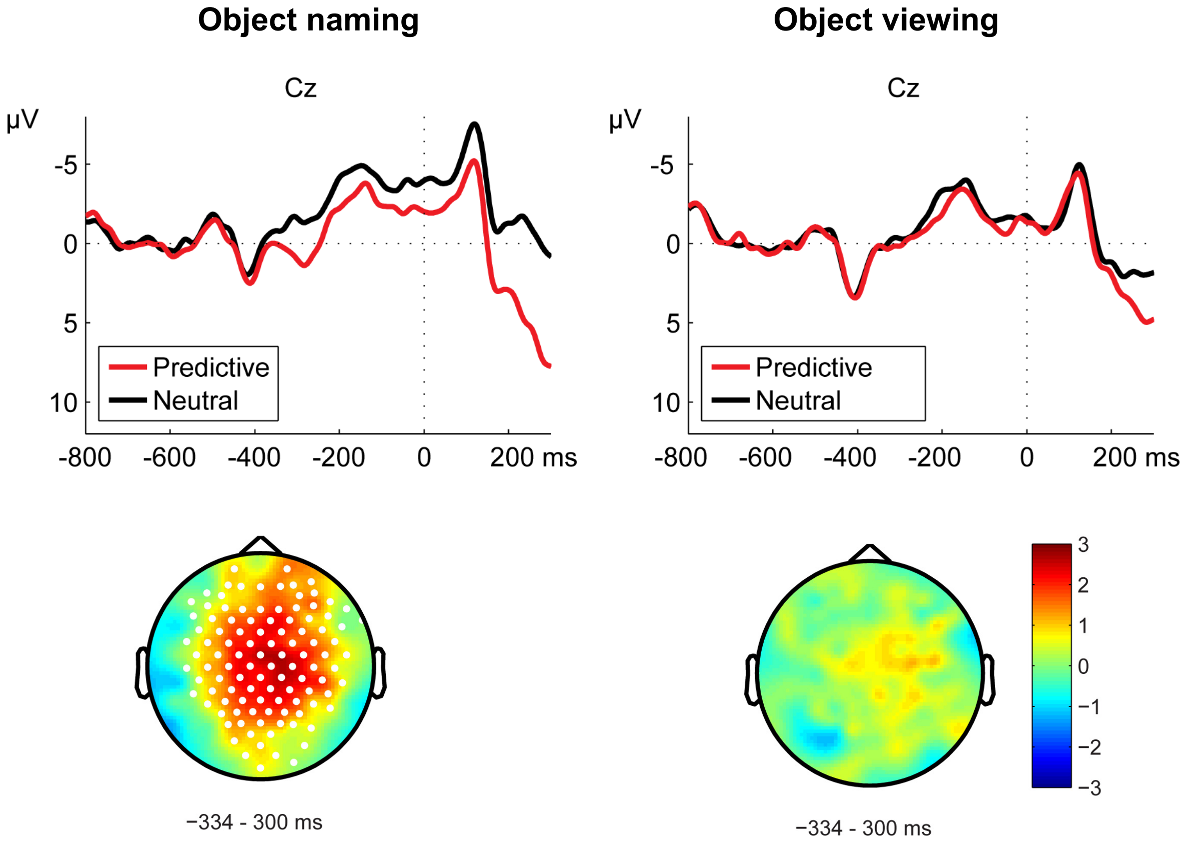


Figure 3. Event-related potentials prior to object onset. The vertical dashed line at time zero indicates object onset. The onset of the last word occurred at -600 ms. Waveforms are shown until the onset of the fastest responses (300 ms). The head maps show the difference between conditions (Predictive - Neutral). Electrodes that were part of a cluster are marked as white dots.

(note that two of the participants had *average* naming latencies shorter than 400 ms in the Predictive condition), these power increases likely reflect artifactual motor activity from mouth movements involved in articulation and are therefore not discussed further.

In contrast, in the object viewing task, the sustained positivity was not present in the ERPs (Figure 3), and no reliable clusters were detected in the comparison between the Predictive and Neutral sentence context (all *ps* > .159). Similarly, the time-frequency analysis (Figure 4) showed that the power decreases and increases observed in the object naming task were absent in the object viewing task, where no clusters were detected, all *ps* > .793.

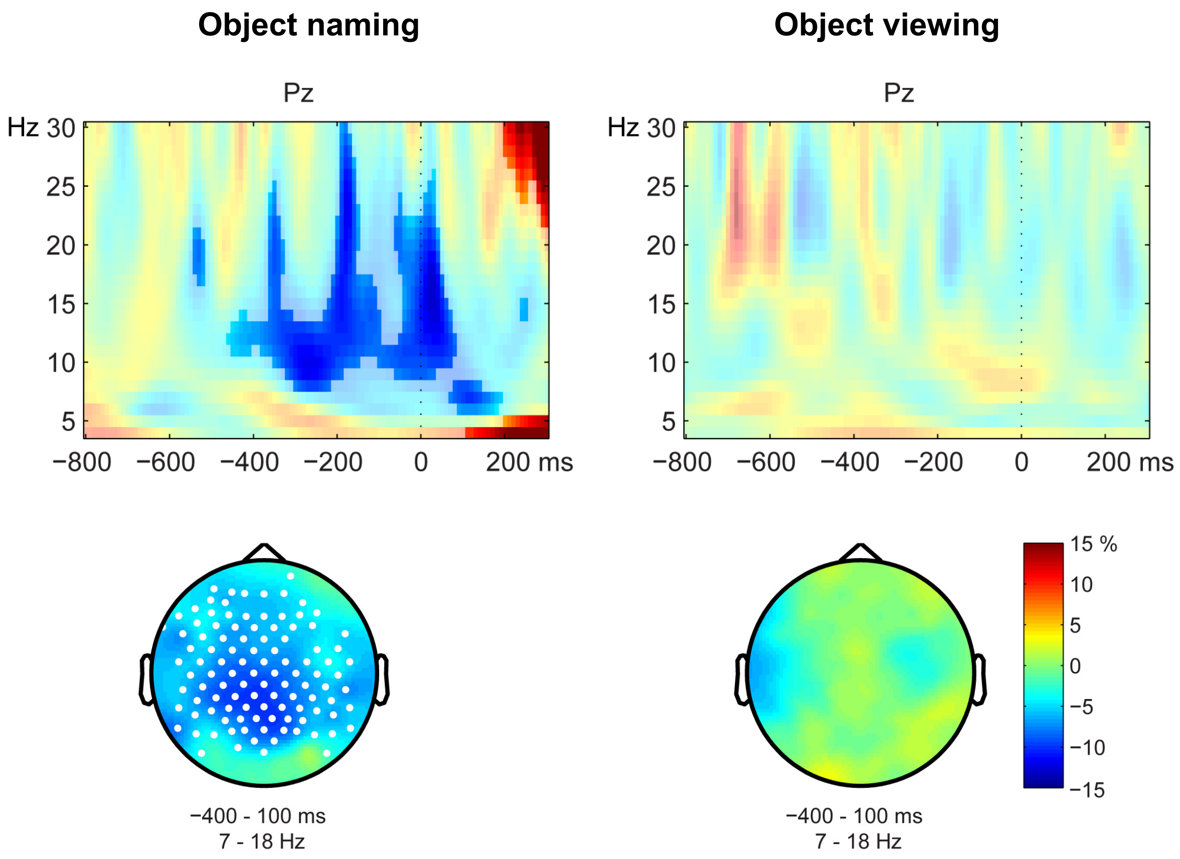


Figure 4. Time-frequency representations of power. The vertical dashed line at time zero indicates picture onset. The onset of the last word of the sentence occurred at - 600 ms. The top panels show the average difference as percentage increase/decrease in the Predictive relative to the Neutral condition with transparent colors, with the statistically significant cluster overlaid in opaque colors. The head maps show the distribution of the difference between the conditions in the time- and frequency-window indicated below the head maps, with channels that formed part of the significant cluster highlighted in white.

