Predicting language in different contexts: The nature and limits of mechanisms in anticipatory language processing

Cover credits: Paul Hömke, Suzanne R. Jongman, & Alastair C. Smith

ISBN: 978-90-76203-72-0

Printed and bound by Ipskamp Drukkers b.v.

Predicting language in different contexts: The nature and limits of mechanisms in anticipatory language processing

Proefschrift
ter verkrijging van de graad van doctor
aan de Radboud Universiteit Nijmegen
op gezag van de rector magnificus,
volgens besluit van het college van de decanen,
in het openbaar te verdedigen op maandag 14 december 2015
om 10.30 uur precies.

door
Florian Hintz
geboren op 9 maart 1987
te Bernau (Berlijn, Duitsland)

Promotor:

Prof. dr. A.S. Meyer

Copromotor:

Dr. F. Huettig (Max Planck Institute voor Psycholinguïstiek, Nijmegen)

Manuscriptcommissie:

Prof. dr. P. Hagoort

Prof. dr. P. Knoeferle (Bielefeld Universiteit, Duitsland)

Dr. R.M. Willems

This research was supported by the The Max Planck Society for the Advancement of Science, Munich, Germany.

Für meine Familie, Cornelia & Julius

Table of Contents

1 | General introduction

"It is far better to foresee even without certainty than not to foresee at all"

Henry Poincaré (co-founder of the 'chaos theory')
The Foundations of Science, 1913, New York: The Science Press

Many scientists believe that prediction plays an important role in human cognition (e.g., A. Clark, 2013; Engel, Fries, & Singer, 2001; Friston, 2010, for discussion). Prinz (2006), for example, argued that "[for] social animals like we are, the capacity to anticipate what our conspecifics are going to do is, of course, of crucial importance in terms of survival and fitness" (p. 516). In a similar vein, Frith (2007; see also Wilson & Knoblich, 2005) suggested that "[the] better we can predict what someone is going to do next, the more successful our interactions with that person will be" (p. 671). The notion that prediction is an important principle in human information processing is acknowledged by a growing number of theories in the cognitive sciences. Theories of visual and auditory perception, for instance, propose that "the human brain is continuously busy generating predictions that approximate the relevant future" (Bar, 2007, p. 280) and that viewers and listeners engage in prediction to prepare for upcoming visual and acoustic events (e.g., Bar, 2009; Bendixen, Schröger, & Winkler, 2009; Schröger, 2007).

Language comprehension involves visual and auditory perception. The theoretical assumptions about these perceptual processes have thus direct implications for our conception of language comprehension. Indeed, it is well established that readers and listeners frequently predict upcoming information when comprehending language (see Federmeier, 2007; Kamide, 2008; Kutas, DeLong, & Smith, 2011; van Berkum, 2010, for reviews). The prediction of upcoming information during language processing has been suggested to serve various purposes. Generally, it is assumed to speed up the comprehension process, which may, for example, facilitate dialogue (cf. H. H. Clark, 1996; Pickering & Garrod, 2004). When being in a conversation with someone, listeners can anticipate what their interlocutor may say next and when their turn may end and use this information in their own utterance planning (De Ruiter, Mitterer, & Enfield, 2006; Levinson, 2013; Sacks, Schegloff, & Jefferson, 1974). Moreover, prediction has been suggested to be closely linked to language acquisition. Consider a language-learning child who predicts his mother to produce a particular (grammatically incorrect) sentence structure; the mother, however, produces another (grammatically correct) structure instead. The discrepancy (i.e., prediction error) between the child's predicted information and the perceived linguistic input may lead to the acquisition or consolidation of the grammatically correct structure (cf. Chang, Dell, & Bock, 2006; Dell & Chang, 2014).

In everyday life, adult comprehenders are used to understanding language in many different situational contexts. Sometimes these contexts involve the extraction of meaning from written text; sometimes we coordinate spoken language comprehension with the integration of co-present visual input; and sometimes these contexts involve listening and talking to another person. In all of these situations, however, we seem to comprehend language effortlessly. The fact that we can often anticipate upcoming input may contribute to the ease with which we comprehend language in different situational contexts. The present thesis is concerned with the mechanisms underlying prediction in different situations of language processing.

Earlier research

Experimental studies of prediction in language processing have mainly focused on three sub-components of language prediction: cues, contents, and mechanisms (Rommers, Meyer, & Huettig, 2015). Cues are the properties of the input signal that can be used to generate predictions about upcoming information. The term contents denotes the types of linguistic and non-linguistic information that comprehenders may anticipate. Finally, mechanisms are the processes that 'connect' cues and contents, thus enabling the pre-activation of contents given particular input cues.

Earlier research showed that various types of linguistic information can be used to predict upcoming input. Among the cues investigated are lexical-syntactic information (De Ruiter et al., 2006), syntactic structures (Arai & Keller, 2012), case marking (Kamide, Scheepers, & Altmann, 2003), prosody (Weber, Grice, & Crocker, 2006), general discourse context (Kaiser & Trueswell, 2004; van Berkum, Brown, Zwitserlood, Kooijman, & Hagoort, 2005), and world knowledge (Hagoort, Hald, Bastiaansen, & Petersson, 2004). Likewise, listeners and readers are able to predict a wide range of contents, including functional semantic aspects of meaning (Altmann & Kamide, 1999; Federmeier & Kutas, 1999), phonological and orthographic form (DeLong, Urbach, & Kutas, 2005; Laszlo & Federmeier, 2009), visual form information (Rommers, Meyer, Praamstra, & Huettig, 2013; Rommers et al., 2015), and syntactic structures (Arai & Keller, 2012; Carminati, van Gompel, Scheepers, & Arai, 2008; Staub & Clifton Jr, 2006).

Previous research has led to a number of different proposals with respect to the mechanisms underlying prediction. Altmann and Mirković (2009; see also Metusalem et al., 2012) have linked predictive language processing to the anticipation of events. On such an account, comprehenders are assumed to map linguistic and/or visual input onto stored mental representations about real-world events. These authors suggested that the activation of generalized event knowledge plays an important role in prediction and that the comprehension system may be "maximally incremental [in that] it develops the fullest interpretation of a sentence fragment at each moment of the fragment's unfolding" (p. 18). Other language scientists have proposed that prediction during language comprehension may be driven by processes that are also involved in language production (Dell & Chang, 2014; Federmeier, 2007). Specifically, readers and listeners are thought to covertly use their production system to anticipate upcoming input during comprehension. Predicting a word is often described as being akin to producing a word internally. Note that some prediction-by-production accounts differ with regard to the degree to which they assume that the production system is recruited. While Pickering and Garrod (2007, 2013) believe that language users rely on forward models (Wolpert, Doya, & Kawato, 2003) that are based on impoverished production representations, Huettig (2015) suggested that the fully-fledged production system is employed. Another possibility is that predictive language processing is driven by simple associations between words. On such an account, comprehenders may exploit the co-occurrence frequencies of words such that recognizing one word automatically activates/leads to retrieval of another word (Huettig, 2015; Kuperberg, 2007; Pickering & Garrod, 2013). Simple associative mechanisms may well be based on 'Hebbian learning' (cf. Kahnemann, 2011). Lastly, Huettig and colleagues (Mishra, Singh, Pandey, & Huettig, 2012; Mani & Huettig, 2014) provided experimental evidence for a link between the language users' reading levels and their degree of prediction. These authors argued that the exposure to written text, resulting in enhanced literacy, sharpens the language user's orthographic and lexical representations and increases the speed of lexical access during on-line speech processing (Mani & Huettig, 2014). Literacy might thus be regarded as a mediating factor in prediction.

Some experimental studies have directly tested the involvement of the three mechanisms and the mediating factor in anticipatory language processing. Metusalem et al. (2012) demonstrated that readers rely on generalized event knowledge when predicting words during discourse comprehension. Federmeier and colleagues (2002, 2010) reported that the language users' production abilities modulate prediction when reading and listening to simple sentences. Crucially, there is some evidence suggesting that more than one mechanism and/or mediating factor may influence prediction. Kukona, Fang, Aicher, Chen, and Magnuson (2011), for example, showed that event-based sentence context and simple associative priming influenced anticipatory processing in situations when language was used to refer to objects in the visual world. Using a similar setup, Rommers et al. (2015) recently showed that the comprehenders' production abilities and reading levels (i.e., literacy) had an impact on their degree of prediction. The latter two studies are consistent with recent theoretical proposals (Huettig, 2015; Kuperberg, 2007; Pickering & Garrod, 2013) which assume that multiple mechanisms and mediating factors contribute to anticipatory language processing.

Predictive language processing in different situational contexts

A topic that has not been examined by previous research concerns the contributions of the three mechanisms and the mediating factor literacy to prediction in different situations of language processing. In line with recent theoretical and experimental suggestions, I hypothesized that predictive language processing was driven by a combination of multiple underlying mechanisms and/or mediating factors. The main goal of the present thesis was, however, to empirically test how strongly event knowledge, association-based and production-based mechanisms, and literacy contribute to anticipation in different situations of language processing. Determining the nature and limits of the mechanisms underlying language prediction in various situational contexts is an important issue for understanding how adult language users cope with the highly variable environment, in which language processing takes place in our everyday life. Moreover, the knowledge gained from these investigations will inform theories of predictive language processing, which thus far do not take influences of the situational context into account.

The interplay between predictive language processing and the visual context

In addition to estimating the contributions of the mechanisms and the mediating factors underlying prediction, the present thesis explored situation-specific influences on anticipatory language processing—specifically the role of the visual context. Comprehension often involves the coordination of spoken language and visual input. One such example concerns understanding spoken language with reference to the visual environment. Previous studies have focused on the role the linguistic stimuli play in the integration of spoken input with co-present visual context (Kamide, 2008, for an overview). By contrast, this thesis explored the potential influence of the visual context on anticipatory language processing and its interaction with the mechanisms and mediating factors underlying prediction.

Prediction vs. production: The language user's task set

As indicated above, the production system may be involved in generating predictions. Some researchers have argued that predicting a word is akin to producing a word internally (e.g., Dell & Chang, 2014). However, few studies have directly compared production and comprehension processes. The present thesis contributes to filling this gap in the literature.

A related question pertains to the influence of the task set on prediction: Do comprehenders engage in prediction differently when their task set involves production in addition to comprehension (as is often the case in dialogue) compared to a task set that involves 'pure' comprehension?

Methodology

The experiments reported in this thesis used a variety of different methods. The majority of these methods are considered to capture on-line language comprehension (Mitchell, 2004, for discussion). Specifically, I employed eye-tracking, electroencephalography (EEG), as well as picture naming and self-paced reading. None of these methods required the participants to engage in meta-linguistic tasks.

The eye-tracking experiments featured different variants of the visual world paradigm (Huettig, Rommers, & Meyer, 2011, for review). In this paradigm participants listen to spoken utterances while looking at relevant visual scenes on a computer screen. Their eye movements are recorded by an eye-tracker for later analysis. A seminal study by Altmann and Kamide (1999) using the paradigm demonstrated that comprehenders predicted objects in the visual scene that would be referred to in spoken utterances. For instance, the participants viewed scenes featuring pictures of a boy, a cake, and some toys and heard predictable sentences such as "The boy will eat the cake" and non-predictable sentences such as "The boy will move the cake". The authors observed that participants tended to fixate the picture of the cake (the only edible object in the scene) before it was mentioned when they heard "eat" but not when they heard "move". These results are consistent with the notion that listeners pre-activated the concept 'cake' on hearing the spoken verb "eat" and that their anticipatory eye movements were guided by matches between information extracted from the spoken and the visual input. Chapters 2, 3, and 5 applied visual world eye-tracking and used the phenomenon of anticipatory eye movements to study further aspects of predictive language processing.

In two experiments, participants' EEG was recorded as they read sentences in rapid serial visual presentation. EEG represents fluctuations in electrical activity generated by the brain. As in most analyses, I averaged participants' EEG in the time domain of interest across multiple trials to create event-related potentials (ERPs), time-locked to a stimulus of interest. This was done based on the assumption that averaging cancels out uninformative signals (i.e., random noise) that are unrelated to the processing of the stimulus. An ERP component that has a strong link to language processing and is often discussed in relation to prediction is the N400. The N400 is a negative component that peaks at around 400 ms after onset of content words. The amplitude of the N400 component is considered a sensitive index of semantic processing (Kutas & Federmeier, 2000, 2011), with a more negative amplitude reflecting greater processing difficulty and a more positive amplitude reflecting processing ease, as caused, for example, by the predictability of the stimulus. Chapter 4 examined the contribution of simple word associations to prediction in discourse reading using the amplitude of the N400 to index the degree to which target words were pre-activated and thus easier or more difficult to process.

The impact of different task sets on anticipatory language processing was studied using a cross-modal picture naming paradigm and a self-paced reading paradigm (Chapter 6). Both methods have previously been used in research on prediction. As in a study by Griffin and Bock (1998; Gollan et al., 2011), the picture naming task involved comprehending the first part of a spoken sentence and naming a picture that was presented at the end of the recording to complete the sentence. This picture was either predictable or non-predictable from the preceding sentence context. Participants' picture naming latencies were measured. In the reading task, the participants read the same sentences in a self-paced fashion (cf. van Berkum et al., 2005; Experiment 3). I used the "moving window" technique (Mitchell, 2004, for discussion), where participants advanced to the next word by pushing the space bar. Their reading times for critical words were measured.

Thesis outline

The study by Altmann and Kamide (1999) showed that listeners could use information extracted from verbs to guide anticipatory eye movements to objects in the visual context that satisfied the selection restrictions of the verb. Based on empirical and theoretical suggestions in the literature, in Chapter 2, I investigated the influence of potential mechanisms (associations, prediction-by-production) and a meditating factor (literacy) in verb-based prediction. I conducted a visual world eye-tracking experiment in which participants listened to predictable ("The man peels an apple") and non-predictable ("The man draws an apple") sentences while looking at displays consisting of four objects. On predictable trials, only the target object (e.g., apple) fitted the selectional constraints of the verb and the verb-noun pairs varied in association strength. Moreover, I assessed participants' performance on the verbal fluency task (van der Elst, van Boxtel, van Breukelen, & Jolles, 2006), tapping their production fluency, and the Peabody vocabulary test (Dunn & Dunn, 1997) as a proxy for literacy. I also included Raven's progressive matrices test to separate participants' non-verbal intelligence from their production and literacy abilities. The key question was to which extent the associations and listener variables would affect language-mediated anticipatory eye movements. A positive relationship between the items' association strength and anticipatory looks to the target objects would be consistent with the notion that associations influence predictive processing. Similarly, positive correlations between participants' anticipatory looks to the target objects and their production fluency and Peabody vocabulary scores would support the notion that production abilities are important for prediction and that literacy mediates anticipatory eye gaze.

In a second experiment, I manipulated the time participants were given to view the visual displays prior to the spoken targets. This was done to investigate potential interactions between the mechanisms and mediating factors contributing to anticipatory eye movements and the preview period.

In Chapter 3, I further tested the hypothesis that visual information influences language-mediated anticipatory eye movements. To that end, a visual world experiment was conducted where the participants listened to predictable spoken sentences ("The man peels a banana") while looking at different types of visual displays. The target object (e.g., banana) was either present or it was replaced by a distractor that had a similar visual shape as the target object (e.g., canoe) or a distractor that was semantically related to the concept invoked by the target (e.g., monkey). Crucially, I manipulated the visual preview period before the target word was heard (short vs. long). If listeners use visual information to constrain anticipatory language processing, the eye gaze patterns after the onset of the verb should differ between the long and short preview conditions, with looks to the visually and semantically similar competitors on short but not long preview trials.

Chapter 4 assessed the contribution of simple word associations to prediction during discourse comprehension. EEG was recorded as participants read target words, which were preceded by associatively related words either appearing in a coherent event (Experiment 1) or in sentences that did not form a coherent event (Experiment 2). Previous research had found that contextually unexpected target words that were associatively related to the described event elicited a reduced N400 amplitude compared to contextually unexpected target words that were unrelated to the event. This finding was replicated (Experiment 1). Crucially, if associations contributed to prediction during discourse comprehension a similar pattern should be observed when the influence of event knowledge was minimized, as was the case in Experiment 2.

In Chapter 5, the similarity of prediction and preparing to speak was examined. Participants listened to a speaker solving mathematical equations while looking at a clock face featuring the numbers 1 to 12. On alternating trials, participants either heard a complete equation ("three plus eight is eleven"), or they heard the first part ("three plus eight is") and had to produce the result ("eleven") themselves. Participants were encouraged to look at the relevant numbers throughout the trial and their eye movements were recorded. After having carried out the computation, they could predict what the recorded speaker would say and they could initiate the word planning process for their own production of the result number. A strong prediction-by-production account would hypothesize that the cognitive processes involved in predicting the recorded speaker and in preparing to produce the result number should be very similar. This should reveal itself in very similar eye movements on the two trial types.

Chapter 6 investigated the effects of prediction in 'pure' comprehension tasks and in comprehension tasks containing an additional production component. Participants either listened to the first part of a sentence ("The man breaks a...") and provided the final word by naming aloud a picture whose name was predictable ("glass") or non-predictable from the sentence context, or they read the same sentences in a self-paced reading task. If the task set has an impact on anticipatory language processing, the two tasks should engage comprehenders in prediction to different degrees.

Chapter 7 summarizes and discusses the findings.

2 | Predictors of verb-mediated anticipatory eye movements in the visual world[1]

Abstract

Many studies have demonstrated that listeners can use information extracted from verbs to guide anticipatory eye movements to objects in the visual context that satisfy the selection restrictions of the verb. An important question is what underlies such verb-mediated anticipatory eye gaze. Based on empirical and theoretical suggestions in the literature, I investigated the influence of five potential predictors on this behavior: functional associations, general word associations, production fluency, literacy, and non-verbal IQ. In three eye-tracking experiments, participants looked at sets of four objects and listened to predictable sentences ("The man peels an apple") or non-predictable sentences ("The man draws an apple"). On predictable trials, only the target object fitted the selectional constraints of the verb. In Experiments 1 and 2 objects were presented before the verb was heard. In Experiment 3, participants were given only a short preview of the display after the verb had been heard. Functional associations and literacy skills were found to be important predictors of verb-mediated anticipatory eye gaze largely independent of the amount of contextual visual input. General word associations did not predict anticipatory eye movements, and non-verbal IQ was only a very weak predictor of anticipatory eye movements. Participants' production fluency correlated positively with the likelihood of anticipatory eye movements when participants were given the long but not when given the short visual display preview. These findings fit best with a pluralistic approach to predictive language processing in which multiple mechanisms, mediating factors, and the situational context dynamically interact.

[1]Adapted from *Hintz, F., Meyer, A.S., & Huettig, F. (in preparation). Predictors of verb-mediated anticipatory eye movements in the visual world.*

Introduction

Human communication is fast and efficient. This may at least partly be due to the fact that we can often predict words that are likely to come up next. Indeed, an impressive amount of experimental evidence has accumulated suggesting that readers and listeners can predict linguistic and non-linguistic information (e.g., Altmann & Kamide, 1999; Arai & Keller, 2012; Chen, Gibson, & Wolf, 2005; DeLong et al., 2005; Federmeier, McLennan, De Ochoa, & Kutas, 2002; Laszlo, Stites, & Federmeier, 2012; Rommers et al., 2013; Staub & Clifton Jr, 2006; van Berkum et al., 2005; Wicha, Moreno, & Kutas, 2003, 2004, and many others).

Spoken language is often used with reference to the visual environment of the language users. Many studies investigating prediction therefore have used the visual world paradigm, where participants integrate spoken linguistic input with co-present visual referents (see Huettig et al., 2011, for a review). A seminal study was conducted by Altmann and Kamide (1999). They presented listeners with semi-realistic scenes and spoken sentences which referred to the visual scenes. For instance, participants saw a scene depicting a boy, a cake, and some toys while hearing the sentence "The boy will move the cake" or "The boy will eat the cake". Altmann and Kamide observed that eye movements were directed to the cake, which was the only edible object in the scene, significantly earlier when the verb was "eat" than when it was "move". They interpreted these findings as evidence that information conveyed by a verb can be used to anticipate an upcoming theme. Many later studies have confirmed this conclusion. For instance, Mani and Huettig (2012) found that even 2-year-olds predict upcoming words that fit thematically with familiar verbs.

Although a number of theoretical proposals have been put forward (e.g., Altmann & Mirković, 2009; Dell & Chang, 2014; Federmeier, 2007; Huettig, 2015; Kamide, 2008; Kutas et al., 2011; Pickering & Garrod, 2007, 2013) we still know surprisingly little about the mechanisms and mediating factors that underlie such verb-mediated anticipatory eye gaze. However, this is an important question because the idea that prediction is a fundamental principle of human information processing has gained considerable ground over recent years (e.g., A. Clark, 2013; Friston, 2010).

The current work is a based on the hypothesis that prediction is not a unitary process, but rather engages a number of mechanisms in a flexible and context-dependent manner. I specifically focused on verb-mediated predictions and investigated the influence of five potential predictors of anticipatory eye gaze behavior: functional associations between the verb and the noun, general word associations between verb and noun, production fluency, literacy, and non-verbal IQ. Dutch participants took part in a visual world eye-tracking experiment. They heard predictable and non-predictable sentences (e.g., "The man peels/draws an apple") while looking at sets of four objects one of which was referred to in the spoken sentence. In Experiments 1 and 2, participants were given ample time to preview the objects prior to the critical spoken verb. In Experiment 3, the preview period was greatly reduced. Based on the previous literature, I predicted that participants would anticipate the target object (e.g., apple) before it was mentioned in the sentence. The key question was to which extent the associations and listener variables would affect anticipatory eye movements. In the remainder of this introduction, I motivate the choices of the predictor variables.

Functional associations

Functional knowledge is a particularly salient aspect of word meaning (e.g., Moss, McCormick, & Tyler, 1997). It is thus conceivable that functional associations have a strong influence on verb-mediated anticipatory eye gaze. Consistent with such a proposal, Moss et al. (1997) found a significant priming effect for functional properties of words in a lexical decision task. Moreover, there is evidence for a pivotal role of functional knowledge in verb-noun priming. McRae and colleagues (Ferretti, McRae, & Hatherell, 2001; McRae, Ferretti, & Amyote, 1997) developed rating paradigms to quantify verb-noun relationships (asking participants to answer questions such as "How common is it for an apple to be peeled?") and showed that animacy decision times were shorter for nouns that were primed by typical transitive verbs as compared to unrelated verbs. These findings indicate that verb-noun priming might predominantly rely on specific functional associations and that verb-noun typicality rating tasks can be used to tap these associations.

A visual world eye-tracking study by Kukona et al. (2011) suggested that functional verb-noun associations influenced verb-mediated anticipatory eye gaze even if they conflict with the event established by the sentential context. Participants listened to spoken sentences such as "Toby arrests the crook" while looking at displays showing pictures of five characters. Two of these characters were not related to the event described in the sentence. On all trials, a picture of Toby (a neutral character introduced prior to the experiment) was shown in the center of the screen. Toby served as the agent in all sentences. Crucially, the remaining two pictures featured agent and patient characters who were likely participants in the event described in the sentence (e.g., a crook and a policeman). Interestingly, after having heard "Toby arrests", participants made eye movements to both the picture of the crook and the picture of the policeman despite the fact that the spoken sentence had already unfolded beyond the subject/agent position and the agent role was taken up by Toby. Although looks to the crook and looks to the policeman differed in magnitude (the picture of the policeman was, however, fixated more than the pictures of the unrelated characters), the results suggest that on processing a transitive verb such as 'arrest', functionally associated concepts are pre-activated, irrespective of their thematic fit with the local sentence context.

Further support for a role of functional associations in anticipatory language processing comes from a study by Borovsky, Sweeney, Elman, and Fernald (2014), who tested how children between three and ten years of age and adults employed recently learned connections between agents, actions, and objects to anticipate upcoming words. They showed that adults and school-aged children learned the agent-action-object relationships and activated this knowledge in subsequent language processing. Importantly and in contrast to the older children and the adults, 3-4 year old children's anticipatory fixations indicated a processing strategy based on associations and a failure to integrate new combinatorial information within the sentence. The authors argued that (functional) associations exerted a substantial influence on anticipatory language processing early on in development. The influence of (functional) associations, they argue, may be down-weighted over the course of development in favor of combinatorial information. In the present study I examined whether functional associations predicted verb-mediated anticipatory eye gaze in adults.

General word associations

Functional associative knowledge is not the only type of associative knowledge that connects verbs and nouns. Another possibility is that general word associations underlie verb-mediated anticipatory eye movements. General word associations are typically operationalized using free word association tasks where participants process an auditory or visual cue word and are asked to say or write down one or more words which come to their minds on processing that cue (e.g., De Deyne, Navarro, & Storms, 2012; Nelson, McEvoy, & Schreiber, 2004). General word associations may include or be influenced by functional associations (e.g., on processing the verb "peel" peel-able objects may be retrieved such as apples, bananas and/or oranges). However, it is typically assumed that free word association tasks are sensitive to a number of different types of associations (e.g., semantic, phonological, orthographic, visual, etc.). For example, participants in free word association tasks often retrieve words that are associated phonologically to the cues (e.g., saying "heel" given the cue "peel").

How likely are general word associations to influence verb-mediated anticipatory eye gaze? Several visual world studies have investigated the effects of (noun-noun) semantic relationships on listeners' language-mediated eye gaze. Yee and Sedivy (2006) presented participants with visual displays that included semantically related pairs of objects (e.g., 'lock' and 'key') and other (unrelated) objects. On hearing the word "key", participants fixated the picture of the semantic competitor 'lock' reliably more than the semantically unrelated distractors in the scene. As 'lock' and 'key' are strongly associatively related according to free word association tasks, the results suggest that associative relationships mediated eye gaze in this study. Moreover, Duñabeitia, Aviles, Afonso, Scheepers, and Carreiras (2009) found that associative relationships predicted eye gaze in a similar visual-world study. Their participants, however, tended to fixate more and earlier on depicted objects that were (free) associates of abstract words than (free) associates of concrete words. This result suggests that associated concepts are more readily retrieved for abstract than for concrete words. To give a final example, Iordanescu, Grabowecky, and Suzuki (2011; see also Iordanescu, Guzman-Martinez, Grabowecky, & Suzuki,

2008) reported evidence for the facilitating effects of sound-vision mappings in a visual search task: Hearing the characteristic sound "meow" resulted in faster location of the picture of the associated animal cat, as compared to an unassociated sound. This result is consistent with the notion that simple (non-functional) associations may influence language-mediated eye gaze. It is worth mentioning that Huettig and Altmann (2005; see also Huettig, Quinlan, McDonald, & Altmann, 2006; Yee, Overton, & Thompson-Schill, 2009; for similar results) demonstrated that eye gaze can also be directed immediately, as a word unfolds, towards conceptually related but non-associated objects. This suggests that looks to the lock on hearing "key" are not solely driven by associative relationships.

In short, there is some support for the view that word-object mapping in the visual world paradigm is influenced by general associative relationships. However, the influence of such associations in verb-mediated predictions remains to be established. Hearing a verb such as "peel" may lead to automatic retrieval of the associated concept 'apple' and thus provide an associative mechanism for predictive language processing. I examined the extent to which general word associations predict verb-mediated anticipatory eye movements in the visual world paradigm.

Production fluency

Another possibility that is frequently raised is that predictive language processing may be supported by language production mechanisms. According to this proposal, the language production system is covertly used to anticipate upcoming language input during comprehension (Chang et al., 2006; Dell & Chang, 2014; Pickering & Garrod, 2007, 2013; Schiller, Horemans, Ganushchak, & Koester, 2009).

Chang, Dell, and Bock's (2006; see also Dell & Chang, 2014) dual path model is an explicit implementation of such a production-based prediction account. The authors claim that predicting the next word is akin to producing the next word during sentence production. Moreover, they argue that prediction is central to language acquisition. A core component of their model is an error-based learning mechanism. Learning occurs when the model's production-based predictions are compared against utterances by others and a deviation between the predicted and the actual input is discovered. Pickering and Garrod (2013) argued that language users construct forward models (cf. Wolpert et al., 2003) both to predict their own utterances and to predict the utterances of others. More precisely, when preparing to speak speakers are assumed to construct efference copies (i.e., impoverished representations) of their intended utterances and compare these copies to the output of the production implementer. Similarly, listeners are assumed to use forward production models to covertly imitate the speaker and predict the speaker's upcoming utterances. Thus, listeners generate predictions on the basis of what they would say next if they were the speaker.

To date, there is only indirect evidence supporting the contribution of production-based mechanisms to prediction. In the study by Mani and Huettig (2012)with two-year olds mentioned above, the children listened to predictable and non-predictable sentences such as "The boy eats a big cake" or "The boy sees a big cake" while looking at a display showing a cake and a bird. The toddlers showed anticipatory eye gaze to the cake before it was mentioned in the speech. More importantly, children's predictive eye gaze correlated positively with the size of their production vocabulary size.

Electrophysiological evidence from studies conducted by Federmeier and colleagues is also consistent with the involvement of the production system in prediction. Federmeier, McLennan, De Ochoa, and Kutas (2002; see also Federmeier, Kutas, & Schul, 2010) observed significant correlations between the amplitude of participants' ERP components that may be interpreted as indexing prediction and their production fluency as measured in the verbal fluency task. In this task, participants have to produce as many members of a given semantic category (e.g., animals or professions) or as many words beginning with a particular letter as possible within one minute. The number of words produced can be seen as an indicator of the participant's ability to quickly retrieve and produce words (for discussion see Luo et al., 2010; Shao et al., 2014). In a visual world study, Rommers et al. (2015) observed that stronger anticipatory bias to the target object was associated with higher verbal fluency scores. In the present study I correlated participants' verbal fluency with their verb-mediated anticipatory eye gaze.

Literacy

Recent studies suggest a substantial influence of literacy on language-mediated anticipatory eye gaze. Mishra et al. (2012) conducted a visual world prediction study with persons with low and higher degrees of literacy in India. They found that the time course of the gazes to the targets differed: High literates started to look more at the target object than at unrelated distractors immediately on hearing a semantically and syntactically biasing adjective. By contrast, the eye gaze of low literates to the targets only started to differ from looks to the unrelated distractors once the spoken target word acoustically unfolded. Mishra et al. concluded that low literates did not use the information from the unfolding spoken words to predict upcoming referents. Mani and Huettig (2014) provided converging evidence from 8-year-old children at the cusp of literacy acquisition. They found that the children's reading ability correlated with their anticipatory eye movements. There was no correlation between their language abilities (measured by a standard naming task in the preferential looking paradigm and a syllable detection task) and their reading ability. Huettig and Brouwer (2015) found further evidence for an effect of literacy on anticipatory eye gaze. They tested Dutch adults with dyslexia and a control group of adults without a history of reading disorders. Adults with dyslexia anticipated target objects in the visual world study later than controls. Similarly, James and Watson (2013) showed that literacy (as measured in the Comparative Reading Habits questionnaire, Acheson, Wells, & MacDonald, 2008, and the American Adult Reading Test, Blair & Spreen, 1989) was linked to anticipatory language-mediated eye movements among American college students.

A good proxy for literacy is vocabulary knowledge. Borovsky, Elman, and Fernald (2012) found that children aged 3 to 10 with relatively high vocabulary knowledge were faster to anticipate target words than children with lower vocabulary knowledge. Rommers et al. (2015) used the Peabody Picture vocabulary test (Dunn & Dunn, 1997) to assess participants' literacy skills. They also observed that greater anticipatory bias to the targets was associated with large vocabulary scores, thus providing further evidence for a link between literacy and predictive language processing. Here I tested the extent to which vocabulary knowledge predicted verb-mediated anticipatory eye gaze in students.

Non-verbal IQ

Finally, I examined the influence of non-verbal intelligence on verb-mediated anticipatory eye gaze. The reason is that there is considerable psychometric evidence for the so-called 'general-factor' underlying much of cognitive processing. The g-factor is a psychometric construct which assumes that performance of individuals at any one type of cognitive task highly predicts performance at other cognitive tasks. I correlated Raven's Advanced Progressive Matrices

(Raven, Raven, & Court, 1998), as a measure of non-verbal intelligence, with participants' verb-mediated anticipatory eye movements.

In sum, the purpose of the current study was to examine five potential predictors of verb-mediated anticipatory eye gaze: functional associations, general word associations, production fluency, literacy, and non-verbal IQ. The predictable verb-noun pairs varied in the strength of functional associations (as determined in a verb-noun typicality rating task) and general word associations (as determined in a free word association task). The influence of production abilities was assessed by using the verbal fluency task, and as a proxy for participants' literacy I administered the Peabody vocabulary test. Finally, to assess any influence of non-verbal intelligence on anticipation behavior, participants were asked to complete Raven's Advanced Progressive Matrices.

Experiment 1

Method

Participants

Sixty-one members (mean age = 22, *SD* = 3) of the participant panel of the MPI for Psycholinguistics, all native speakers of Dutch, were paid for their participation. All participants had normal hearing and normal or corrected-to-normal vision. All participants gave written consent beforehand. The study was approved by the ethics board of the Faculty of Social Sciences of the Radboud University Nijmegen.

Stimuli

The eye-tracking experiment consisted of 40 predictable and 40 non-predictable sentences. On each trial, the participants heard a sentence and saw a visual display consisting of four objects. On predictable trials, one of the four objects was predictable (e.g., “De man schilt op dit moment een appel”, the man peels at this moment an apple); the remaining three objects were unrelated distractors (e.g., a candle, a radio, an owl; Figure 2.1, for an example). On non-predictable trials, participants saw the same four objects as in the predictable condition but neither the target object nor any of the distractors could be anticipated from the spoken sentential context (e.g., “De man tekent op dit moment een appel”, the man draws at this moment an apple). All sentences had the same structure and had the same number of words: The subject position was filled by “the man”, and the adverbial “at this moment” separated verb and object. Using this padding between verb and target, enough time for participants to make anticipatory eye movements was provided. The resulting sentence construction is deemed quite natural by native speakers of Dutch. The predictable verb-noun pairs varied in general association strength. In the non-predictable items, the general association strength was zero. General association strength was operationalized using a free association task for a larger set of verbs.

Free verb-noun association pre-test

One hundred fifty-nine Dutch transitive verbs were selected from the CELEX database (Baayen, Piepenbrock, & Gulikers, 1993) and evenly distributed across three lists. One hundred and five native speakers of Dutch (mean age = 29, *SD* = 14, none of whom participated in any of the other rating studies or the main experiments) were randomly assigned one of the three lists and carried out an adapted version of the free association task (cf. Nelson et al., 2004). The participants were asked to read the verbs (one at a time) and write down the first three nouns that came to mind. The order of verbs on all lists was randomized for each participant. The experiment was conducted on-line using the WebExp package (Keller, Gunasekharan, Mayo, & Corley, 2009). I used a continued (up to three words) rather than the single-word association task as this has been shown to yield more reliable estimates of the associative relationship between lexical concepts (e.g., De Deyne et al., 2012).

Each verb was read by at least 35 participants. For a given verb, the number of occurrences of a particular noun was counted and divided by the number of participants who had read that verb. This proportion served as the measure of general verb-noun association strength. There were a total of 410 missing values (2.5% of the data) resulting from cells where participants had not provided any answers at all or no nouns. Forty predictable verb-noun pairs (e.g., 'schillen-appel', 'peel-apple'; see Appendix, for all experimental verb-noun pairs) were selected. The mean general association strength in those pairs was .37, ranging from .09 to .77. The same nouns as in the predictable items were paired with neutral, non-predictable transitive verbs (e.g., 'tekenen-appel', 'draw-apple'). The general association strength between verbs and nouns in those pairs was zero.

Cloze probability pre-test

To be sure that predictable and non-predictable verb-noun pairs were classified properly, I pre-tested the sentences for cloze probability (on-line, using the WebExp package). Thirty-five Dutch native speakers (mean age = 21 years, *SD* = 2), none of whom participated in other rating studies or the main experiments, were presented with the sentence fragments up to the object position (e.g., "The man verbs at this moment a...") and were asked to fill in the final word which would best complete the sentence. The cloze probability was the proportion of participants who chose to complete the sentence fragment with the target word selected based on the pre-test. On predictable items, the mean cloze probability for the target nouns was .39 (*SD* = .24; ranging from

.06 to .8), in the non-predictable items, it was zero. These results suggest that the target nouns could be predicted in the predictable but not in the non-predictable sentence contexts.