Given the clear effects of predictability in object naming and their disappearance in object viewing, one may wonder whether interactions between Naming and Predictability back up the pattern of results. We could not compute the interaction using the cluster-based permutation tests, as the randomization procedure is designed to compare only two conditions at a time. For completeness we report Analyses of Variance computed on the data averaged across all pre-object data points corresponding to each cluster detected in the object naming task. For the TFRs, we also averaged over the frequencies participating in the cluster. These yielded the expected

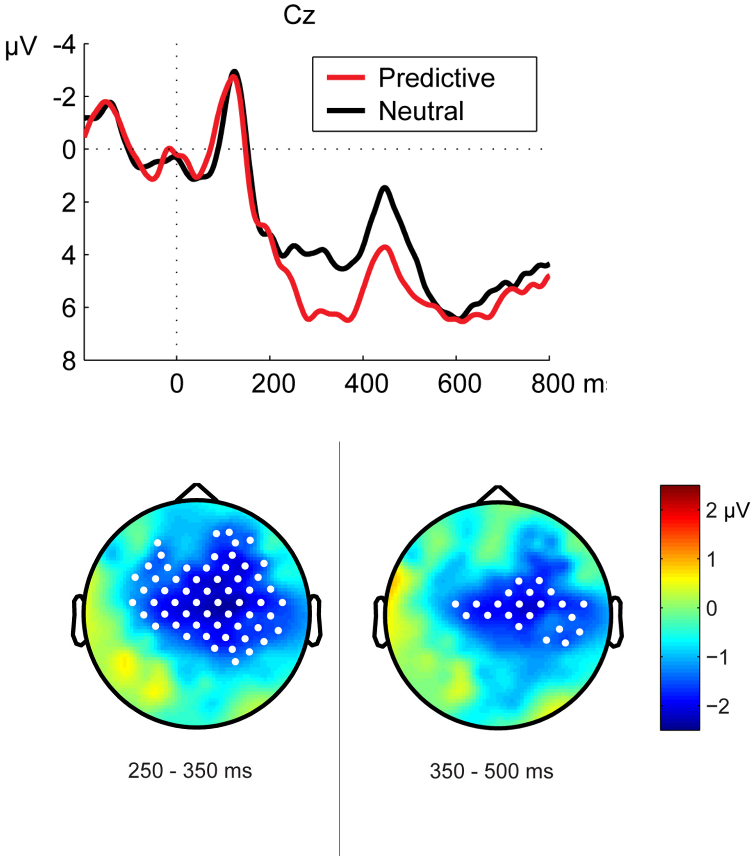


Figure 5. Post-object event-related potentials from the object viewing task. Time zero indicates object onset. The head maps indicate the distribution of the difference between conditions. Note that for the calculation of the difference wave, the waveforms of the Predictive condition were subtracted from those of the Neutral condition, as the N300 and N400 are usually plotted as negativities.

interaction, both on the sustained positivity in the ERPs, $F(1,21) = 5.251$, $p = .032$, and on the alpha/beta power decreases, $F(1,21) = 11.146$, $p = .003$.

The post-object-onset results from the object viewing task are shown in Figure 5. As expected, we observed an N300 that was larger in the Neutral than the Predictive condition, $p = .005$, and an N400 in the same direction, $p = .029$. The N300 effect had a frontocentral distribution, while the N400 effect had a central distribution and was relatively strongly right-lateralized.

Discussion

We presented objects in a naming and a viewing task and analyzed pre-stimulus electrophysiological effects induced by predictive, compared to neutral, lead-in sentences. In the object naming task, ERPs revealed a sustained positivity which reached statistical significance more than 300 ms before object onset. Time-frequency analyses revealed alpha and beta suppression in predictive relative to the neutral sentences, which began well before object onset (at around 400 ms). Alpha and beta suppression have previously been observed to precede naming responses (Laaksonen et al., 2012) as well as motor responses in nonverbal tasks (e.g., Pfurtscheller, Brunner, Schlögl, & Lopes da Silva, 2006). The fact that the effects were observed before object onset suggests that participants assembled a motor plan in an anticipatory fashion, based on the preceding lead-in sentences.

Importantly, the sustained positivity and alpha and beta suppression prior to object onset seen in the object naming task were not observed in the object viewing task. The results thus highlight the electrophysiological signatures of language production processes that are not shared with comprehension. Since articulation itself is by definition not shared between object naming and silent reading of sentences, it is perhaps not too surprising that motor *preparation* is not shared between the tasks either. However, from a language comprehension point of view it is important to note that reading predictable sentences is clearly different from preparing to complete sentences out loud. As discussed in the Introduction, alpha and beta suppression in nonverbal motor preparation have been shown to extend to motor imagery (McFarland, Miner, Vaughan, & Wolpaw, 2000; Neuper & Pfurtscheller, 1999; Neuper, Wörtz, & Pfurtscheller, 2006). Moreover, imagery has been implicated in prediction as a general mechanism (Moulton & Kosslyn, 2009) and might also play a role in language comprehension: Perhaps listeners imagine what they would say next if they were the speaker (e.g., Pickering & Garrod, in press). In that sense, alpha and beta suppression were important candidates for reflecting production processes involved in comprehension. Contrary to this hypothesis, alpha and beta suppression were found only in the production task, not in the comprehension task.

As discussed in the Introduction, there is evidence that listeners and readers anticipate specific upcoming words, and several authors have suggested that the production system could generate such predictions. The differences between the results from both tasks show that reading predictable sentences is clearly different from preparing to complete sentences out loud. This constrains the production processes that may be involved in language comprehension to stages before motor preparation. Before discussing further implications of this result, we briefly address three potential concerns.

First, one might argue that the results differed in the two tasks because participants paid less attention to the sentences during object viewing than during object naming. At least two aspects of our data argue against this possibility. First, participants' accuracy in answering the comprehension questions was equally high in both tasks, indicating that they did pay attention in both tasks. Second, the analysis of the ERPs after object onset yielded clear N300 and N400 effects in the comprehension task. The presence of these effects is consistent with previous studies (e.g., Federmeier & Kutas, 2001) and inconsistent with the hypothesis of lack of attention during the object viewing task.

A second potential concern is that, as no pre-stimulus effects were observed in the object viewing task, participants were predicting during the object naming task, but they were not doing so in the object viewing task. However, this seems unlikely for several reasons. First, many previous studies have observed prediction in non-naming situations at similar degrees of predictability. For instance, the average cloze probability of our sentences (83%) is similar to that in the studies by van Berkum et al. (2005; 86%) and Wicha, Moreno, and Kutas (2003; 80%).

Both of these studies observed pre-stimulus effects of prediction in comprehension through a gender match/mismatch manipulation. Furthermore, our materials were counterbalanced across tasks, and the prediction-induced motor preparation in the object naming task implies that our sentences were successful in eliciting predictive behavior from the participants. Finally, the post-object analysis elicited an N400 effect, and at least part of the amplitude of the N400 component is explained by prediction (Lau, Holcomb, & Kuperberg, in press).

A third potential concern is that anticipatory alpha suppression has not only been observed during motor preparation but also for predictable (relative to unpredictable) sensory stimuli (e.g., Rohenkohl & Nobre, 2011; Thut, Nietzel, Brandt, & Pascual-Leone, 2006). Therefore, the alpha suppression observed in the present study may reflect the participants' sensory predictions of the upcoming object rather than motor preparation resulting from such predictions. However, such sensory predictions typically elicit alpha suppression with a focal occipital distribution (see Rohenkohl & Nobre, 2011; Thut et al., 2006). As shown in Figure 4, the alpha suppression in the Predictive relative to the Neutral condition in the object naming task had a rather widespread distribution and was strikingly absent on the channels over occipital cortex. This leads us to favor an interpretation of the observed alpha suppression in terms of motor preparation.

Although it is likely that there is some involvement of production processes in prediction given the observed correlations between production and prediction skills (e.g., Federmeier, McLennan, De Ochoa, and Kutas, 2002; Mani & Huettig, 2012), there are limits to this involvement. For instance, the results of Brunnellière and Soto-Faraco (2013), who observed regional accent-specific phonetic anticipation, could be explained by assuming that production is involved in prediction including the level of, in Levelt, Roelofs, and Meyer’s (1999) terms, phonetic encoding. (Note that Brunnellière and Soto-Faraco themselves do not take a stance on this issue) Our study suggests that the effects reported by Brunnellière and Soto-Faraco (2013) are unlikely to stem from the production system.