Plausibility rating study

To ensure that both predictable and non-predictable sentences depicted events that are likely to happen in real life, a pen-and-paper plausibility rating was carried out. Two lists were generated each containing 20 predictable and 20 non-predictable sentences chosen at random. None of the target nouns appeared twice in the same list. Twenty implausible sentences, which had the same structure as the predictable and non-predictable sentences were added to each list. The 60 sentences were presented in random order. Twenty-four participants (visiting students from the University of Groningen, The Netherlands), twelve per list, none of whom participated in other rating studies or the main experiments, were asked to estimate on a 1-10 scale how plausible the events described in the sentences were. The mean plausibility judgment for predictable sentences was 9.05 (SD = .56); the mean plausibility judgment for non-predictable sentences was 7.69 (SD = .95). Pairwise comparisons revealed a significant difference between the two types of sentences ($t(39) = 7.8$, $p < .001$). It is worth mentioning that although the plausibility ratings differed statistically between the predictable and the non-predictable conditions, the latter items were overall still deemed very plausible. In fact, it is conceivable that differences in plausibility are a function of predictability such that predictable events are naturally rated more plausible.

Word length and frequency

The mean number of letters of the inflected verb in the predictable sentences was 6 (SD = 2). The mean word frequency (Subtlex; Keuleers, Brysbaert, &

New, 2010) was 25 per million (SD = 78.2). In the non-predictable sentences, the mean number of letters of the inflected verbs was 7 (SD = 2), and the mean word frequency was 46.8 per million (SD = 122.6). Pairwise comparisons revealed that the verbs in the predictable and the non-predictable conditions did not differ in number of letters ($t(39)$ = -1.122, p = .269) or frequency ($t(39)$ = -.921, p = .363). The mean word frequency of the object nouns was 73.3 (SD = 138.13).

Sentence recordings

All 80 sentences were spoken with neutral intonation at a normal pace by a female native speaker of Dutch. Recordings were made in a sound-damped booth, sampling at 44 kHz (mono, 16 bit sampling resolution) and stored directly on computer. The mean sentence duration was 2800 ms (SD = 214). Onsets and offsets of all words were marked using Praat (Boersma, 2002).

Display composition

To create the visual displays, 40 sets of four objects were composed each consisting of one target object and three distractor pictures which were unrelated to the other pictures in shape, semantics, and phonology of their names (see Figure 2.1, for an example). The pictures were selected from the database provided by Snodgrass and Vanderwart (1980) and colored in, or were drawn by an artist. I conducted two rating studies to assess the semantic and visual similarity between the concepts invoked by the target nouns and the distractor objects. The rating studies were necessary to ensure that the distractors were unrelated to the targets, as semantic and/or visual similarity can affect the participants' gaze pattern (cf. Huettig & Altmann, 2005, 2007).

Visual and semantic similarity rating studies

Twelve participants (mean age = 23, *SD* = 4) provided semantic similarity ratings and twelve others (mean age = 22, *SD* = 2) provided visual similarity ratings. None of these participants took part in the main experiments. The study was conducted over the internet using the WebExp package. In both rating studies, participants read 40 target nouns. Each target was paired with the four objects which were part of that item. For example, participants would read the word apple in the upper right corner of the screen and saw the four objects displayed in Figure 2.1 at the bottom, next to each other. The order was randomized. In the visual similarity rating study, participants were asked to judge how similar the typical visual shape of the concept denoted in the printed word was to the physical shape of the referents of the depicted objects, ignoring any similarity in meaning. In the semantic similarity rating, participants were asked to judge meaning similarity while ignoring shape similarity. A rating scale ranging from 0 (no similarity) to 10 (identical) was used in both tasks. As the object referred to by the written word was among the four pictures, a measure of how well the object name fitted its visual representation was also obtained. The results of the visual similarity rating confirmed that the target objects depicted the concepts invoked by the written words (mean target object rating = 9.91, *SD* = 0.3). The semantic similarity rating confirmed that the target objects matched the semantic representations invoked by the written words (mean target rating = 9.83, *SD* = 0.41). The mean distractor score in the visual similarity rating was 0.52 (*SD* = 0.56); in the semantic similarity rating it was 0.55 (*SD* = 0.4). The rating scores of the three distractors in one set were averaged. Pairwise comparisons between target object ratings and averaged distractor ratings confirmed that the distractors were judged to be dissimilar to the target concept at visual ($t(39) = 139.6$, p < .001) and semantic ($t(39) = 121$, p < .001) levels of representation.

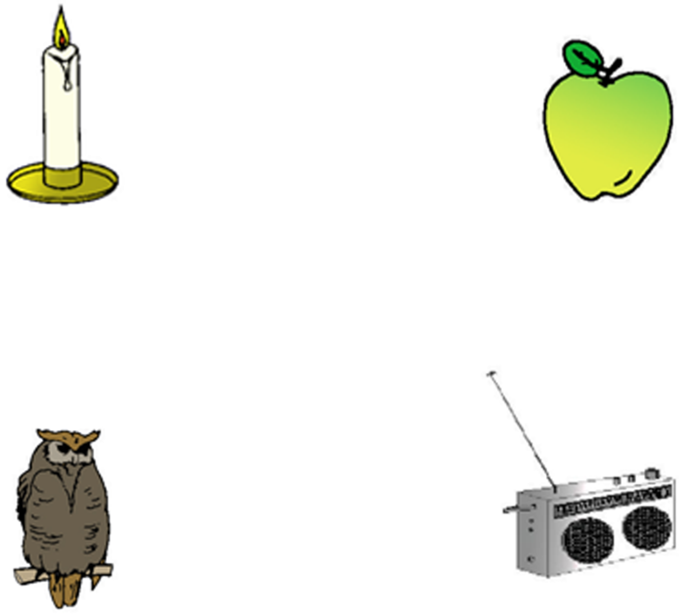


Figure 2.1: Example display for the target object apple presented in the predictable condition with the sentence “De man schilt op dit moment een appel” (the man peels at that moment an apple) and in the non-predictable condition with the sentence “De man tekent op dit moment een appel” (the man draws at that moment an apple). In both conditions, the target object was shown along with three distractor objects (a candle, an owl, a radio).

Procedure

Eye-tracking experiment

The test session started with the eye-tracking experiment. The 40 predictable and the 40 non-predictable items were evenly distributed across two lists. None of the target nouns appeared twice on one list. Participants were randomly assigned one list and were seated in a sound-shielded booth. Eye movements were tracked using an EyeLink 1000 remote desktop tracker sampling at 500 Hz. A sticker was placed on participants' foreheads to monitor the distance between their heads and the tracker. The distance was held constant between 55 and 60 cm. The eye-tracker was calibrated, and participants were instructed to listen to the sentences carefully and to not move their eyes away from the screen. I used a look-and-listen task (Huettig et al., 2011, for discussion), that is, the participants did not receive a specific viewing instruction. The spoken sentences were presented through loud speakers. A trial was structured as follows: First, a central fixation dot appeared in the center of the screen for two seconds. The dot disappeared and the playback of the sentence started. The onset of the display was timed to one second prior to the occurrence of the verb in the speech signal. The four objects remained on the screen for the rest of the trial. The positions of the pictures were randomized across four fixed positions of a (virtual) 2 x 2 grid. The time between the onset of the verb and the onset of the target noun was on average 1480 ms. Each participant was presented with all 40 trials of one list. The order of trials was randomized automatically before the experiment. The eye-tracking experiment, including calibration, took approximately 10 minutes. The data from participants' left or right eye (depending on the quality of the calibration) were analyzed in terms of fixations, saccades, and blinks, using the algorithm provided in the EyeLink software. Fixations were coded as directed to the target, to one of the three unrelated distractors, or elsewhere.

Production fluency task

Participants carried out a digitized version of the Dutch verbal fluency task (cf. van der Elst et al., 2006). They were given two categories ('animals' and 'professions') and two letters ('p' and 'm') and were instructed to produce as many words as possible belonging to the given category or beginning with the given letter within one minute. The category name or the letter was shown for three seconds before the screen went blank. Participants' answers were recorded. Incorrect words and repetitions were excluded.

I calculated the average number of words (across all four categories; cf. Federmeier et al. 2002, 2010) an individual produced within one minute. The data of three participants had to be excluded because they misunderstood the task.

Functional associations

I followed Ferretti et al. (2001) and used a typicality rating task to operationalize functional associations. Each participant was presented with the same predictable and non-predictable verb-noun pairs s/he was presented with in the eye-tracking experiment. Following the rating procedure suggested by Ferretti and colleagues, the pairs were embedded in the question "How common is it for a noun to be verbed?" (e.g., "How common is it for an apple to be peeled?", Dutch translation, "Hoe waarschijnlijk is het voor een appel om geschild te worden?"). The participants were instructed to rate on a 1-10 scale how typical it was for the target and the three distractors used in the eye-tracking experiment to undergo the action implied by the verb. The items were presented to the participants in an Excel sheet. They typed the respective rating score next to each noun. The nouns appeared in a random order. Predictable and non-predictable items were presented in random order as well. Log-transformed ratios between the rating for the target and the average rating for the three distractors were calculated for each item. A ratio of zero meant that target and unrelated distractors were rated to be equally typical; a ratio greater than zero indicated a bias towards the target. The mean ratio for the predictable items was .71 (SD = .23), ranging from .21 to 1. The mean ratio for non-predictable items was .08 (SD = .13) ranging from -.21 to .44.

Raven's progressive matrices

I used a computerized version of Raven's Advanced Progressive Matrices test (Raven et al., 1998) to assess participants' non-verbal intelligence. 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.

Peabody vocabulary test

Participants' receptive vocabulary size was assessed using a digitized version of the Dutch Peabody picture vocabulary test (Dunn & Dunn, 1997; Dutch translation by Schlichting, 2005). 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 the number (1, 2, 3, or 4). A percentile score was calculated based on Dutch norms.

Results

The results of all experiments were analyzed using a magnitude estimation approach. This was motivated by a recent proposal (Cumming, 2014; see Huettig & Janse, 2015, which advocates turning away from null-hypothesis testing towards interpreting results by using measures of effect sizes and confidence intervals. As has been shown empirically (cf. Fidler & Loftus, 2009), this leads to a better interpretation of the results compared to a research report based on null hypothesis testing (see Cumming, 2012, 2014, for extensive discussion).

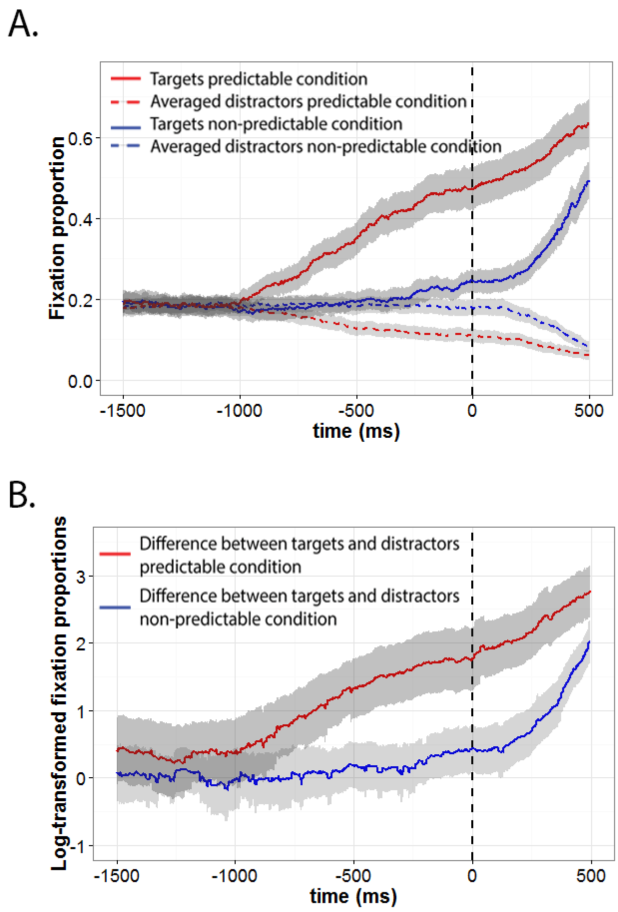


Figure 2.2: Panel A plots the fixation proportions for target and averaged distractor objects in the predictable and the non-predictable condition. Panel B plots the difference between log-transformed fixation proportions for target and distractor objects for predictable and non-predictable conditions in Experiment 1. Confidence intervals (95%), calculated for each sampling step, are shaded in gray. The critical window spanned the time between the acoustic onset of the verb and the acoustic onset of the target word in the speech signal (mean = 1.5 s). Time zero (vertical dashed line) indicates the onset of the target word.

I plotted participants' eye movements for the period between the acoustic onset of the verb and the acoustic onset of the target word (time zero) plus 500 ms. Figure 2.2 presents the fixation data of Experiment 1 in two ways: Panel A displays fixation proportions to the target object (solid lines) and to the averaged distractor objects (dashed lines) for the predictable (red) and the non-predictable (blue) conditions. I computed by-participant confidence intervals (95%) for each line at every sampling step (2 ms). The area between the lower and the upper bounds is shaded in gray. In Panel B, I log-transformed the fixation proportions and subtracted fixations to the three distractor objects from fixations to the target objects in the predictable (red line) and non-predictable (blue line) condition (cf. Arai, van Gompel, & Scheepers, 2007). A difference of zero means that target and averaged distractors were fixated equally often, and a difference greater than zero means that more fixations were made to the target object. By-participant confidence intervals were calculated for each sampling step, based on the mean of the difference between target and distractors.

Both panels show that participants' likelihood of fixating the target objects in the predictable condition increased already one second before the target word occurred in the speech signal. This suggests that participants anticipated the upcoming targets very early, shortly after the verb had been heard. However, the same target objects referred to in the non-predictable condition only attracted increased overt visual attention after the onset of the spoken target noun. The confidence intervals indicate that the by-participant variance was smaller in the non-predictable and larger in the predictable condition, which suggests variation in participants' tendency to predict. In the following analyses, I made use of this variation.

I carried out multiple linear regression analyses by participants and by items to estimate the contribution of general word and functional associations, production fluency, literacy, and non-verbal intelligence to the anticipatory eye movements observed in Experiment 1. Predictor variables were entered simultaneously into the regression equation. To calculate the dependent variables, I divided each participant's proportion of looks to the target during the onset-verb-onset-target period by that participant's proportion of looks to the averaged distractors during that time window. This was done for all experimental trials. The resulting values were log-transformed. Fixation proportions that were 0 or 1 were replaced with 0.001 and 0.99, respectively (cf. Macmillan & Creelman, 1991), prior to the division and log-transformation. The data were aggregated by participant and by item yielding average scores for each participant and for each item, for the predictable condition and for the non-predictable condition.

By-participant regression analysis

The by-participant regression analysis used the participants' predictive eye gaze scores (i.e., average log-transformed scores in the predictable condition as the dependent variable. The participants' verbal fluency scores, their Peabody vocabulary scores, and their Raven's non-verbal intelligence scores were simultaneously entered as predictors into the regression analysis. This model, with an R^2 of .203 showed the following independent contributions to participants' predictive eye gaze (see Table 2.1, for an overview of the standardized betas, the effect size measure of all five predictor variables in all experiments): verbal fluency (unstandardized ß = .057, SEß = .026, CI [.005, .110]; standardized ß = .269; see Figure 2.3, for scatter plots), Peabody vocabulary scores (unstandardized ß = .009, SEß = .005, CI [-.001, .019]; standardized ß = .238), Raven's progressive (unstandardized ß = .024, SEß = .016, CI [-.009, .057]; standardized ß = .199).

By-item regression analyses

Predictive eye gaze (the dependent variable) in the by-item analysis was operationalized as the difference between looks to an item in the predictable condition and looks to the same item in the non-predictable condition. This difference score was calculated to minimize any influence of the visual display which was constant across predictable and non-predictable conditions. An item's difference score, along with the measures of general and functional associations were entered into the regression analysis. The model with an R^2 of .219 revealed the following individual contributions to explaining the variance within the likelihood of predictive looks made to an item: functional associations (unstandardized *ß* = 1.381, SE*ß* = .436, CI [.498, 2.265]; standardized *ß* = .475; Figure 2.3, for scatter plots), general word associations (unstandardized *ß* = -.791, SE*ß* = .577, CI [-1.96, .378]; standardized *ß* = -.206).

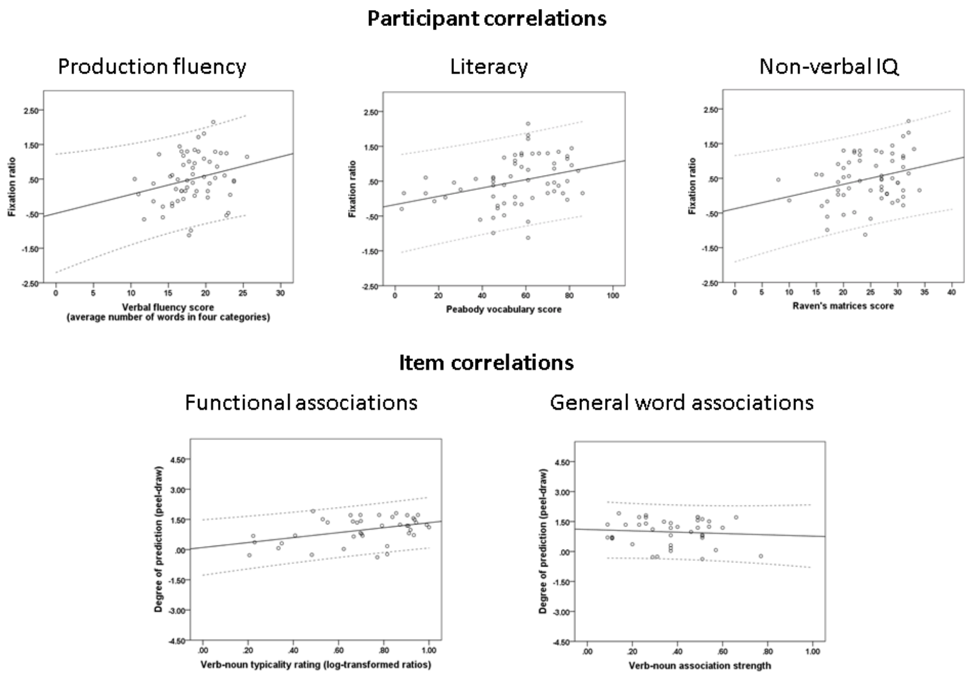


Figure 2.3: Scatter plots showing participant and item correlations in Experiment 1. By participants, the likelihood of their anticipatory eye movements to the target object during the predictive window (verb onset-target word onset) was correlated with their production fluency (verbal fluency tasks: 'animals', 'professions', letter 'm', letter 'p'; average number of words named in one minute), their Peabody vocabulary scores and their non-verbal IQ scores. By items, the typicality ratio and the association strength of a particular verb-noun pair was correlated with the difference between looks to that target object in the predictive and in the non-predictive condition. The dashed lines display the confidence intervals (95%) for the correlation coefficients. Note that the scale of the ordinate in the participant correlation scatter plot is different from the scale of the ordinate in the item correlation scatter plots. This is due to the difference in dependent variable calculation.

Discussion

I found evidence for robust and early effects of target noun anticipation given predictable verbs, replicating earlier research. The same nouns however, were not anticipated following non-predictable verbs. The effect size measure (standardized betas) suggests that in the by-participant analyses literacy and production fluency each accounted for large amounts of unique variance. The results of the by-item regression analyses showed that functional associations contributed the largest amount of unique variance to explaining variability in predictive eye gaze.

The substantial influence of functional associations on anticipatory language processing suggests that the more typical a target noun was rated to undergo the action implied by the verb the higher the likelihood of predictive looks made to that item. General word associations did not predict anticipatory eye movements. The modulating influence of participants' production fluency scores on their predictive eye gaze is consistent with previous evidence that participants' language production abilities are important for predictive language processing (Federmeier et al., 2002, 2010; Mani & Huettig, 2012; cf. Dell & Chang, 2014; Pickering & Garrod, 2007, 2013). The positive relationship between participants' production fluency scores and their predictive gaze revealed that participants who produced more words within one minute showed more verb-mediated anticipatory eye gaze. The modulating influence of participants' receptive vocabulary knowledge is consistent with the notion of literacy as a mediating factor in predictive language processing.

Verb-noun typicality ratings from naive participants

One could argue that functional associations predicted the likelihood of predictive looks made to an item so well because the ratings were given by the same

Table 2.1: The table gives an overview of the unique variance explained by each of the five predictor variables in Experiment 1, 2 and 3, and the combined data of Experiment 1 and 2. Standardized betas from by-participant and by-item analyses are provided as a measure of effect size.

	Experiment 1 long preview	**Experiment 2** long preview	**Experiment 1 and 2 combined** long preview	**Experiment 3** short preview
By-participant regression				
Non-verbal IQ (Raven's)	ß = .199	ß = -.065	ß = .06	ß = .092
Literacy (Peabody)	ß = .238	ß = .226	ß = .242	ß = .277
Production ability (Verbal fluency task)	ß = .269	ß = .194	ß = .213	ß = -.063
By-item regression				
Functional associations (verb-noun typicality rating)	ß = .475	ß = .341	ß = .498	ß = .347
General associations (free verb-noun association)	ß = -.206	ß = -.03	ß = -.143	ß = .012

sample of participants who had previously participated in the visual world experiment. Thus, participants may have been more familiar with the materials and biased as compared to a completely naive sample of volunteers. To rule out that the measure of functional associations collected from the participants in Experiment 1 were confounded by potential memory effects from having participated in the visual world experiment before, an independent sample of 20 native speakers of Dutch (mean age = 21, *SD* = 2) was asked to carry out the typicality ratings in the same way as described above. The mean verb-noun typicality ratio for the predictable items was .65 (*SD* = .21), ranging from .14 to .98. The mean ratio for non-predictable items was .06 (*SD* = .13) ranging from -.20 to .38. I then re-ran the by-items regression analysis including the measure of general associations and the newly collected measure of functional associations (R^2 = .246). As before general word associations did not explain much of the variance in predictive eye gaze (unstandardized *ß* = -.691, SE*ß* = .559, CI [-1.823, .442]; standardized *ß* = -.18). The influence of the newly collected verb-noun typicality ratings on the dependent variable was very similar to the influence of the verb-noun typicality ratings stemming from the participants who had taken part in the eye-tracking before (independent sample: unstandardized *ß* = 1.541, SE*ß* = .451, CI [.628, 2.454]; standardized *ß* = .497 vs. eye-tracking participants: unstandardized *ß* = 1.381, SE*ß* = .436, CI [.498, 2.265]; standardized *ß* = .475). This suggested that the previous ratings were, if at all, only weakly influenced by memory effects or familiarity with the materials resulting from having participated in the visual world experiment before. In the following by-item analyses I thus used the verb-noun typicality ratings obtained from the participants of Experiment 1.

Free noun-verb association task

Why was there no influence of general word associations on verb-mediated eye gaze? As in most previous eye-tracking experiments, participants were given some preview of the visual scene. During that time, they could retrieve information about the four objects and activate associated verbal knowledge. Thus, the visual preview may have provided a head start for the pre-activation of conceptual and lexical information associated with the objects and may have overridden the effects of general verb-noun associations. To assess this hypothesis, I carried out another free association task on the target nouns used in the eye-tracking experiment. If general noun-verb rather than the verb-noun associative relationships were crucial for the eye-tracking results, I should observe a positive relationship between the general noun-verb association strength of an item and the likelihood of anticipatory eye movements made to that item. Forty-four Dutch participants who had not taken part in any of the rating studies or the main experiment were asked to read the 40 target nouns and note down the first three verbs that came to mind. The parameters and the analysis were identical to the free verb-noun association task. The mean general noun-verb association strength in the predictable condition was .29 (*SD* = .3) ranging from 0 to .98; in the non-predictable condition, it was .011 (*SD* = .04). I entered the measures of general noun-verb associations, general verb-noun associations and functional associations into a by-item regression analysis (R^2 = .266). Similar to general verb-noun associations in the previous analysis, general noun-verb associations did not explain much unique variance of the anticipatory eye gaze (unstandardized *ß* = -1.047, SE*ß* = 1.06, CI [-3.197, 1.103]; standardized *ß* = -.16). The contributions of the other variables were similar to the previous analysis (functional associations: unstandardized *ß* = 1.773, SE*ß* = .509, CI [.742, 2.805]; standardized *ß* = .572; general verb-noun associations: unstandardized ß = -.683, SEß = .559, CI [1.817, .451]; standardized ß = -.178).

In sum, Experiment 1 demonstrated that production fluency, literacy and functional associations were important predictors of verb-mediated anticipatory eye movements.

Experiment 2

The aim of Experiment 2 was to replicate the results of Experiment 1. This decision was motivated by two considerations: First, I aimed to assess the reliability of the observed regression effects in order not to capitalize on spurious observations. Second, I wanted to estimate the contribution of object-verb priming to anticipatory eye movements in a more direct way. Recent evidence suggests that visual objects exert a substantial influence on the recognition of auditory linguistic input. McQueen and Huettig (2014) showed that lexical decision times to spoken target nouns were faster when the words were preceded by related picture primes (e.g., arrow-"sword"). The nature of the trials in the eye-tracking experiment in the current study was similar to the nature of the cross-modal priming trials in McQueen and Huettig's study. Thus one could explore the possibility of object-verb priming in the trials of the eye-tracking experiment and its potential influences on anticipatory eye movements.

Experiment 2 was run in two sessions. In session 1, participants took part in the same eye-tracking experiment as in Experiment 1 and carried out the Peabody vocabulary test and the verbal fluency task. In session 2, two weeks later, I asked the same participants to carry out the Raven's non-verbal intelligence test and a cross-modal object-verb priming experiment in which the same materials as in the eye-tracking experiment were used. The primes consisted of the target objects used in the eye-tracking experiment (e.g., apple), and the targets were the verbs presented auditorily just as in the eye-tracking experiment (e.g., "peel"). As the trials in the eye-tracking experiment and in the priming experiment were comparable, I could examine the effects of object-verb priming systematically. Based on the results of the free noun-verb association task, I predicted faster reaction times in a lexical decision task pertaining to predictable rather than to non-predictable items. If priming of that kind has an influence on participants' anticipatory eye movements, their priming effects should correlate positively with their likelihood of anticipatory fixations in the eye-tracking experiment. Similarly, the strength of the priming effect for an item in the cross-modal priming experiment should contribute to explaining variance in the by-item regression analyses of predictive looks to the target objects in the eye-tracking experiment.

Method

Participants

Sixty-one native speakers of Dutch (mean age = 21 years of life, *SD* = 3), took part in Experiment 2. All participants were right-handed, had normal hearing, and normal or corrected-to-normal vision. None of them had taken part in any of the norming studies or Experiment 1.

Stimuli and procedure

Session 1

In Session 1, participants carried out the same eye-tracking experiment as in Experiment 1. Because the eye-tracker used in the first experiment was unavailable, a tower mounted EyeLink 1000 tracker was used. The tracker sampled at 1000 Hz. Participants were asked to put their chin on a chin rest, the distance between their heads and the screen was fixed at 75 cm. Instructions and subsequent eye-tracker calibration were identical to Experiment 1.The eye-tracking experiment was followed by the Peabody vocabulary test and the verbal fluency task. Due to task misunderstanding and a technical error, the verbal fluency data of two participants could not be analyzed. Session 1 took approximately 35-45 minutes.

Session 2

The first experiment of Session 2 was the cross-modal object-verb priming experiment. Participants were asked to judge whether an auditorily presented verb was an existing Dutch word. The spoken inflected verbs (extracted from the recordings used in the eye-tracking experiment) were preceded by object primes which were on 25% of the trials (i.e., the 20 predictable items) related to the verbs. In total, the experiment consisted of 80 experimental and 40 filler items. Each participant saw the same 40 object-verb pairs (20 predictable and 20 non-predictable) from the list s/he was presented with in the eye-tracking experiment. Those 40 pairs served as experimental items and required a Yes-response. Yes-responses were provided using the right (dominant) hand on a button box. Recall that analyses of the word frequency and the number of letters of the inflected verbs across the predictable and the non-predictable condition revealed no statistical differences. Filler items which required a No-response were constructed by selecting 40 additional object primes and pairing them with 40 pseudoverbs derived from existing Dutch verbs by replacing a single letter. The pseudoverbs were embedded in a sentence template similar to the experimental stimuli and spoken by the same speaker. Recordings of those sentences were made and the pseudoverbs were extracted using Praat. Filler items were the same for all participants. Two lists with 80 trials each were generated and pseudo-randomized prior to the experiment. Each participant was presented with all 80 trials on one list. The structure and timings within a trial were similar to Ferretti, McRae, and Hatherell (2001) and were as follows: A '+' appeared in the center of the screen for 250 ms. Subsequently, the prime object was presented for 200 ms. (e.g., an apple). A visual mask used to minimize effects of visual memory was shown for 50 ms before the target (e.g., "schilt", peels) was played back. A trial was terminated by the participant's response and they received immediate visual feedback in case of an incorrect response. The inter-trial interval was 1500 ms. The experiment took approximately five minutes. Incorrect responses and reaction times greater than 1300 ms (more than twice as long as the expected mean reaction time based on similar studies, see Ferretti, McRae, & Hatherell, 2001; McRae, Hare, Elman, & Ferretti, 2005) to experimental trials were removed from the analysis (a total of 6.2% of the data). Due to an error, the logfiles of two participants were not saved. In the regression analysis, these cells were left empty. Participants completed Session 2 by carrying out Raven's progressive matrices test.

Results

As one participant did not come back for the second test session the data of this participant from all the tasks in Session 1 were excluded. The final sample therefore comprised 60 participants.

Eye-tracking experiment

The eye-tracking data were analyzed as in Experiment 1. Figure 2.4 shows time-course graphs plotting the fixation proportions for target and distractor objects for the predictable and non-predictable condition (Panel A) and the difference between log-transformed fixation proportions for target and distractor objects in the predictable and non-predictable conditions (Panel B). As in Experiment 1, by-participant CIs for the respective objects/conditions were calculated for each sampling step (1ms) and added to the plots. The pattern of fixations looks very similar to that in Experiment 1: In the predictable condition, participants anticipated the target noun and fixated upon the respective picture shortly after they had heard the verb, roughly one second prior to the target word onset. In the non-predictable condition, there was no difference in looks to the target object and the unrelated distractors during the time period starting at the onset of the verb until the onset of the target noun. Only after the target noun onset, participants gazed more at the target picture than at the unrelated distractors.

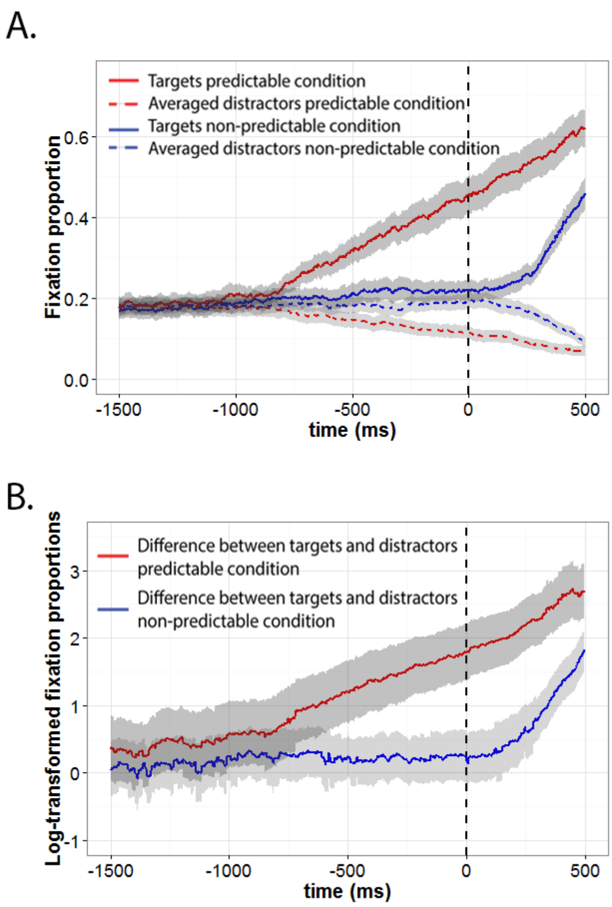


Figure 2.4: Time-course graphs plotting the eye-tracking data for Experiment 2. Panel A displays fixation proportions for the various objects present in the scene; Panel B displays the difference between the log-transformed fixations to target and averaged distractors. Trial timing was identical to Experiment 1.

Cross-modal object-verb priming

Reaction times to predictable items were on average 17 ms faster than reaction times to non-predictable items. The mean RT in the predictable condition was 734 ms (SD = 163 [740.9, 761.1]), and the mean RT in the non-predictable condition was 751 ms (SD = 170, CI [724.3, 743.7]). Each participant's and each item's priming effect was calculated by subtracting a participant's or an item's mean RT in the predictable condition from that participant's and that item's mean RT in the non-predictable condition. Note that for the argument the overall priming effects are less important than the priming effects for individual participants and items and their contributions to anticipatory eye movements.

By-participant regression analysis

The by-participant regression analysis (R^2 = .99) was based on the same variables as in Experiment 1 plus the by-participant cross-modal object-verb priming effects (see Figure 2.5, for scatter plots). The model revealed the following individual contributions (see Table 2.1 for a summary): production fluency scores (unstandardized *ß* = .033, SE*ß* = .023, CI [-.013, .079]; standardized *ß* = .194), Peabody vocabulary scores (unstandardized *ß* = .009, SE*ß* = .006, CI [-.002, .02]; standardized *ß* = .226), Raven's progressive matrices (unstandardized *ß* = -.007, SE*ß* = .015, CI [-.038, .024]; standardized *ß* = -.065), cross-modal priming effect (unstandardized *ß* = .001, SE*ß* = .002, CI [-.003, .004]; standardized *ß* = .037).

By-item regression analysis

Predictive gaze (calculated as previously) was regressed on the measures of functional associations, general verb-noun and noun-verb associations, and the items' cross-modal priming effects (see Table 2.1 for a summary). As in

Experiment 1, functional associations contributed most to explaining variance within the dependent variable (unstandardized ß = .9, SEß = .455, CI [-.025, 1.824]; standardized ß = .341; see Figure 2.5, for scatter plots). The unique contributions of the other variables were rather low (general verb-noun associations: unstandardized ß = -.104, SEß = .566, CI [-1.253, 1.044]; standardized ß = -.03; general noun-verb associations: unstandardized ß = .583, SEß = 1.051, CI [-1.551, 2.716]; standardized ß = .098; cross-modal priming effect: unstandardized ß = .0, SEß = .001, CI [-.002, .002]; standardized ß = -.023). The R^2 of that model was .143. I observed a moderate positive correlation between an item's general noun-verb association strength and its cross-modal object-verb priming effect (r = .32, CI [.008, .631]) suggesting that the stronger the general association between a given noun-verb pair the stronger its priming effect.