Current theories can accommodate a lack of involvement of production processes in comprehension by posing multiple mechanisms supporting prediction. For instance, Pickering and Garrod (in press) proposed that in in addition to a prediction route that uses the production system, there is also a “prediction-by-association” route that builds on comprehension knowledge. It has yet to be shown which types of situations emphasize which types of predictive mechanisms. One might, for instance, expect that articulatory involvement in language comprehension is more likely when the stimuli are presented as speech compared to written presentation. Our study employed visual presentation as a strict test of whether articulatory involvement occurs by default.

Further research into the involvement of production processes in language comprehension should specify the aspects of production that are involved in comprehension and should explain why motor preparation apparently is not one of them. In addition, a further functional specification of alpha and beta desynchronization during language production is desirable. At present, it is unclear whether these effects represent only a final pre-articulatory stage, as a motor program is prepared for articulation, or also represent earlier processes such as phonological encoding. The latter is conceivable given how far ahead of the participants' responses the desynchronization began. If more stages turn out to be represented by alpha and beta desynchronization, the present results would impose further constraints on theories of which production processes are involved in comprehension.

In summary, the present study showed preparatory motor effects in object naming before the objects were presented. These effects were not observed in object viewing. The results constrain the stages of language production that might be involved in comprehension to stages before motor preparation. The question which stages are shared between production and comprehension provides an exciting agenda for future research.

Acknowledgments

The authors thank Mira Cladder-Micus and Esther Kroese for assistance with data collection.

References

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

Altmann, G. T. M., & Mirković, J. (2009). Incrementality and Prediction in Human Sentence Processing. *Cognitive Science, 33*(4), 583-609.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*, 390–412.

Bar, M. (2007). The proactive brain: Using analogies and associations to generate predictions. *Trends in Cognitive Sciences, 11(7),* 280-289.

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

Bastiaansen, M. C. M., Mazaheri, A., & Jensen, O. (2012). Beyond ERPs: Oscillatory neuronal dynamics. In S. J. Luck & E. S. Kappenman (Eds.), *The Oxford Handbook of Event-related Potential Components*. New York: Oxford University Press.

Boersma, P., & Weenink, D. (2009). Praat: Doing phonetics by computer (Version 5.1).

Brunnellière, A., & Soto-Faraco, S. (2013). The speakers' accent shapes the listeners' phonological predictions during speech perception. *Brain & Language, 125,* 82-93.

Chang, F., Dell, G. S., & Bock, K. (2006). Becoming syntactic. *Psychological Review, 113(2),* 234-272.

Chatrian, G.E., Peterson, M.N.C., Lazarte, J.A. (1959). The blocking of the rolandic wicket rhythm and some central changes related to movement. *Electroencephalography and Clinical Neurophysiology, 11,* 497– 510.

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences, 36(3),* 181-204.

Clark, E. V., & Hecht, B. F. (1983) Comprehension, production, and language acquisition. *Annual Review of Psychology, 34,* 325-349.

Clark, H. H., & Malt, B. C. (1984). Psychological constraints on language: A commentary on Bresnan and Kaplan and on Givon. In W. Kintsch, J. R. Miller, and P. Polson (Eds.), *Method and tactics in cognitive science* (pp. 191-214). Hillsdale, NJ: Erlbaum.

DeLong, K. A., Urbach, T. P., & Kutas, M. (2005). Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. *Nature Neuroscience, 8*(8), 1117-1121.

Federmeier, K. D. (2007). Thinking ahead: The role and roots of prediction in language comprehension. *Psychophysiology, 44*(4), 491-505.

Federmeier, K. D., & 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.

Federmeier, K. D., & Kutas, M. (2001). Meaning and modality: influences of context, semantic memory organization, and perceptual predictability on picture processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27(1),* 202-224.

Federmeier, K. D., & Kutas, M. (2005). Aging in context: Age-related changes in context use during language comprehension. *Psychophysiology, 42,* 133-141.

Federmeier, K. D., Kutas, M., & Schul, R. (2010). Age-related and individual differences in the use of prediction during language comprehension. *Brain and Language, 115,* 149-161.

Federmeier, K. D., McLennan, D. B., De Ochoa, E., & Kutas, M. (2002). The impact of semantic memory organization and sentence context information on spoken language processing by younger and older adults: An ERP study. *Psychophysiology, 39,* 133-146.

Gagnepain, P, Henson, R. N., & Davis, M. H. (2012). Temporal predictive codes for spoken words in human auditory cortex. *Current Biology, 22(7),* 615-622.

Ganis, G., Kutas, M., & Sereno, M. I. (1996). The search for ‘common sense’: an electrophysiological study of the comprehension of words and pictures in reading. *Journal of Cognitive Neuroscience, 8*, 89–106.

Ganushchak, L. Y., Christoffels, I., & Schiller, N. (2011). The use of electroencephalography (EEG) in language production research: A review. *Frontiers in Psychology, 2,* 208.

Garrett, M. F. (1980). Levels of processing in sentence production. In B. Butterworth (Ed.), *Language production, Vol. 1: Speech and talk* (pp. 177-220). London: Academic Press.

Goncharova, I. I., McFarland, D. J., Vaughan, T. M., Wolpaw, J. R. (2003). EMG contamination of EEG: Spectral and topographical characteristics. *Clinical Neurophysiology, 114 (9),* 1580 – 1593.

Griffin, Z. M., & Bock, K. (1998). Constraint, word frequency, and the relationship between lexical processing levels in spoken word production. *Journal of Memory and Language, 38,* 313–338.

Jescheniak, J. D., & Levelt, W. J. M. (1994). Word frequency effects in speech production: Retrieval of syntactic information and of phonological form. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*, 824-843.

Junghöfer, M., Elbert, T., Tucker, D. M., & Rockstroh, B. (2000). Statistical control of artifacts in dense array EEG/MEG studies. *Psychophysiology, 37*(4), 523-532.

Kamide, Y. (2008). Anticipatory processes in sentence processing. *Language and Linguistics Compass, 2*, 647-670.

Kutas, M., DeLong, K. A., Smith, N. J. (2011). A look around at what lies ahead: Prediction and predictability in language processing. In: M. Bar (Ed.), *Predictions in the Brain: Using Our Past to Generate a Future* (pp. 190-207). Oxford University Press.

Kutas, M., & Federmeier, K. D. (2011). Thirty Years and Counting: Finding Meaning in the N400 Component of the Event-Related Brain Potential (ERP). *Annual Review of Psychology, 62*, 621-647.

Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science, 207(4427),* 203-205.

Kutas, M. and Hillyard, S.A. (1984) Brain potentials during reading reflect word expectancy and semantic association. *Nature, 307,* 161–163.

Laaksonen, H., Kujala, J., Hultén, A., Liljeström, M., & Salmelin, R. (2012). MEG evoked responses and rhythmic activity provide spatiotemporally complementary measures of neural activity in language production. *Neuroimage, 60,* 29-36.

Lau, E., Holcomb, P. J., & Kuperberg, G. R. (in press). Dissociating N400 effects of prediction from association in single word contexts. *Journal of Cognitive Neuroscience*.

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences, 22,* 1-38.

Mani, N., & Huettig, F. (2012). Prediction during language processing is a piece of cake - but only for skilled producers. *Journal of Experimental Psychology: Human Perception and Performance, 38(4),* 843-847.

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods, 164*(1), 177-190.

McCauley, S. M. & Christiansen, M. H. (2011). Learning simple statistics for language comprehension and production: The CAPPUCCINO model. In L. Carlson, C. Hölscher, &

T. Shipley (Eds.), *Proceedings of the 33rd Annual Conference of the Cognitive Science Society* (pp. 1619-1624). Austin, TX: Cognitive Science Society.

McFarland, D. J., Miner, L. A., Vaughan, Th. M., & Wolpaw, J. R. (2000). Mu and beta rhythm topographies during motor imagery and actual movements. *Brain Topography, 12(3),* 177-186.

McPherson, W. B., & Holcomb, P. J. (1999). An electrophysiological investigation of semantic priming with pictures of real objects. *Psychophysiology,36(1),* 53-65.

Moulton, S. T., & Kosslyn, S. M. (2009). Imagining predictions: Mental imagery as mental emulation. *Philosophical Transactions of the Royal Society B, 364,* 1273-1280.

Nigam, A., Hoffman, J., & Simons, R. F. (1992). N400 to semantically anomalous pictures and words. *Journal of Cognitive Neuroscience, 4(1),* 15-22.

Neuper, C., & Pfurtscheller, G. (1999). Motor imagery and ERD. In: G. Pfurtscheller and F. H. Lopes da Silva (Eds.), *Event-related desynchronization: Handbook of electroencephalography and clinical neurophysiology* (pp.303-325). Revised edition Vol. 6. Amsterdam: Elsevier.

Neuper, C., Wörtz, M., & Pfurtscheller, G. (2006). ERD/ERS patterns reflecting sensorimotor activation and deactivation. *Progress in Brain Research, 159,* 211–222.

Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J.-M. (2011). FieldTrip: Open Source Software for Advanced Analysis of MEG, EEG, and Invasive Electrophysiological Data. *Computational Intelligence and Neuroscience, 2011*, 9.

Oostenveld, R., & Praamstra, P. (2001). The five percent electrode system for high-resolution EEG and ERP measurements. *Clinical Neurophysiology, 112*(4), 713-719.

Pfurtscheller, G., Brunner, C., Schlögl, A., & Lopes da Silva, F. H. (2006). Mu rhythm (de)synchronization and EEG single-trial classification of different motor imagery tasks. *Neuroimage, 31(1)*, 153-159.

Pickering, M. J., & Garrod, S. (2007). Do people use language production to make predictions during comprehension? *Trends in Cognitive Sciences, 11*(3), 105-110.

Pickering, M. J., & Garrod, S. (in press). An integrated theory of language production and comprehension. *Behavioral and Brain Sciences*.

Rohenkohl, G., & Nobre, A. C. (2011). Alpha oscillations related to anticipatory attention follow temporal expectations. *The Journal of Neuroscience, 31(40),* 14076-14084.

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.

Szekely, S., et al. (2004). A new on-line resource for psycholinguistic studies. *Journal of Memory and Language, 51,* 247-250.

Szewczyk, J., & Schriefers, H. (2013). Prediction in language comprehension beyond specific words: An ERP study on sentence comprehension in Polish. *Journal of Memory and Language, 68,* 297-314.

Thut, G., Nietzel, A., Brandt, S. A., Pascual-Leone, A. (2006). α-Band electroencephalographic activity over occipital cortex indexes visuaspatial attention bias and predicts visual target detection. *The Journal of Neuroscience, 26(37),* 4949-9502.

van Berkum, J. J. A. (2004). Sentence comprehension in a wider discourse: Can we use ERPs to keep track of things? In M. Carreiras & C. Clifton (Eds.), *On-Line Study of Sentence Comprehension: Eyetracking, ERPs and Beyond* (pp. 229-270). Hove: Psychology Press.

van Berkum, J. J. A. (2010). The brain is a prediction machine that cares about good and bad - Any implications for neuropragmatics? *Italian Journal of Linguistics, 22,* 181-208.

van Berkum, J. J. A., Brown, C. M., Zwitserlood, P., Kooijman, V., & Hagoort, P. (2005). Anticipating upcoming words in discourse: Evidence from ERPs and reading times. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31(3),* 443-467.

Wicha, N.Y., Moreno, E.M., & Kutas, M. (2003). Expecting gender: An event related brain potential study on the role of grammatical gender in comprehending a line drawing within a written sentence in Spanish. *Cortex, 39(3),* 483-508.

Wicha, N. Y., Moreno, E. M., and Kutas, M. (2004). Anticipating words and their gender: an event-related brain potential study of semantic integration, gender expectancy, and gender agreement in Spanish sentence reading. *Journal of Cognitive Neuroscience,16(7),* 1272-1288.

Willems, R. M., Özyürek, A., & Hagoort, P. (2008). Seeing and Hearing Meaning: ERP and fMRI Evidence of Word versus Picture Integration into a Sentence Context. *Journal of Cognitive Neuroscience, 20(7),* 1235-1249.

Chapter 6

Summary and discussion

Summary and discussion

This chapter summarizes the main findings from the previous chapters. In addition, the findings are put in a somewhat broader perspective by proposing generalizations and highlighting potential avenues for future research.

Summary of the results

Many studies suggest that visual representations, such as the orientation and shape of objects, become activated during language processing (e.g., Stanfield & Zwaan, 2001; Zwaan, Stanfield, & Yaxley, 2002). Chapter 2 investigated whether such representations influence performance in a routine or task-dependent manner. Participants read sentences that implied a particular shape or orientation of an object. After each sentence, they were presented with a picture of that object, which either matched or mismatched with the implied orientation or shape. The effects of these manipulations were examined in three tasks: participants named the object (Experiment 1), they decided whether the object had been mentioned in the sentence (Experiment 2), or they visually imagined the situation described in the sentence before naming the object (Experiment 3). No reliable effects of orientation were observed, suggesting that orientation representations do not play a crucial role in language processing. Ratings from an independent group of participants (Experiment 4) showed that the lack of orientation effects did not stem from a relative weakness in the orientation manipulation in the materials. Clear effects of shape were observed only when participants had to compare the sentences to the pictures (Experiment 2) or when they had been instructed to use visual imagery (Experiment 3). This suggests that implied shape influences performance in a task-dependent manner. These results are noteworthy because they cast doubt on earlier claims that orientation representations are routinely activated (Wassenburg & Zwaan, 2010). Furthermore, they suggest that imagery is worth further investigation as a possible mediating factor between language processing and the activation of shape representations.

While Chapter 2 investigated the influence of visual information on subsequent task performance, Chapter 3 investigated the role of object shape representations in a strong form of incremental sentence comprehension: the anticipation of references to objects. The main question was whether object shape representations can be activated in an anticipatory fashion. In one experiment, participants listened to predictable words (e.g., "moon") in sentences (e.g., "In 1969 Neil Armstrong was the first man to set foot on the moon"). Five-hundred milliseconds before the onset of the predictable word, they were presented with a visual display with several objects, while their eye movements were tracked. The displays contained, besides three distractor objects that were unrelated to the sentence, a critical object, which was either the target object (e.g., a moon), an object with a similar shape (e.g., a tomato), or an unrelated control object (e.g., rice). In a time window before the spoken word (e.g., "moon") could influence eye movements, participants already fixated both the target object and the object with a similar visual shape more often than the other objects. This demonstrates that object shape information had been retrieved in an anticipatory fashion. An ERP experiment where the same materials were used without visual displays revealed analogous effects on N400 amplitude. Words that were semantically anomalous given the sentence context elicited larger N400 amplitudes than words that did fit in the sentence, but the effect was attennuated when their referents' shape was similar to the expected word's referent (e.g., "In 1969 Neil Armstrong was the first man to set foot on the tomato"). This result provides converging evidence for the conclusions from the eye-tracking experiment and in addition suggests that a visual environment is not necessary to obtain effects of object shape. While earlier studies had reported effects of shape similarity on eye movements (Dahan & Tanenhaus, 2005; Huettig & Altmann, 2007) and ERPs (Kellenbach, Wijers, &

Mulder, 2000), these studies did not employ highly predictive sentence contexts. The findings from Chapter 3 suggest an additional anticipatory locus for these effects, implying a link between vision and language processing at very early processing stages. Listeners can activate object shape representations very rapidly, even before words that refer to the objects are uttered.