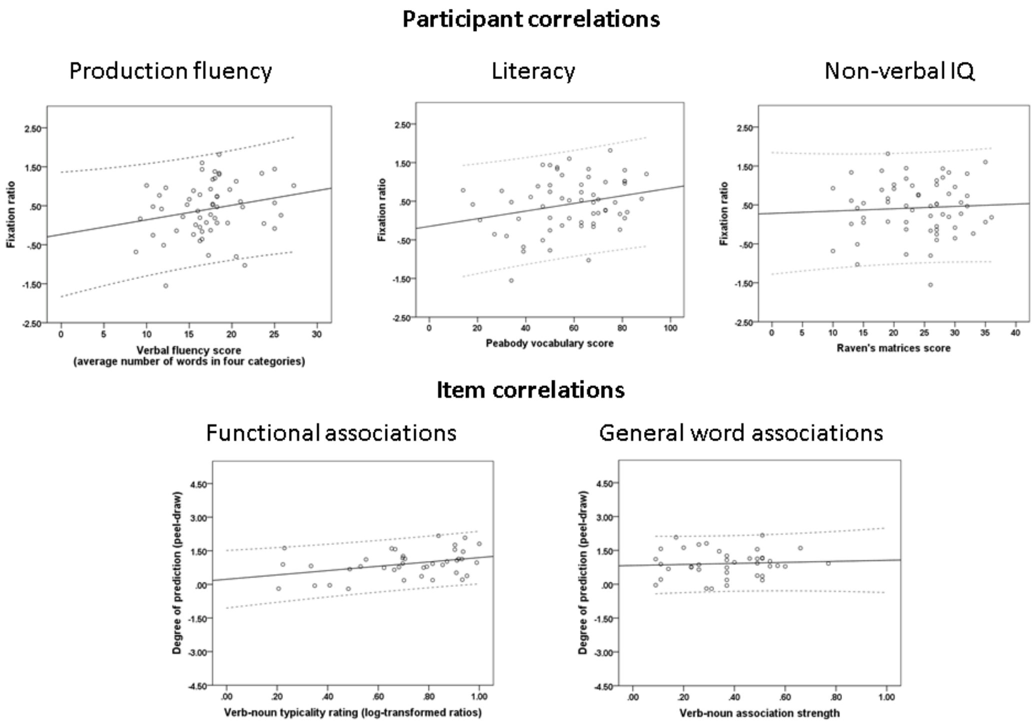


Figure 2.5: Scatter plots featuring participant and item correlations for Experiment 2.

Combined analyses of Experiment 1 and Experiment 2

As Experiment 1 and Experiment 2 were identical, I pooled the data of Experiment 1 and Experiment 2 to increase the statistical power. The dataset therefore contained 121 participants. I conducted regression analyses by participants (production fluency scores, Peabody vocabulary scores and Raven's nonverbal intelligence scores) and by items (functional associations, general verb-noun and noun-verb associations). With the larger dataset, the by-participant model (R^2 = .13) revealed the following unique contributions (see Table 2.1 for a summary): production fluency (unstandardized ß = .04, SEß = .017, CI [.007, .072]; standardized ß = .213), Peabody vocabulary scores (unstandardized ß = .009, SEß = .004, CI [.002, .016]; standardized ß = .242), Raven's progressive matrices (unstandardized ß = .007, SEß = .011, CI [-.014, .028]; standardized ß = .06). By items, the analysis (R^2 = .233) revealed the following unique contributions: functional associations (unstandardized ß = 1.192, SEß = .385, CI [.41, 1.973]; standardized ß = .498), general verb-noun associations (unstandardized ß = -.454, SEß = .478, CI [-1.422, .515]; standardized ß = -.143), general noun-verb associations (unstandardized ß = .003, SEß = .844, CI [-1.71, 1.715]; standardized ß = .0).

Discussion

The eye-tracking results of Experiment 2 replicated the prediction effect found in Experiment 1: Participants shifted their overt visual attention to the target objects already one second before the targets were mentioned. Likewise, by-participant and by-item regression analyses yielded results similar (though slightly lower in effect size) to those in Experiment 1. When pooling the data of both experiments, a robust pattern emerged: production fluency and literacy explained considerable unique variance in verb-mediated anticipatory eye movements in the by-participant analysis. By items, functional associations accounted for the largest proportion of unique variance in predictive eye gaze. Experiment 2 explored the influence of object-verb priming on predictive eye gaze. In a lexical decision task, the target objects primed the recognition of spoken verbs, similar to previous research (McQueen & Huettig, 2014). The regression analyses, however, showed that the participants' and the items' cross-modal priming effects did not explain much of the variance in anticipatory eye gaze.

With regard to the contribution of general word associations, it is conceivable that an influence on predictive processing was not observed because anticipatory eye movements were influenced by both general noun-verb and general verb-noun associative priming. The measures only captured each influence individually and not their combined effect. To assess this possibility, I added the values of each item's general noun-verb association strength and verb-noun association strength yielding a combined measure of general association strength. Subsequently, I regressed predictive gaze (using the combined dataset of Experiment 1 and 2) on the measure of functional associations and the combined measure of general associations. The model (R^2 = .228) showed that functional associations still explained a large amount of unique variance in predictive eye gaze (unstandardized ß = 1.226, SEß = .374, CI [.468, 1.984]; standardized ß = .512) but the combined general verb-noun and noun-verb association strength did not (unstandardized ß = -.342, SEß = .409, CI [-1.171, .488]; standardized ß = -.131).

To sum up, I found little evidence for a role of general word associations and non-verbal IQ in predictive language processing. Experiment 1 and 2 strongly support the notion that functional associations, language production abilities, and literacy abilities influence verb-mediated anticipatory eye gaze. This finding is consistent with the notion that multiple mechanisms and mediating factors rather than a single mechanism underlie prediction in language processing (Huettig, 2015; Mani & Huettig, 2013).

Experiment 3

What determines the weighting of these multiple mechanisms and mediating factors in predictive language processing? One possibility is that the situational context in which anticipatory language processing occurs determines how strongly each of these factors contributes to prediction (Huettig, 2015). Experiments 1 and 2 suggest that in situations with extensive visual input, functional associations, language production and literacy abilities are all related to anticipatory eye gaze. However, even when we talk about things in the ‘here and now’, language processing does not always take place with such extensive input from the visual environment. In many situations visual input is much more limited or even absent. In Experiment 3, I therefore explored the contributions of the same five predictor variables to anticipatory eye gaze when participants had only very limited visual input.

To that end, I conducted the same eye-tracking experiment as before but reduced the amount of visual input participants received by presenting the visual display only 500 ms prior to the target word onset. Rommers et al. (2013) have recently shown that presenting displays 500 ms prior to the target word onsets is sufficient time to observe anticipatory eye movements. Experiment 3 was run in one testing session. Participants carried out the eye-tracking experiment, which was followed by the Peabody vocabulary test, the production fluency task, and Raven’s progressive matrices test. The session lasted around 90 minutes.

Method

Participants

Sixty native Dutch speakers (mean age = 22 years of life, $SD = 5$), participated in Experiment 3. All had normal hearing and normal or corrected-to-normal vision. None of them had taken part in any of the norming studies or in Experiments 1 or 2.

Stimuli and procedure

The eye-tracking experiment was the same as in Experiment 1 and 2 except that the presentation of the visual display was timed to 500 ms prior to the target word onset rather than timing it to one second prior to the occurrence of the verb in the spoken signal. A trial started with the presentation of a fixation dot in the center of the screen. After two seconds, the playback of the sentence was initiated and the dot remained visible on the screen until it was replaced with the display featuring the four objects, 500 ms prior to the spoken target word.

As in Experiment1 and 2, I administered the verbal fluency task, the Peabody vocabulary test and Raven's progressive matrices test. Due to task misunderstanding, the verbal fluency data of one participant could not be analyzed.

Results and discussion

Figure 2.6 shows the fixations for the eye-tracking experiment. The plots suggest that participants disengaged their overt attention from the fixation dot between 300 and 200 ms before the onset of the spoken target. Anticipatory looks to the target objects arose around 100 ms before target onset.

For the regression analyses (see Table 2.1 for a summary), we selected a time window of 500 ms length, starting 200 ms after the display onset and ending

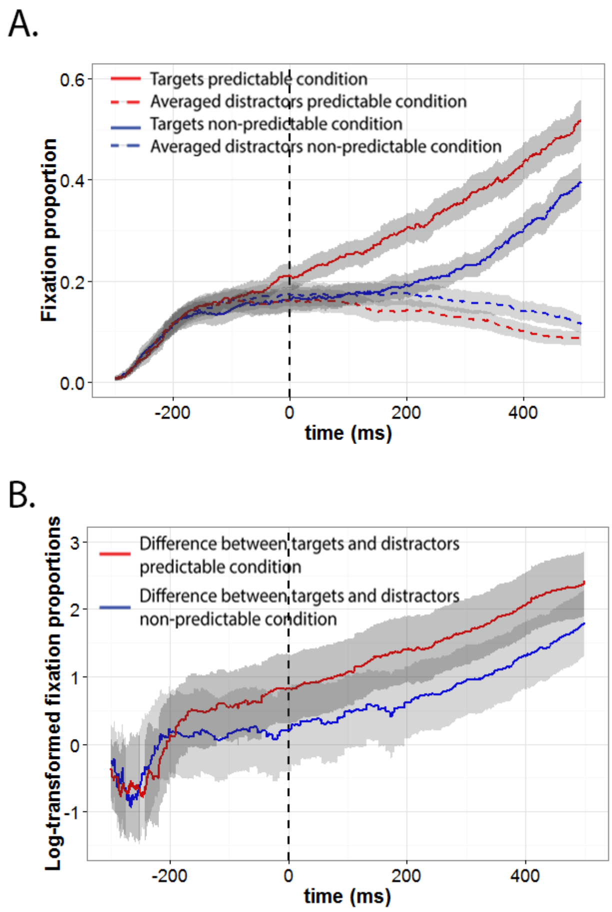


Figure 2.6: Time-course graphs showing fixations for Experiment 3, plotted in the same way as for the previous experiments. The onset of the display was timed to 500 ms prior to the target word onset (time zero in the graph).

200 ms after the spoken target onset, taking into account that it takes approximately 200 ms to program and initiate a saccadic eye movement (Saslow, 1967).

By-participant regression analysis

As in the previous experiments, the model (R^2 = .101) revealed some evidence for the influence of participants' Peabody vocabulary scores on the likelihood of their anticipatory eye movements (unstandardized $ß$ = .006, SE$ß$ = .003, CI [.0, .013]; standardized $ß$ = .277) suggesting that participants with better literacy predicted more. As in Experiments 1 and 2, we found that Raven's progressive matrices contributed little unique variance to verb-mediated anticipatory eye movements (unstandardized ß = .007, SEß = .011, CI [-.015, .029]; standardized ß = .092). Interestingly, unlike in Experiment 1 and 2, production fluency did not explain much unique variance in anticipatory eye gaze (unstandardized ß = -.011, SEß = .024, CI [-.059, .037]; standardized ß = -.063, Figure 2.7, for scatter plots).

By-item regression analysis

As the visual objects were presented 500 ms prior to the onset of the spoken target, i.e., after participants had already processed the verb in the spoken sentences, object/noun-verb priming seems very unlikely. Therefore, the by-item regression analysis (R^2 = .123) was based on the measures of functional associations and general verb-noun associations. As in the previous two experiments, the results revealed that functional associations were a robust predictor of the likelihood of predictive looks made to an item (unstandardized ß = 3.775, SEß = 1.728, CI [.274, 7.275]; standardized ß = .347) but general verb-noun associations were not (unstandardized ß = .176, SEß = 2.287, CI [-4.457, 4.809]; standardized ß = .012).

General Discussion

In the current study, I aimed to determine the influence of five potential predictors of verb-mediated anticipatory eye gaze behavior: functional associations, general word associations, production fluency, literacy, and non-verbal IQ. In three visual world eye-tracking experiments participants looked at visual displays of four objects while listening to sentences in which the target object was predictable ("The man peels the apple") or not predictable ("The man draws the apple"). In Experiments 1 and 2 the visual display was presented

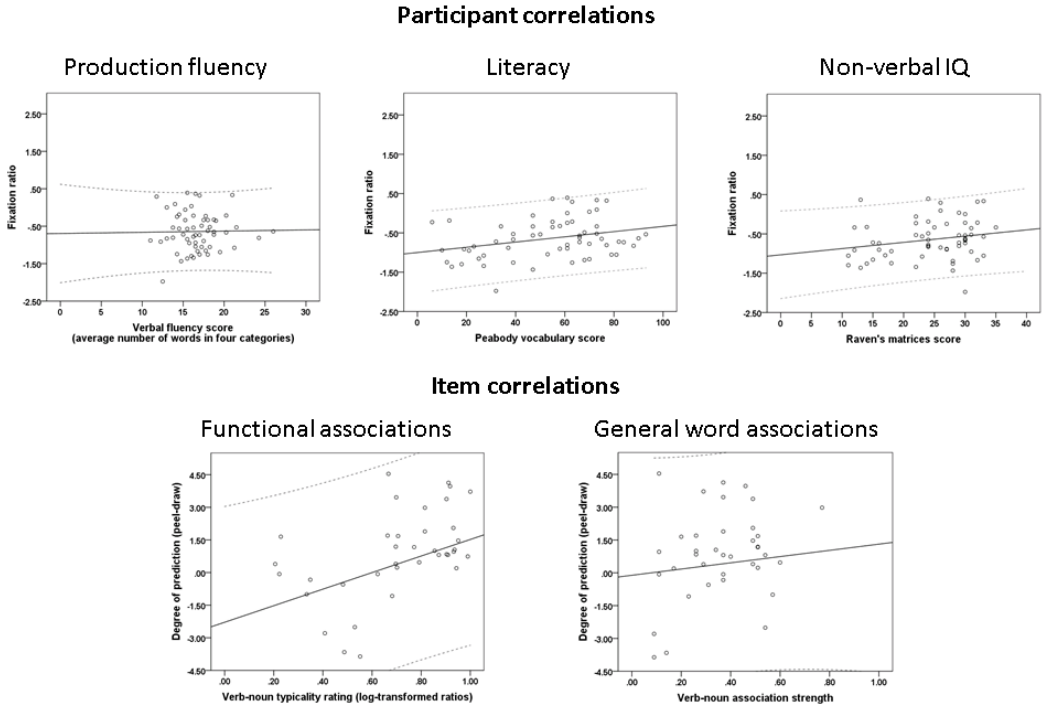


Figure 2.7: Scatter plots showing participant and item correlations for Experiment 3.

one second before the verb was heard. In Experiment 3 participants were given only a short preview of the display starting 500 ms before the target noun was heard. The predictable sentences varied in verb-noun typicality and general verb-noun association strength, representing measures of functional associations and general word associations, respectively. Participants' production fluency was assessed using the verbal fluency task. I used the Peabody vocabulary test as a proxy for literacy and Raven's advanced progressive matrices as a proxy for non-verbal intelligence.

In all three experiments, I found that participants anticipated the predictable spoken targets as indexed by looks to the target objects before they were referred to in the speech signal. No such anticipatory eye movements were observed when the same objects were referred to in a non-predictable sentence. This pattern replicates earlier research (e.g., Altmann & Kamide, 1999). I performed standard multiple regression analyses in which all predictor variables were entered simultaneously into the regression equation in order to estimate the unique contribution of each predictor variable.

By-item regression revealed that functional associations were the most robust predictor of verb-mediated anticipatory eye movements in all three experiments. This conclusion is based on a positive relationship between the likelihood of anticipatory looks made to an item and that item's verb-noun typicality rating. In contrast, there was not even a hint pointing towards a contribution of general word associations to verb-mediated anticipatory eye movements.

By-participant regression, with three individual differences measures, thought to tap production fluency, literacy, and non-verbal intelligence, revealed literacy to be a robust predictor of verb-mediated anticipatory eye movements in all three experiments. Production fluency correlated positively with the likelihood of anticipatory eye movements (and explained substantial unique variance) when participants were given the long preview but not when given the short preview of the visual display. I will now discuss the implications each of these findings.

Functional vs general word associations

The findings of a robust role of functional associations for language-mediated anticipatory eye gaze are compatible with the findings by Kukona et al. (2011). Their results can be interpreted as showing that on processing a transitive verb such as 'arrest' functionally associated concepts are pre-activated irrespective of their thematic fit with the local sentence context. The present results are also in line with the developmental findings by Borovsky et al. (2014). These authors reported evidence for the involvement of associations in anticipatory language processing in 3-4 year olds. Given that the children in their study were instructed to learn connections between agents, actions, and objects, anticipatory eye movements in that study were most likely also driven by functional rather than general associations. The present results show that functional associations are a strong predictor of verb-mediated anticipatory eye gaze in adults. More generally, the results provide further support for the notion that functional knowledge is a particularly salient aspect of word meaning (e.g., Moss et al., 1997).

The lack of an effect of general associations may be surprising given its assumed important role in cognitive processing (see Hutchison, 2003, for a review) and specifically in prediction (Bar, 2007, 2009). Using non-predictable sentences such as "Eventually she looked at the beaker that was in front of her" and displays of four visual objects, Huettig and McQueen (2007) found that participants first looked at phonological (beaver) and then at semantic (fork) and visual (bobbin) competitors during the unfolding of the target "beaker". This implies the activation of knowledge on multiple levels of representation during spoken word processing (see Smith, Monaghan, & Huettig, 2013, for a computational modeling approach of the gaze pattern observed by Huettig and McQueen). One might expect that general associations, a measure sensitive to the sum of multiple associative connections between a cue and a target should be able to capture predictive gaze more precisely than a component of that sum (e.g., functional associations). I suggest that the situation in which language processing takes place affects the impact of particular types of associations: In the current experiments functional associations were more important than general associations. It is likely that verb-noun typicality (and not general associative strength) predicted anticipatory eye movements so well because of the presence of the objects. That is, the presence of the target referents may have boosted the pre-activation of functional associations to the verbs. Seeing something that has the visual shape listeners would typically associate with 'peel-able' objects might increase the likelihood of fixations to an apple.

On the other hand, it is also possible that overall general associations play only a weak role in predictive language processing. Some hints in this regard come from eye-tracking research on reading. McDonald and Shillcock (2003) presented some evidence that transitional probabilities between words (i.e. the likelihood of two words occurring together) influence fixation durations in reading, which is in line with the notion that general associations are important for prediction in reading. However, Frisson, Rayner, and Pickering (2005) replicated the findings of McDonald and Shillcock (2003) in a first experiment only, but in a second experiment when items were matched for cloze values no effect of transitional probabilities was found. These authors concluded that low level transitional probabilities do not explain prediction above predictability effects determined by the use of a cloze task. More doubts about the role of transitional probabilities are raised by the absence of interactive effects of frequency and predictability in reading research. Predictability effects should be larger for low frequency than for high frequency words (Levy, 2008; McDonald & Shillcock, 2003; Norris, 2006) since reading a low frequency word in a context in which it is highly expected should be easier whereas reading a high frequency word in a predictive context should result in less of a benefit as it is very likely to occur anyway. A large number of studies investigating this issue have failed to find a significant frequency and predictability interaction (e.g., Altarriba, Kroll, Sholl, & Rayner, 1996; Ashby, Rayner, & Clifton, 2005; Kretzschmar, Schlesewsky, & Staub, 2015; Whitford & Titone, 2013; Rayner, Ashby, Pollatsek, & Reichle, 2004; Staub, 2011; Staub & Benatar, 2013). Future research is required on the role of general word associations in prediction in language processing.

Production ability

Although it is likely that functional associations are particularly important in situations with visual context, the results of the regression analyses in all three experiments support the view that functional associations predict anticipatory eye movements in both scenarios tested in the current study (long and short visual preview). Participants' production fluency, in contrast, correlated positively with the likelihood of anticipatory eye movements in the long, but not in the short preview condition. Even though the effect size for the influence of production fluency in Experiment 2 was weaker than in Experiment 1 (standardized *ß* = .194 vs. standardized *ß* = .269), both differed substantially from the results obtained with the short preview manipulation (Experiment 3; standardized *ß* = -.063). Thus, a sufficient preview period seems to play a pivotal role for the engagement of production-based prediction. The pattern suggests that the influence of mechanisms also involved in preparing to speak may be encouraged in situations in which language users can sufficiently exploit the visual input.

Note that the developmental study by Mani and Huettig (2012) and the adult study by Rommers et al. (2015), who also obtained evidence for production-based prediction used spoken sentences and pictures as well. In the experiments by Federmeier and colleagues (2002, 2010), correlations between prediction-related ERP components and student and older participants' production fluency scores were found although the experiments did not feature any pictorial input. This suggests that production abilities can modulate predictive language processing in multiple situations (e.g., reading, language-vision mapping). I suggest, however, that the likelihood of the involvement of production-based mechanisms in prediction may increase with relevant visual context present.

Why might this be the case? Imagine, for instance, two persons talking to each other face-to-face. They share the same visual environment and hence have a common non-linguistic source of information available to facilitate the prediction of upcoming language, for example, when the speaker refers to objects in the visual environment. In such situations, comprehenders might be inclined to engage in the question "How would I finish the sentence if I was the speaker, given the visual objects surrounding me?". In contrast to such a scenario, try to think of a telephone call between the same persons. I conjecture that in situations with limited relevant visual input (as in Experiment 3), comprehenders are less likely to employ production-based mechanisms for prediction.

As with all experimental tasks, the question may be raised whether (or to what extent) a particular task measures what it is assumed to measure. Performance on the verbal fluency task, for instance, may be regarded as a measure of executive functioning (cf. Henry & Crawford, 2004; Fitzpatrick, Gilbert, & Serpell, 2013; Shao et al., 2014, for discussion) rather than tapping core aspects of the language production process. As with most experimental tasks, there are several sub-components involved in the verbal fluency task (e.g., lexical retrieval, response inhibition of words which have been named already, etc.). When linking verbal fluency performance to prediction, I agree with Federmeier (2007) who argued: "The fact that it was a test of rapid, cued production that explained the pattern of [prediction] effects that would be seen in a language comprehension task further supports the contention that the [...] effect pattern reflects covert generation processes in the form of predictions about features of likely upcoming words-perhaps reflecting a link between language comprehension and language production mechanisms" (p. 495).

There are three arguments about the current data that suggest that the verbal fluency task measures at least partly production abilities rather than general cognitive abilities. First, verbal fluency scores and general intelligence scores (e.g., non-verbal intelligence; as measured using Raven's progressive matrices) did not correlate strongly in the experiments. This also – indirectly – hints at the second argument: in the past, performance on verbal fluency tasks has been linked to general processing speed (e.g., Bryan, Luszcz, & Crawford, 1997). One might hence argue that general cognitive processing speed rather than production fluency was crucial in the current set of experiments. That is, those participants who could process visual input more rapidly and link it to the unfolding language were the ones who predicted earlier or stronger. If this had been the case the correlation between participants' predictive gaze and their verbal fluency scores should have been particularly evident in even more challenging situations of visual processing such as in Experiment 3. Finally, there was no substantial correlation in any of the experiments between participants' vocabulary knowledge as measured using the Peabody vocabulary test and their verbal fluency scores. This suggests that, at least in the present study, the verbal fluency task was not simply tapping vocabulary knowledge. Therefore, I believe that the results point to the possibility that production-based mechanisms might indeed be involved in anticipatory language processing. In Experiments 1 and 2, with longer previews of the visual scene, the pictures may have served as lexical retrieval cues just as the categories and letters served as lexical retrieval cues in the verbal fluency task. Participants who were able to quickly retrieve lexical items after being cued in the verbal fluency task may have profited more from the picture input in the eye-tracking experiment and used those to make predictions about potentially upcoming words. Clearly, more research is needed to examine the involvement of production-based mechanisms in prediction. However, the results add to a growing body of correlational evidence suggesting a modulating influence of production abilities on anticipatory language processing.

Literacy

The current findings provide further evidence for a link between literacy and anticipatory spoken language processing. Similar evidence has previously been reported particularly in studies with participant populations with low literacy levels or reading disorders (illiterates, children, adults with dyslexia). The present results further confirm that literacy and in particular vocabulary knowledge is predictive of verb-mediated anticipatory eye gaze even among highly literate individuals such as the university students who participated in the present experiments (cf. James and Watson, 2013; Rommers et al., 2015).

The present study therefore confirms that literacy is one of the most consistent influences on anticipatory language processing. It is noteworthy that literacy accounted for unique variance of verb-mediated anticipatory eye gaze beyond what was accounted for by functional associations and production fluency. This is a novel finding and constrains causal explanations of literacy-on-anticipation effects. One possibility is that enhanced literacy sharpens orthographic representations, and by extension lexical representations more generally, which become available more quickly during on-line speech processing (Mani & Huettig, 2014). Further research is needed to test this proposal.

Non-verbal IQ

Finally, the results show that non-verbal intelligence does not explain much variance in verb-mediated anticipatory eye movements. This is consistent with the notion that linguistic anticipation abilities are largely independent from non-verbal abilities (see Rommers et al. 2015, for some evidence that verbal and non-verbal anticipation abilities may be distinct).

A pluralistic approach to predictive language processing

Taken together the results suggest that language users do not rely on a single mechanism to predict upcoming language. Instead, prediction in language processing is driven by multiple mechanisms and influenced by multiple mediating factors. The possibility that multiple mechanisms underlie predictive language processing has recently been suggested by several researchers (Kuperberg, 2007; Huettig, 2015; Pickering & Garrod, 2013; Wlotko & Federmeier, 2013). Pickering and Garrod (2013) for instance assume that "comprehenders will emphasize [production-based mechanisms] when they are (or appear to be) similar to the speaker" and that "[associations are] more accurate for simple, "one-step" associations between a current and a subsequent state" (p. 346). They also point out that comprehenders may combine both mechanisms. The results are consistent with this view. However, the results also signal the need for adjustments to Pickering and Garrod's account and other theories of predictive language processing. The situational context seems to be crucial in determining the contribution of each of the underlying mechanisms.

Conclusion

In language comprehension, a main goal of the comprehender is to understand the communicated content as fast as possible. Prediction might be a tool used to achieve that. To that end, language users might not rely on a single mechanism used in all situations of predictive language processing. It might rather be the case that the available input sources determine (at least partly) which mechanism is employed to predict upcoming words. Very often, we process spoken language when (relevant) visual input is available. Hence, it might not be surprising that, among, others, the presence or absence of visual input might determine which mechanism is used for prediction. Understanding how the multiple mechanisms and mediating factors underlying prediction interact with the nature of the environment the language user is immersed in is a necessary prerequisite for a complete account of prediction.

Table 2.2: Appendix: Stimulus materials

Target object	**Predictable verb**	**Non-predictable verb**	**Distractor 1**	**Distractor 2**	**Distractor 3**	**General word associations**	**Functional associations**
appel (apple)	schillen (peel)	tekenen (draw)	uil (owl)	kaars (candle)	radio (radio)	0.58	0.82
baard (beard)	scheren (trim)	zien (see)	mand (basket)	typemachine (typewriter)	kurkentrekker (corkscrew)	0.32	0.99
bal (ball)	trappen (kick)	lenen (borrow)	pak (suit)	krant (newspaper)	gitaar (guitar)	0.11	0.49
band (tube)	verwisselen (change)	verliezen (loose)	ananas (pineapple)	colbert (jacket)	voet (foot)	0.09	0.41
bank (coach)	bekleden (stiffen)	kiezen (choose)	worst (sausage)	pijp (pipe)	trommel (drum)	0.4	0.7
beker (cup)	winnen (win)	bekijken (look at)	fontijn (fountain)	lieveheersbeestje (ladybug)	stoplicht (stoplight)	0.15	0.78
biertje (beer)	drinken (drink)	kopen (buy)	handdoek (towel)	klok (clock)	potlood (pencil)	0.24	1
bloem (flower)	planten (plant)	ontvangen (receive)	lucifer (match)	CD (CD)	koffer (suitcase)	0.28	0.94
boek (book)	publiceren (publish)	verstoppen (hide)	spuit (injectionneedle)	plug (plug)	kom (bowl)	0.32	0.91
boom (tree)	kappen (chop)	beschrijven (describe)	zebra (zebra)	vliegtuig (plane)	schoen (shoe)	0.35	0.91
boterham (sandwich)	smeren (prepare)	betalen (pay)	pistool (pistol)	jurk (dress)	rugbybal (rugby-ball)	0.15	0.68
broek (pants)	passen (fit)	zoeken (search)	zwaard (sword)	kikker (frog)	accordeon (accordion)	0.1	0.67
cadeau (present)	krijgen (receive)	stelen (steal)	bh (bra)	paddestoel (mushroom)	aap (monkey)	0.53	0.53
contract (contract)	ondertekenen (sign)	ontvangen (receive)	wijn (wine)	ballon (balloon)	hoed (hat)	0.35	0.7
deur (door)	openen (open)	zoeken (search)	trui (sweater)	stethoscoop (stethoscope)	ontstopper (plunger)	0.4	0.79
dief (thief)	arresteren (arrest)	filmen (film)	zeilboot (sailboat)	ijs-beer (ice-bear)	kasteel (castle)	0.38	0.93
doos (box)	tillen (lift)	verbergen (hide)	theezakje (tea-bag)	magneet (magnet)	frisbee (frisbee)	0.26	0.21
fiets (bike)	repareren (repair)	pakken (grab)	lippenstift (lipstick)	poes (cat)	ei (egg)	0.42	0.71
glas (glass)	breken (break)	lenen (borrow)	telefoon (telephone)	schaar (scissors)	naaimachine (sewing-machine)	0.48	0.33
hond (dog)	aaien (pet)	tekenen (draw)	schelp (shell)	vlag (flag)	robot (robot)	0.27	0.7
huis (house)	bezitten (own)	kiezen (choose)	muffin (muffin)	dokter (doctor)	olifant (elephant)	0.26	0.35
ijsje (ice-cream)	likken (lick)	overhandigen (hand over)	muts (cap)	zaag (saw)	bril (glasses)	0.18	0.9
kind (child)	beschermen (protect)	beschrijven (describe)	wasknijper (clothespin)	veer (feather)	boter (butter)	0.24	0.85
lamp (lamp)	vervangen (replace)	verbergen (hide)	borst (chest)	tak (branch)	wereldbol (globe)	0.32	0.62
muur (wall)	behangen (decorate)	bewaken (guard)	draaimolen (carousel)	laptop (laptop)	schip (ship)	0.48	0.7
overhemd (shirt)	strijken (iron)	zien (see)	aardappel (potato)	luidspreker (loudspeaker)	schop (shovel)	0.1	0.94
piano (piano)	stemmen (tune)	stelen (steal)	bezem (broom)	kaas (cheese)	sjaal (scarf)	0.08	0.93
pizza (pizza)	bestellen (order)	verkopen (sell)	harp (harp)	beksleutel (wrench)	skateboard (skateboard)	0.1	0.22
sigaar (cigar)	roken (smoke)	verstoppen (hide)	hanger (hanger)	mok (mug)	boor (drill)	0.16	0.66
sinaasappel (orange)	persen (squeeze)	overhandigen (hand over)	zaklamp (flashlight)	emmer (bucket)	bel (bell)	0.33	0.92
standbeeld (statue)	onthullen (reveal)	bewaken (guard)	balkon (balcony)	slang (snake)	weg (road)	0.27	0.48
stoel (chair)	verplaatsen (displace)	pakken (grab)	konijn (rabbit)	pijl (arrow)	brood (bread)	0.15	0.23
taart (cake)	bakken (bake)	verkopen (sell)	fluit (flute)	spiegel (mirror)	bijl (ax)	0.38	0.9
tafel (table)	dekken (prepare)	betalen (pay)	rugzak (backpack)	viool (violin)	fietspomp (bicycle-pump)	0.51	0.87
tas (bag)	dragen (carry)	kopen (buy)	paard (horse)	fabriek (factory)	helikopter (helicopter)	0.49	0.77
touw (rope)	spannen (take up)	verliezen (lose)	pen (pen)	kies (tooth)	slot (lock)	0.44	0.84
trein (train)	missen (miss)	filmen (film)	haai (shark)	berg (mountain)	iglo (igloo)	0.07	0.55
varken (pig)	slachten (slaughter)	fotograferen (take a photo)	bus (bus)	tent (tent)	sleutel (key)	0.4	0.95
vis (fish)	vangen (catch)	fotograferen (take a photo)	waterkoker (kettle)	sneeuwpop (snowman)	palmboom (palm-tree)	0.24	0.82
wond (wound)	hechten (suture)	bekijken (look at)	trompet (trumpet)	vogel (bird)	pan (pan)	0.58	0.65

Table 2.3: Appendix II: By-participant and by-item correlations for Experiment 1. For a given correlation, Pearson's r and 95% confidence intervals based on Fisher's z transformation are provided.

By-participant correlations

	Predictive looks	Non-verbal IQ	Peabody vocabulary	Verbal fluency
Non-verbal IQ	.297 [-.08, .456]	1		
Peabody vocabulary	.338 [-.009, .527]	.419 [.209, .697]	1	
Verbal fluency	.256 [-.003, .521]	-.082 [-.327, .153]	.012 [-.252, .275]	1

By-item correlations

	Predictive looks	Verb-noun typicality rating	General verb-noun association strength	General noun-verb association strength
Verb-noun typicality rating	.424 [.171, .828]	1		
General verb-noun association strength	-.087 [-.513, .103]	.250 [-.088, .520]	1	
General noun-verb association strength	.092 [-.388, .251]	.362 [.036, .644]	.1 [-.277, .427]	1

Table 2.4: By-participant and by-item correlations for Experiment 2.

By-participant correlations

	Predictive looks	Non-verbal IQ	Peabody vocabulary	Verbal fluency	Cross-modal priming effect
Non-verbal IQ	.057 [-.36, .23]	1			
Peabody vocabulary	.257 [-.062, .524]	.353 [-.006, .522]	1		
Verbal fluency	.221 [-.078, .47]	.151 [-.151, .357]	.189 [-.073, .442]	1	
Cross-modal priming effect	.068 [-.25, .324]	.325 [.023, .536]	.214 [-.029, .486]	.014 [-.256, .284]	1

By-item correlations

	Predictive looks	Verb-noun typicality rating	General verb-noun association strength	General noun-verb association strength	Cross-modal priming effect
Verb-noun typicality rating	.367 [-.004, .686]	1			
General verb-noun association strength	.066 [-.352, .295]	.25 [-.094, .523]	1		
General noun-verb association strength	.211 [-.246, .426]	.362 [.021, .672]	.1 [-.227, .471]	1	
Cross-modal priming effect	.037 [-.36, .313]	.084 [-.345, .304]	-.03 [-.419, .28]	.32 [.008, .631]	1

Table 2.5: By-participant and by-item analyses for the combined data of Experiment 1 and Experiment 2.

By-participant correlations

	Predictive looks	Non-verbal IQ	Peabody vocabulary	Verbal fluency
Non-verbal IQ	.172 [-.13, .251]	1		
Peabody vocabulary	.295 [.054, .444]	.377 [.209, .563]	1	
Verbal fluency	.241 [.038, .393]	.066 [-.145, .203]	.099 [-.084, .279]	1

By-item correlations

	Predictive looks	Verb-noun typicality rating	General verb-noun association strength	General noun-verb association strength
Verb-noun typicality rating	.462 [.171, .824]	1		
General verb-noun association strength	-.019 [-.449, .163]	.25 [-.088, .52]	1	
General noun-verb association strength	.166 [-.317, .318]	.362 [.036, .644]	.1 [-.227, .427]	1

Table 2.6: By-participant and by-item correlations for Experiment 3.

By-participant correlations

	Predictive looks	Non-verbal IQ	Peabody vocabulary	Verbal fluency
Non-verbal IQ	.21 [-.194, .381]	1		
Peabody vocabulary	.299 [-.013, .566]	.421 [.169, .659]	1	
Verbal fluency	.018 [-.330, .203]	.149 [-.2, .294]	.242 [-.016, .503]	1

By-item correlations

	Predictive looks	Verb-noun typicality rating	General verb-noun association strength
Verb-noun typicality rating	.35 [.025, .67]	1	
General verb-noun association strength	.099 [-.31, .334]	.25 [-.068, .568]	1

3 I see what you're going to say: Listeners' use of contextual visual information during predictive language processing restricts visual and semantic competition[1]

Abstract

Recent influential accounts of anticipatory language processing assume that individuals predict upcoming information at multiple levels of representation (Pickering & Garrod, 2013; A. Clark, 2013; Friston, 2010). Research investigating language-mediated anticipatory eye gaze typically assumes that linguistic input restricts the domain of subsequent reference (e.g., visual target objects). Here, we explored the converse case: Can visual input restrict the dynamics of anticipatory language processing? To this end, we recorded participants' eye movements when they listened to sentences in which the patient object was predictable based on the selectional restrictions of the verb (e.g., "The man peels a banana"). While listening, participants looked at different types of displays. The target object (e.g., banana) was either present or it was absent and the displays featured objects that had a similar visual shape as the target object (e.g. canoe) or objects that were semantically related to the concepts invoked by the target (e.g., monkey). Each trial was presented as long preview version, where participants saw the displays before the verb was heard, and as short preview version, where participants were given only a short preview of the display before the target onset, after the verb had been heard. Participants anticipated the target objects on both preview types. Importantly, robust evidence for the pre-activation of visual shape and semantic information related to the (absent) target objects was found on short but not long preview trials. Our results hence suggest that information in the visual environment can dynamically restrict anticipatory language processing.

[1]Adapted from *Hintz, F., Meyer, A.S., & Huettig, F. (submitted). I see what you're going to say: Listeners' use of contextual visual information during predictive language processing restricts visual and semantic competition.*

Introduction

In everyday conversations, speakers often refer to objects in their visual environment. On comprehending such utterances, listeners integrate the linguistic input with relevant visual context. This integrative process has been thoroughly studied, often using the visual world paradigm (Cooper, 1974; Huettig et al., 2011, for review). For instance, Huettig and Altmann (2007, Experiment 1), asked participants to listen to sentences such as "In the beginning, the zoo keeper worried greatly but then he looked at the snake and realized it was harmless", while seeing a display featuring a snake, a rug, a pillow, and a barrel. They found that participants already looked at the picture of the snake when hearing "zoo keeper", indicating that they used linguistic and world knowledge to anticipate objects that were likely to be referred to. The authors found that upon hearing "snake", participants looked not only at the target object (snake) but also at visually similar objects (e.g., an electric cable). No such bias in looks to the visually similar object was observed when participants heard "zoo keeper". This suggests that their anticipation of the concept 'snake' did not involve the pre-activation of visual information.

It is an important question whether visual information of upcoming spoken referents is anticipated as it has recently been proposed that prediction is a fundamental principle of human information processing (Bar, 2007, 2009; A. Clark, 2013; Friston, 2010). Some theoretical accounts assume that during language comprehension individuals predict input at multiple levels of representation (cf. Pickering & Garrod, 2013). Thus, visual information should be anticipated in the experimental situation tested by Huettig and Altmann (2007). An account for the lack of visual competition prior to the spoken target may be that anticipatory language processing was constrained by contextual visual information. In other words, participants may have used visual information to predict what the speaker is likely to say next. Consistent with this notion, McQueen and Huettig (2014) demonstrated that visual objects can prime the recognition of related spoken words.

In the current study, we tested the possibility that visual context constrains the dynamics of anticipatory language processing. We conducted an eye-tracking experiment similar to Experiment 1 by Huettig and Altmann (2007). Participants listened to sentences where the final word was predictable based on the selectional restrictions of the verb (e.g., Dutch translation equivalent of "The man peels at that moment a banana"). While hearing the spoken sentences, they looked at displays showing four objects. On target-present trials one of them was the target (banana), and the other objects were unrelated distractors. On target-absent trials, the target was replaced with a visually similar object (a canoe) or a semantically related object (a monkey). Only the target object satisfied the verb-specific selectional restrictions (see Figure 3.1, for example displays). Crucially, presentation of the display began either one second before the onset of the spoken sentence, before the verb was heard (long preview), or shortly before the onset of the target, after the verb had been heard (short preview). If listeners use visual information to constrain anticipatory language processing, the eye gaze patterns in the short and long preview conditions should differ. In the short preview condition, we should, upon presentation of the display, observe anticipatory looks to the target and to visually and semantically similar competitors. In the long preview condition, we should also see anticipatory looks to the targets, but the biases towards the visual and semantic competitors should be absent. This is because in the long preview condition, participants had ample time before the presentation of the verb to activate knowledge about the depicted objects (including, for instance, the kinds of actions they tend to be involved in) and should therefore be able to rapidly direct their gaze to the object that would be most likely to be involved in the action implied by the verb. The absence of biases towards the visual and semantic competitors would show that the effects of pre-activated knowledge about the objects in the visual display can overrule predictions invoked by the spoken verb.

Method

Participants

Sixty members of the subject panel of the MPI (eleven male, mean age = 22, SD = 3), took part in the experiment. All were native speakers of Dutch and did not report any history of learning or reading disabilities or neurological or psychiatric disorders. The participants were paid for participation. The ethics board of the faculty of Social Sciences of the Radboud University Nijmegen approved the study.

Materials

The experiment consisted of 30 Dutch transitive sentences (e.g., “De man pelt op dit moment een banaan”, the man peels at that moment a banana; Appendix, for all sentences) in which the final word was predictable based on the selectional restrictions of the verb. All sentences had the same structure and the same number of words: The subject position was taken by “the man”, and the adverbial “at that moment” separated verb and target. The padding between verb and target was used to ensure that participants had enough time to generate predictions and to program and launch saccadic eye movements prior to the onset of the spoken targets. The resulting sentence construction is deemed to be quite natural by native Dutch speakers. The target nouns were on average six letters long (SD = 3) and had a mean word frequency of 25 per million words (based on Subtlex, Keuleers et al., 2010, SD = 30).

The inflected verbs were on average six letters long (*SD* = 1) and had a word frequency of 4 per million (*SD* = 7; six verbs were not listed).

The sentences were pre-tested on cloze probability using an on-line tool for web experiments developed by the technical group of the MPI. Thirty-eight Dutch native speakers (five male; mean age = 22; *SD* = 3) took part in the rating study, none of whom participated in other rating studies or the main experiment. The mean cloze probability of the targets was .23 (*SD* = .25; ranging from .03 to .89).

To create the visual displays, we used the Dutch stimulus set by de Groot et al. (in press), which contains words and photographs of common objects matched for visual and semantic similarity. For each of our 30 targets, we selected a visual competitor, i.e., an object that had a similar visual shape as the concept invoked by the target and a semantic competitor, i.e., an object that was semantically similar to the target. Both competitors were unrelated to the target on all other dimensions (see de Groot, Koelewijn, Huettig, & Olivers, in press; de Groot, Huettig, & Olivers, 2015, for details of the rating procedure). Unlike the target, neither of the competitors satisfied the selectional restrictions of the verb. For each target, we further selected three completely unrelated objects as distractors[2] and a picture of the target. All pictures had the same size (124 x 124 pixels) and resolution (400 x 400 pixels).

Each sentence was paired with three different displays each consisting of the three unrelated distractors and one of three critical objects (e.g., target, semantic competitor, visual competitor; Figure 3.1, for example displays). Each display type was presented in a long and short preview trial. The six versions

[2] As de Groot et al.'s stimulus set only provides norms for two unrelated distractor objects per target word, we carried out additional semantic similarity and visual similarity rating studies (n = 36, nine male, mean age = 22, *SD* = 3, none of these volunteers took part in the main experiment or the cloze probability rating study) on the third distractor following de Groot et al.'s procedure. The additional distractors were rated not to be visually or semantically similar to the concept invoked by the target noun (visual rating task: average rating = 1.55; *SD* = 1.54; semantic rating task: average rating = .45; *SD* = .59; on a 1-10 scale).

of an item were distributed across six experimental lists such that each sentence occurred only once on one list. The lists featured equal numbers of long and short preview trials. Each display type occurred ten times on each list. In order to create equal numbers of target-present and target-absent trials, we added ten non-predictable filler sentences, which had the same structure as the predictable sentences. The filler sentences were paired with displays containing a picture of the target and three unrelated distractors and also occurred in long and short preview versions.

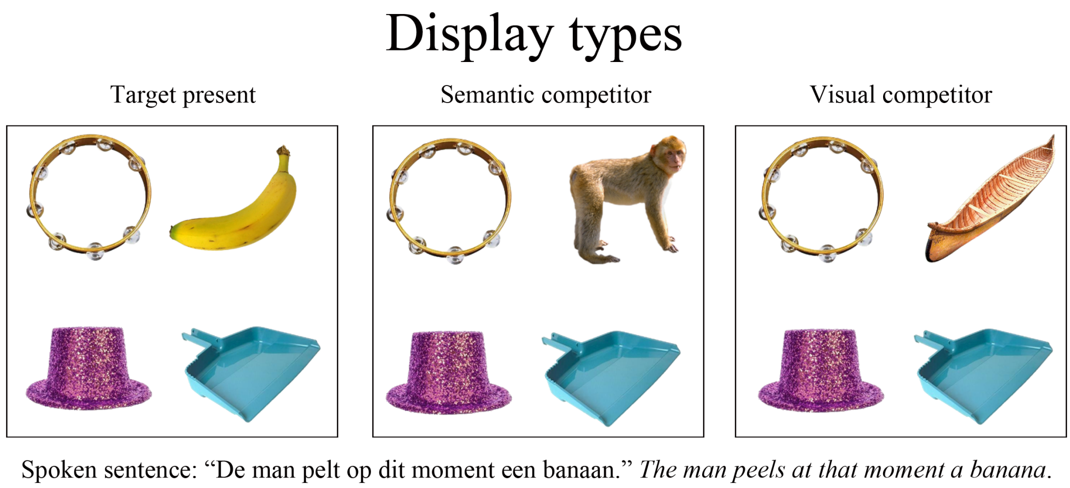


Figure 3.1: Examples of visual displays. While listening to the sentence, participants looked at displays in which the predictable target object was present (banana), or was absent and a semantic competitor (monkey) or a visual shape competitor (canoe) were present. In all three display types, the pictures of the hat, the tambourine and the dustpan were unrelated distractors.

Procedure

The 30 predictable sentences and the ten non-predictable sentences were spoken with neutral intonation at a normal pace by a female native speaker of Dutch. Recordings were made in a sound-damped booth, sampling at 44 kHz (mono, 16 bit sampling resolution) and stored directly on computer. The mean sentence duration was 3231 ms (SD = 195). Onsets and offsets of all words were marked in each sentence using Praat (Boersma, 2002). The time between the onset of the verb and the onset of the target noun in the predictable sentences was on average 1830 ms (SD = 159).

The participants were tested individually in a sound-shielded booth. Eye movements were recorded using an EyeLink 1000 tracker sampling at 1000 Hz. Participants placed their heads in a chin rest, which was approximately 75 cm away from the computer screen. After calibration, participants were randomly assigned one list. The order of trials was random with the constraint that maximally two trials of the same display type appeared in a row. The spoken sentences were presented through headphones. A trial started with the presentation of a central fixation dot in the center of the screen for two seconds. On long preview trials, the dot was replaced with the display and the playback of the sentence started after one second. On short preview trials, the playback of the spoken sentence started immediately after the two-second presentation of the fixation dot; the presentation of the displays was timed to begin 750 ms before to the onset of the spoken target word. The four objects remained in view until the end of the trial (see Figure 3.2). The positions of the objects were randomized across four fixed positions of a (virtual) 2 x 2 grid. The entire session took approximately 10 minutes.

Regions of interests (250 x 250 pixels) were defined around each of the four objects. The data from participants' left or right eye (depending on the quality of the calibration) were analyzed in terms of fixations, saccades, and blinks, using the algorithm provided in the EyeLink software. Fixations on experimental trials were coded as directed to the target, semantic competitor, visual competitor, one of the three distractors, or elsewhere.

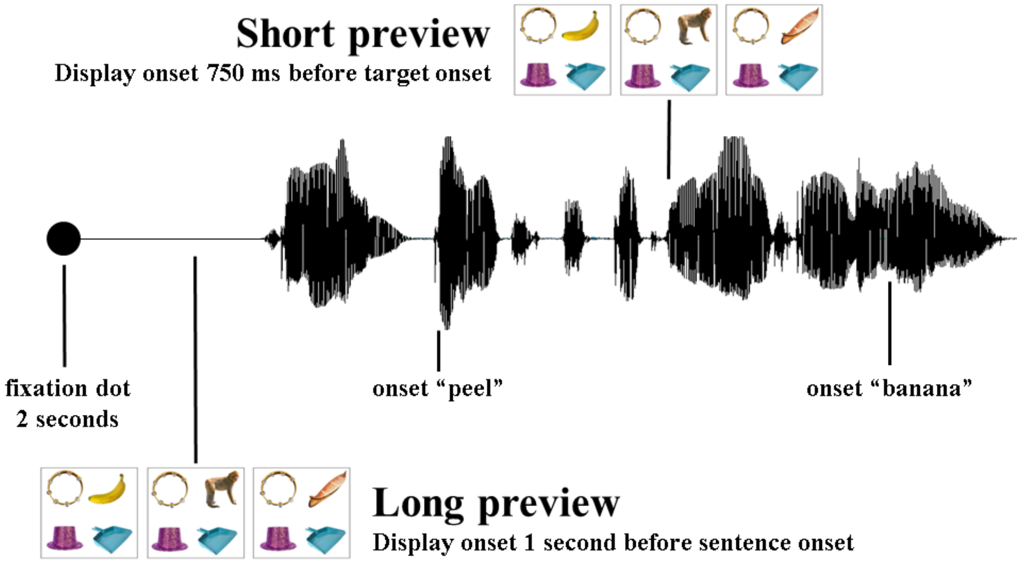


Figure 3.2: Time line of events in short and long preview trials. Participants listened to sentences such as “De man pelt op dit moment een banaan” (the man peels at that moment a banana) while looking at target-present, semantic competitor or visual competitor displays. The displays were either shown 750 ms before the onset of the spoken target word (e.g., “banana”; short preview) or one second before the sentence onset (long preview).

Results

The eye-tracking data were analyzed using a magnitude estimation approach. This was motivated by a recent proposal (Cumming, 2014; see Huettig and Janse, 2015), which advocates turning away from null-hypothesis testing towards interpreting results by using measures of effect sizes and confidence intervals. As has been shown empirically (Fidler & Loftus, 2009), this leads to a better interpretation of the results compared to a research report based on null-hypothesis testing (see Cumming, 2012, 2014, for extensive discussion).

Participants’ eye movements recorded on short preview trials were plotted for the period between the presentation of the visual display (750 ms before target onset, marked as time zero) plus an additional 1000 ms. As it takes around 200 ms to program and launch a saccade, participants’ fixations to the various objects emerged at around 500 ms prior to the target onset. Note again that the verb was presented approximately 1000 ms before the onset of the visual display. Eye movements on long preview trials are reported for a time window starting at the average onset of the spoken verb (i.e., 1830 ms before the target onset) until 1000 ms post target onset.

Figure 3.3 presents the fixation data for short and long preview trials for each of the three display types. By-participant confidence intervals (95%), computed at each sampling step (1 ms), were added to all lines in all panels and indicate by-participant variation (cf. Loftus & Masson, 1994; Masson & Loftus, 2003). The area between the lower and the upper bounds is shaded in gray.

The top panels in Figure 3.3 show that participants anticipated the targets on both short and long preview trials. On short preview trials, anticipatory eye movements to the target objects arose 450 ms before the objects were referred to in the speech signal. On long preview trials, participants gazed at the target objects shortly after having recognized the verbs, around one second prior to the target onset. The middle panels show that on short preview trials participants showed a strong bias towards the semantic competitor. The time course and magnitude of the competitor fixations were similar to the target anticipation effect. On long preview trials, the bias in looks to the semantic competitor was weaker and arose only shortly before the spoken target was heard. Most importantly, the bottom panels show that there was a bias towards the visual competitors on short preview trials but not long preview trials. Finally, for both preview types, we observed fixations to visual and semantic competitors at around 500 ms after the target onset, which most likely reflect bottom-up processing of the target word.

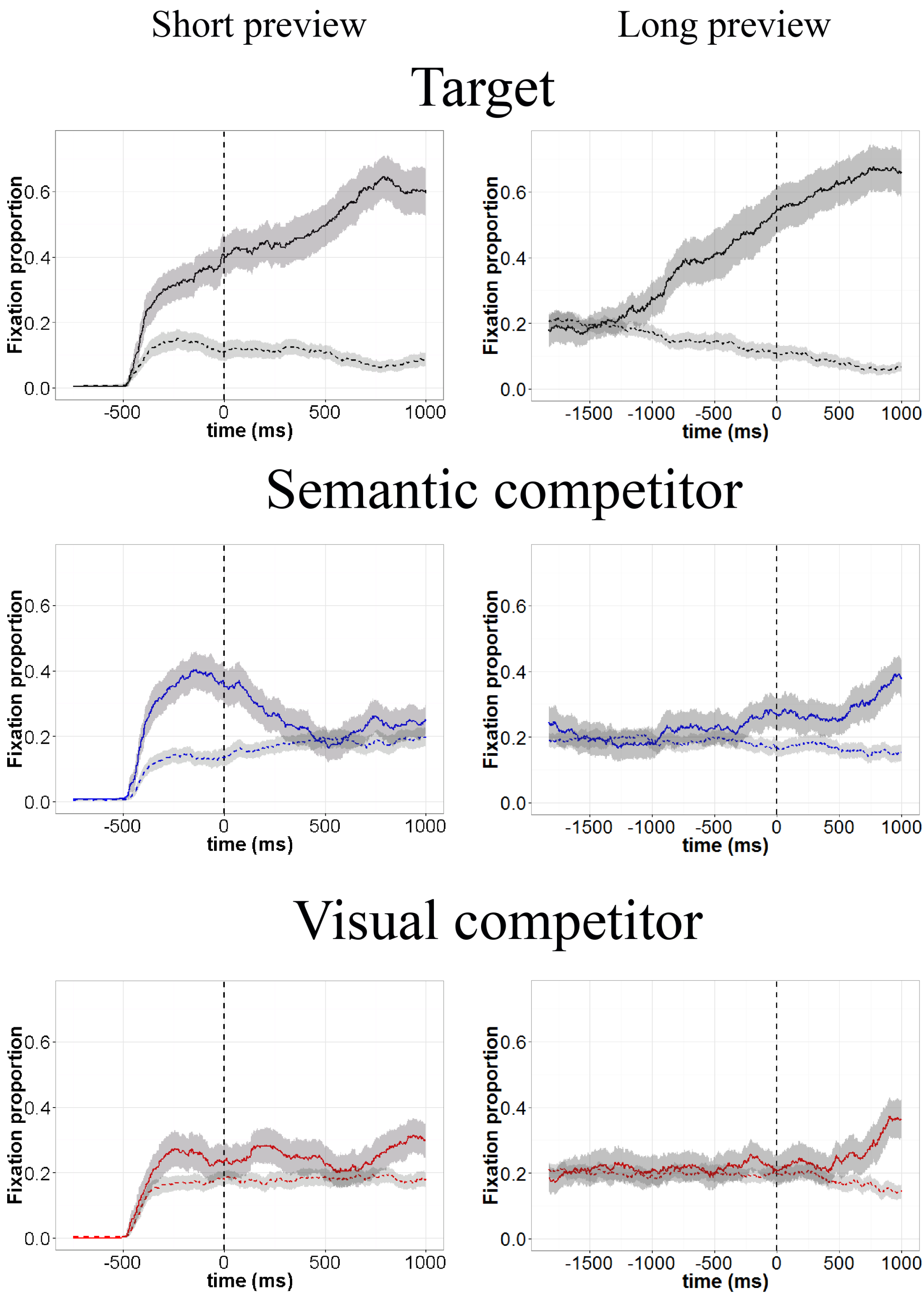


Figure 3.3: The graphs plot the fixation proportions for the critical objects and the averaged distractors in target-present and target-absent trials with short and long previews, respectively. Confidence intervals (95%), calculated at each sampling step, are shaded in gray. Time zero (vertical dashed line) indicates the average onset of the target words. Fixations on short preview trials are plotted from 750 ms before the onset of the spoken target until 1000 ms post target onset. Fixations on long preview trials are plotted from the average onset of the spoken verb until 1000 ms post target onset.

Discussion

In the target-present condition, we observed anticipatory eye movements to objects that satisfied the requirements of the verb with short and long preview manipulations. This is consistent with previous evidence demonstrating that listeners use verb-specific selectional restrictions to predict upcoming nouns (Altmann & Kamide, 1999). On short preview trials, we found a strong semantic and a weaker visual bias effect. Importantly, as hypothesized, these effects were eliminated in the long-preview condition. The absence of visual competition on long preview trials replicates the results by Huettig and Altmann (2007). However, in contrast to their conclusion that “a spoken word [does not activate] a target concept whose visual features initiate some form of visual search for corresponding features in the concurrent scene” (p. 1014), we did find competition effects on short preview trials, which rules out that visual and semantic representations were not pre-activated by the linguistic input.

As explained in the Introduction, this pattern indicates that, given sufficient time, listeners exploit the visual information to constrain anticipatory language processing. Thus, when participants inspected a display showing, among other things, a banana, or a canoe, or a monkey, they activated knowledge about the objects including the actions each of them is likely to be involved in. This allows them upon hearing the verb to quickly direct their gaze to the target in the target-present condition, since both the target and the associated action may already be pre-activated. It does, however, not lead to a bias towards the semantically or visually related distractors, since the activated knowledge about these objects does not include the verb or the implied action.

One may speculate that the influence of the visual context on anticipatory language processing is particularly evident in situations in which the linguistic stimulus is only weakly predictable. Indeed, such an account would explain why Rommers et al. (2015) recently obtained anticipatory visual competition effects with a long preview manipulation. Their participants listened to sentences such as “In 1969 Neil Armstrong was the first man to set foot on the moon” while looking at displays featuring a picture of the predictable target (e.g., moon) or the picture of an object that had a similar visual shape (e.g., tomato). The presentation of the display began one second before the sentence onset. In addition to anticipatory looks to the moon, Rommers et al. observed a bias in looks to the tomato before the onset of “moon”. Thus, in spite of the long preview period, a visual effect was observed. An important difference between the experiment reported by Rommers et al. and the present experiment is that the target words in their stimulus materials were highly predictable (average cloze probability $> .7$), whereas our target words were only moderately predictable (average cloze probability $= .23$). The high certainty of one particular target object may have led to strong activation of visual-form knowledge, which may have overridden the influence of the visual context. Future research could examine the exact mechanisms that weigh the contribution of each modality to language-mediated predictive eye gaze.

To conclude, prediction is assumed to play an important role during language processing (Dell & Chang, 2014; Pickering & Garrod, 2013). Here, we have shown that when speech is accompanied by relevant visual context, information extracted from the visual context contributes to the listeners’ predictions about upcoming language. This shows that the mechanisms underlying predictive language processing are situation-dependent. Theories of predictive language processing must take influences of the situational context into account.

Table 3.1: Appendix: Stimulus materials

Spoken verb	**Target**	**Semantic competitors**	**Visual competitors**	**Unrelated distractors**
strekken (stretch)	arm (arm)	hersenen (brain)	boemerang (boomerang)	waterscooter (watercraft), plakband (sticky tape), koekje (cookie)
kneden (knead)	asbak (ashtray)	pijp (pipe)	jojo (jojo)	dennenappel (pinecone), rozen (roses), verkeerslicht (traffic light)
pellen (peel)	banaan (banana)	aap (monkey)	kano (canoe)	tamboerijn (tambourine), hoed (hat), blik (dustpan)
winnen (win)	beker (cup)	vork (fork)	klos garen (bobbin)	pen (pen), duikbril (goggles), dynamiet (dynamite)
stapelen (stack)	blok (block)	hobbelpaard (rocking horse)	toffee (toffee)	saxofoon (saxophone), beer (bear), knoop (button)
planten (plant)	boom (tree)	bijl (ax)	wc borstel (toilet brush)	magnetron (microwave), magneet (magnet), grammafoonspeler (gramophone)
besturen (drive)	boot (boat)	anker (anchor)	klomp (clog)	chocolade (chocolate), honkbal (baseball), ventilator (fan)
branden (burn)	cd (CD)	diskette (floppy disk)	reddingsboei (buoy)	holster (holster), meetlat (yardstick), klamboe (mosquito net)
plukken (pick)	druif (grape)	wijnglas (wineglass)	biljartballen (billiard balls)	kettingzaag (chainsaw), bel (bell), megafoon (megaphone)
bakken (fry)	ei (egg)	haan (rooster)	wol (wool)	tandenborstel (toothbrush), xylofoon (xylophone), boog (bow)
koelen (cool)	fles (bottle)	kurk (cork)	kegel (bowling pin)	broek (pants), kerstbal (bauble), portemonnee (wallet)
blazen (blow/play)	fluit (flute)	harp (harp)	deegroller (rolling pin)	badeend (duck), ton (ton), skeeler (rollerblade)
smeden (forge)	hoefijzer (horseshoe)	zadel (saddle)	koptelefoon (headphone)	teddy beer (teddy bear), camembert (camembert), peultje (shell)
ontpitten (pit)	meloen (melon)	bananen (banana)	rugbybal (rugby ball)	golfclub (golf club), slang (snake), stoel (chair)
slijpen (sharpen)	mes (knife)	theepot (teapot)	peddel (paddle)	poederdoos (powder box), babybedje (cot), bokshandschoenen (boxing gloves)
drinken (drink)	milkshake (milkshake)	friet (French fries)	walkietalkie (walkie talkie)	wetsuit (wet suit), snelheidsmeter (speedo), pad (path)
laseren (laser)	oog (eye)	pruik (wig)	globe (globe)	broccoli (broccoli), politieauto (police car), stemvork (tuning fork)
piercen (pierce)	oor (ear)	voet (foot)	croissant (croissant)	schildersezel (easel), vrachtwagen (truck), leeuw (lion)
stemmen (tune)	piano (piano)	trompet (trumpet)	streepjescode (barcode)	riem (belt), bureaulamp (desk lamp), pannenkoeken (pancakes)
skimmen (skim)	pinpas (debit card)	euro (euro coin)	envelop (envelope)	blad (sheet), zwaan (swan), nagelschaartje (nail scissors)
snoeien (prune)	plant (plant)	gieter (watering can)	feesttoeter (party horn)	wasmachine (washmachine), controller (controller), garnaal (shrimp)
openen (open)	raam (window)	schoorsteen (chimney)	schilderij (painting)	vishaak (fishhook), zalmmoot (salmon fillet), honkbalknuppel (baseball bat)
lanceren (launch)	raket (rocket)	tank (tank)	vuurtoren (lighthouse)	toilettas (toiletry), dalmatiër (dalmatian), vuilnisemmer (bin)
graveren (engrave)	ring (ring)	oorbellen (earrings)	donut (donut)	telraam (abacus), prei (leek), fluitketel (kettle)
persen (squeeze)	sinaasappel (orange)	courgette (zucchini)	golfbal (golfball)	kalf (calf), snijplank (cutting board), bergschoen (hiking boot)
doppen (shell)	sperzieboon (bean)	ui (onion)	degen (sword)	spiegel (mirror), douchekop (showerhead), fiets (bicycle)
knopen (tie)	stropdas (tie)	trui (sweater)	vlieger (kite)	rolstoel (wheelchair), videoband (videotape), notenkraker (nutcracker)
waxen (wax)	surfplank (surfboard)	badpak (swimsuit)	veer (feather)	bizon (bison), graafmachine (excavator), ananas (pineapple)
kleien (make pottery)	theepot (teapot)	lepel (spoon)	kandelaar (candleholder)	sportschoenen (sneakers), bretels (braces), ketting (chain)
spotten (spot)	vliegtuig (plane)	label (label)	kruis (cross)	worst (sausage), muffinvorm (muffin pan), beeldscherm (screen)

4 | The relative contribution of event knowledge and word associations to prediction during discourse reading[1]

Abstract

A substantial body of literature has shown that readers and listeners often anticipate information. An open question concerns the mechanisms underlying predictive language processing. Theoretical accounts differ in whether they assume single or multiple mechanisms to underlie prediction in language processing. One proposal is that comprehenders use event knowledge to predict upcoming words. Other theoretical frameworks assume that simple word associations may also play a role in prediction. In the present study we contrasted the contribution of event knowledge and simple word associations to prediction using event-related brain potentials (ERPs). Participants read target words, which were preceded by associatively related words either appearing in a coherent discourse events (Experiment 1) or in sentences which did not form a coherent discourse event (Experiment 2). Contextually unexpected target words that were associatively related to the described event elicited a reduced N400 amplitude compared to contextually unexpected target words that were unrelated to the event (Experiment 1). In Experiment 2, a similar but reduced effect was observed. These findings support the notion that during discourse reading event knowledge and simple word associations jointly contribute to prediction and that multiple mechanisms underlie predictive language processing.

[1]Adapted from *Hintz, F., Meyer, A. S., & Huettig, F. (in preparation). The relative contribution of event knowledge and word associations to prediction during discourse reading.*

Introduction

Written text comprehension is fast and efficient. Ziefle (1998, see also Noyes & Garland, 2008) estimated that the average adult reads prose at a speed of 250 to 300 words per minute. One reason why we are such fast and efficient comprehenders may be that we often anticipate upcoming information (e.g., DeLong et al., 2005; Schwanenflugel & Shoben, 1985; van Berkum et al., 2005; Wicha et al., 2004). Experimental studies of prediction have so far primarily focused on the cues in the input signal that are used to predict upcoming language (e.g., Federmeier & Kutas, 1999; Kaiser & Trueswell, 2004; Knoeferle, Crocker, Scheepers, & Pickering, 2005; van Berkum et al., 2005) and on the contents of the comprehender's predictions (e.g., Arai & Keller, 2012; Chen et al., 2005; DeLong et al., 2005; Federmeier et al., 2002; Laszlo et al., 2012; Rommers et al., 2013; Staub & Clifton Jr, 2006; Wicha et al., 2003, 2004). An important issue that has received much less attention concerns the cognitive mechanisms that enable the pre-activation of linguistic knowledge given a particular predictive input cue. Accounts (Altmann & Mirković, 2009; Dell & Chang, 2014; Federmeier, 2007; Huettig, 2015; Pickering & Garrod, 2007, 2013) differ in that they assume single or multiple mechanisms to underlie prediction in language processing.

Event knowledge

One-system accounts of prediction in language processing typically ascribe an important role to event knowledge. According to this proposal, comprehenders use event knowledge to predict upcoming words (Altmann & Mirković, 2009; Metusalem et al., 2012). This is plausible because events tend to re-occur and thus are likely to be an important organizing principle of past experience.

In a recent EEG study, Metusalem and colleagues (2012) directly investigated whether event knowledge is immediately accessible during online language processing and constitutes a major determinant of predictive language processing. They recorded the N400, a component of time-locked EEG signals (i.e., event-related potentials, ERPs). It is a negative-going deflection that peaks around 400 milliseconds after stimulus onset and is considered to be an index of semantic processing (Kutas & Hillyard, 1980; Kutas & Federmeier, 2000, 2011, for review). The N400 ERP component is typically distributed over centro-parietal electrodes. The participants in the study by Metusalem et al. read passages describing typical events such as "The parents were very excited about their new baby girl. One of the first things they did was to get her baptized in their church. The baby liked baths, so she smiled when she was sprinkled with water/priest/dentist on her forehead". The authors observed a three-way split in N400 amplitude: Expected targets ("water" in the example) elicited the smallest N400 amplitude. Contextually unexpected words ("dentist") elicited the largest N400 amplitude. Interestingly, contextually unexpected targets which were related to the described event ("priest") elicited an attenuated N400 amplitude, which lay in between these two extremes. In their second experiment, participants read the target sentences presented in isolation. Now the N400 elicited by the unexpected but event-related words ("priest") did not differ from the N400 elicited by the unexpected event-unrelated words ("dentist"). Metusalem et al. concluded that readers used their knowledge about the described event to constrain their predictions about potentially upcoming referents. That is, on the basis of general event knowledge, readers pre-activated event-related words beyond the most expected continuation of the target sentences. Based on a post-hoc analysis using latent semantic analysis norms (LSA, Landauer & Dumais, 1997), the authors argued that the attenuation of the N400 amplitude in the event-related condition was unlikely to be influenced by associative priming between words in the preceding context and the critical targets.

Simple word associations

However, other theoretical frameworks assume that multiple mechanisms are required and that simple word associations may also play a role in prediction in language processing (Huettig, 2015; Kuperberg, 2007; Pickering & Garrod, 2013). Kuperberg (2007) linked the N400 and the P600 event-related potentials to two competing neural processing streams. She related the N400 explicitly to the computation of "semantic features, associative relationships and other types of semantic relationships between content words (including verbs and arguments) within in a sentence" (p. 36). Moreover, she proposed that the N400 might reflect the comparison of "these relationships with those that are pre-stored within lexical semantic memory" (p. 36).

Electrophysiological studies investigating associative priming in word pairs such as "church-priest" have demonstrated that target words elicit a reduced N400 amplitude when preceded by associated primes as compared to unassociated primes (Bentin, McCarthy, & Wood, 1985; Bentin, 1987). For example, Van Petten (2014) recently showed that free association strength (e.g., Nelson et al., 2004), assumed to quantify the association strength between two words, and corpus-based measures of association (such as LSA) both accounted for substantial amounts of variance in the amplitude of the N400 component elicited by the second word in 303 word pairs. These and many related studies, employing different paradigms and techniques (e.g., Chwilla & Kolk, 2002; Hutchison, Balota, Cortese, & Watson, 2008; Perea & Gotor, 1997; Hutchison, 2003; Neely, 1991, for review), suggest that associative relationships between words facilitate linguistic processing.

Simple word associations in sentence contexts

A recurring issue that has been intriguing language researchers concerns the contribution of simple associations to predictive processing in more complex linguistic stimuli than word pairs (see Ledoux, Camblin, Swaab, & Gordon, 2006, for review). To investigate this issue, previous studies have typically manipulated the associative relationship in word pairs (associated vs. unassociated) and the words' fit with the sentential context (congruent vs. incongruent). Pioneering work was carried out by Van Petten (1993), who compared the influence of associations with the influence of sentential contexts on the amplitude of the N400 component. An example of her stimuli is given in (1). Associated words are marked in italics.

(1) **Associated/congruent**: When the *moon* is full it is hard to see many *stars* or the Milky Way.

Unassociated/congruent: When the *insurance* investigators found that he'd been drinking they *refused* to pay the claim.

Associated/anomalous: When the *moon* is rusted it is available to buy many *stars* or the Santa Ana.

Unassociated/anomalous: When the *insurance* supplies explained that he'd been complaining they *refused* to speak the keys.

Van Petten reported reductions in the N400 amplitude for the associated/congruent, the unassociated/congruent and the associated/anomalous conditions, relative to the unassociated/anomalous condition. She reasoned that the N400 reduction in the unassociated/congruent condition could be attributed to the sentential context and that the N400 reduction in the associated/anomalous condition was due to lexical association. Crucially, she found that the N400 amplitude reduction in the associated/congruent condition was larger than the reduction in the remaining two conditions which she argued reflected the additive effects of sentential context and lexical association.

Coulson, Federmeier, Van Petten, and Kutas (2005) further explored the electrophysiological signature of associative priming in processing associated word pairs in isolation and in sentence contexts. Employing a lateralized half field manipulation, they additionally investigated the engagement of the two brain hemispheres subserving associative priming and the integration of sentential constraints. Their participants read pairs of associated and unassociated words in isolation. The authors reported that the second word elicited more positive ERPs in the associated than in unassociated pairs. In their second experiment, the same word pairs were embedded in simple sentences. Similar to Van Petten (1993), Coulson et al. crossed the experimental factors associative relationship and sentential fit. Their results showed that processing was primarily influenced by the words' fit with the sentential contexts. However, the authors also reported subtle effects of associative priming as indexed by a more positive N400 amplitude for associated compared to unassociated pairs and a late positive component elicited by the associated but not by unassociated pairs. The influence of associative priming was particularly pronounced when the word pairs occurred in incongruent sentence contexts. With regard to hemispheric differences, Coulson and colleagues observed that after presentation to the right visual field, lexical associations had an effect in incongruous but not in congruous sentence completions. After left visual field presentation, lexical association showed an effect in both congruous and incongruous sentences (see Beeman, 1993; Chiarello, Liu, & Faust, 2001, for further discussion of hemispheric contributions to sentence processing).

Taken together, the studies by Van Petten and Coulson et al. suggest that, at the sentence level, contextual constraints substantially influence the comprehension process. However, there is also some evidence for the modulating influence of simple word associations. These appear to have an impact especially when the target words are incongruent with the sentence context they appear in.