In Chapter 4, the eye-tracking experiment from Chapter 3 was successfully replicated with a longer preview time of the visual displays. This time, the displays were presented from one second before the sentence until two seconds after sentence offset. In addition, participants performed several other tasks, and their performance on these tasks was compared to their performance in the eye-tracking experiment. Most notably, participants performed a spatial cueing task where they indicated the location of an X symbol on the screen (left or right) after viewing a briefly presented cue. The cues were valid (an arrow pointing in the direction where the X would appear), invalid (an arrow pointing in the other direction), or neutral (a plus sign). Participants' facilitation from the predictive arrow cues was quantified by subtracting response times in the valid condition from those in the neutral condition. Individual differences in this measure predicted anticipatory eye movements to targets (e.g., the moon; albeit statistically marginally) and to shape competitors (e.g., the tomato). Participants who were most strongly facilitated by predictive arrow cues also showed the strongest effects of predictable language input on their anticipatory eye movements. Individual differences in performance on the other tasks – a predictable-action observation task, a fluid intelligence task, a vocabulary task, and a verbal fluency task – did not form statistically reliable predictors of the anticipatory eye movements. The results support the idea that some anticipatory attentional mechanisms may be shared across verbal and non-verbal tasks. Because in previous research different types of predictability effects have been studied in different experiments, even in different fields, these results represent a first step in exploring similarities and differences between predictive mechanisms in language and other cognitive domains.

Chapter 5 compared language production and comprehension. Participants were presented with objects (e.g., a broom) after predictive lead-in sentences (e.g., “He swept the floor with a…”) or neutral lead-in sentences (e.g., “He saw the drawing of a…”). Participants named the objects or just viewed them, while answering occasional comprehension questions. Their EEG was recorded. In the object naming task, naming latencies were shorter for objects in predictive than in neutral sentence contexts. ERP analyses yielded a sustained positivity in Predictive relative to Neutral sentence contexts, starting before object onset. In addition, time-frequency analyses revealed a suppression of power in predictive relative to neutral sentence contexts in frequencies typically associated with motor preparation, which also started before the object was presented. This suggests that participants assembled a motor plan in an anticipatory fashion. In contrast, in the object viewing task, these pre-stimulus effects were not observed. Thus, at the level of motor preparation, reading predictable sentences is clearly different from preparing to complete sentences out loud. The results constrain the possible involvement of language production in comprehension to stages before motor preparation. Whereas several recent theories and models treat language production and comprehension as highly similar or tightly linked (Chang, Dell, & Bock, 2006; McCauley & Christiansen, 2011; Pickering & Garrod, in press), these results urge caution in linking production and comprehension.

The contents of predictions

The results from Chapters 3 and 4 provided one piece of the puzzle regarding the contents of predictions. Not only functional semantic aspects of meaning (e.g., edibility; Altmann & Kamide, 1999), but also information regarding an object’s physical shape can become pre-activated during sentence comprehension. This process might enable language-vision interactions to proceed smoothly and efficiently in real-world situations where language is often used to guide visual attention towards relevant objects.

The fact that this finding adds to accumulating evidence demonstrating that many types of information can become pre-activated (as discussed in Chapter 1) invites a tempting generalization: Can all types of information that can become *activated* when processing words or concepts also become *pre-activated*? This appears to be a viable working hypothesis. One challenge for the future lies in delineating the time course with which these different representations become activated. For instance, is semantic, syntactic, and phonological information pre-activated in parallel or serially? And if serially, then in what order? Of course, the answer to such questions may depend strongly on the (linguistic and non-linguistic) context. The current thesis used sentence contexts that allowed listeners to predict upcoming linguistic input with a high degree of certainty. It may not be very efficient to pre-activate shapes of the wide range of objects that might be referred to under more uncertain conditions. Ideally then, a model of predictive language processing would not only specify the contents of predictions, the cues that predictions are based on, and the mechanisms involved, but also the role that context may play in dynamically promoting or inhibiting ongoing hypotheses regarding what may come next. In addition, future studies could investigate whether interactions exist between the cues, contents and mechanisms, in the sense that certain cues enable the preactivation of certain contents. As a hypothetical example, semantic cues might be used to predict *which* words are likely to come up, whereas prosodic cues might be more relevant to predicting *when* a word will occur.

Task-dependence and different types of visual representations

In Chapters 3 and 4 it was observed that object shape representations were activated in an anticipatory fashion. These representations influenced eye movements both with a short preview of the visual display (Chapter 3) and with a long preview (Chapter 4), and they had an effect on event-related brain potentials during a task where no visual displays were presented at all (Chapter 3). None of these experiments required participants to engage in meta-linguistic decisions. One may wonder why then in Chapter 2, where such decisions were required, clear evidence for the activation of object shape representations was only observed during meta-linguistic tasks, but not when participants simply read sentences and named objects.

Because the studies described in the different chapters were conducted using different participants, materials, tasks, and experimental techniques, there is a large number of factors that the difference in results could be attributed to. Apart from methodological considerations, however, an important conceptual difference between the chapters in terms of the shape information involved should not be overlooked. In Chapters 3 and 4, object shape representations of certain objects (e.g., a moon) were found to influence how other objects (e.g., a tomato) and references to those objects were processed. The object shapes involved generally represented the canonical forms of the objects, which are likely to be the first shapes that come to mind when thinking about concepts. In contrast, in Chapter 2 the manipulation involved certain objects (e.g., an eagle) of which a particular variant of their shape (e.g., flying or sitting) was contextually implied. Any effect of such a manipulation requires a relatively more elaborate integration of the object with its context, perhaps at the level of a situation model. This might make the activation of such representations more task-dependent. The effects of implied visual representations on task performance are likely to be only as strong as the weakest link in all the processes occurring between receiving linguistic input and responding to the presented stimuli.

At the same time, the fact that the shape representations in Chapters 3 and 4 were activated under conditions where participants were simply asked to listen for comprehension does not warrant the complementary conclusion, namely that the activation of canonical shapes would be task-independent. It is worth noting that the effects in all three chapters, though statistically reliable, were relatively small. Average effects of implied shape on naming latencies (Chapter 2) were no larger than 20 milliseconds. Anticipatory biases toward shape competitor objects (Chapters 3 and 4) occurred on a relatively small proportion of trials. Finally, the shape similarity effect on ERPs (Chapter 3) occurred relatively late. In addition to the factors discussed in the relevant chapters, the categories from which the objects were taken might be relevant, although this was not investigated. For instance, in a functional magnetic resonance imaging (fMRI) study, Chao, Haxby, and Martin (1999) observed a distinction between objects in the category of four-legged animals, which differ mainly in visual form, and objects in the category of tools, which among other things can be differentiated by the motion associated with their use. Greater activity in the medial occipital cortex (associated with visual perception) was observed in response to photographs and written names of four-legged animals than in response to tools, suggesting that categories differ in the extent to which they elicit activation of visual form. In summary, although for both canonical and contextually implied shape effects it is important to demonstrate that they exist, neither of the two should necessarily be regarded as an automatic consequence of receiving linguistic input.

Relationships across domains

In Chapter 4 a positive relationship was observed between individual differences in response facilitation by predictive arrow cues and language-mediated anticipatory eye movements. Future investigations should investigate the robustness of this relationship across different contexts and tasks. Such research should further specify which processes and/or representations underlying anticipatory behavior may be shared across verbal and nonverbal domains.

This relationship between anticipatory behavior in a spatial cueing task - widely used in attention research - and the visual world paradigm - widely used in psycholinguistic research - also implies that researchers in different areas might profit from each other's achievements. For instance, predictive mechanisms that gain substantial empirical support in attention, perception, or action research should also be tested in language comprehension. The present thesis is not alone in endorsing this view. For instance, Kutas, DeLong, and Smith (2011) emphasize that language researchers investigating prediction should take "a look around at what lies ahead", that is, they should let themselves be informed by developments in other fields when plotting future directions (see also van Berkum, 2010). A useful first step in making such crosstalk effective would be to devise common definitions of the core concepts. For instance, "prediction" and "anticipation" are often used more or less interchangeably (unfortunately, this thesis is no exception). It would obviously be beneficial to spell out the commonalities and differences between these terms. Fortunately, in the domain of neuroscience recently Bubic, von Cramon, and Schubotz (2010) proposed definitions for the terms "prediction" (general orientation towards the future), "expectation" (representations of what is predicted to occur in the future; similar to what has been called "contents" in this thesis), "anticipation" (formulating and communicating expectations to sensory and motor areas), and "prospection" (consideration of potential distant future events). These definitions might not be perfect, but further discussion could facilitate a more broadly integrated investigation of predictive behavior.