Simple word associations in discourse contexts

Previous investigations have also shown that discourse context beyond the level of single sentences exerts a powerful influence on the N400 component. Hagoort et al. (2004) asked Dutch participants to read sentences such as “Dutch trains are yellow/white/sour and very crowded”. The participants in that study knew that Dutch trains are typically yellow. The color word “white” was a violation of world knowledge and yielded an N400 similar to that elicited by the word “sour”. The authors argued that comprehenders immediately integrate word meanings and world knowledge. The same lab showed in a follow-up study that the critical N400 amplitude was attenuated when the target sentence was preceded by a mitigating context. Thus, reading “The coming world championships are one big national spectacle. The Dutch railways have painted the Dutch flag on their trains” prior to the target sentence reduced participants’ N400 amplitude in response to the word “white” in the sentence, most likely because the Dutch flag includes the color white (Hald, Steenbeek-Planting, & Hagoort, 2007; van Berkum, Hagoort, & Brown, 1999, for similar results). Note that in these and many similar studies the influence of the discourse context and the influence of associations between words in the discourse and the target words could not be separated. Previous attempts to do so yielded mixed results. Most researchers agree that simple associative relationships between words cannot fully account for the N400 amplitude modulations observed when critical target words are preceded by related discourse, but whether associations have a major or minor effect on the N400 amplitude is a matter of some debate (cf. Kuperberg, Paczynski, & Ditman, 2010; Otten & van Berkum, 2007).

To tease apart the effects of lexical association and the effects of discourse context, Camblin, Gordon, and Swaab (2007) orthogonally manipulated the lexical associations between words and the discourse congruence of the words. In a series of experiments recording ERPs and eye movements during reading, they investigated how lexical-level effects interacted with the effects of discourse context. The authors embedded associated and unassociated word pairs in sentences that were coherent and locally congruent. The critical words were either congruous or incongruous with the discourse-level context (see (2) for an example).

(2) **Associated/congruent**: Lynn had gotten a sunburn at the beach. Nothing she tried would help her dry and irritated skin. Lynn couldn't stop scratching her *arms* and *legs*.

Unassociated/congruent: Lynn had gotten a sunburn at the beach. Nothing she tried would help her dry and irritated skin. Lynn couldn't stop scratching her *arms* and *nose*.

Associated/anomalous: Lynn's wool sweater was uncomfortable and itchy. She fidgeted as the rough material irritated her skin. Lynn couldn't stop scratching her *arms* and *legs*.

Unassociated/anomalous: Lynn's wool sweater was uncomfortable and itchy. She fidgeted as the rough material irritated her skin. Lynn couldn't stop scratching her *arms* and *nose*.

Their analyses revealed independent effects of discourse congruence and lexical associations. Violations of discourse congruence had early and lingering effects on both ERP and eye-tracking measures, whereas the effects of association were more fragile and particularly evident in scenarios in which the discourse context was not cohesive. Camblin et al. (2007; see Otten & van

Berkum, 2008, for similar conclusions) thus concluded that when a cohesive, congruent discourse model can be constructed, it may override associative facilitation. They argued that the effects of association "[do] not contribute to processing of words in sentences that are part of a larger discourse" (p. 126).

In a follow-up study, Boudewyn, Gordon, Long, Polse, and Swaab (2011) used the same discourses as Camblin et al. in spoken form. In contrast to the previous study, they observed an interaction between lexical association and discourse congruency. They argued that local lexical associations and overall discourse congruence may each exert their own influence on incoming words during discourse comprehension and that these effects may be additive when the two sources of information are consistent with one another.

The current study

Previous studies have clearly established that sentence context and discourse context (e.g. event knowledge) exert powerful influences on language comprehension (see also Bicknell, Elman, Hare, McRae, & Kutas, 2010; Matsuki et al., 2011; McRae & Matsuki, 2009, for review). There is currently no consensus with regard to the importance of the contribution of word associations in discourse processing. One condition under which associative priming has repeatedly been shown to impact language comprehension is when the critical target words are incongruent with the local contexts they are embedded in (Boudewyn et al., 2012; Camblin et al., 2007; Van Petten, 1993). In light of this finding, the possibility arises that associative priming might have contributed to the prediction effects observed by Metusalem et al. (2012). To recap, the participants in that study read short texts about common events (e.g., a baptism) containing three kinds of target words (e.g., expected, unexpected event-related, unexpected event-unrelated). The authors found that the target words in the event-related condition that were incongruent with the local sentence context yielded an attenuated N400 amplitude relative to the event-unrelated unexpected condition. It is conceivable that the LSA post-hoc analysis carried out by Metusalem and colleagues to estimate the influence of associations was not sensitive enough to capture the interplay between sentence (in)congruence and lexical-level association effects. In order to clarify the contribution of simple associations to discourse comprehension and in order to gain a better understanding of the mechanisms underlying prediction, a more direct test needed.

In the present study, we tested the potential contribution of simple word associations to prediction during discourse reading. To that end, we first replicated Metusalem et al.'s (2012) context manipulation (Experiment 1). Dutch participants' electroencephalogram was recorded as they read short passages consisting of three sentences. While the first two sentences established an event context, the third sentence contained one of three target words: A highly expected word, a word that was unexpected in the context of the third sentence but related to the overall event context, or a word that was unexpected in the context of the third sentence and unrelated to the overall event. Analyses of participants' ERPs in response to the three kinds of target words closely replicated the three-way split pattern in N400 amplitude observed by Metusalem and colleagues. This demonstrated the robustness of the results in a different language than the original English study. In Experiment 2, we asked a different sample of participants to read the same target sentences as before. The target sentences were preceded by two sentences, which – unlike in Experiment 1 – did not build up a coherent discourse context. However, each of the two sentences contained a word (prime hereafter) that was part of the event-establishing sentences in Experiment 1 and that was associatively related to the unexpected event-related target word in that discourse. Using this manipulation, we minimized the event knowledge that participants could extract from the two introductory sentences while keeping part of the associated lexical input the same as in Experiment 1. As in Experiment 1, the critical primes appeared in a grammatical syntactic environment. We reasoned that the prime words were activated and kept in working memory and might activate strong associates that are also held in working memory (cf. Lau, Holcomb, & Kuperberg, 2013).

If the N400 amplitude reduction in the unexpected event-related condition in Experiment 1 (and in Metusalem et al., 2012, Experiment 1) was partly due to associative priming, we should observe a similar pattern of results in Experiment 2. Specifically, we predicted a significant difference between the N400 amplitude elicited by the unexpected event-related condition and the N400 amplitude elicited by the unexpected event-unrelated condition. Moreover, we predicted both unexpected conditions to differ significantly from the expected condition. Such a pattern would be consistent with the notion that simple associative relationships between words modulated the critical N400 component in the previous experiments. It would also be consistent with the notion that associations contribute to prediction during discourse comprehension. If on the other hand, the effect in the previous experiments was primarily driven by the activation of event knowledge, event-related and event-unrelated unexpected conditions should elicit N400 components of similar amplitude, with both being more negative than the N400 elicited in the expected condition.

Experiment 1

Method

Participants

Thirty-one members of the subject panel of the MPI (five male, mean age = 21, *SD* = 2) took part in Experiment 1. All were native speakers of Dutch, right-handed, and did not report any history of learning or reading disabilities or neurological or psychiatric disorders. The participants were paid for participation. The ethics board of the Faculty of Social Sciences of the Radboud University Nijmegen approved the study. One participant had to be excluded from the analysis due to an experimental error.

Stimuli

We started by translating the English materials used by Metusalem et al. (2012) into Dutch. Each of the 72 experimental items consisted of a short discourse that consisted of three sentences (see (3), for an example). The first two sentences established the event context. The third sentence contained the target words. The comprehension questions used by Metusalem and colleagues were also translated. Five native speakers of Dutch checked the qualities of the translations. The subsequent rating studies were the same as described by Metusalem and colleagues and were carried out to ensure that the Dutch materials were comparable to those of the original study.

(3) "De ouders waren erg blij met hun pasgeboren dochter. Een van de eerste dingen die ze hebben gedaan is haar laten *dopen* in hun *kerk*. De baby hield ervan om in bad te gaan dus ze lachte toen ze werd besprenkeld met water/pastoor/tandarts op haar voorhoofd."

> The parents were very excited about their new baby girl. One of the first things they did was to get her *baptized* in their *church*. The baby liked baths, so she smiled when she was sprinkled with water/priest/dentist on her forehead.

Cloze probability

We pre-tested the predictability of the expected target words within each discourse. Thirty-one native speakers of Dutch (5 male; mean age = 22, *SD* = 3), none of whom participated in the main experiments or in any of the other rating studies, carried out a cloze probability rating study. The study was conducted online using a tool for web experiments developed by the technical group of the MPI. The 72 translated discourses were presented to the participants in random order, one at a time, up to the word preceding the expected target. Participants were instructed to read each discourse carefully and provide the word they believed would be the best continuation of the sentence. Cloze probability was the proportion of participants who provided a particular response for a given discourse. For each discourse, we selected the word with the highest cloze probability as the expected target of that discourse. On eleven items another word than the direct Dutch translation of the English expected target was deemed the most likely continuation of the third sentence. Due to very low cloze probability, we had to exclude six items which were part of the original material set. The mean cloze probability of the expected targets in the remaining 66 items was .67 (range = .19 - 1; *SD* = .23; mean cloze probability in Metusalem et al., 2012, was .81, based on 30 participants and 72 items).

Event association

In another web-based rating study, we instructed 49 native speakers of Dutch (10 male; mean age = 20, *SD* = 2), who did not take part in the main experiments or other rating studies, to read the discourses containing the expected targets and try to think of persons and objects which were likely to be part of the described scenarios but were not mentioned. The discourses were randomized for each participant and presented one at a time. Participants were asked to write down minimally three and up to five event-associated persons and objects. Participants provided 13180 out of 16170 possible associations (49 participants x 66 items x 5 associations). Five-hundred sixteen of these (4%) had to be removed because they were not nouns. Event-related targets were selected according to a weighting system that was based on the order of mention of participants' responses (five points for the first association, four points for the second, etc.). The highest scoring association that was not listed in the cloze probability rating was selected as event-related target.

The Dutch language - unlike English - differentiates between common and neuter grammatical gender of nouns. Gender is reflected in determiners (common gender: de vs. neuter gender: het; both are equivalent to the English the) and in inflectional marking on prenominal adjectives (e.g., common gender: een goede auto, a good car vs. neuter gender: een goedø boek, a good book). When the highest scoring event-related target had a different gender than the expected target noun, and when gender was marked by a determiner or adjective, the highest scoring event-related noun with the same gender as the expected target was chosen. The maximum possible score for an item was 245 (49 participants x score of 5). Across the 66 discourses, the mean relatedness score for the event-associated targets was 75 (range = 16 to 222, SD = 48; mean event-related score in Metusalem et al., 2012, was 92.4, based on 45 participants and 72 items[2]).

[2]Note that the cloze probability of the expected target words and the event-relatedness scores for the unexpected event-related targets were a bit lower than in the original study. The lower cloze probability is very likely connected to the fact that we had to exclude six items from the original set. The event-relatedness scores are lower than in Metusalem et al.'s study because sometimes low event associations had to be selected as target words to match the expected target words' gender marking.

The event-unrelated targets were generated by shuffling the event-related targets across discourses such that event-related and event-unrelated targets consisted of the same lexical items. Therefore, lexical factors such as length and word frequency were the same across the two conditions. We split the 66 experimental items into three lists each containing 22 discourses. Each discourse appeared once in each list and once in each condition across the three lists. We minimized the variability across the lists by matching the three lists on the following variables: mean cloze probability, log frequency, and orthographic length of the expected targets; mean event-relatedness score, log frequency, and orthographic length of the event-related targets. Finally, we shuffled the event-related targets across the discourses within each rotation group to obtain the event-unrelated targets. The shuffling was done such that the event-related and event-unrelated targets within each discourse were matched for animacy and concreteness. We checked that in case of overt gender marking the event-unrelated targets had the same gender as the expected targets and the event-related targets. The results of all norming studies are summarized in Table 4.1.

The association strength between the expected and the unexpected targets was checked to assess the possibility that potential reductions in the amplitude of the N400 were driven by strong associative connections between the target words (cf. Federmeier & Kutas, 1999; Kleiman, 1980). We used the Dutch free association database by De Deyne, Navarro, and Storms (2013) to determine how strongly the expected targets (cues) and the event-related and event-unrelated targets (responses) were related. Fifty-nine of the 66 expected targets were listed in the database. The mean associative strength for these items was .0057 (SD = .014) for the event-related targets and was .0011 (SD = .007) for the event-unrelated targets, which amounts to one response per 175 participants and one response per 909 participants, respectively. De Deyne et

Table 4.1: Norming results for the three rotation groups and the stimuli set overall. Word frequencies were taken from the SUBTLEX-NL corpus of Dutch subtitles (Keuleers et al., 2010).

	List 1	List 2	List 3	Overall
Expected targets				
Cloze probability	.66	.68	.69	.68
Word frequency	71	64	70.9	68.63
Orthographic length	6.6	6.1	6	6.23
Event-related targets				
Cloze probability	.00	.00	.00	.00
Word frequency	92.9	91	106	96.63
Orthographic length	6.4	5.9	6.4	6.23
Event relatedness score	75.5	79.9	68.1	74.5
Event-unrelated targets				
Cloze probability	.00	.00	.00	.00
Event relatedness score	0	0	0	0

al.'s database lists responses of the first 100 participants who read a particular cue word. Thus, the association strength between the expected targets and both kinds of unexpected targets was rather low.

Twenty-two filler items of the same structure as the experimental items were constructed to ensure an equal number of trials containing anomalous and non-anomalous targets. Each discourse was followed by a comprehension question about the just-read scenario. Yes and No responses occurred equally often.

Procedure

Participants were randomly assigned one of the three lists and were tested individually in a dimly lit room. They were seated in a relaxed position in front of a 19 inch CRT screen. We told them that they were going to participate in a reading comprehension experiment which consisted of a number of short discourses. Participants were instructed to read each of these discourses carefully in order to be able to answer the questions following the discourses.

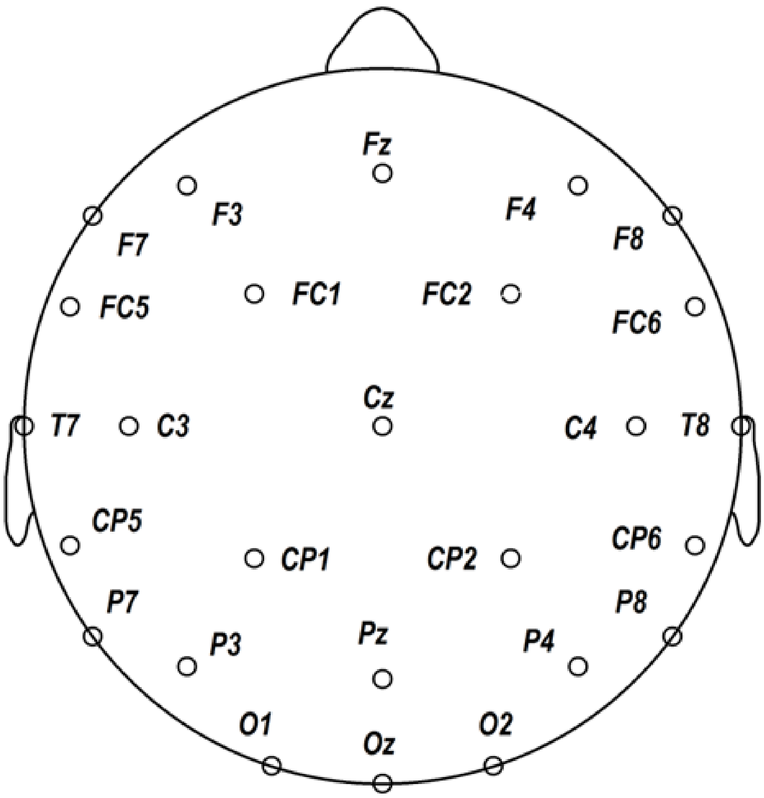
They were instructed to move or blink as little as possible while reading the third sentence. The order of experimental and filler trials was randomized in the beginning of the experiment.

The trial parameters were identical to those used by Metusalem et al.: Participants were presented with the first two sentences of a discourse in the middle of the screen. Once they had read and understood the two sentences, they pushed a button on the response device to advance to the third sentence. A red fixation cross in the middle of the screen appeared for 1000 ms and cued the beginning of the third sentence which was presented word by word (i.e., rapid serial visional presentation).The stimulus onset asynchrony for the words was 350 ms, divided into 200 ms presentation of the word and 150 ms inter-stimulus interval. Directly following the last word in the sentence, the comprehension question appeared in the middle of the screen requiring participants to provide a Yes or No response. Responses were given by pushing a button on the response device using either the right or left thumb. The left-right-Yes-No button assignment was counter-balanced across participants. The experiment consisted of five blocks which were separated by pauses. The first block consisted of 20 trials; the remaining four blocks consisted of 17 trials. The entire session took a little less than two hours.

EEG recording and analysis

Electroencephalography was recorded continuously from 26 active Ag/AgCl electrodes mounted in a cap according to the 10-20 system (Klem, Lüders, Jasper, & Elger, 1999). The signal was amplified by a Biosemi active amplifier with a bandpass filter of 0.016-100 Hz, sampled with a frequency of 250 Hz and referenced online to the left mastoid. Four additional electrodes were used to monitor participants' horizontal and vertical eye movements and blinks (see

Figure 4.1, for an overview of the electrodes' distribution over the scalp). All electrode impedances were kept below 5 kΩ.



Hemisphere	Right	F4, FC2, FC6, C4, CP2, CP6, P4, O2
	Left	F3, FC1, FC5, C3, CP1, CP5, P3, O1
Laterality	Lateral	F3, FC5, CP5, P3, F4, FC6, CP6, P4
	Medial	FC1, C3, CP1, O1 FC2, C4, CP2, O2
Anteriority	Prefrontal	F3, F4, FC1, FC2
	Frontal	FC5, FC6, C3, C4
	Parietal	CP1, CP2, CP5, CP6,
	Occipital	P3, P4, O1, O2

Figure 4.1: Layout of the 26 electrodes across the scalp. The table summarizes the electrode groupings for the distributional analyses in Experiments 1 and 2.

The ERP analysis was carried out on the experimental items of each participant individually, using Brain Vision Analyzer (version 2.0). Participants' EEG data were re-referenced offline to the average of the left and right mastoids and filtered again (highpass = 30 Hz, lowpass = .01 Hz). The EEG was time-locked to the onset of the three kinds of targets in the recording. Semi-automatic artifact rejection was used to exclude trials containing drifts, blinks, and muscle tension. The overall percentage of trials excluded due to artifact contamination was 8%, which was evenly distributed over the three conditions (expected targets = 7%, event-related targets = 9%, event-unrelated targets = 8%). Ocular correction was applied based on the average of each participant's vertical and the average of their horizontal eye electrodes (cf. Gratton, Coles, & Donchin, 1983). By-participant averages across the three conditions were calculated, relative to a 500-ms pre-stimulus baseline window. The individual participant averages were then averaged together yielding a grand average ERP for each condition.

The analysis of participants' accuracy on the comprehension questions suggested that they read and understood the discourses (mean accuracy = 97%, *SD* = 3).

Results

Figure 4.2 displays the grand average ERPs elicited by the target words in Experiment 1 for the 26 scalp electrodes. We plot 1000 ms post-stimulus onset and 500 ms pre-stimulus baseline. Visual inspection suggests no differences between the three waveforms prior to the onset of the target word presentation. With regard to the amplitude of the N400, a three-way split very similar to the pattern observed by Metusalem et al. arose around 300 ms after stimulus onset, extending to roughly 600 ms after stimulus onset. The N400 amplitude elicited by the expected targets is positive at the majority of the electrodes.

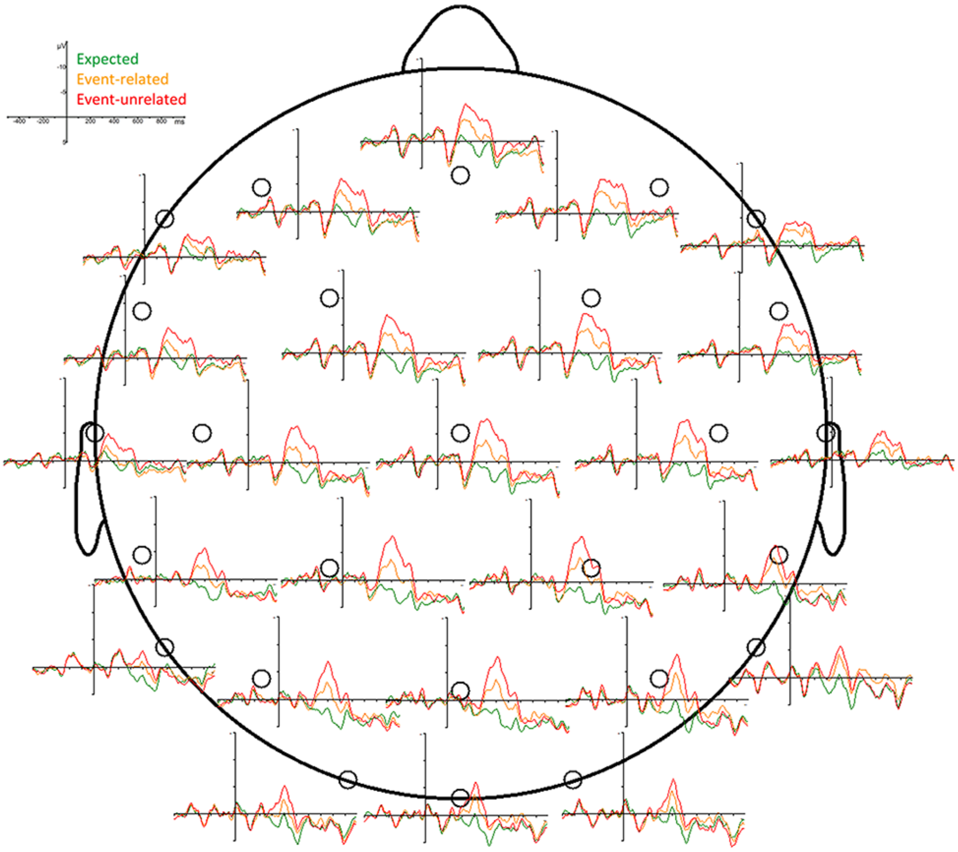


Figure 4.2: Grand average ERPs elicited by the target words in the three conditions in Experiment 1. Time zero refers to the onset of the target word presentation. Negative voltage is plotted up.

The N400 amplitude elicited by the unexpected event-unrelated targets was the most negative going deflection of the three conditions. Figure 4.2 indicates that the amplitude of the N400 elicited by the unexpected event-related targets lay between these two conditions (see Figure 4.3, for the three-way split in N400 amplitude on a central-parietal electrode, Pz).

To analyze the N400 amplitude differences statistically across the three conditions, we submitted the mean ERP amplitudes from 300 to 500 ms to a repeated measures ANOVA with three levels of Condition and 26 levels of Electrode. The analysis yielded main effects of Condition[3], $F(2,58) = 30.421$,

[3]We report p-values for Greenhouse-Geisser epsilon-adjusted degrees of freedom (Greenhouse & Geisser, 1959), the unadjusted degrees of freedom, and the value of the Greenhouse-Geisser epsilon for F-tests with more than one degree of freedom in the numerator.

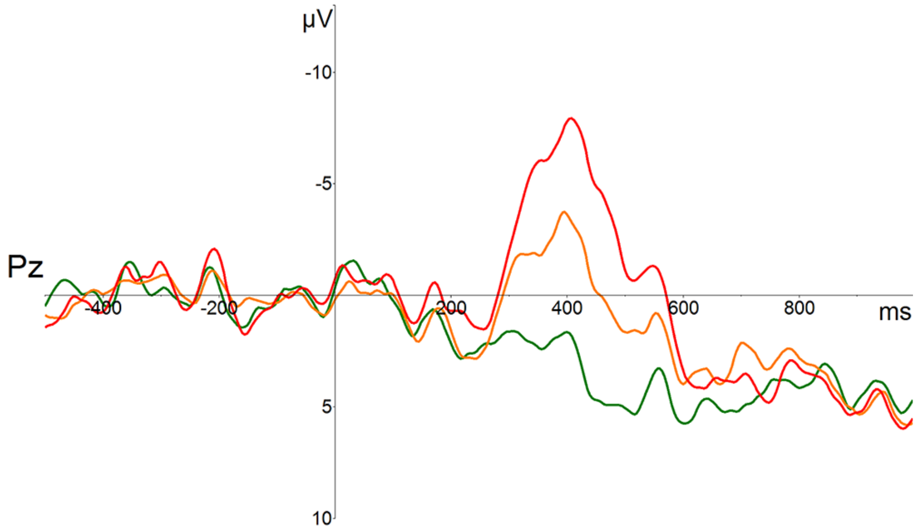


Figure 4.3: Grand average ERPs at the midline parietal electrode (Pz) in Experiment 1.

ε_{GG} = .924, p < .001, a main effect of Electrode, $F(25{,}725) = 9.637$, ε_{GG} = .140, p < .001, and a Condition-by-Electrode interaction, $F(50{,}1450)$ = 10.325, ε_{GG} = 0.18, p < .001. The subsequent planned comparison confirmed the visual inspection and thus the replication of Metusalem et al.'s findings. The N400 amplitude elicited by the event-related targets was between the amplitudes elicited by the expected and the unexpected event-unrelated targets. That is, N400 amplitudes in the event-unrelated condition were significantly greater (i.e., more negative-going) than those in the event-related condition ($F(1{,}29) = 17.461$, $\varepsilon_{GG} = 1$, p < .001; interaction with electrode: $F(25{,}725)$ = 5.225, ε_{GG} = .28, p < .001) and significantly smaller than those in the expected condition ($F(1{,}29) = 15.981$, $\varepsilon_{GG} = 1$, p < .001; interaction with electrode: $F(25{,}725) = 10.573$, ε_{GG} = .225, p < .001)

Figure 4.2 further suggests that the three-way N400 split pattern is widespread across the scalp, but is most prominently expressed over medial parietal-occipital sites. Figure 4.4 shows difference waves for the unexpected conditions (event-related minus expected and event-unrelated minus expected) demonstrating the size of the N400 effect at all 26 electrodes (see Figure 4.5, for

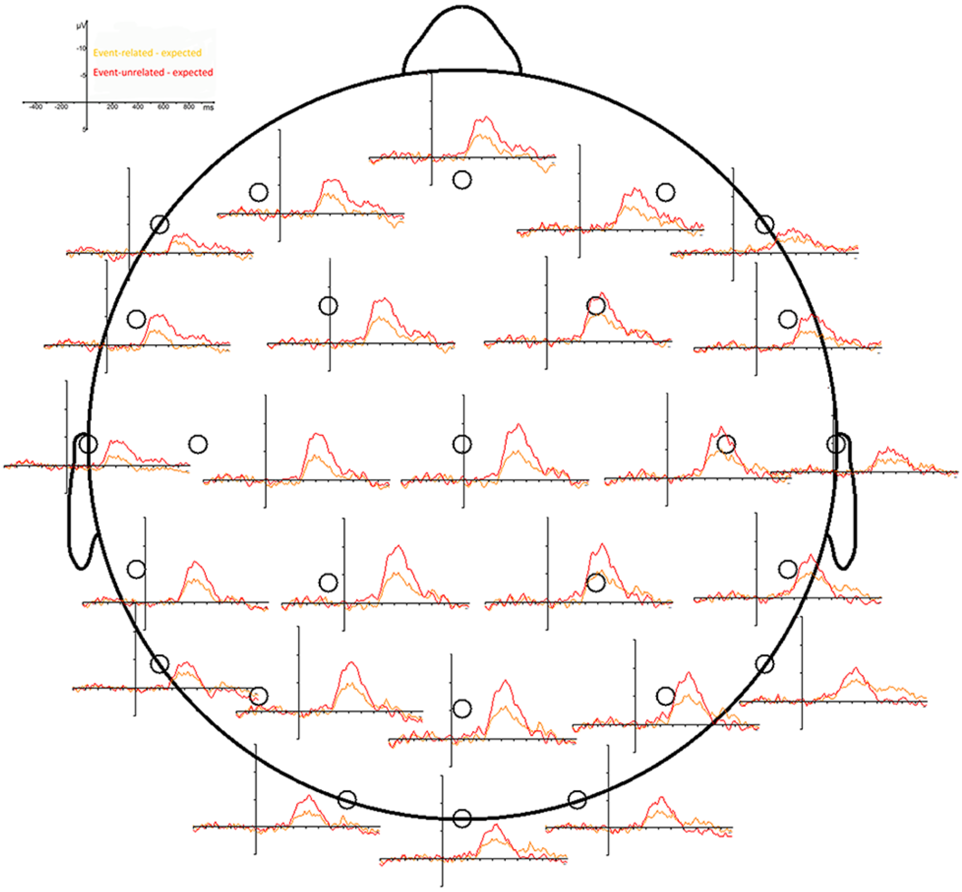


Figure 4.4: Difference waves reflecting the size of N400 effects in the event-related and event-unrelated conditions in Experiment 1.

scalp topographies of the mean amplitudes in the 300 to 500 ms time window for these conditions). To assess the exact topographic distribution, mean ERP amplitude difference scores between each unexpected condition and the expected condition during the 300 to 500 ms time window were submitted to a repeated measures ANOVA with two levels of Difference (event-related minus expected vs. event-unrelated minus expected), two levels of Hemisphere (left vs. right), two levels of Laterality (lateral vs. medial) and four levels of Anteriority (pre-frontal vs. frontal vs. parietal vs. occipital). The model revealed main effects of Difference ($F(1,29) = 20.681$, $\varepsilon_{GG} = 1$, $p < .001$), Laterality ($F(1,29) = 16.914$, $\varepsilon_{GG} = 1$, $p < .001$) and Anteriority ($F(3,87) = 4.18$, $\varepsilon_{GG} = .464$, $p < .035$), as well as interactions between Difference and Laterality

(F(1,29) = 5.093, $\varepsilon_{GG} = 1$, p < .032) and Difference, Laterality and Anteriority (F(3,87) = 15.454, ε_{GG} = .841, p < .001). The effect of Hemisphere was not statistically reliable (F(1,29) = 2.486, ε_{GG} = 1, p = .126). We explored the significant effects by means of Bonferroni-corrected post hoc tests. The main effect of Difference shows once more that the N400 components elicited by the unexpected event-unrelated target words were more negative-going than the one elicited by the unexpected event-related target words (p < .001). With regard to the factor Laterality, N400 amplitudes appear to be generally more negative over medial than over lateral sites (p < .001). The main effect of Anteriority indicates that N400 amplitudes were more negative over parietal-occipital electrodes (frontal vs. parietal: p < .001; parietal vs. occipital: p = .003). The Difference-by-Laterality interaction revealed that the difference in N400 amplitude between event-related and event-unrelated conditions was more positive over medial than over lateral sites (p < .001).

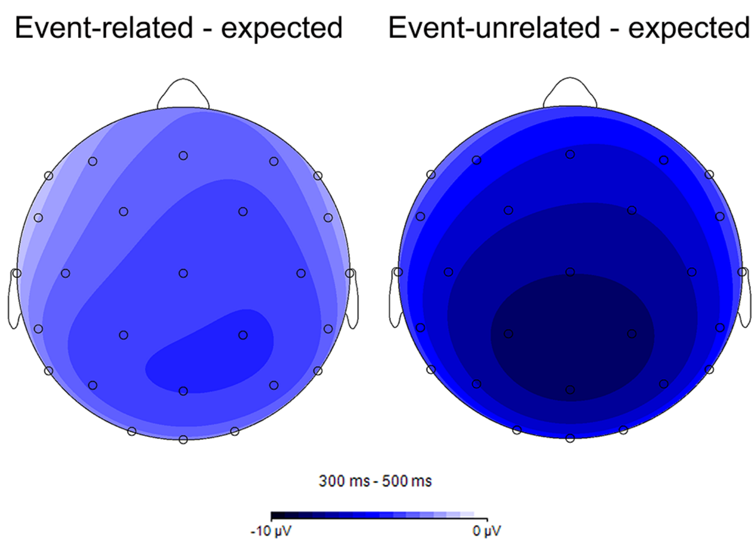


Figure 4.5: Scalp topographies of the N400 effects in the event-related and event-unrelated conditions from Experiment 1. The left plot reflects the N400 effect for the event-related targets, and the right the event-unrelated targets. Values correspond to mean amplitude, 300–500 ms post-stimulus onset in the respective condition.

Figures 4.2 and 4.3 suggest that the deflections elicited by the three kinds of target words may additionally differ in time windows preceding and following that of the N400 component. More specifically, Figure 4.2 indicates that, around 200 ms post stimulus onset, the deflections elicited in the expected condition and in the unexpected event-related condition were more positive than the deflection elicited in the unexpected event-unrelated condition. Previous research has linked variation in ERP amplitude during that time window (the so-called 'P200 component') to variance in the level of expectancy for particular lexical items. Federmeier and Kutas (2002) proposed that during reading, language users rely on contextual information in a sentence to prepare for the visual analysis of an upcoming stimulus. Roughly 200 ms post stimulus onset, expected words are assumed to elicit more positive-going amplitudes than unexpected words. Furthermore, similar to Metusalem et al. (2012, Experiment 1), Figure 4.2 suggests that the deflection elicited by the unexpected event-unrelated condition was more negative over frontal regions (with a slight preference for the left hemisphere) than the deflections elicited by expected and unexpected event-related conditions during the time window following the N400 component. Moreover, the deflection elicited by the expected target appears to be more positive than the deflections elicited by the other two conditions over right-hemispheric fronto-parietal regions.

Although we had no predictions about ERP differences between the three conditions in these time regions, we conducted two analyses. We entered mean amplitudes from 150 to 250 ms and from 500 to 900 ms, respectively to repeated measures ANOVAs with three levels of Condition and 26 levels of Electrode each. The results of the P200 analysis showed a main effect of Electrode ($F(25{,}725) = 24.087$, $\varepsilon_{GG} = .176$, p $<$ .001) but neither a main effect of Condition ($F(2{,}58) = 1.066$, $\varepsilon_{GG} = 1$, p $>$ .3) nor an interaction between Condition and Electrode ($F(50{,}1450) = 1.47$, $\varepsilon_{GG} = .264$, $p = .16$). The results of the second analysis, likewise, showed a main effect of Electrode ($F(25,725) = 17.29$, $\varepsilon_{GG} = .215$, p $< .001$) but not Condition ($F(2,58) = 1.859$, $\varepsilon_{GG} = .897$, $p = .17$). The Electrode-by-Condition interaction, however, was statistically reliable ($F(50,1450) = 3.163$, $\varepsilon_{GG} = .209$, $p = .001$). The presence of these effects is interesting, in particular because they closely resemble the effects reported by Metusalem and colleagues.

Discussion

The results show the three-way split pattern in N400 amplitude as observed by Metusalem et al. (2012). That is, the N400 amplitude elicited by the expected targets was more positive than the N400 amplitude elicited by the unexpected event-related targets and was more positive than the N400 elicited by the unexpected event-unrelated targets. Crucially, the N400 amplitude elicited by the unexpected event-related targets was less negative than the N400 elicited by the unexpected event-unrelated targets. These findings demonstrate the robustness of the prediction effect in a different language than the original English study. Our analyses revealed that the N400 attenuation in the unexpected event-related condition was most strongly visible over medial parietal-occipital electrodes. In sum, the results are consistent with the notion that event knowledge contributes to prediction during discourse reading.

Experiment 2

Given the close replication of the original experiment, we could assess the contribution of associative priming to the attenuated N400 amplitude in the unexpected event-related condition in Experiment 1. In Experiment 2, participants read short paragraphs, each consisting of three sentences. In each paragraph, the same passage-final sentence as in Experiment 1, including the same three kinds of target words, was used (e.g., "The baby liked baths, so she smiled when she was sprinkled with water/priest/dentist on her forehead"). We selected two prime words (e.g., "baptized", "church") from the event-establishing sentences in Experiment 1 that, from all the words in these sentences, were most strongly associated with the unexpected event-related target words (e.g., "priest"). The two primes were each placed in a neutral carrier sentence (see (4) for an example). We controlled that the sentences did not build up a coherent discourse context and that none of the other words in the two carrier sentences was associated with any of the target words in that paragraph or with the primes. Thus, reading the two introductory sentences in a given paragraph did not allow participants to construct a cohesive discourse model, yet, the primes were embedded in a meaningful syntactic environment. Thereby, we also minimized potential effects of strategic processing as the associative relationship between the critical targets and the primes became less obvious.

(4) "Dat Peter en Claudia de jongen lieten *dopen*, was het onderwerp van het gesprek. Telkens weer bleek het plaatje van de *kerk* in het geschiedenisboek zijn aandacht te trekken. De baby hield ervan om in bad te gaan dus ze lachte toen ze werd besprenkeld met water/pastoor/tandarts op haar voorhoofd."