Constraints on production-comprehension links

An inspiring theory implicates the production system in comprehension, especially in predictive language comprehension. However, in Chapter 4, individuals' anticipatory eye movements did not vary systematically as a function of their verbal fluency performance in a task that involves production. In Chapter 5, it was found that the electrophysiological signal recorded while participants silently read predictable sentences was very different from that recorded while participants completed the sentences out loud. Of course, these results may well depend on the aspects of production tapped by the measures employed. Furthermore, as mentioned previously in Chapter 4, the participants were mainly students, and the production system might be especially involved in comprehension at ages when comprehension develops or declines (cf. Federmeier, Kutas, & Schul, 2010; Federmeier, McLennan, De Ochoa, & Kutas, 2002; Mani & Huettig, 2012). Nevertheless, the involvement of the production system in comprehension did not receive support in this thesis. Instead, constraints on production-comprehension links were highlighted.

Conclusions

Our ability to communicate through language is usually seen as very special, because it is a uniquely human skill. However, this does not entail that language should be studied as isolated from other cognitive functions that we happen to share with animals. Language use necessitates the coordination of linguistic operations with other processes such as visual perception and attention. This thesis examined interactions between language and other cognitive functions as listeners and readers processed and anticipated language referring to objects.

One issue that was addressed are the contents of predictions. The reported experiments revealed a rapid linking process between language and vision: as participants listened to sentences, their eye movements showed that they began mapping words onto objects at a visual level before these words had been heard. Event-related brain potentials were also consistent with a role for visual representations in the contents of predictions. In some sense, predictive sentence contexts thus allow listeners to "see" what is coming next. On the other hand, another set of experiments showed that visual representations do not influence language processing performance to the same extent in all tasks. This highlighted the influence of context on the activation of visual information during sentence comprehension.

The *mechanisms* underlying predictive language processing were also addressed. Analyses of individual differences showed that participants who strongly used non-linguistic predictive arrow cues also strongly used predictive linguistic context. Thus, predictive language processing appears to be supported at least in part by mechanisms that are not unique to language processing.

Furthermore, while some theories have proposed that predictions during comprehension are generated by covert language production, the results reported here did not provide support for this view. For instance, participants who are good at producing as many words as possible within a certain amount of time are not necessarily better predictors. Furthermore, the brain activity underlying the process of reading predictable sentences is quite distinct from that of completing sentences out loud, especially at the level of motor preparation. Earlier studies that have observed relationships between production skills and predictive language comprehension can be reconciled with these results by the suggestion that prediction is likely to rely on multiple mechanisms. In summary, this thesis underscored context-dependent links between language, vision, and attention.

References

Bubic, A., von Cramon, D. Y., & Schubotz, R. I. (2010). Prediction, cognition and the brain. *Frontiers in Human Neuroscience, 4,* 25.

Chang, F., Dell, G. S., & Bock, K. (2006). Becoming syntactic. *Psychological Review, 113(2),* 234-272.

Chao, L. L., Haxby, J. V., & Martin, A. (1999). Attribute-based neural substrates in temporal cortex for perceiving and knowing about objects. *Nature Neuroscience, 2*(10), 913-919.

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.

Federmeier, K. D., Kutas, M., & Schul, R. (2010). Age-related and individual differences in the use of prediction during language comprehension. *Brain and Language, 115,* 149-161.

Federmeier, K. D., McLennan, D. B., De Ochoa, E., & Kutas, M. (2002). The impact of semantic memory organization and sentence context information on spoken language processing by younger and older adults: An ERP study. *Psychophysiology, 39,* 133-146.

Kutas, M., DeLong, K. A., Smith, N. J. (2011). A look around at what lies ahead: Prediction and predictability in language processing. In: M. Bar (Ed.), *Predictions in the brain: Using our past to generate a future* (pp. 190-207). Oxford University Press.

Huettig, F., & Altmann, G. T. M. (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.

Kellenbach, M. L., Wijers, A. A., & Mulder, G. (2000). Visual semantic features are activated during the processing of concrete words: event-related potential evidence for perceptual semantic priming. *Cognitive Brain Research, 10*(1-2), 67-75.

Mani, N., & Huettig, F. (2012). Prediction during language processing is a piece of cake - but only for skilled producers. *Journal of Experimental Psychology: Human Perception and Performance, 38(4),* 843-847.

McCauley, S. M. & Christiansen, M. H. (2011). Learning simple statistics for language comprehension and production: The CAPPUCCINO model. In L. Carlson, C. Hölscher, & T. Shipley (Eds.), *Proceedings of the 33rd Annual Conference of the Cognitive Science Society* (pp. 1619-1624). Austin, TX: Cognitive Science Society.

Pickering, M. J., & Garrod, S. (in press). An integrated theory of language production and comprehension. *Behavioral and Brain Sciences*.

Stanfield, R.A., & Zwaan, R.A. (2001). The effect of implied orientation derived from verbal context on picture recognition. *Psychological Science, 12*(2), 153-156.

van Berkum, J. J. A. (2010). The brain is a prediction machine that cares about good and bad - Any implications for neuropragmatics? *Italian Journal of Linguistics, 22,* 181-208.

Wassenburg, S. I., & Zwaan, R. A. (2010). Readers routinely represent implied object rotation: The role of visual experience. *Quarterly Journal of Experimental Psychology, 63*(9), 1665-1670.

Zwaan, R.A., Stanfield, R.A., & Yaxley, R.H. (2002). Language comprehenders mentally represent the shapes of objects. *Psychological Science, 13*(2), 168-171.

Nederlandse samenvatting

Volwassen lezers en luisteraars activeren en integreren de betekenis van binnenkomende taal op basis van eerder vergaarde kennis. Veel van deze kennis is van abstracte aard (bijvoorbeeld de betekenis van het woord “betekenis”), maar een deel van de kennis heeft te maken met direct waarneembare fysieke aspecten van de wereld om ons heen, zoals hoe objecten er uitzien. Het lezen van het woord “appel” bijvoorbeeld kan kennis van hoe een appel er uitziet activeren. Dit proefschrift behandelt de mechanismen en representaties die het verwerken van referenties naar objecten in zinscontext ondersteunen. Hieronder worden kort enkele bevindingen aangestipt.

Eerder onderzoek heeft laten zien dat visuele representaties zoals de orientatie en vorm van objecten wordt geactiveerd tijdens taalverwerking. Hoofdstuk 2 onderzocht of zulke representaties gedrag op een routinematige manier beïnvloeden of op een taak-afhankelijke manier. Wanneer je zinnen over objecten leest, bedenk je dan altijd (bewust of onbewust) hoe ze er uitzien, of hangt dit af van de taakomgeving? In drie verschillende reactietijdentaken had de oriëntatie van objecten geen effect. Dit suggereert dat orientatierepresentaties wellicht niet cruciaal zijn voor taalverwerking. Duidelijke effecten van visuele vorm werden alleen gevonden wanneer proefpersonen zinnen en plaatjes met elkaar moesten vergelijken of wanneer ze zich een in een zin beschreven situatie eerst visueel hadden ingebeeld. Samengenomen kunnen op basis van deze resultaten claims over een routinematige activatie van visuele representaties in twijfel worden getrokken.

Wanneer we woorden lezen of horen wordt er veel informatie geactiveerd, en dit gebeurt ook nog eens heel snel. Sterker nog, voordat een zin is afgelopen voorspellen we vaak al hoe hij verder zal gaan en activeren we informatie die nog niet genoemd is. Hoofdstuk 3 onderzocht of deze gepreactiveerde informatie ook visueel van aard kan zijn. Proefpersonen luisterden naar voorspelbare woorden (bijvoorbeeld “maan”) in zinnen (“In 1969 zette Neil Armstrong als eerste mens voet op de maan") terwijl ze naar een set objecten keken op een beeldscherm. Ondertussen werden hun oogbewegingen gemeten, omdat luisteraars die vooruitlopen op de taal soms al naar objecten kijken voordat die genoemd worden. Uit dit experiment bleek dat luisteraars die een woord voorspelden niet alleen veel naar voorspelbare objecten zelf keken (bijvoorbeeld naar een plaatje van een maan), maar ook naar andere objecten met een vergelijkbare vorm (zoals een tomaat, die ook rond is). Dit suggereert dat luisteraars inderdaad op een voorspellende manier visuele informatie kunnen activeren. Een ander experiment waarbij hersengolven uit het electro-encephalogram (EEG) werden gemeten ondersteunde de resultaten. Dergelijke visuele voorspellingen zouden nuttig kunnen zijn voor luisteraars in dagelijkse situaties waarin sprekers verwijzen naar objecten die ter plaatse aanwezig zijn.

In Hoofdstuk 4 werden de mechanismen van dit voorspellingsgedrag verder onderzocht. Is er een speciaal taalvoorspellingssysteem, of maakt taal gebruik van voorspellingsmechanismen die ook in andere cognitieve domeinen worden gebruikt? In dat geval zouden mensen die veel voorspellen in talige taken ook veel moeten voorspellen in nonverbale taken. Het oogbewegingsexperiment werd succesvol herhaald in combinatie met andere taken. In een van de taken reageerden proefpersonen op stimuli (X-vormige symbolen) die rechts of links op een beeldscherm werden getoond, voorafgegaan door een pijltje dat met 80% zekerheid aangaf waar een stimulus zou verschijnen. Proefpersonen wiens gedrag in deze nonverbale taak sterk gefaciliteerd werd door zulke voorspellende pijltjes vertoonden ook de sterkste neiging tot het maken van voorspellende oogbewegingen tijdens het luisteren naar zinnen. Dit suggereert dat bepaalde voorspellende aandachtsprocessen wellicht gebruikt worden in zowel verbale als nonverbale taken, en er dus niet (alleen) een uniek taalvoorspellingssysteem is.

Sommige theorieën stellen dat voorspellende taalverwerking gebruikmaakt van het taalproductiesysteem. Volgens deze hypothese maken lezers of luisteraars zinnen in hun hoofd af door zich als het ware af te vragen wat zij zouden zeggen wanneer ze de spreker waren. Dit kan betekenen dat het stil lezen van voorspelbare zinnen vergelijkbaar is met het hardop afmaken van zinnen en zelfs tot een voorspellende activatie van motorische spraakprocessen zou kunnen leiden. In Hoofdstuk 5 lazen proefpersonen stil voorspelbare en neutrale zinnen die eindigden met een afbeelding van een object, of ze maakten deze zinnen hardop af. Ondertussen werd hun hersenactiviteit gemeten met EEG. Het bleek dat de hersenactiviteit duidelijk verschilde tussen de twee taken, vooral op het niveau van het voorbereiden van motorische spraakprocessen. Het lezen van voorspelbare zinnen is op dit niveau dus niet erg vergelijkbaar met het hardop afmaken van zinnen. Dit beperkt de mogelijke rol van taalproductie in voorspellende taalverwerking tot processen voorafgaand aan motorische voorbereiding.

Samengevat zijn met dit proefschrift enkele puzzelstukjes gelegd met betrekking tot de verwerking en voorspelling van taal die naar objecten verwijst. De resultaten laten zien hoe specifiek de informatie die wordt geactiveerd kan zijn en hoe snel deze activatie kan plaatsvinden. Ook geven ze steun voor het idee dat talige voorspellingsprocessen ten minste ten dele worden ondersteund door aandachtsmechanismen die niet uniek zijn voor taal en benadrukken ze de belangrijke modulerende rol van context tijdens taalverwerking.

Acknowledgements

Of course every PhD project has its ups and downs, but looking back I see mainly ups. There are many people who have contributed to this. Unfortunately I cannot mention all of you, but I am grateful to everybody.

First of all I would like to thank my supervisors, Falk Huettig and Antje Meyer. Working with you both is a great experience. Falk, thank you for always being available for questions and brainstorming sessions that shaped my thinking, for valuing my ideas as if I was a fellow researcher, for providing feedback at the speed of light, and for all the humor. Antje, thank you for your excellent support, for insisting on theoretically relevant and simple experiments, for trusting me with freedom to pursue my interests, and for also giving feedback at an amazing speed.

I am grateful to Prof. James McQueen, Prof. Herbert Schriefers, and Dr. Christian Dobel for reading and evaluating this thesis. Although he will sadly not be able to read this anymore, I also want to thank Prof. Rob Schreuder for agreeing to be on my reading committee.

Peter Praamstra, thank you for showing me around your lab, for your knowledge of EEG, and for being very helpful during the reviewing process. Vitória Piai, thank you for your natural generosity with time, resources, and friendship. I hope we will continue to collaborate.

I enjoyed the occasional helpful chat about matters of statistics and programming with Matthias Sjerps and with Holger Mitterer. I don't (want to) know how I would have done things without the basics of programming Holger taught in his course. Here I would also like to mention Laurie Stowe, Marcel Bastiaansen and Ton Dijkstra who guided me on my first steps into the world of psycholinguistics.

Thank you to the technical group, especially Johan Weustink, Ronald Fisher and Alex Dukers, for keeping my experiments running and my computer up to date. Thanks to the library and the administration for keeping things going in the background. I also want to thank the IMPRS coordinators Rachel Sheer, Els den Os and Dirkje van der Aa for organizing nice events and enrolling me in helpful courses. I am grateful to the many research assistants of our department for their help with some tedious tasks, like annotating sound files, and to Tilman Harpe for drawing many pictures for one of the studies.

My paranymphs, Zeshu Shao and Maartje van de Velde, I am glad to have you standing by my side at the defense. Zeshu, dear office mate, thank you for sharing the ups and downs and for our nice chats about all kinds of scientific and nonscientific matters. Dear Maartje, thank you too for our science- and non-science chats and your always cheerful and enthusiastic presence.

Thanks also to everybody else who tolerated my presence in the same office. I very much enjoyed the sarcastic jokes and life stories of Matthias and Marijt as well as Janise in 362, talking about cultures and careers with Loan in 232, and the laughs and fun stories first with Wencui and then of course with Conny and Carmen (and for a while Isabelle) in 362. Thanks to Francesca and Florian for being like officemates too.

I am grateful to the members of and everyone connected to the Psychology of Language department over the years: Agnieszka, Alastair, Alma, Annelies, Aoju, Chu, Dan, Esther, Evelyn, Florian, Francesca, Jana, Jannie, Johanna, Johanne, Leonie, Lesya, Linda, Neil, Odette, Shiri, Suzanne, Svetlana, and Will. Thanks for your help, thoughts, comments, and for all the fun times we had during group activities, conferences, and simply in and around the office as co-workers and friends. With respect to this thesis I also appreciate Neil's reading of my general introduction and discussion. Thanks to the members of Anne Cutler's comprehension department, who provided me with additional feedback, fun lunches, and exciting meetings to take part in when

Antje’s department was still in its startup phase. Also thanks to the MPI football team for the active Friday evenings.

To my friends outside work (you know who you are!): it is always fun to hang out even though we sometimes don’t see each other for a while. I am sure we will continue to enjoy nice parties, activities, game nights and holidays, which were great ways to take my mind off work for a bit. Thank you to the members of 't Haasje for the sports and the fun during my first year when I was still a regular on the track.

Thank you to my family. Pap, mam, Jelle, Wieke, Daan, you always support me in whatever I do, even if the research as explained by me must sometimes sound somewhat obscure. Thank you for being there.

Last but not least, Suzanne, consider this an honorary last authorship. I am very grateful for your love and support, as well as for your tolerance of my absence during busy evenings and weekends. Thank you for the great times we had and for those ahead of us.

Curriculum Vitae

Joost Rommers was born in 1985 in Rijswijk, the Netherlands. In 2007, he obtained a bachelor's degree in General Linguistics (cum laude) with minors in Psychology from the University of Groningen. This was followed by a master's degree in Cognitive Neuroscience (cum laude) from the Radboud University Nijmegen in 2009. In that same year he started his PhD project at the Max Planck Institute for Psycholinguistics. The results of his PhD research in the Psychology of Language Department are described in this thesis. He is now a postdoctoral researcher in the same department.