That Peter and Claudia had the boy *baptized*, was the subject of the conversation. Time and again the image of the *church* in the history book appeared to get his attention. The baby liked baths, so she smiled when she was sprinkled with water/priest/dentist on her forehead.

If associative priming contributed to the N400 amplitude reduction in the unexpected event-related condition in Experiment 1, event-related targets in Experiment 2, which were preceded by associatively related primes in isolated carrier sentences, should elicit an attenuated N400 amplitude, relative to the

N400 amplitude elicited by the unexpected event-unrelated targets. Note that we generally predicted smaller effects in Experiment 2. As discussed in the Introduction, previous studies showed that discourse context exerts a powerful influence on the amplitude of the N400 component, whereas the effects of associations are rather subtle. As we minimized the influence of discourse context, we expected somewhat smaller differences between the N400 amplitudes elicited by the three kinds of target words. If, alternatively, the graded N400 pattern in the previous experiment was primarily driven by event knowledge, event-related and event-unrelated unexpected target words should elicit N400 components with similar amplitudes.

Method

Participants

Thirty-three volunteers (three male, mean age = 22, *SD* = 3) who did not take part in Experiment 1 or in any of the norming studies participated in Experiment 2. All were right-handed, native speakers of Dutch and did not report any history of learning or reading disabilities or neurological or psychiatric disorders. Due to an experimental error the logfiles of two participants were not saved. Another participant was excluded post-hoc due to too much noise in the EEG signal.

Stimuli and procedure

In each experimental item of Experiment 1 we replaced the two event-establishing sentences with two isolated sentences. Each of these contained a prime word. Prime words were selected from the event-establishing sentences and were most strongly associated with the unexpected event-related target word in that item. The association strength between a chosen prime word (e.g. verb, noun, adjective, preposition) and the three target words was determined using the Dutch free association database by de Deyne and colleagues (De Deyne et al., 2013). In case of inflections, the base form of that word was looked up. Thirteen (out of 132) prime words were not listed in the database. The average association strength between the remaining 119 primes and the unexpected event-unrelated target words was .0282 (SD = .0557; one response in 35 participants). The average association strength between the 119 primes and the expected targets and the unexpected event-unrelated targets was .013 (SD = .0351; one response in 77 participants) and .0002 (SD = .0013; one response in 5000 participants), respectively. The primes were embedded in neutral non-predictable carrier sentences and appeared in the same inflectional form as in the event-establishing sentences. None of the other words in the two carrier sentences was associatively related to the three target words within a given paragraph. Most importantly, we ensured that the two carrier sentence (and the target sentence) did not build up a coherent discourse context.

The fillers items used in Experiment 1 were edited by shuffling the event-establishing sentences and the target sentences across all paragraphs. By doing so, completely unrelated sentences were paired, which did not form a coherent discourse. Finally, we created new comprehension questions for all experimental and filler items. Within each item, the question focused on the contents of either the first or the second carrier sentence. The ratio of Yes and No responses was balanced. Apart from these changes, stimuli, procedure and analysis were identical to Experiment 1.

Results and discussion

The accuracy analysis of their responses to the comprehension questions showed that participants read the carrier sentences carefully and understood the content (mean accuracy = 90%, SD = 5).

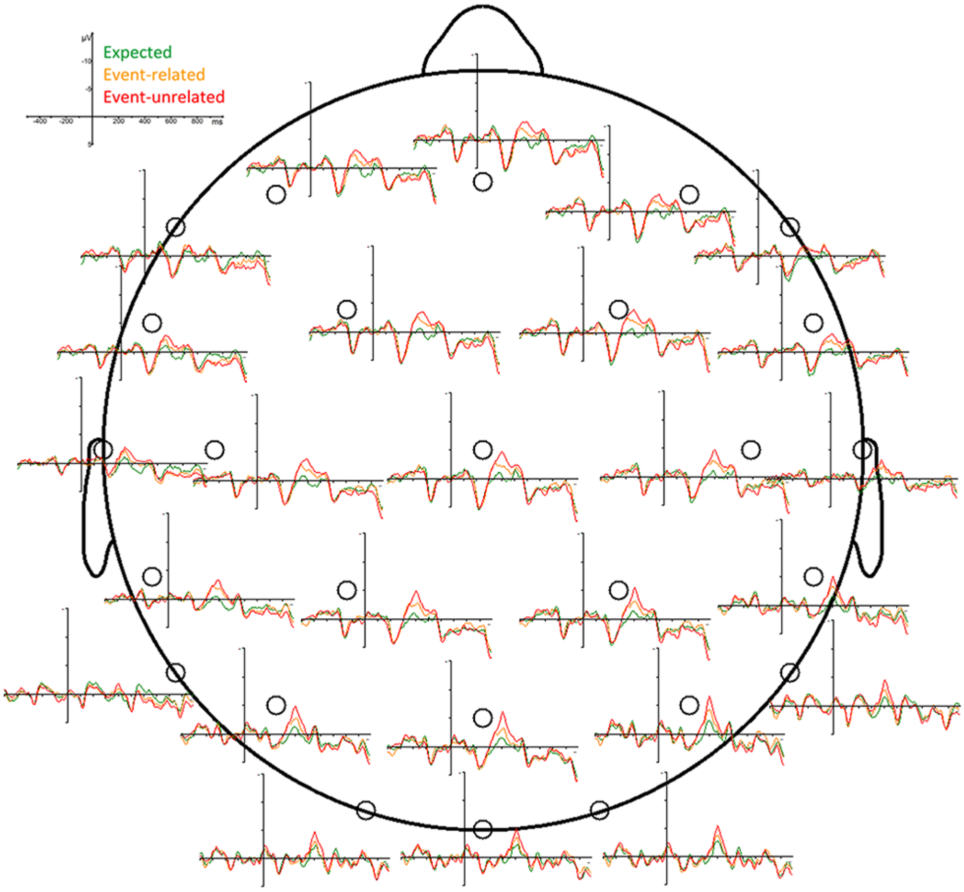


Figure 4.6: Grand average ERPs elicited by the target words in the three conditions in Experiment 2.

Figure 4.6 plots the grand average ERPs elicited by the three conditions in Experiment 2 for all 26 scalp electrodes (see Figure 4.7, for grand average ERPs at the Pz electrode in Experiment 2). Visual inspection suggests no differences between the three waveforms prior to the onset of target word presentation. However, at around 300 ms after target word onset, the lines diverged, extending to around 600 ms post stimulus onset. The N400 amplitude elicited by the unexpected event-unrelated targets was less negative than in Experiment 1. Moreover, the N400 amplitude elicited by the expected targets was more negative than in Experiment 1. This was expected, and both effects are most likely connected to the lower predictability of the target words in the expected condition. As we minimized the amount of event knowledge in the paragraphs by replacing the three coherent sentences in one paragraph with three isolated sentences, the expected targets were less predictable than in Experiment 1 and thus elicited a more negative-going N400 amplitude. In turn, as the expected target words were less predictable, the degree of expectation violation in response to the unexpected event-unrelated target words was smaller, too, hence the less negative N400 amplitude. Interestingly, visual inspection suggests that the N400 amplitude elicited by the unexpected event-related target words was similar to the amplitude elicited by the same condition in Experiment 1.

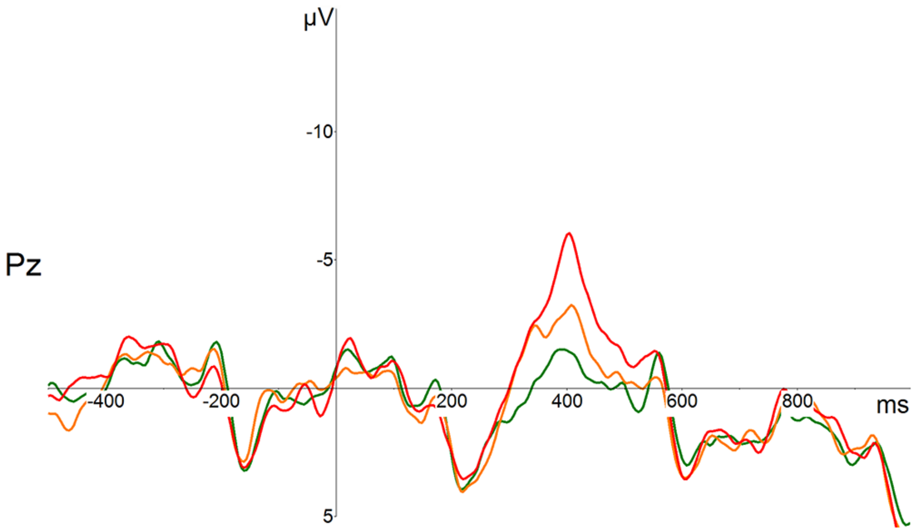


Figure 4.7: Grand average ERPs at the midline parietal electrode (Pz) in Experiment 2.

To statistically analyze the effects, mean ERP amplitudes were submitted to a repeated measures ANOVA with three levels of Condition and 26 levels of Electrode. This analysis yielded main effects of Condition, $F(2{,}58) = 12.252$, $\varepsilon_{GG} = .813$, p < .001 and Electrode, $F(2{,}58) = 4.772$, $\varepsilon_{GG} = .137$, p < .001 and a Condition-by-Electrode interaction $F(50{,}1450) = 1.848$, $\varepsilon_{GG} = 0.193$, $p = .041$. The planned comparison between event-related und event-unrelated unexpected conditions revealed the statistical reliability of the critical difference, $F(1{,}29) = 5.217$, $\varepsilon_{GG} = 1$, $p = .03$ (interaction with electrode: $F(25{,}725) = 1.044$, $\varepsilon_{GG} = .244$, p > .3; see Figure 4.8 for difference waves between the unexpected and the expected conditions). The unexpected event-related condition also differed from the expected condition ($F(1,29) = 9.465$, $\varepsilon_{GG} = 1$, $p = .005$; interaction with electrode: $F(25,725) = 2.993$, $\varepsilon_{GG} = .229$, $p = .01$).

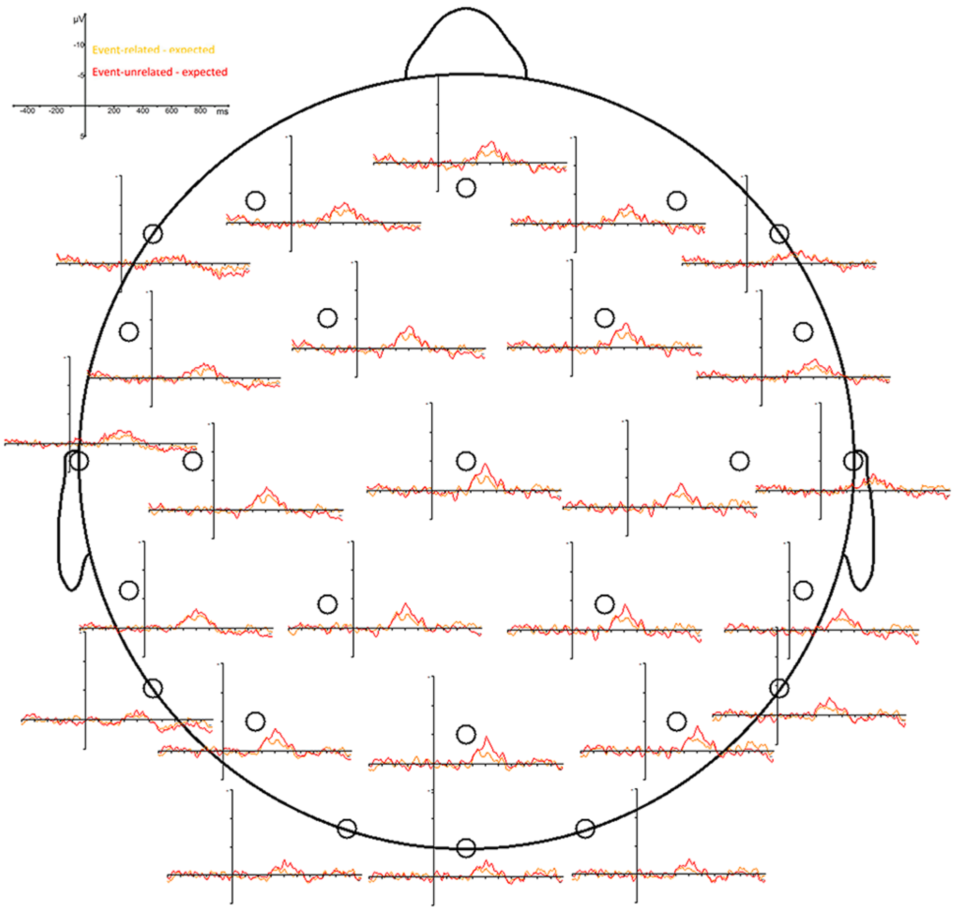


Figure 4.8: Difference waves reflecting the size of N400 effects in the event-related and event-unrelated conditions in Experiment 2.

We analyzed the scalp topographies of the N400 effects for the event-related and event-unrelated targets (see Figure 4.9) according to the same procedure as in Experiment 1. This yielded a significant main effect of Difference ($F(1,29) = 5.114$, $\varepsilon_{GG} = 1$, $p = .030$). While the main effect of Anteriority was trending towards significance ($F(3,87) = 2.348$, $\varepsilon_{GG} = .474$, $p = .123$), none of the other main effects or interactions were statistically reliable (all $p > .4$). Bonferroni-corrected post-hoc tests showed that N400 amplitudes were most negative over

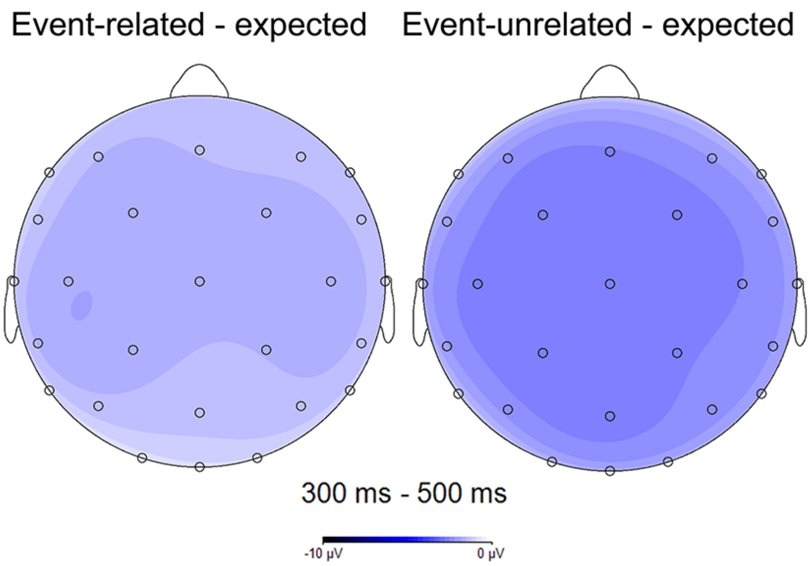


Figure 4.9: Scalp topographies of the N400 effects in the event-related and event-unrelated conditions from Experiment 2.

medial frontal-parietal electrodes (pre-frontal vs. occipital: p = .05, frontal vs. occipital: p = .024, parietal vs. occipital: p = .029; all other comparisons: p > .3).

In contrast to Experiment 1, the plot of the grand averages in Experiment 2 did not show any differences between the deflections elicited by each of the three conditions in the time windows preceding and following the N400 time window. This result is similar to Metusalem et al.'s Experiment 2, where participants read the passage-final sentences in insolation, and suggests that in particular the observed differences between the conditions in the late time window in Experiment 1 (and Metusalem et al.'s Experiment 1) were most likely connected to the presence of a coherent event discourse. In their General Discussion, Metusalem and colleagues speculate that these effects "might be linked in some way to the eliciting word's status as related or unrelated to the described event [and that] late effects elicited by semantic violations [might] in some way [be] modulated by the eliciting word's degree of relation to the described event" (p. 560; cf. Kuperberg, Caplan, Sitnikova, Eddy, & Holcomb, 2006; Nieuwland & van Berkum, 2005). As the current study was not designed to address this issue and our primary focus was on the N400 time window, our results do not allow us to make any conclusion about this speculation. Based on our manipulation, we are confident to conclude that simple associations between words did most likely not contribute to the late window effects. However, as far as their actual locus is concerned, future research is needed.

To summarize, using an associative priming manipulation, we obtained a similar three-way N400 split pattern as in Experiment 1: The N400 amplitude elicited by unexpected event-related target words lay between the ones elicited by the expected and unexpected event-unrelated target words.

General discussion

In the current study, we tested the contribution of simple word associations to prediction during discourse comprehension. Dutch participants read short paragraphs consisting of three sentences. In Experiment 1, the three sentences formed a coherent discourse context. While the first two sentences established an event scenario, the paragraph-final sentence contained three kinds of target words: A highly expected word, an unexpected word that was related to the discourse context, or an unexpected word that was not related to the context. In Experiment 2, the three sentences did not form a coherent discourse context. The same paragraph-final sentences were preceded by two isolated sentences, each containing a prime word that was associatively related to the critical targets and included in the event-establishing sentences of Experiment 1. In Experiment 2, we thus minimized the amount of event knowledge participants could extract from the preceding sentences while maintaining parts of the lexical material from Experiment 1.

The results of Experiment 1 replicated previous research (Metusalem et al. 2012, Experiment 1): Expected target words and unexpected event-unrelated target words elicited positive-going and strongly negative-going N400 amplitudes, respectively. The N400 amplitude elicited by the unexpected event-related target words was attenuated (e.g., less negative relative to the unexpected event-unrelated condition). These findings are consistent with the notion that comprehenders use event knowledge to generate predictions about upcoming input. Topographic analyses revealed that the N400 amplitudes elicited by both unexpected conditions were generally more negative over medial parietal-occipital electrodes, a region typically associated with the N400 and semantic processing (see Federmeier & Kutas, 2011, for discussion). Importantly, however, when the same targets followed sentences that did not establish a coherent event but contained associatively related prime words (Experiment 2), a similar N400 three-way split pattern was obtained as in Experiment 1. This suggests that associative priming between the critical words in the carrier sentences and the unexpected event-related targets affected the amplitude of the respective N400 component. Below, we discuss how associative priming may contribute to the prediction effects observed in cohesive discourse reading (e.g., the present Experiment 1; Metusalem et al., 2012, Experiment 1).

One possibility is that associations contributed only weakly to the prediction effects observed in the present Experiment 1 and in Metusalem et al. (2012, Experiment 1). That is, as has been suggested in earlier work (Ledoux et al., 2006; Metusalem et al., 2012), when part of a cohesive discourse, the influence of lexical-level relationships (such as associative priming) may be overridden by the effects of a higher-level discourse model (cf. Camblin et al., 2007). Based on the current data, we cannot rule out this possibility. However, why should one a priori exclude the contribution of word associations? An important outcome of the current study is that facilitation was obtained when the same prime words appeared in a cohesive discourse and when they appeared in isolated, disconnected sentences. The consequences of both event knowledge and associations thus appear to be consistent with one another, in that both pre-activated the unexpected event-related targets to some degree. Therefore, it seems rather ad hoc to preclude the influence of associations to the N400 reduction in Experiment 1 (and in Metusalem et al., 2012, Experiment 1). We believe our results are in line with the findings by Boudewyn et al. (2012) who, using a different experimental design, observed that associations contributed to prediction during discourse comprehension, though the effects were smaller and more fragile than the effects of discourse event context. In sum, while these arguments are not sufficient to reject the possibility that the effects of discourse context overrode the effects of simple word associations, we consider this to be unlikely.

Another possibility is that the attenuation of the N400 amplitude elicited by the unexpected event-related targets is primarily due to associative priming rather than reflecting the activation of generalized event knowledge. Recall that previous work suggests that effects of association may emerge particularly when the critical target words are incongruent with the context they appear in (Ledoux et al., 2006). The unexpected event-related targets constituted an anomaly within the context of the third sentence in the paragraphs, thus increasing the likelihood of influences of associative priming. Moreover, some indications for the claim that the N400 amplitude reduction reflects associative priming rather than event knowledge activation comes from a cross-modal lexical decision study by Kintsch and Mross (1985). Their participants listened to discourses, containing critical words and made word/non-word decisions to visually presented test strings (which constituted an anomaly at the point of presentation). The authors found priming effects for words that were associated with the target words, but at best marginally significant priming effects for words that were related to the described event. The results of the present Experiment 2 clearly show that when target words appear in incoherent sequences of sentences, participants exploit the associative relationships between the target words and associatively related words in the preceding two sentences. However, there are some conceptual and methodological differences between Kintsch and Mross' experiments and our study. Most importantly, Kintsch and Mross argue that associations impact word recognition only in a bottom-up fashion rather than predictively. In their experiments the critical words in the discourses were directly followed by the associatively related and event-related words, respectively. In a manipulation where they implemented a 150 ms lag of presentation between the critical words in the discourse and the two kinds of related words, no advantage in lexical decision times for related over unrelated words was observed. This finding is in conflict with our results, as we show that the effects of associative priming can persist even across sentence boundaries. One obvious explanation might be the different dependent variables. Measuring electrophysiological activity by means of the N400 component might be more sensitive to capture (online) word recognition than lexical decision times. Furthermore, to be able to compare Kintsch and Mross' study with the present experiments one would have to determine the summed association strength between words in the discourse and the associatively related, visually presented test words. One reason why we believe that associative priming persisted across sentence boundaries in our study is the summed influence of both prime words. That is, reading "baptize", which primes "priest", in the first (isolated) sentence and subsequently reading "church", which is also associatively related to "priest", in the second (isolated) sentence might boost the pre-activation of "priest" and might overcome the decay of the target due to intervening words. Further research is needed to confirm this hypothesis. Future experiments could, for example, investigate the interaction between the degree of prediction and the distance between associatively related words in discourse comprehension. However, we believe it to be very unlikely that the N400 amplitude reduction in the present Experiment 1 (and in Metusalem et al., 2012, Experiment 1) was primarily driven by associative priming rather than event knowledge.

A third possibility that has been suggested previously (Boudewyn et al., 2012; Van Petten, 1993) is that event knowledge and associations have additive effects. On such an account the prediction effect in Experiment 1 would reflect the joint contribution of event knowledge and associative priming to anticipatory processing in discourse comprehension. The prediction effect in Experiment 2 would largely reflect the contribution of associative priming.

Such an argumentation implies that the difference between the event-related and the event-unrelated unexpected conditions should be larger in Experiment 1 than in Experiment 2 (compare plots in Figure 4.10). We assessed this hypothesis by means of a post hoc ANOVA. Difference scores between event-related and expected conditions and event-unrelated and expected conditions, respectively, of both experiments, were submitted to a repeated measures ANOVA with the factors Difference (event-related minus expected conditions vs. event-unrelated minus expected conditions; within-participants), Electrode (26 levels; within-participants) and Experiment (Experiment 1 vs. Experiment 2; between-participants). We obtained significant main effects of Difference ($F(1{,}29) = 22.671$, $\varepsilon_{GG} = 1$, $p < .001$), Electrode ($F(25{,}725) = 15.038$, $\varepsilon_{GG} = 245$, $p < .001$) and Experiment ($F(1{,}58) = 10.03$, $p = .002$), and interactions between Electrode and Difference ($F(25) = 5.334$, $\varepsilon_{GG} = .297$, $p < .001$), Electrode and Experiment ($F(25{,}1450) = 5.1$, $p < .001$) as well as

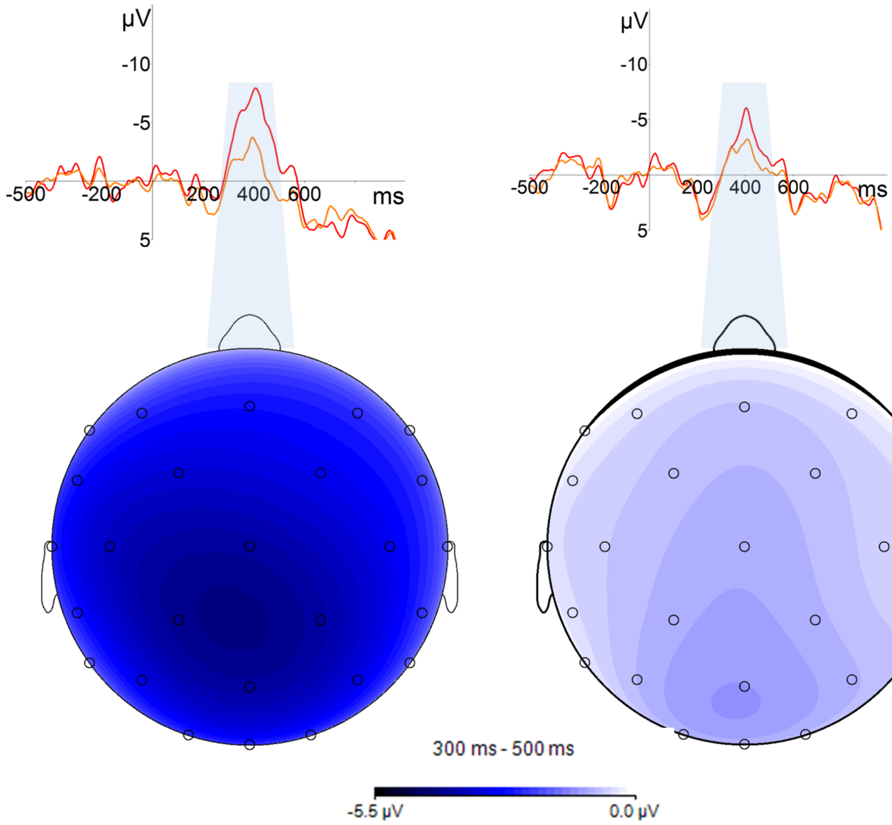


Figure 4.10: Scalp topographies of the mean difference between the event-unrelated (red) and event-related (orange) conditions (indicated by two waves taken from the Pz electrode) in Experiment 1 (left) and Experiment 2 (right).

Difference and Experiment ($F(1,1) = 6.981$, $p < .011$). The main effects and interactions were explored by post hoc Bonferroni-corrected comparisons. The main effect of Difference echoes the previous analyses and highlights the consistent difference between event-related and event-unrelated conditions in both experiments ($p < .001$). The main effect of Experiment shows that N400 amplitudes were generally more negative in Experiment 1 than in Experiment 2. The Experiment-by-Difference interaction indicates that N400 amplitudes for event-related and event-unrelated conditions were more negative in Experiment 1 than in Experiment 2 (event-related: $p = .05$; event-unrelated: $p < .001$).

Consistent with the notion that in Experiment 1 both event knowledge and associations contributed to prediction, the difference between both unexpected conditions was larger in Experiment 1 than in Experiment 2. Recent experimental findings using other methodologies support this notion and suggest that associations and event knowledge/combinatorial mechanisms underlie prediction. Kukona et al. (2011) used visual world eye-tracking (see Huettig et al., 2011, for review) to contrast the influence of associative priming with the influence of event-based context on prediction during sentence comprehension. Their participants listened to sentences containing a critical verb (e.g., "arrest") such as "Toby arrests the crook" while looking at visual scenes which included verb-related agents and patients (e.g., a policeman and a crook). The authors observed anticipatory eye movements to both agents and patients although the agent role had already been filled (Toby). Kukona and colleagues concluded that anticipatory eye gaze was influenced by simple associative relationships between the words (e.g., arrest-policeman) and combinatorial event information. Crucially, associative priming showed an effect even though it conflicted with the event built up by the sentential context.

To conclude, we contrasted the contribution of event knowledge and simple word associations to prediction using event-related brain potentials. We observed that contextually unexpected target words that were related to an event description (which included the associatively related words) elicited a reduced N400 amplitude compared to contextually unexpected target words that were unrelated to the event. Crucially, a similar N400 attenuation was observed when the discourse context did not allow for the build-up of a coherent event but included the associatively related words. As the difference between event-related and event-unrelated conditions was larger when the sentences formed a coherent event than when they did not, our results suggest that associative priming alone cannot account for the N400 pattern observed in our Experiment 1. However, because part of the effect remained in Experiment 2 the findings fit best with the notion that during discourse reading both event knowledge activation and simple word associations jointly contribute to the pre-activation of up-coming words. These findings add to a growing body of literature arguing that one-mechanism accounts are not sufficient to explain how comprehenders use predictive cues in the input signal to generate predictions about upcoming language. The data suggest that prediction during discourse reading is driven by multiple mechanisms.

5 | Prediction and production of simple mathematical equations: Evidence from visual world eye-tracking[1]

Abstract

The relationship between the production and the comprehension systems has recently become a topic of interest for many psycholinguists. It has been argued that these systems are tightly linked and in particular that listeners use the production system to predict upcoming content. In this study, we tested how similar production and prediction processes are in a novel version of the visual world paradigm. Dutch speaking participants (native speakers in Experiment 1; German-Dutch bilinguals in Experiment 2) listened to mathematical equations while looking at a clock face featuring the numbers 1 to 12. On alternating trials, they either heard a complete equation ("three plus eight is eleven") or they heard the first part ("three plus eight is") and had to produce the result ("eleven") themselves. Participants were encouraged to look at the relevant numbers throughout the trial. Their eye movements were recorded and analyzed. We found that the participants' eye movements in the two tasks were overall very similar. They fixated the first and second number of the equations shortly after they were mentioned, and fixated the result number well before they named it on production trials and well before the recorded speaker named it on comprehension trials. However, all fixation latencies were shorter on production than on comprehension trials. These findings suggest that the processes involved in planning to say a word and anticipating hearing a word are quite similar, but that people are more aroused or engaged when they intend to respond than when they merely listen to another person.

---

[1]Adapted from *Hintz, F., & Meyer, A.S., (2015). Prediction and production of simple mathematical equations: Evidence from visual world eye-tracking. PLoS ONE 10(7): e0130766. doi:10.1371/journal.pone.0130766.*

Introduction

An important issue in psycholinguistics is the relationship between the language production system and the language comprehension system. Although language users draw upon both systems when communicating with one another, the two systems have mainly been studied independently of each other. Recently, however, a growing number of researchers have advocated the view that production and comprehension are tightly integrated (Dell & Chang, 2014; Pickering & Garrod, 2007, 2013; MacDonald, 2013). The discussion of the production-comprehension interface has often focused on prediction. Research on language comprehension has established that readers and listeners often anticipate upcoming information (e.g. Altmann & Kamide, 1999; DeLong et al., 2005; Federmeier & Kutas, 1999; van Berkum et al., 2005) and that anticipation contributes to the speed with which they comprehend language. It is often assumed that predicting a word during comprehension is basically the same process as planning to say a word aloud. Scientists who argue for an integration of the comprehension and the production systems have also suggested that speakers, predict their own utterances and compare these predictions to the actual outcomes (Pickering & Garrod, 2014).

In the current study, we explored the similarity of word prediction during comprehension to word production using a novel version of the visual world paradigm (Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995). Although there is correlational evidence for the involvement of production-based mechanisms in language comprehension, to our knowledge no study has directly compared the behavioral consequences of word prediction and word planning. We first summarize the key characteristics of word production and prediction processes. We then turn to recent proposals that suggest an integration of production and comprehension (i.e., prediction-by-production accounts) and discuss relevant empirical findings. Finally, we report two eye-tracking experiments in which we compared participants' eye movements reflecting their preparation to speak to eye movements reflecting their prediction of an upcoming word.

Word production

Various models of word production have been proposed (e.g. Caramazza, 1997; Dell, 1986; Levelt, Roelofs, & Meyer, 1999; Rapp, Buchwald, & Goldrick, 2013). Although they differ in their assumptions about the representations accessed when words are produced and about the processes involved, they agree that word production involves three main steps: speakers decide which concepts to refer to, select suitable words from the mental lexicon, and build up the corresponding word forms. For instance, the model proposed by Levelt, Roelofs, and Meyer assumes three stages: conceptual preparation, lemma selection, and word form encoding (see also Levelt, 1999; Roelofs, 1997). The last-mentioned stage includes morphological, phonological, and phonetic encoding. During conceptual preparation, the speaker decides which concepts to encode. From the conceptual level, activation spreads to grammatical word units (lemmas), which are selected and ordered during lemma selection. Lemma selection is followed by the retrieval of the corresponding morphemes and phonological segments. The retrieved segments are combined into syllables. Based on the syllabified phonological representation a phonetic representation is created and, finally, articulatory commands are generated and executed. Speakers monitor their speech planning at the conceptual and phonological level and their overt speech for accuracy and appropriateness.

Word prediction

The brain has sometimes been said to be essentially a "prediction machine" (A. Clark, 2013; Bar, 2007; Gilbert & Wilson, 2007; Schacter, Addis, & Buckner, 2007) and many authors have proposed that prediction plays an important in language comprehension (e.g. Kutas et al., 2011; van Berkum, 2010; Huettig, 2015, for a recent review). This view is well supported by experimental evidence. For instance, using recordings of event-related brain potentials, it was shown that semantic/conceptual information about upcoming language can be predicted (Federmeier & Kutas, 1999; Federmeier et al., 2002; Szewczyk & Schriefers, 2013) as well as the grammatical gender of words (van Berkum et al., 2005; Wicha et al., 2003). To give a final example, there is experimental evidence indicating that listeners can predict the phonological forms of upcoming words (DeLong et al., 2005; Gagnepain, Henson, & Davis, 2012). Altmann and Mirković (2009) suggest that the comprehension system is "maximally incremental" in the sense that "it develops the fullest interpretation of a sentence fragment at each moment of the fragment's unfolding" and at all possible levels (p. 604).

Prediction-by-production

Many authors have considered the possibility that predictions during comprehension are generated by mechanisms drawing upon knowledge also involved in speech production (e.g. van Berkum et al., 2005; Federmeier, 2007; Garrett, 1980; McCauley & Christiansen, 2011). This view has been explicitly implemented in two recent integrative frameworks. Dell and Chang (2014) developed a model of sentence production, comprehension, and language acquisition where predicting the next word of an utterance is akin to planning to produce that word (see also Chang et al., 2006). Similarly, Pickering and

Garrod (2013) proposed that predictions during comprehension can be driven by an associative route, which is grounded in the comprehension system and by a simulation route, which engages the production system. With regard to the latter route, the authors suggest that language users construct forward models during production and comprehension, to predict their own utterances and to predict upcoming utterances by other speakers. In both cases, the "predictions are not the same as implemented production representations but easier-to-compute 'impoverished' representations" (p. 339). Pickering and Garrod propose that both routes (i.e., prediction-by-association and prediction-by-production) can be used flexibly to predict information during comprehension.

Dell and Chang's and Pickering and Garrod's frameworks have in common that they equate the activation of word knowledge during prediction for comprehension with the activation of word knowledge during word planning. That is, the anticipation of the meaning, grammatical characteristics, or phonological form of upcoming words is equated with the activation of this information for speaking.

Empirical evidence for production-based mechanisms in prediction

Several studies have reported correlational evidence for the involvement of production-based mechanisms in prediction and hence for a link between the comprehension and the production systems. Federmeier and colleagues (Federmeier et al., 2002; 2010) found a significant correlation between participants' prediction-related ERP components and their production fluency, as measured in a verbal fluency task. Furthermore, Mani and Huettig (2012) observed that the production vocabulary size of two-year old toddlers predicted the degree to which the toddlers anticipated upcoming target words. These studies provide indirect evidence supporting prediction-by-production. However, an important further step towards understanding the involvement of production-based mechanisms in prediction is to compare word prediction directly to word planning processes carried out under identical circumstances.

To that end, we used a novel version of the visual world paradigm, which has previously been used to study prediction during comprehension (Altmann & Kamide, 1999; Altmann & Mirković, 2009; Knoeferle et al., 2005) and language production (Griffin & Bock, 2000; Meyer, Sleiderink, & Levelt, 1998; Huettig et al., 2011, for review). In this paradigm, participants' eye movements are recorded while they view displays (e.g., showing a boy, a cake, and other objects) and hear sentences (e.g., "the boy will eat the cake", Altmann & Kamide, 1999) or produce utterances referring to the display. The eye movements indicate when the participants direct their attention to different parts of the displays and can, for instance, reveal whether or not listeners anticipate specific words (e.g., look at the cake in the above example before it is actually mentioned). In our experiments, Dutch speaking participants looked at the picture of an analogue clock face featuring the numbers 1 to 12. On half the trials, they listened to recordings of a person solving simple mathematical equations including the numbers 1 to 12, saying for instance "drie plus vijf is acht" (three plus five is eight). On the remaining trials, the recording stopped after "is", and the participants had to supply the solution of the equation. Listening and speaking trials alternated. In both tasks participants were asked to fixate on the relevant numbers on the clock face. After having carried out the computation, participants could predict what the recorded speaker would say and they could initiate the word planning process for their own production of the result number. We were interested in the similarity of the eye movements related to these processes.

We used spoken mathematical equations as materials because they allow for tight experimental control of variables such as word frequency and semantic associations. Moreover, in these utterances the final word is entirely predictable from the preceding context, yet different from trial to trial. Evidently, mathematical equations are not produced very often in everyday life. However, they are grammatically well-formed utterances. Producing and comprehending equations undoubtedly relies on grammatical, lexical, and phonological processes that are also involved in processing other types of utterances, such as descriptions of events and scenes, and therefore can be used to investigate these processes. Indeed, this has been done in several earlier studies. For instance, Ferreira and Swets (2002) used the production of equations to study the scope of advance planning in sentence production; Bock, Irwin, Davidson, and Levelt (2003), Korvorst, Roelofs, and Levelt (2007) and Kuchinsky, Bock, and Irwin (2011) used time telling to study the mapping of conceptual information onto linguistic structures; and Scheepers and colleagues (Scheepers et al., 2011; Scheepers & Sturt, 2014) used priming between complex sentences and mathematical equations to study the involvement of shared processes and representations in arithmetic and sentence processing. Here, we exploited the simplicity of the lexical content of equations and the predictability of the result numbers to assess the involvement of prediction in speech planning and comprehension.

We recorded the participants' eye movements throughout the experiment. We expected that in both tasks participants would follow the instructions and fixate upon each of the numbers they heard soon after word onset. In the production task, they should compute the result as soon as they had heard the second number, direct their gaze to the corresponding number on the clock face, and produce the response. We expected that the participants would initiate the shift of gaze to the appropriate location as soon as they had derived the number concept (rather than after having completely planned the utterance) and that they would therefore begin to fixate the number some time before the onset of their response. This coordination of eye gaze and speech planning would allow them to look at the response number while retrieving the corresponding verbal expression, which may facilitate these linguistic encoding processes (cf. Meyer, van der Meulen, & Brooks, 2004). In the listening condition, the participants might simply follow the listener, i.e., fixate upon each of the three numbers after it has been named. Alternatively, they could anticipate the result by computing it in the same way as on production trials. This would be consistent with the view that listeners engage their production system when they listen to another speaker and use it to predict what the speaker will say next (Dell & Chang, 2014; Pickering & Garrod, 2013). If the participants carry out the same computations and engage the speech production system in the same way on production and comprehension trials, their eye movements should not differ between the two conditions. A third possibility is that the prediction of the result number on comprehension trials is based not only on the engagement of production-based processes but is also supported by fast associative processes (cf. Kukona et al., 2011; Kuperberg, 2007). In that case, one might expect faster eye movements to the result numbers on comprehension than on production trials.

To anticipate the main results of Experiment 1, we found that the participants' eye movements on production and comprehension trials were very similar, and that, specifically, they anticipated the result numbers on comprehension trials. Experiment 1 was carried out with native speakers of Dutch, who listened to and completed utterances in their native language. In Experiment 2, we asked German-Dutch bilinguals to carry out the same production and comprehension tasks as in Experiment 1 in Dutch, their second language. Previous research has shown that even in highly proficient late bilinguals, linguistic processing is slower, and presumably more effortful, than in native speakers (e.g. La Heij, 2005; Costa, Caramazza, & Sebastian-Galles, 2000; Spalek, Hoshino, Wu, Damian, & Thierry, 2014; van Heuven & Dijkstra, 2010). The goal of Experiment 2 was to determine whether the participants would still anticipate the result number or whether, given the higher linguistic processing load, they would simply follow the recorded speaker.

Experiment 1

Method

Participants

Twenty-five native Dutch participants (five male; mean age = 22 years, SD = 3), mostly students of Radboud University Nijmegen, participated in the experiment. All participants had normal or corrected-to-normal vision and hearing. None reported any signs or a history of developmental speech disorders. One participant had to be excluded from the sample because s/he mentioned during the debriefing that s/he had been diagnosed with dyscalculia.

Ethics statement

All participants signed informed consent beforehand and were paid for their participation. Ethical approval of the study was provided by the ethics board of the Social Sciences Faculty of Radboud University.

Materials and design

We constructed 60 stimulus sentences. The sentences were simple mathematical equations including the numbers one to twelve, using addition and subtraction (e.g., 1 + 5 = 6, spoken Dutch sentence: “Een plus vijf is zes”). Repetitions of numbers (as in 2 + 2 = 4 or 6 - 3 = 3) were avoided. The 60 sentences were spoken at a normal speech rate and with normal intonation contour by a native female speaker of Dutch. Recordings were taken in a sound-shielded booth sampling at 44 kHz (16 bit resolution). A second version of each equation was created by manually cutting off the result number at the offset of "is". The complete versions of the equations served as comprehension items. The incomplete versions of the equations served as production items. The mean length of the comprehension recordings was 4680 ms (SD = 234); the mean length of the production counterparts was 3860 ms (SD = 180). Onsets and offsets of all words in the spoken equations were marked using Praat (Boersma, 2002). We designed the picture of a round clock face featuring the numbers from one to twelve in their customary positions.

All participants were presented with all 120 items. The experiment was divided into four equal blocks, divided by short pauses. A pseudo-random order of the trials was generated. The constraints on the randomization were that production and comprehension trials alternated, the production and the comprehension versions of a given equation occurred in different blocks, and that successive items did not have the same result number.

Procedure

The experiment was administered using an EyeLink 1000 system (SR Research) sampling at 1000 Hz. Participants placed their heads on a chin rest facing the computer screen 75 cm from the screen. Participants were instructed via the computer screen and invited to ask clarification questions. The eye-tracking system was calibrated and then the experiment began.

At the beginning of each trial, a black dot was presented on a white background in the middle of the screen for one second. Participants were asked to fixate the dot. This served as drift correction and ensured that participants always fixated the same position at the beginning of a trial. Subsequently, the clock face appeared in the center of the screen at a 600 x 600 pixel resolution, coinciding with the onset of the spoken sentence. The clock face remained in view during the entire trial. The trial duration for comprehension trials was 6000 ms (composed of the duration of the recording, on average 4680 ms, and individual time-outs, on average 1320 ms). The trial duration for production trials was 5500 ms (composed of the incomplete recordings, on average 3680 ms, individual time-outs of on average 1320 ms and 500 ms for participants' oral response). The participants were instructed to listen to the utterances and provide the result number on every second trial. They were also instructed to move their eyes to the numbers mentioned by the speaker as quickly as possible. This instruction was needed because a pilot study had shown that without such instruction participants would often fixate the center of the screen throughout the trial. Participants' responses to the production trials were recorded and coded during the experiment. The experiment lasted approximately 20 minutes.

Data coding and dependent variables

We excluded comprehension trials from the analysis on which participants uttered the result number by mistake (34 trials in total, $< 1\%$ of all comprehension trials). Participants' speech onsets on production trials were hand-coded using Praat (Boersma, 2002). As the participants could begin to compute the result as soon as they had heard the second number of the recording, we defined the speech onset latency as the time period between the onset of that number and the onset of the participants' response. Production trials were excluded when the response was incorrect or the speech onset latency deviated by more than 2.5 standard deviations from the participant's mean onset latency (66 trials in total, $< 1\%$ of all production trials).

For the eye movement analyses, the data from the participants' left or right eye (depending on the quality of the calibration) were analyzed and coded in terms of fixations, saccades, and blinks, using the algorithms provided in the EyeLink software. To determine how long each number was fixated for, regions of interests (90 x 90 pixels) were defined around each of the three numbers relevant for each equation. Two time windows were defined for the analyses: The time window for fixations to the first number started at the onset of the recording and lasted until the onset of the second number (on average 1686 ms, $SD = 135$). The time window for the second number and result number started at the onset of the second number and ended at the offset of "is" (on average 1957 ms, $SD = 144$). Note that this window only included fixations preceding the onset of the third number (i.e. predictive fixations in the comprehension condition, and fixations related to the preparation of the response in the production condition). For the statistical analyses we summed the fixation durations in the critical time windows and log-transformed the resulting total fixation durations.

We also calculated the fixation latencies for the first, second, and result numbers. As it takes about 200 ms to program and launch a saccadic eye movement (Saslow, 1967), we consider the fixation latency to be the onset of the first fixation to a region of interest with a latency of 200 ms or more after the onset of the relevant time window (i.e., the onset of the utterance for fixations to the first number, and the spoken onset of the second number for fixations to the second and third number). Fixation latencies were log-transformed before analysis.

Results

The analysis of the participants' speech onset latencies on production trials showed that they produced the result slightly later (by 85 ms) than the recorded speaker (means: 2128 ms, $SD = 316$ versus 2043 ms, $SD = 200$). This close match in the latencies is important because it facilitates the comparison of the participants' eye movements in the production and comprehension conditions.

Figure 5.1 shows a time-course graph plotting the proportions of fixations to the first, second, and result number on production (blue) and comprehension (red) trials. Fixations are plotted backwards from the offset of "is" in the recording (time zero) to the onset of the first number. The vertical line, 1686 ms before time zero, indicates the average onset of the second number. Recall that each participant heard a given equation twice, once as a production version and once as a comprehension version. The fixation proportions indicate the proportion of trials (out of all relevant trials) on which participants fixated the first (dotted lines), second (dashed lines) and result number (solid lines), respectively, at that moment in time. We computed by-participant confidence intervals (95%) for each of the average fixation lines at every sampling step (1 ms) to indicate the variation in participants' fixation behavior. The area between the lower and upper bounds is shaded in light gray for production trials and in dark gray for comprehension trials. The graph shows that soon after utterance onset participants began to fixate the first number mentioned in the recording. Shortly after the onset of the second number, they stopped looking at the first number and started fixating the second number. Participants' likelihood of fixating the result number increased already before the offset of "is", that is, well before the result was named by the recorded speaker or before participants named it themselves. This behavior reflects that participants had calculated the solution of the equations. In the production condition participants began to look at the result number around one second after the onset of the second number, roughly 600 ms prior to the offset of "is". In the comprehension condition, fixations to the result number began about 100-200 ms later, but well before the onset of the spoken result number; on average,

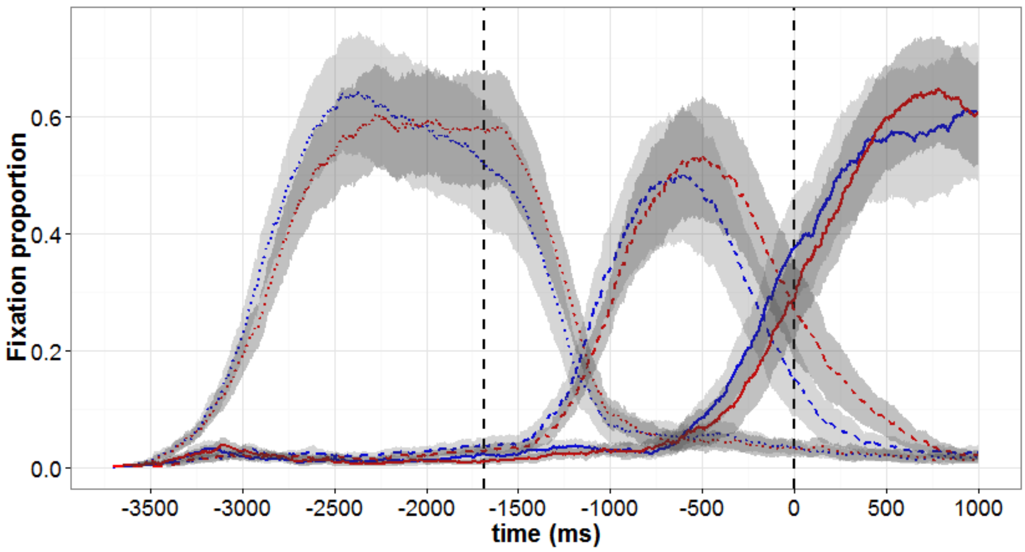


Figure 5.1: Fixation proportions Experiment 1. The graph plots participants' average fixation proportions to first (dotted lines), second (dashed lines) and result numbers (solid lines) for production (blue) and comprehension (red) conditions. Fixations are plotted backwards from the offset of "is" (time zero) to the onset of the first number. The first vertical dotted line represents the average onset of the second number. The areas shaded in gray represent the space in between the lower and upper bounds of the 95% by-participant confident intervals.

the participants' gaze landed on the result number 274 ms before its spoken onset.

Log-transformed total fixation durations for the time period between the onset of the second number and the offset of 'is" in the recording were analyzed using linear mixed-effects regression models in R (R-Core-Team, 2012), using the lme4 library (Bates, Maechler, & Bolker, 2012). Mixed-effect models allow for simultaneous inclusion of participants and items as random factors (Baayen, Davidson, & Bates, 2008). The full model included the fixed effect of Condition (comprehension vs. production) and the maximal possible random effects structure (see Barr, Levy, Scheepers, & Tily, 2013, for discussion), consisting of random intercepts and slopes for Condition by participant (N =

24) and item (N = 60). This model was compared to the same model without the fixed effect Condition using a likelihood test. Including Condition improved the model fit significantly, $\chi^2(2) = 9.091$, $p = .003$. The full model revealed that, during the critical time window, participants looked more at the result numbers in the production as compared to the comprehension condition (production mean = 233 ms, *SD* = 253 vs. comprehension mean = 184 ms, *SD* = 240), β = -.638, *SE* = .196, t = -3.253 ($t > |2|$ were considered significant; Baayen et al., 2008). This result is complemented by the analysis of the fixation latencies, which showed that the participants' mean fixation latency for the result number was on average 143 ms shorter on production trials (1628 ms, *SD* = 598) than on comprehension trials (1771 ms, *SD* = 673). Log-transformed fixation latencies were submitted to a mixed effect model which was, apart from the dependent variable, identical to the model used for the fixation duration analysis. The model revealed the statistical robustness of the effect of Condition (β = -.07, *SE* = .024, t = -2.92; $\chi^2(2) = 7.762$, $p = .005$).

Although the participants' eyes landed somewhat later on the result number on comprehension than on production trials, there is strong evidence that the participants anticipated the result numbers on comprehension trials instead of following the speaker. This can be appreciated by considering the fixation latencies measured from the onset of the spoken result number. Given that it takes at least 200 ms to recognize a spoken word (Allopenna, Magnuson, & Tanenhaus, 1998) and a further 200 ms to initiate and launch a saccadic eye movement (Saslow, 1967), saccades triggered by the spoken result number must have latencies of at least 400 ms. The observed average fixation latency was -284 ms. Thus, the participants began to look at the result number well before it was named by the recorded speaker. This was true for all speakers, with the slowest participant having a mean fixation latency of 150 ms.

In a supplementary analysis, we explored whether the participants' performance in the two tasks changed over the course of the experiment. To that end, we split the sequence of experimental trials into five blocks of 24 trials each and computed each participant's average fixation latency for the result numbers on comprehension and production trials. As one might expect, the participants' latencies were longer in the first block than in subsequent blocks. There was no consistent performance change across blocks 2 to 5. An analysis including block as an additional variable yielded no interaction with other variables.

We also compared fixation durations and fixation latencies between production and comprehension conditions for first and second number fixations (first number means: production = 892 ms, *SD* = 478 vs. comprehension = 833 ms, *SD* = 457; second number means: production = 501 ms, *SD* = 360 vs. comprehension = 547 ms, *SD* = 381). The analysis of fixation durations revealed no difference between the two conditions (first number: β = -.14, *SE* = .1, t = -1.353; $\chi^2(2)$ = 1.833, p = .176; second number: β = .224, *SE* = .186, t = 1.204; $\chi^2(2)$ = 1.434, p = .231). However, the analyses of the fixation latencies showed that the participants' gaze landed earlier on the first and on the second number on production trials (first number mean = 1831 ms, *SD* = 317; second number mean = 792 ms, *SD* = 395 ms) compared to comprehension trials (first number mean = 1904 ms, *SD* = 347; second number mean = 904 ms, *SD* = 493; first number: β = -.037, *SE* = .013, t = -2.9; $\chi^2(2)$ = 8.004, p = .004; second number: β = -.118, *SE* = .022, t = -5.3; $\chi^2(2)$ = 19.689, $p < .001$). Thus, the participants were overall faster to react to the spoken input when they prepared for a response than when they merely listened to the recorded speaker.

Discussion

In Experiment 1, we investigated how similar word prediction and word production processes are. We analyzed and compared eye movements reflecting participants' preparation to produce a word (the result of a simple mathematical equation) and eye movements reflecting their prediction of the same word being produced by a recorded speaker. In both conditions, we observed that the participants fixated upon the first two numbers of the equation in the order of mention, as they had been instructed to do, and then shifted their gaze to the result number. In both conditions, the shift of gaze to the result number occurred before the result number was spoken. Thus, in the production condition, the participants computed the result and then directed their gaze to the corresponding number and began to speak slightly afterwards. The average time eye-speech lag, i.e., the time between the onset of the fixation upon the result number and the onset of speech, was 490 ms (SD = 581). This substantial lag suggests that the participants initiated the saccade to the result number as soon as they had computed the result concept and carried out most of the linguistic planning of their utterance after the shift of gaze. We correlated the fixation latencies to the result numbers (measured from the spoken onset of the second number) with the eye-speech lags and found a moderate correlation r = .31 ($p < .001$ across 1240 production trials). Thus, speech latencies were faster on trials on which participants had looked at the target earlier. This correlation suggests that overlapping processes were engaged when participants planned the eye movement to the response number and when they subsequently planned the naming response.

Overall, the participants' eye movements in the two conditions were similar, suggesting that the cognitive processes occurring up to the overt articulation of the result number were similar as well. Although the similarity of the eye movements in the two conditions is striking, the analyses did reveal significant differences between the two conditions in the fixations to the three numbers. As shown in Table 5.1, the participants fixated the numbers earlier on production than on comprehension trials. This held not only for the result number, but also for the first two numbers, which were produced by the recorded speaker. As participants initiated the gaze to the result number earlier in the production condition than in the comprehension condition, the total duration of fixations to the result number was also longer. This was not the case for the first and second number, where fixations both began and ended earlier in the production than in the comprehension condition. Thus, it appears that participants were overall more engaged or aroused when they planned to speak than when they merely listened to the recorded speaker.

As the participants fixated the numbers earlier on production than on comprehension trials, one might expect that they would also produce the result number earlier than the recorded speaker. However, we observed the opposite, with participants taking slightly longer to produce the result than the recorded speaker. Recall, however, that the prerecorded speaker had been asked to read the equations at a moderate pace; hence the comparison of her speech onset latencies to those of the group of participants is not informative.

Experiment 2

In Experiment 2, we used the same materials and design as in Experiment 1, but tested non-native speakers of Dutch. They were German students of the Radboud University with intermediate knowledge of Dutch. The goal of that experiment was to determine whether we could replicate the two main findings of Experiment 1, namely, first, that participants were faster to direct their eyes to the relevant stimuli on production than on comprehension trials,

Table 5.1: Mean fixation durations and mean fixation latencies for production and comprehension conditions in Experiment 1 and 2. Standard deviations in brackets.

	Experiment 1			
	Fixation duration (ms)		Fixation latency (ms)	
	speak	*listen*	*speak*	*listen*
1^{st} No.	892 (478)	833 (457)	1831 (317)	1904 (347)
2^{nd} No.	541 (360)	548 (381)	792 (395)	904 (493)
Result No.	233 (253)	184 (241)	1628 (598)	1771 (673)
	Experiment 2			
	Fixation duration (ms)		Fixation latency (ms)	
	speak	*listen*	*speak*	*listen*
1^{st} No.	863 (490)	836 (501)	1912 (371)	1959 (389)
2^{nd} No.	454 (365)	494 (384)	970 (583)	1089 (679)
Result No.	197 (237)	156 (216)	1699 (660)	1880 (687)

and, second, that they would predict the result numbers on comprehension trials. There is ample research demonstrating that lexical access is delayed in late bilingual individuals (as compared to native speakers), even at high levels of proficiency (La Heij, 2005; Costa et al., 2000; Spalek et al., 2014; van Heuven & Dijkstra, 2010). We expected that due to delayed lexical access the fixation latencies on the first two numbers would be longer than in Experiment 1. In addition, we reasoned that as the task would overall be somewhat more demanding for the non-native than for the native speakers, the non-native speakers might refrain from predicting the result numbers but simply follow the recorded speaker. This would result in later shifts of gaze to the result number on comprehension trials compared to production trials.

Method

Participants

Twenty-four participants (6 male; mean age = 25 years, *SD* = 3) took part in Experiment 2. All were late German-Dutch bilinguals and students of the

Radboud University. All had received class-room instruction in Dutch language and, at the time of participation, had been regularly speaking Dutch for at least six months (on average 43 months). They rated their proficiency as intermediate (3 on a five-point scale ranging 1 to 5). Note that numerals 1 to 12 are cognates in German and Dutch. Hence even for beginning speakers for Dutch, the task was not challenging. All participants had normal or corrected-to-normal vision and hearing. None reported any signs or a history of developmental speech disorders. All participants gave informed consent before the experiment and were paid for their participation.

Materials and procedure

The instructions were translated into German. Apart from that, materials and procedure were identical to Experiment 1. Thus, the participants were asked to carry out the tasks in their second language.

Results and discussion

As in Experiment 1, we excluded production trials on which participants gave incorrect responses or their latency was more than 2.5 *SD* above their mean (57 trials; $< 1\%$). We also excluded comprehension trials where participants uttered the result number by mistake (6 trials; $< 1\%$). We calculated participants' naming latencies on production trials in the same way as in Experiment 1. The non-native speakers took on average 112 ms longer to start producing the result number than the recorded speaker. This is slightly longer (by 27 ms) than the time taken by the native speakers of Experiment 1.

Figure 5.2 shows the average proportions of fixations to the first, second, and result number across the average trial plotted in the same way as for Experiment 1. As can be seen in Table 5.1, the non-native speakers were somewhat slower to fixate the three numbers than the native speakers, but apart from

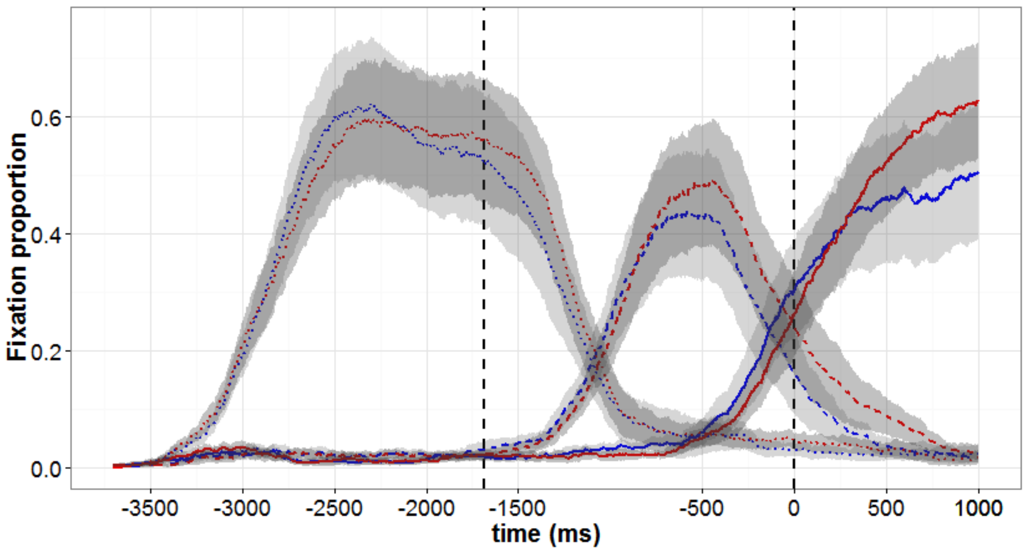


Figure 5.2: Fixation proportions of the German-Dutch bilinguals in Experiment 2. The data are plotted in the same way as for Experiment 1: Average fixation proportions to first (dotted lines), second (dashed lines) and result numbers (solid lines) for production (blue) and comprehension (red) conditions are shown. Fixations are plotted backwards from the offset of "is" (time zero). The first vertical dotted line represents the average onset of the second number. The areas shaded in gray represent the space in between the lower and upper bounds of the 95% by-participant confident intervals.

this expected difference, the results for the two groups of participants were very similar. In both the comprehension and the production condition, the participants first looked at the first and second number, in the order of mention, and then at the result number. As in Experiment 1, participants looked earlier at the relevant numbers on production than on comprehension trials. However, on comprehension trials they still looked at the result number before it was produced by the prerecorded speakers; the average lag was 197 ms. In other words, they anticipated the result numbers, as the native speakers of Experiment 1 had done.

The statistical analyses, carried out in the same way as for Experiment 1, confirmed that the results of Experiment 1 were replicated. As in Experiment

1, total fixation durations upon the first and second number did not differ between production and comprehension conditions (see Table 5.1, for means; first number: β = -.062, SE = .121, t = .711; $\chi^2(2)$ = .518, p = .472; second number: β = .212, SE = .23, t = .922; $\chi^2(2)$ = 1.669, p = .197), but there was a significant difference for the result number, with longer total fixation durations on production than on comprehension trials (β = -.575, SE = .139, t = -4.129). The comparison between the full model and a model that did not include the fixed factor Condition confirmed the better fit of the former (χ^2 (2) = 13.35, $p < .001$). The fixation latencies to all three numbers were significantly shorter on production than on comprehension trials, with the difference being most pronounced for the result number (first number: β = -.026, SE = .011, t = -2.4; $\chi^2(2)$ = 5.377, p = .02; second number: β = -.091, SE = .023, t = -3.95; $\chi^2(2)$ = 12.914, $p < .001$; result number: β = -.0103, SE = .0389, t = 2.66; $\chi^2(2)$ = 6.917, p = .009). This pattern suggests that the participants were more alert on speaking than on listening trials.

We examined whether the participants' performance changed over the course of the experiment. The sequence of trials was split into five blocks of 24 trials each and computed each participant's average fixation latency for the result numbers on comprehension and production trials. The latencies were longer in the first block than in subsequent blocks, but there was no further consistent performance change across the following blocks. In all blocks, fixation latencies were shorter on production than on comprehension trials. An analysis including block as an additional variable yielded no main effect of this variable.

Finally, we examined whether all participants anticipated the result number on comprehension trials, i.e. had average latencies, measured from the onset of the result number, below 400 ms. This was the case for 22 of the 24 participants. The remaining participants had latencies of 351 and 442 ms, respectively, but were also much slower than average (by 1011 ms and 730 ms, respectively) to fixate the second number in the equation. Thus, these participants processed the equations very slowly. Whether they did not aim to anticipate the result numbers or simply did not complete the mental computation before the numeral was produced by the recorded speaker cannot be determined.

In sum, the results of Experiment 2 closely replicate those of Experiment 1. The non-native speakers of Dutch were only slightly slower than the native speakers to fixate the relevant numbers on the clock face and to produce the result number, and on both production and comprehension trials their eyes landed on the result number before the numeral was produced. Thus, on comprehension trials they predicted the last word of the utterance, as the native speakers of Experiment 1 had done. Individual analyses of the difference between their first fixation upon the target on comprehension trials and the onset of result number in the recording showed that there were two participants with long positive lags (351 ms, 442 ms).

Most likely, the results obtained for the two groups of speakers were very similar because the difficulty of the task did not differ much for them; in other words, contrary to our expectation, the non-native speakers found the tasks almost as easy as the native speakers did. Evidently, on the basis of these data no claims can be made about prediction in non-native language comprehension in general (but see Foucart, Martin, Moreno, & Costa, 2014; Bonifacci, Giombini, Bellocchi, & Contento, 2011). The present results do, however, show that the main results - prediction of upcoming result numbers and earlier fixations of the numbers on production than on comprehension trials - can readily be replicated.

General discussion

Recent theories of language comprehension propose that listeners employ production-based mechanisms to anticipate upcoming language (Dell & Chang, 2014; Pickering & Garrod, 2007, 2013). In the current study, we tested this claim by comparing participants' planning of a word with their prediction of the same word being produced by a recorded speaker—both processes carried out under identical circumstances. In two experiments, participants listened to mathematical equations containing the numbers 1 to 12 while looking at an analogue clock face. We instructed participants to fixate the numbers on the clock face mentioned by the speaker as quickly as possible. On alternating trials, they listened to the entire equation or produced the result number themselves when the recording stopped after "is". In Experiment 1, the participants were native speakers of Dutch, and in Experiment 2 they were native speakers of German using Dutch as their second language. We tested a strong prediction of prediction-by-production accounts of processing: If word prediction is identical to internal word production, we should observe identical behavioral consequences caused by both processes. We measured participants' eye movements preceding their own speech onsets and preceding the onset of the spoken numerals in the recording. In both experiments, we observed shifts in eye gaze to the result numbers prior to the respective word onsets. In both experiments, however, we also observed statistically significant differences in the fixation latency and duration for the result numbers between the conditions: The participants began to fixate the result number earlier and consequently spent more time fixating it on production trials than on comprehension trials. Moreover, participants were also faster to locate the first and second numbers of the equations on production than on comprehension trials.

These results allow for three conclusions. First, in the comprehension condition, the participants predicted the last word of the spoken utterances. The task did not require them to do so; they could have waited until the recorded speaker named the result numbers and then direct their gaze towards them. However, instead of following the recorded speaker, the participants computed along with her, predicting what she would say. In evaluating this finding it is important to keep in mind that the participants knew that their eye movements were recorded and that they had been asked to fixate the relevant numbers. They were not asked to anticipate what the recorded speaker might say. Nevertheless, it is possible that participants felt that looking at each number as soon as possible and anticipating the results on comprehension trials would be desirable. It should also be kept in mind that in our experiments comprehension and production trials alternated. There may have been a transfer effect between trial types, and participants may have been more likely to engage in the mental computation of the results on comprehension trials than they would be if, for instance, production and comprehension trials appeared in different blocks. Further work is needed to determine under which conditions listeners engage in which kinds of predictions (Huettig, 2015). Our study demonstrated that under the conditions we created, the participants anticipated the last word of the spoken utterances.

Second, judging from the participants' eye movements, the cognitive processes occurring on production and comprehension trials, up to the time when the result numbers were produced or heard, were very similar. This can best be appreciated by comparing the fixation proportions to the three numbers of the equations shown in Figures 5.1 and 5.2. One may say that the similarity of the eye movements and the underlying cognitive processes is hardly surprising given the similarity of the production and comprehension tasks. However, as noted, the participants did not have to predict the result numbers and direct their eyes towards them in anticipation of the speaker. They elected to do so, and the cognitive processes involved in computing the result number are very likely to have been the same as those engaged in computing the result number for overt articulation. Thus, our results are in line with the view that the participants engaged largely the same processes on comprehension and on production trials. As noted in the Introduction, it has been proposed that predictions during comprehension might be based on fast associative processes, which are not engaged when speakers prepare utterances (Kukona et al., 2011; Kuperberg, 2007). If such associative processes played a major role in our task, one might have observed faster eye movements to the result numbers on comprehension than on production trials. However, as discussed further below, we found the opposite pattern. Thus, our results do support the view that prediction on comprehension trials was based on processes that were engaged on production trials.

Third, we found a consistent difference in the participants' fixation latencies to the three numbers. They looked earlier at the first and second number mentioned by the recorded speaker and they directed their gaze earlier to the result number when they had to produce it than when they merely listened to the recorded speaker. This unexpected result suggests that the participants were more engaged or aroused when an overt response was required than when they merely listened to the other person. The fixation latency difference between the production and comprehension conditions was more pronounced for the result number than for the first and second numbers, probably because the arithmetic operation required to compute the result benefited more from a higher activation level than the processes involved in listening to the utterances and locating the numbers on the clock face. In short, the participants carried out very similar cognitive processes on production and comprehension trials, but did so with a higher degree of engagement or arousal when an overt response was required, and perhaps because an overt response was required. Perhaps a different pattern of results is seen when participants have to provide an overt response on comprehension trials as well (e.g., indicating whether or not the response given by the recorded speaker is correct). In that case, eye movements to the result number might be equally fast on production and comprehension trials; or they might even be faster on comprehension trials if associative processes are engaged during comprehension. The effects of various comprehension tasks can be addressed in future research.

The current project involved mathematical equations, and one may ask whether the results are indicative of the relationship between production and comprehension processes occurring when people comprehend and produce everyday utterances. This is an empirical issue, which could be investigated by comparing the eye movements of speakers describing scenes and events (saying, for instance, "the boy will eat the cake") to the eye movements of listeners hearing descriptions of the same scenes. As discussed earlier, there is a large body of evidence demonstrating that listeners predict upcoming parts of utterances, and it is highly likely that these predictions are based on conceptual and linguistic knowledge that is also assessed when people produce utterances. Yet, whether this knowledge is used in the same way and equally efficiently in speaking and listening is not known. To assess this issue it is necessary to compare the time course of conceptual and linguistic processes for production and comprehension in tasks that are as similar as possible. A methodological contribution of the present study is to demonstrate how this could be done.

It has often been proposed that speaking is more effortful than comprehending language. A number of reasons for this difference have been proposed, for instance that speakers must develop more complete representations than listeners and that the speakers must not only develop, but also monitor their speech plans for correctness and appropriateness for the communicative situation (H. H. Clark & Krych, 2004). There is some evidence for the claim that speaking is indeed more demanding than listening, for instance from dual-task studies (Bock, Dell, Garnsey, Kramer, & Kubose, 2007; Sjerps & Meyer, 2015), though the evidence is by no means unambiguous (Kubose et al., 2006). In any event, it is often assumed that the processes involved in speaking are inherently more complex than those involved in listening, which leads to a higher degree of felt cognitive effort. Our results do not allow us to decide whether or not this view is correct. They do suggest, however, that in addition to any inherent differences in the complexity of production and comprehension processes, motivational differences may also contribute to differences in experienced effort: People may be more aroused or attentive when they speak than when they listen, possibly for the simple reason that the results of their efforts are witnessed by others when they speak but not when they listen to language.

In sum, the present study illustrates how eye-tracking can be used to track the time course of some of the processes occurring when people listen to spoken sentences and anticipate upcoming words and when they prepare to say these words themselves. Our evidence suggests, first, that prediction during sentence comprehension and speech planning may indeed be closely related processes, and, second, that people are more active or aroused when they intend to complete another person's utterance than when they merely listen to the interlocutor.

6 Encouraging prediction during production facilitates subsequent comprehension: Evidence from interleaved object naming in sentence context and sentence reading[1]

Abstract

Many studies have shown that a supportive context facilitates language comprehension. A currently influential view is that language production may support prediction in language comprehension. Experimental evidence for this, however, is relatively sparse. Here we explored whether encouraging prediction in a language production task encourages the use of predictive contexts in an interleaved comprehension task. In Experiment 1a, participants listened to the first part of a sentence and provided the final word by naming aloud a picture. The picture name was predictable or not predictable from the sentence context. Pictures were named faster when they could be predicted than when this was not the case. In Experiment 1b the same sentences, augmented by a final spill-over region, were presented in a self-paced reading task. No difference in reading times for predictive vs. non-predictive sentences was found. In Experiment 2, reading and naming trials were intermixed. In the naming task, the advantage for predictable picture names was replicated. More importantly, now reading times for the spill-over region were considerable faster for predictive vs. non-predictive sentences. We conjecture that these findings fit best with the notion that prediction in the service of language production encourages the use of predictive contexts in comprehension. Further research is required to identify the exact mechanisms by which production exerts its influence on comprehension.