Publications

Rommers, J., Meyer, A. S., & Huettig, F. (in press). Object shape and orientation do not routinely influence performance during language processing. *Psychological Science.*

Rommers, J., Dijkstra, T., & Bastiaansen, M. C. M. (2013). Context-dependent semantic processing in the human brain: Evidence from idiom comprehension. *Journal of Cognitive Neuroscience, 25(5),* 762-776.

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.

Huettig, F., Rommers, J., & Meyer, A. S. (2011). Using the visual world paradigm to study language processing: A review and critical evaluation. *Acta Psychologica, 137,* 151-171.

MPI Series in Psycholinguistics

1. The electrophysiology of speaking: Investigations on the time course of semantic, syntactic, and phonological processing. *Miranda van Turennout*
2. The role of the syllable in speech production: Evidence from lexical statistics, metalinguistics, masked priming, and electromagnetic midsagittal articulography. *Niels O. Schiller*
3. Lexical access in the production of ellipsis and pronouns. *Bernadette M. Schmitt*
4. The open-/closed-class distinction in spoken-word recognition. *Alette Haveman*
5. The acquisition of phonetic categories in young infants: A self-organising artificial neural network approach. *Kay Behnke*
6. Gesture and speech production. *Jan-Peter de Ruiter*
7. Comparative intonational phonology: English and German. *Esther Grabe*
8. Finiteness in adult and child German. *Ingeborg Lasser*
9. Language input for word discovery. *Joost van de Weijer*
10. Inherent complement verbs revisited: Towards an understanding of argument structure in Ewe. *James Essegbey*
11. Producing past and plural inflections. *Dirk Janssen*
12. Valence and transitivity in Saliba: An Oceanic language of Papua New Guinea. *Anna Margetts*
13. From speech to words. *Arie van der Lugt*
14. Simple and complex verbs in Jaminjung: A study of event categorisation in an Australian language. *Eva Schultze-Berndt*
15. Interpreting indefinites: An experimental study of children's language comprehension. *Irene Krämer*
16. Language-specific listening: The case of phonetic sequences. *Andrea Weber*
17. Moving eyes and naming objects. *Femke van der Meulen*
18. Analogy in morphology: The selection of linking elements in Dutch compounds. *Andrea Krott*
19. Morphology in speech comprehension. *Kerstin Mauth*
20. Morphological families in the mental lexicon. *Nivja H. de Jong*
21. Fixed expressions and the production of idioms. *Simone A. Sprenger*

22. The grammatical coding of postural semantics in Goemai (a West Chadic language of Nigeria). *Birgit Hellwig*
23. Paradigmatic structures in morphological processing: Computational and crosslinguistic experimental studies. *Fermín Moscoso del Prado Martín*
24. Contextual influences on spoken-word processing: An electrophysiological approach. *Daniëlle van den Brink*
25. Perceptual relevance of prevoicing in Dutch. *Petra M. van Alphen*
26. Syllables in speech production: Effects of syllable preparation and syllable frequency. *Joana Cholin*
27. Producing complex spoken numerals for time and space. *Marjolein Meeuwissen*
28. Morphology in auditory lexical processing: Sensitivity to fine phonetic detail and insensitivity to suffix reduction. *Rachèl J. J. K. Kemps*
29. At the same time...: The expression of simultaneity in learner varieties. *Barbara Schmiedtová*
30. A grammar of Jalonke argument structure. *Friederike Lüpke*
31. Agrammatic comprehension: An electrophysiological approach. *Marlies Wassenaar*
32. The structure and use of shape-based noun classes in Miraña (North West Amazon). *Frank Seifart*
33. Prosodically-conditioned detail in the recognition of spoken words. *Anne Pier Salverda*
34. Phonetic and lexical processing in a second language. *Mirjam Broersma*
35. Retrieving semantic and syntactic word properties. *Oliver Müller*
36. Lexically-guided perceptual learning in speech processing. *Frank Eisner*
37. Sensitivity to detailed acoustic information in word recognition. *Keren B. Shatzman*
38. The relationship between spoken word production and comprehension. *Rebecca Özdemir*
39. Disfluency: Interrupting speech and gesture. *Mandana Seyfeddinipur*
40. The acquisition of phonological structure: Distinguishing contrastive from noncontrastive variation. *Christiane Dietrich*
41. Cognitive cladistics and the relativity of spatial cognition. *Daniel B.M. Haun*
42. The acquisition of auditory categories. *Martijn Goudbeek*
43. Affix reduction in spoken Dutch. *Mark Pluymaekers*
44. Continuous-speech segmentation at the beginning of language acquisition: Electrophysiological evidence. *Valesca Kooijman*

45. Space and iconicity in German Sign Language (DGS). *Pamela Perniss*
46. On the production of morphologically complex words with special attention to effects of frequency. *Heidrun Bien*
47. Crosslinguistic influence in first and second languages: Convergence in speech and gesture. *Amanda Brown*
48. The acquisition of verb compounding in Mandarin Chinese. *Jidong Chen*
49. Phoneme inventories and patterns of speech sound perception. *Anita Wagner*
50. Lexical processing of morphologically complex words: An information-theoretical perspective. *Victor Kuperman*
51. A grammar of Savosavo, a Papuan language of the Solomon Islands. *Claudia Wegener*
52. Prosodic structure in speech production and perception. *Claudia Kuzla*
53. The acquisition of finiteness by Turkish learners of German and Turkish learners of French: Investigating knowledge of forms and functions in production and comprehension. *Sarah Schimke*
54. Studies on intonation and information structure in child and adult German. *Laura de Ruiter*
55. Processing the fine temporal structure of spoken words. *Eva Reinisch*
56. Semantics and (ir)regular inflection in morphological processing. *Wieke Tabak*
57. Processing strongly reduced forms in casual speech. *Susanne Brouwer*
58. Ambiguous pronoun resolution in L1 and L2 German and Dutch. *Miriam Ellert*
59. Lexical interactions in non-native speech comprehension: Evidence from electroencephalography, eye-tracking, and functional magnetic resonance imaging. *Ian FitzPatrick*
60. Processing casual speech in native and non-native language. *Annelie Tuinman*
61. Split intransitivity in Rotokas, a Papuan language of Bougainville. *Stuart Robinson*
62. Evidentiality and intersubjectivity in Yurakaré: An interactional account. *Sonja Gipper*
63. The influence of information structure on language comprehension: A neurocognitive perspective. *Lin Wang*
64. The meaning and use of ideophones in Siwu. *Mark Dingemanse*
65. The role of acoustic detail and context in the comprehension of reduced pronunciation variants. *Marco van de Ven*
66. Speech reduction in spontaneous French and Spanish. *Francisco Torreira*

67. The relevance of early word recognition: Insights from the infant brain. *Caroline Junge*
68. Adjusting to different speakers: Extrinsic normalization in vowel perception*.* *Matthias J. Sjerps*
69. Structuring language: contributions to the neurocognition of syntax. *Katrien R. Segaert*
70. Infants' appreciation of others' mental states in prelinguistic communication: a second person approach to mindreading. *Birgit Knudsen*
71. Gaze behavior in face-to-face interaction. *Federico Rossano*
72. Sign-spatiality in Kata Kolok: how a village sign language of Bali inscribes its signing space. *Connie de Vos*
73. Who is talking? Behavioural and neural evidence for norm-based coding in voice identity learning. *Attila Andics*
74. Lexical processing of foreign-accented speech: Rapid and flexible adaptation. *Marijt Witteman*
75. The use of deictic versus representational gestures in infancy. *Daniel Puccini*
76. Territories of knowledge in Japanese conversation. *Kaoru Hayano*
77. Family and neighbourhood relations in the mental lexicon: A cross-language perspective. *Kimberley Mulder*
78. Contributions of executive control to individual differences in word production. *Zeshu Shao*
79. Hearing and seeing speech: Perceptual adjustments in auditory-visual processing. *Patrick van der Zande*
80. High pitches and thick voices: The role of language in space-pitch associations. *Sarah Dolscheid*
81. Seeing what's next: Processing and anticipating language referring to objects. *Joost Rommers*