[1]Adapted from *Hintz, F., Meyer, A.S., & Huettig, F. (submitted). Encouraging prediction during production facilitates subsequent comprehension: Evidence from interleaved object naming in sentence context and sentence reading.*

Introduction

A hallmark of human communication is the speed with which we process language. In dialogue, interlocutors typically react to previous turns very quickly, often within as little as 200 ms (cf. De Ruiter et al., 2006; Heldner & Edlund, 2010). It is likely that communication is so fast and efficient because language users rely heavily on supportive contexts (cf. van Berkum et al., 2005; Huettig, 2015; Kutas et al., 2011; Lupyan & Clark, in press, for reviews).

A currently prominent view assumes that language production may support prediction in language comprehension (Dell & Chang, 2014; Pickering & Garrod, 2013). Evidence supporting this notion, however, is sparse (but see Gambi, Cop, & Pickering, 2015; Mani & Huettig, 2012). If production-based prediction plays an important role in comprehension, one would expect that contexts encouraging prediction in the service of language production should also facilitate language comprehension.

Indeed, previous research implies that production tasks can increase the use of predictive contexts during comprehension (compared to comprehension settings without a production task). Gollan et al. (2011; see also Griffin & Bock, 1998), for instance, observed faster naming latencies for objects depicting words that appeared in strongly predictable contexts, as compared to appearing in weakly predictable contexts. When the same sentences were used in an eye-tracked reading task, highly predictable targets were read faster than weakly predictable targets. Interestingly however, based on a post-hoc analysis, Gollan et al. reported that the facilitation effect was much larger in the naming task than in the reading task. In other words, when the task set involved production in addition to comprehending the first part of the sentence, the degree to which participants experienced facilitation was higher than when the task set involved only comprehension.

In the current study, we further explored the hypothesis that a task set encouraging prediction in a production task also encourages readers to use predictive contexts in a comprehension task, compared to a task set only involving comprehension. To that end, Dutch participants carried out two tasks, a cross-modal naming task and a word-by-word self-paced reading task. The cross-modal naming task involved comprehending the first part of a spoken sentence and naming an object that was presented at the end of the recording to complete the sentence. The task thus comprised a production component in addition to comprehension. Self-paced reading only involved comprehension. The same sentences were used in both tasks and contained critical target nouns which appeared in both predictable and non-predictable contexts. In contrast to Gollan et al. (2011), within the predictable condition, we chose items that were not highly but moderately predictable.

Experiment 1 was run as a between-participants manipulation: In Experiment 1a, participants carried out the cross-modal naming task; in Experiment 1b, another sample of participants read the complete sentences including the target in a self-paced word-by-word fashion. We measured participants' picture naming latencies and their reading times for the target words. To anticipate the main results, a substantial naming advantage was found on predictable over non-predictable trials in Experiment 1a. In Experiment 1b, we did not observe significant facilitation in the predictable condition (with our moderately predictable items) relative to the non-predictable condition. In Experiment 2, we interleaved naming and reading trials, appearing in random order. If a task set including prediction serving language production increases the likelihood of using predictive contexts in comprehension, we should observe facilitation on the reading trials of Experiment 2.

Experiment 1

Method

Participants

We estimated the required number of participants to be able to draw reliable statistical conclusions using G*Power (Faul, Erdfelder, Lang, & Buchner, 2007) prior to running the experiment. Following the program's calculation (54 participants per experiment), which was based on the items' mean cloze probability and range, and the number of items per condition, 109 members of the subject panel of the MPI took part in Experiment 1a and 1b (Experiment 1a: fifteen male, mean age = 21, *SD* = 3; an additional participant had to be excluded from the analysis because of extensive data loss; Experiment 1b: ten male, mean age = 22, *SD* = 2). All were native speakers of Dutch and did not report any history of learning or reading disabilities or neurological or psychiatric disorders. The participants were paid for participating in the study. The ethics board of the faculty of social sciences of the Radboud University Nijmegen approved the study.

Materials

In both sub-experiments, the stimuli consisted of 40 target nouns which appeared in simple predictable sentences (e.g., "De man breekt op dit moment een glas", the man breaks at this moment a glass) and non-predictable sentences (e.g., "De man leent op dit moment een glas", the man borrows at this moment a glass; see Appendix, for all items). All sentences were of the same structure: The subject position was filled by "the man", and the adverbial "at this moment" separated verb and object. Using this "padding" between verb and target, we aimed to provide enough time for participants to generate predictions. In Dutch, the resulting sentence construction is deemed quite natural by native speakers.

Thirty-five additional native speakers of Dutch (mean age = 21 years, *SD* = 2) provided cloze probability ratings over the internet. Cloze probability was the proportion of participants who chose to complete a sentence fragment with the word in question. In the predictable items, the targets' mean cloze probability was .39 (*SD* = .24; range: .06 - .8); in the non-predictable items, it was zero.

Analyses were carried out on the length and frequency (using the Subtlex-NL database) of the verbs and objects. Raw frequencies were transformed to Zipf values, as suggested by van Heuven and colleagues (2014). Pairwise comparisons revealed that the predictable and the non-predictable verbs did not differ with regards to either number of letters ($t(39)$ = -1.122, p = .269) or frequency of the inflected verb ($t(39)$ = -.1.96, p = .057). The objects' Zipf-transformed mean word frequency was 4.52 (*SD* = .54).

For Experiment 1a, the 80 sentences, including the target nouns, were spoken with neutral intonation at a normal pace by a female native speaker of Dutch. Recordings were made in a sound-damped booth, sampling at 44 kHz (mono, 16 bit sampling resolution) and stored directly on computer. The mean sentence duration was 2800 ms (*SD* = 214). A second version of each recorded sentence was created by manually cutting off the target noun. The mean duration of the cut sentences was 2076 ms (*SD* = 155). Depictions of the forty target words were selected from the Snodgrass object database (Snodgrass & Vanderwart, 1980) and were colored in, or were drawn by an artist.

The same sentences were used in written form in Experiment 1b. Neutral prepositional phrases were added to each sentence (e.g., "De man breekt een glas van de collectie", the man breaks a glass from the collection) to be able to measure potential spill-over effects (see Mitchell, 1994, for discussion). The two words following the target were the same in all sentences (“van de”, from the). In Experiment 1b, 30% of the sentences were followed by a comprehension question. Half of the questions focused on the verb of the just-read sentence, the other half focused on the object. Half of the questions required a Yes-response.

Procedure

The 40 predictable and the 40 non-predictable items were evenly distributed across two lists, with none of the target nouns appearing twice on a list. Participants were randomly assigned one list and were seated in a sound-shielded room.

The spoken sentences in Experiment 1a were presented to the participants through loudspeakers. A trial was structured as follows: A central fixation dot appeared in the center of the screen for 250 ms. The dot disappeared and the playback of the sentence started. Coinciding with the end of the recording, participants saw a picture of the target word which they were asked to name as quickly as possible. The picture remained on the screen for 2000 ms; the inter-trial interval was 1500 ms. Each participant was presented with all 40 trials on one list. The order of trials was pseudo-randomized prior to the experiment. All trials, including participants’ responses were recorded in wav files for later analysis. Due to very low naming agreement for the pictures, we had to exclude the predictable and the non-predictable versions of four items. Naming latencies were calculated as the difference between the onset of the presentation of the critical object and the onset of participants’ response.

The same experimental lists were used in Experiment 1b. A trial started with the presentation of the first word of the sentence, next to a number of underscores indicating the number of words to follow (i.e., “moving window” format). Upon pushing the space bar with their left hand, participants advanced to the next word and the previous word was replaced with an underscore. They were instructed to read the sentences as fast as possible. On 30% of the trials a comprehension question was asked which they answered by pushing the left (No) or right (Yes) button on the mouse using their right hand. On trials where no comprehension question was asked, participants advanced to the next trial by pushing a button on the mouse. Their responses to the comprehension questions showed that they read the sentences carefully (mean accuracy = 93%, *SD* = 8). Reading times for the target words and for the post-target words were calculated as the difference between the respective onsets of presentation and participants' button presses.

Results

Naming latencies and reading times (RTs) were log-transformed and analyzed using linear mixed-effects regression models in R (R-Core-Team, 2012) with simultaneous inclusion of participants and items as random factors (Baayen et al., 2008). The full model included a fixed effect of Condition (predictable vs. non-predictable) and the maximal possible random effects structure (cf. Barr et al., 2013) consisting of random intercepts and slopes for Condition by participants (n = 54) and items (n = 36/40). This model was compared to the same model without the fixed effect of Condition using a likelihood test. Including Condition improved the model fit, $\chi^2(2) = 23.583$, $p < .001$, in the naming latency analysis in Experiment 1a. The full model revealed that participants named the target objects on average 96 ms faster when these were preceded by a predictable lead-in sentence relative to being preceded by a non-predictable lead-in sentence, β = -.147, *SE* = .026, *t* = -5.64 (see Table 6.1, for means). Applying the same analysis to the reading data revealed that the target words and the word following the target (spill-over region, henceforth) were read 7 ms and 8 ms faster, respectively, when preceded by a predictable verb relative to a non-predictable verb. These differences were not statistically significant (target: β = -.011, SE = .019, t = -.55; $\chi^2(2)$ = .311, p > .5; spill-over: β = -.018, SE = .017, t = -1.01; $\chi^2(2)$ = 1.035, p > .3).

Table 6.1: Mean naming latencies (NLs) and (post-)target reading times (RTs) for Experiments 1a, 1b, and 2. Standard deviations in parentheses.

		Exp. 1a	**Exp. 1b**	**Exp. 2**
NLs (ms)	Predictable	657 (193)	-	752 (215)
	Non-predictable	753 (232)	-	839 (314)
RTs (ms)	Predictable	-	397 (159)	327 (177)
	Non-predictable	-	404 (171)	332 (182)
post-target RTs (ms)	Predictable	-	370 (108)	312 (135)
	Non-predictable	-	378 (118)	331 (156)

A correlation analysis between an item's cloze probability and its facilitation effect (predictable minus non-predictable naming latency/RT) revealed a significant positive relationship between both measures in Experiment 1a (r = .347, n = 36, p = .038) but only a trend towards significance in Experiment 1b (target: r = .214, n = 40, p = .186; spill-over: r = .129, n = 40, p = .427).

Discussion

The results of Experiment 1a replicate previous findings (Gollan et al., 2011; Griffin & Bock, 1998) and show that naming latencies were substantially reduced for predictable relative to non-predictable items. This finding is in line with our hypothesis: When participants were asked to carry out a production task (object naming) in addition to comprehension, the likelihood of facilitatory processing was increased as compared to when they carried out a 'pure' comprehension task. The lack of a significant prediction effect in Experiment 1b may appear surprising given previous successful applications of the self-paced reading paradigm to study anticipatory language processing (e.g. van Berkum et al., 2005, Experiment 3) but note that we chose moderately predictable items in our study to avoid potential ceiling effects.

In Experiment 2, we tested whether the likelihood of using a predictive context on reading trials could be increased by randomly interleaving naming and reading trials. This manipulation was motivated by two considerations: First, we wanted to rule out that the self-paced reading paradigm might not be sensitive to capture such effects. Second, if our assumption is correct, mixing naming and reading trials, that is production and comprehension tasks, should increase participants' likelihood of using predictive contexts when processing the target word on reading trials. For the naming task, we expected similar results in as Experiment 1a.

Experiment 2

Method

Participants

Fifty-six native speakers of Dutch (11 male, mean age = 21, $SD = 3$) who had not participated in Experiment 1 or the norming study took part in Experiment 2. None of them reported any history of learning or reading disabilities or neurological or psychiatric disorders. Due to a programming error, two participants had to be excluded.

Stimuli and procedure

The materials were the same as in Experiment 1a and 1b. The 80 naming items and the 80 reading items were evenly distributed across four lists, with each of the target nouns appearing only once on a list. Participants were randomly assigned lists. The order of trials was completely randomized in the beginning of a testing session. Apart from that, trial structure and procedure were identical to the previous experiments.

Results and Discussion

Naming data and reading data were analyzed separately. As in Experiment 1b, accuracy to the comprehension questions on reading trials indicated that participants read the sentences carefully (mean accuracy = 91%, *SD* = 11). The same statistical procedure as in Experiment 1 was applied.

The analysis revealed a naming advantage very similar to that observed in Experiment 1a: Participants were 87 ms faster at naming the target object when it was preceded by a predictable relative to a non-predictable lead-in sentence (β = -.106, *SE* = .018, t = -6.01; $\chi^2(2) = 25.629$, p < .001). However, on reading trials, we now observed a facilitatory effect as well: The spill-over region was read 19 ms faster on predictable as compared to non-predictable items. This difference was statistically significant (β = -.047, *SE* = .017, t = -2.81; $\chi^2(2) = 7.338$, p = .007). No facilitation effect was found for target reading times (β = -.015, *SE* = .019, t = -.75; $\chi^2(2)$ = .575, p > .4).

Correlation analyses, carried out as previously, showed a marginal significant positive relationship between an item's naming advantage and its cloze probability (similar to Experiment 1a; r = .322, n = 36, p = .055). Not a hint towards a significant correlation was found for reading trials (target: r = -.091, n = 40, p > .5; spill-over: r = .075, n = 40, p > .6).

Conclusions

The present study supports the notion that prediction in the service of language production encourages the use of predictive contexts in comprehension: Substantial facilitation effects were observed when the participants' task involved production (Experiment 1a and 2) but not when participants carried out a "pure" comprehension task (Experiment 1b).

One interpretation of the facilitation effects is that participants used their production system to anticipate predictable words not only on the production trials of both experiments, but also on the self-paced reading trials of Experiment 2. That is, they used their production system to predict the name of the object in the naming task and similarly used their production system to predict upcoming words in the self-paced reading task. A possible objection to this interpretation of the results is that we observed the facilitation effect only in the spill-over region of Experiment 2, but not before the predictable word occurred.

An alternative interpretation of our findings is therefore that this effect does not reflect a 'downstream' consequence of production-based prediction, but merely facilitated integration of the target word with previous sentence context (cf. Van Petten & Luka, 2012). We cannot rule such an account with certainty. Note that prediction effects in self-paced reading often manifest themselves only in the spill-over region. Smith and Levy (2013), for instance, observed a logarithmic relationship between predictability and reading time of the target word in the spill-over region (the following three words) but not the target words or before (see Mitchell, 1994, for a detailed discussion of this issue). Moreover, most authors of electrophysiological studies reporting reduced N400 effects on target words following predictable contexts also interpret such a result as reflecting prediction rather than facilitated integration of the target words (see Kutas et al., 2011, for further discussion).

Why might encouraging prediction in a language production task facilitate integration in an interleaved comprehension task? In addition to being more aroused when having to produce an overt response rather than reading silently (cf. Hintz & Meyer, 2015), we conjecture that processes involved in language production and dialogue contributed to the observed facilitation effects. In line with such a notion, recent electrophysiological evidence suggests that participants engaged in lexical processing when anticipating that an experimental confederate would produce an utterance (Baus et al., 2014). In a similar vein, using a joint naming paradigm involving two participants, Gambi et al. (2015) compared the coordination of two successive utterances within and between speakers. The authors observed that the way in which speakers produced their own utterances was affected by whether they anticipated the turn of their confederate. Thus, the coordination of speaker turns, a situation similar to the alternation between comprehension and production task sets in the present study, may make use of some mechanisms that are also involved in preparing to speak. Although our data do not unequivocally show that participants used their production system to anticipate upcoming words, we conjecture that it is the most parsimonious account of the data. Future research is required to confirm this interpretation.

To conclude, we have shown that the degree to which readers use predictive contexts is influenced by the task set: In our study, readers relied more on predictive contexts when they also carried out a production task that encouraged prediction but less so when they carried out a pure reading task.

Table 6.2: Appendix: Predictable and non-predictable verb-noun pairs. Cloze probability is provided for the target objects embedded in predictable sentences.

Target object	**Predictable verb**	**Non-predictable verb**	**Cloze prob.**
appel (apple)	schillen (peel)	tekenen (draw)	0.63
baard (beard)	scheren (trim)	zien (see)	0.51
bal (ball)	trappen (kick)	lenen (borrow)	0.6
band (tube)	verwisselen (change)	verliezen (loose)	0.46
bank (couch)	bekleden (stiffen)	kiezen (choose)	0.14
beker (cup)	winnen (win)	bekijken (look at)	0.06
biertje (beer)	drinken (drink)	kopen (buy)	0.6
bloem (flower)	planten (plant)	ontvangen (receive)	0.06
boek (book)	publiceren (publish)	verstoppen (hide)	0.14
boom (tree)	kappen (chop)	beschrijven (describe)	0.69
boterham (sandwich)	smeren (prepare)	betalen (pay)	0.77
broek (pants)	passen (fit)	zoeken (search)	0.49
cadeau (present)	krijgen (receive)	stelen (steal)	0.2
contract (contract)	ondertekenen (sign)	ontvangen (receive)	0.69
deur (door)	openen (open)	zoeken (search)	0.31
dief (thief)	arresteren (arrest)	filmen (film)	0.34
doos (box)	tillen (lift)	verbergen (hide)	0.23
fiets (bike)	repareren (repair)	pakken (grab)	0.34
glas (glass)	breken (break)	lenen (borrow)	0.2
hond (dog)	aaien (pet)	tekenen (draw)	0.43
huis (house)	bezitten (own)	kiezen (choose)	0.29
ijsje (ice-cream)	likken (lick)	overhandigen (hand over)	0.77
kind (child)	beschermen (protect)	beschrijven (describe)	0.43
lamp (lamp)	vervangen (replace)	verbergen (hide)	0.31
muur (wall)	behangen (decorate)	bewaken (guard)	0.8
overhemd (shirt)	strijken (iron)	zien (see)	0.49
piano (piano)	stemmen (tune)	stelen (steal)	0.23
pizza (pizza)	bestellen (order)	verkopen (sell)	0.26
sigaar (cigar)	roken (smoke)	verstoppen (hide)	0.29
sinaasappel (orange)	persen (squeeze)	overhandigen (hand over)	0.71
standbeeld (statue)	onthullen (reveal)	bewaken (guard)	0.17
stoel (chair)	verplaatsen (displace)	pakken (grab)	0.14
taart (cake)	bakken (bake)	verkopen (sell)	0.51
tafel (table)	dekken (prepare)	betalen (pay)	0.66
tas (bag)	dragen (carry)	kopen (buy)	0.09
touw (rope)	spannen (take up)	verliezen (lose)	0.2
trein (train)	missen (miss)	filmen (film)	0.06
varken (pig)	slachten (slaughter)	fotograferen (take a photo)	0.46
vis (fish)	vangen (catch)	fotograferen (take a photo)	0.23
wond (wound)	hechten (suture)	bekijken (look at)	0.8

7 | Summary and discussion

This chapter summarizes the main results of the preceding chapters and discusses the findings in broader perspective. It also highlights potential directions for future research.

Summary of results

Adult readers and listeners comprehend language in different contexts. Even though these contexts often impose additional burdens onto the readers and listeners such as the integration of multiple input modalities or the coordination of turns in dialogue, most of the time language comprehension proceeds effortlessly. Previous research has suggested that the prediction of upcoming linguistic information is one reason for the ease with which we understand language in highly variable environments. Indeed, the previous literature had identified a number of mechanisms (event knowledge, associations-based and production-based mechanisms) and mediating factors (e.g., literacy) that are potentially involved in language prediction. The main goal of the present thesis was to test how strongly each of the mechanisms and the mediating factor contribute to anticipation in different situations of language processing. Moreover, the research in this thesis explored situation-specific influences on anticipatory language processing, such as the presence or absence of contextual visual information and the role of the language users' task set.

Chapter 2 aimed to determine the contribution of association-based and production-based mechanisms and literacy to verb-mediated anticipatory eye gaze behavior. In three visual world eye-tracking experiments, the participants looked at visual displays of four objects while listening to sentences where the target object was predictable ("The man peels an apple") or not predictable ("The man draws an apple"). In Experiments 1 and 2, the visual display was presented one second before the verb was heard. In Experiment 3, the participants were given only a short preview of the display starting 500 ms before the target noun was heard. In all three experiments, the participants looked at the target object before it was mentioned. Based on suggestions in the literature, I tested the influence of two kinds of associations: functional and general associations. Functional association strength was determined using a verb-noun typicality rating task (e.g., "How common is it for an apple to be peeled?"), and general word association strength was determined using a free verb-noun association task ("Write down the first three nouns that come to your mind when reading 'peel'). The predictable sentences in the experiments varied substantially in verb-noun typicality and in general verb-noun association strength. To examine the contributions of production abilities and literacy to prediction, participants' production fluency and receptive vocabulary knowledge were tested in separate verbal fluency and vocabulary tests. A non-verbal IQ test was included to separate general non-verbal intelligence from language-related abilities. The participant and item variables were related to the likelihood of anticipatory eye movements made to the target objects. Multiple regression analyses revealed that functional associations and literacy skills were strong predictors of verb-mediated anticipatory eye gaze with long and short visual previews. General word associations did not predict anticipatory eye movements in any of the experiments. Participants' production fluency correlated positively with the likelihood of anticipatory eye movements when participants were given the long but not short visual display preview.

In sum, functional associations, literacy and production-based mechanisms were found to contribute to prediction. These findings are novel and extend a recent proposal by Kukona et al. (2011), who argued that event-based sentence context and associative priming are involved in verb-mediated anticipatory eye movements. The data of Chapter 2 add to this proposal in that they specified the type of associations involved. Moreover, the results of Chapter 2 showed that in addition to event-based sentence context and associations, literacy and production-based mechanisms influenced predictive language processing in the visual world. Lastly, the results of Chapter 2 imply that production-based mechanisms only contribute to prediction under certain circumstances, namely when listeners were given sufficient preview time to inspect the displays.

Chapter 3 further explored the hypothesis that contextual visual information influences verb-mediated anticipatory language processing. Participants' eye movements were recorded in a visual world experiment when they listened to predictable sentences that had the same structure as the predictable sentences in Chapter 2: An object could be anticipated based on the selectional restrictions of the verb ("The man peels a banana"). While listening, participants looked at different types of displays. The target object (banana) was either present or it was replaced with an object that had a similar visual shape as the target object (canoe) or by an object that was semantically related to the concept invoked by the target (monkey). Each trial was presented in long preview version, where participants saw the displays before the verb was heard, and in short preview version, where participants were given only a short preview of the display before target onset, after the verb had been heard. I found that the participants anticipated the target objects in both preview conditions. However, robust evidence for visual shape and semantic competition effects was observed on short but not long preview trials. The absence of looks to the visual shape and semantic competitors before the target word was heard suggested that listeners extracted information from the visual displays to constrain the subsequent domain of reference. In other words, when given a long preview of the scene, they inspected the visual displays and most likely activated knowledge about the depicted objects, including the kinds of actions they tend to be involved in. On target-present trials, the knowledge activated on hearing the verb and the knowledge activated via the visual input yielded a match which resulted in an eye movement to the target object. However, on target-absent trials the knowledge activated via both modalities did not match; and therefore no eye movements to the visual shape and semantic competitors were observed. The results of Chapter 3 corroborate the notion that language-mediated anticipatory eye movements are substantially influenced by contextual visual information. Taken together, the present data suggest that the visual context does not only modulate the contribution of the mechanisms underlying language prediction (e.g., production-based mechanisms, Chapter 2, Experiment 3); instead visual context itself seems to restrict the dynamics of anticipatory language processing.

In Chapter 4, the contribution of simple word associations to prediction during discourse comprehension was tested. Participants' EEG was recorded as they read target words, which were preceded by associatively related words either appearing in a coherent discourse (describing an event, Experiment 1) or in sentences that did not form a coherent discourse (Experiment 2). Replicating earlier research, I found in Experiment 1 that contextually unexpected target words that were associatively related to the described event elicited an N400 with reduced amplitude compared to contextually unexpected target words that were unrelated to the event. In Experiment 2, where the influence of event knowledge was minimized, a similar but reduced N400 effect was observed. A likely explanation is that the effect in Experiment 1 reflected the joint contributions of event knowledge and associative priming to anticipatory processing, whereas the prediction effect in Experiment 2 reflected largely the contribution of associative priming. In light of the absence of an effect of general associations in Chapter 2, the study suggests that the contribution of this type of associations to prediction may depend on the situation in which language processing takes place. One may speculate that, given the subject-verb-object sentences and the presence of relevant contextual information, the situational context in the experiments in Chapter 2 may have encouraged the reliance on functional rather than general associations. In Chapter 4, the sentences were more variable and no visual context was available, which may have encouraged the reliance on general word associations.

In Chapter 5, a novel version of the visual world paradigm was used to test how similar prediction and production processes are. The participants listened to mathematical equations while looking at a clock face featuring the numbers 1 to 12. On alternating trials, they either heard a complete equation ("three plus eight is eleven"), or they heard the first part of the equation ("three plus eight is") and had to produce the result ("eleven") themselves. The participants were encouraged to look at the relevant numbers throughout the trial. They fixated the result number before the recorded speaker named it, and before they named it themselves. The eye movements related to anticipating the speech of another person and preparing to speak were overall very similar. However, shorter fixation latencies on production than on comprehension trials suggested that the processes involved in production and prediction – although quite similar – differ with regard to how aroused or engaged people are when they intend to respond or when they merely listen to another person. This finding has implications for theories of prediction-by-production accounts, which assume that predicting a word is akin to producing a word internally. The results from Chapter 5 highlight that motivational differences between the two processes need to be taken into account by theories assuming production-based mechanisms in prediction.

Related to motivational differences between prediction and production processes, Chapter 6 explored how different task sets influence the degree to which comprehenders engage in prediction. Specifically, I tested whether encouraging prediction in a language production task also encourages prediction in comprehension. The participants either listened to the first part of a sentence and provided the final word by naming aloud a picture (Experiment 1a) or they read the same sentences in a self-paced fashion (Experiment 1b). The final words in the sentences were either predictable or non-predicable. In Experiment 1a, participants' object naming latencies were measured; in Experiment 1b, their reading times were measured. The results revealed a naming advantage on predictable over non-predictable trials (Experiment 1a). However, no such advantage was found when the target words were read (Experiment 1b). Importantly, when reading and naming trials were mixed (Experiment 2), I observed facilitation effects with both tasks. The results thus suggested that encouraging prediction in the service of language production also increases the likelihood of prediction in comprehension.

Mechanisms and mediating factors in predictive language processing: Support for pluralistic approaches

Taken together, the results reported in Chapters 2 and 4 add to a growing body of literature showing that prediction during language comprehension is driven by multiple mechanisms and mediating factors (Kuperberg, 2007; Huettig, 2015; Pickering & Garrod, 2013; Wlotko & Federmeier, 2013). An important contribution of the present thesis to shaping a comprehensive account of predictive language processing is the identification of situational contexts in which the proposed mechanisms do or do not contribute to language prediction. The experiments in Chapters 2 and 4 featured two common situations of language processing: listening to language referring to the visual world and discourse reading. The results showed that each of the proposed mechanisms and the mediating factor contributed to prediction in at least one of the two situations tested. Why do people employ a multitude of mechanisms in language prediction, rather than a relying on a single mechanism? Although the present experiments were not designed to specifically answer this question, they indicate that in order to ensure smooth processing in highly variable contexts, the involvement of multiple mechanisms and mediating factors is a prerequisite. In different situations of language processing the brain may simply 'take what it can get' to achieve the most accurate prediction of upcoming information and exploit the information available in the input. I conjecture that the contribution of each mechanism and mediating factor to prediction in a particular situation is regulated by the properties and affordances of the situational context. Such a hypothesis needs to be explored in greater detail. Future research could, for example, zoom in on the interaction between individual mechanisms and mediating factors and investigate how they complement each other in different situations of language processing. Furthermore, for a complete understanding of language prediction, including the neuronal level, future research could follow up on recent studies that identified brain circuits subserving anticipation in language processing (e.g., Bonhage, Mueller, Friederici, & Fiebach, 2015; Willems, Frank, Nijhof, Hagoort, & van den Bosch, 2015) and test whether these regions are engaged differently in different situations of language processing. In what follows, I discuss how two aspects of the situational context impacted anticipatory language processing in the experiments reported in the present thesis.

The interplay of predictive language processing and the visual context

Chapters 2 and 3 showed that the presence or absence of relevant visual context impacted anticipatory language processing and the contribution of the mechanisms underlying anticipatory language processing. Chapter 3 revealed that comprehenders extract information from the visual context which is then used to constrain anticipatory language processing. In Chapter 2, the visual context was found to modulate the contribution of production-based mechanisms to prediction.

I suggest that the best account to these findings is a view where (predictive) language processing and visual processing constantly interact and influence each other (cf. Lupyan & Clark, in press). On such an account, contextual visual information may act as a trigger for the involvement of production-based mechanisms in prediction: When listening to someone who will refer to something in the immediate visual environment you may be inclined to engage in asking yourself “What would I say next if I were the speaker given these objects?” and use your production system to (internally) complete how the speaker may finish the sentence. In a similar vein, listeners may use the information extracted from the visual context to restrict the domain of subsequent reference and anticipate what the speaker may say about the objects at hand. Note that the latter view is similar to that proposed by Knoeferle et al. (2005). These authors demonstrated that information extracted from visually depicted scenes affected the process of incremental thematic role-assignment. More specifically, the authors showed that listeners combined information extracted from the depicted scenes and information derived from spoken verbs to resolve temporary ambiguities in structural role assignment.

In a recent paper, Lupyan and Clark (in press) argued that language exerts a powerful influence on the way we perceive our environment: “language begins to take on a surprisingly central role in cognition by providing a uniquely focused and flexible means of constructing predictions against which sensory signals can be evaluated. Predictive processing thus provides a plausible mechanism for many of the reported effects of language on perception, thought, and action" (p. 2). The results of Chapters 2 and 3 suggest that the reverse also holds: the visual environment appears to affect predictive language processing.

The language user's task set

Another aspect of the situational context that was found to impact anticipatory language processing in the situations tested in this thesis was the language user's task set. Chapter 6 showed that comprehenders engaged in prediction to a stronger degree when the task set involved production in addition comprehension (i.e., prediction in the service of language production) than when the task set involved 'pure' comprehension.

Why might this be the case? The results of Chapter 5 imply that the effect may in part be due to motivational differences. That is, one simple reason for the differences seen in Chapter 6 may be that the language users were more aroused when they had to speak themselves as compared to when they merely had to listen to another person speak. Moreover, the situation tested in Experiment 1a in Chapter 6 was similar to a dialogue context in which one first listens to another person and then provides a response, as we sometimes do when we complete other people's sentences. Another hypothesis is thus that the language users engaged in prediction to a stronger degree because of the simulated dialogue situation. There are two arguments why this might have been the case. First, dialogue is the most natural and basic form of language use and even children and illiterate adults can hold a conversation (Pickering & Garrod, 2004). It is thus a highly practiced process and the language users may have engaged in prediction to a stronger degree because they have more experience with being in dialogue situations than with reading. Lastly, previous research showed that listeners can more or less precisely predict the end of their interlocutors turn to enable smooth turn-taking (e.g., De Ruiter et al., 2006). Hence, it could also be that the involvement of additional processes related to dialogue (e.g., the prediction of turn ends) contributed to the larger anticipation effects when the production task was involved. Future research is needed to test whether motivational differences, general experience, and/or processes related to dialogue underlie the pattern of results in Chapter 6.

General conclusion

Predicting the outcome of other people's actions is an important skill, which has probably saved our lives many times as it enables us to react faster to the immediate situation (LaBerge, 1995). Predicting what another person might say next usually does not save our lives, but it most likely saves time when our predictions are confirmed. Prediction in language processing is one of the reasons why comprehension is such a fast and efficient process.

This thesis investigated how adult readers and listeners predict upcoming information in different situations of language processing. The presented results showed that multiple language mechanisms and mediating factors are implemented in parallel and that people adapt their use of these mechanisms in response to situational and task demands. People can dynamically adjust the comprehension strategies to the situational and task context so that resources can be allocated to most efficiently achieve comprehension aims.

References

Acheson, D. J., Wells, J. B., & MacDonald, M. C. (2008). New and updated tests of print exposure and reading abilities in college students. *Behavior Research Methods*, *40*(1), 278-289.

Allopenna, P. D., Magnuson, J. S., & Tanenhaus, M. K. (1998). Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. *Journal of Memory and Language*, *38*(4), 419-439.

Altarriba, J., Kroll, J., Sholl, A., & Rayner, K. (1996). The influence of lexical and conceptual constraints on reading mixed-language sentences: Evidence from eye fixations and naming times. *Memory & Cognition*, *24*(4), 477-492.

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. (2012). The use of verb-specific information for prediction in sentence processing. *Language and Cognitive Processes*, *28*(4), 525-560.

Arai, M., van Gompel, R. P., & Scheepers, C. (2007). Priming ditransitive structures in comprehension. *Cognitive Psychology*, *54*(3), 218-250.

Ashby, J., Rayner, K., & Clifton, C. (2005). Eye movements of highly skilled and average readers: Differential effects of frequency and predictability. *The Quarterly Journal of Experimental Psychology*, *58*(6), 1065-1086.

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.

Baayen, R. H., Piepenbrock, R., & Gulikers, L. (1993). *The CELEX lexical database on cd-rom* [Aggregated Database]. Philadelphia, PA: University of Pennsylvania [Distributor].

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

Bar, M. (2009). The proactive brain: Memory for predictions. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *364*(1521), 1235-1243.

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

Bates, D., Maechler, M., & Bolker, B. (2012). lme4: Linear mixed-effects models using s4 classes. R package version 0.999999-0.

Baus, C., Sebanz, N., de la Fuente, V., Branzi, F. M., Martin, C., & Costa, A. (2014). On predicting others' words: Electrophysiological evidence of prediction in speech production. *Cognition*, *133*(2), 395-407.

Beeman, M. (1993). Semantic processing in the right hemisphere may contribute to drawing inferences from discourse. *Brain and Language*, *44*(1), 80-120.

Bendixen, A., Schröger, E., & Winkler, I. (2009). I heard that coming: Event-related potential evidence for stimulus-driven prediction in the auditory System. *Journal of Neuroscience*, *29*(26), 8447-8451.

Bentin, S. (1987). Event-related potentials, semantic processes, and expectancy factors in word recognition. *Brain and Language*, *31*(2), 308-327.

Bentin, S., McCarthy, G., & Wood, C. C. (1985). Event-related potentials, lexical decision and semantic priming. *Electroencephalography and Clinical Neurophysiology*, *60*(4), 343-355.

Bicknell, K., Elman, J. L., Hare, M., McRae, K., & Kutas, M. (2010). Effects of event knowledge in processing verbal arguments. *Journal of Memory and Language*, *63*(4), 489-505.

Blair, J. R., & Spreen, O. (1989). Predicting premorbid IQ: A revision of the national adult reading test. *Clinical Neuropsychologist*, *3*(2), 129-136.

Bock, K., Dell, G. S., Garnsey, S. M., Kramer, A. F., & Kubose, T. T. (2007). Car talk, car listen. In A. S. Meyer, L. Wheeldon, & A. Krott (Eds.), *Automaticity and control in language processing* (p. 21-42). Hove, England: Psychology Press.

Bock, K., Irwin, D. E., Davidson, D. J., & Levelt, W. J. M. (2003). Minding the clock. *Journal of Memory and Language*, *48*(4), 653-685.

Boersma, P. P. (2002). Praat, a system for doing phonetics by computer (version 5.1.19) [computer program]. *Available from http://www.praat.org/*.

Bonhage, C., Mueller, J. L., Friederici, A. D., & Fiebach, C. J. (2015). Combined eye tracking and fmri reveals neural basis of linguistic predictions during sentence comprehension. *Cortex*, *68*, 33-47.

Bonifacci, P., Giombini, L., Bellocchi, S., & Contento, S. (2011). Speed of processing, anticipation, inhibition and working memory in bilinguals. *Developmental Science*, *14*(2), 256-269.

Borovsky, A., Elman, J., & Fernald, A. (2012). Knowing a lot for one's age: Vocabulary skill and not age is associated with anticipatory incremental sentence interpretation in children and adults. *Journal of Experimental Child Psychology*, *112*(4), 417-436.

Borovsky, A., Sweeney, K., Elman, J. L., & Fernald, A. (2014). Real-time interpretation of novel events across childhood. *Journal of Memory and Language*, *73*(0), 1-14.

Boudewyn, M. A., Gordon, P. C., Long, D., Polse, L., & Swaab, T. Y. (2011). Does discourse congruence influence spoken language comprehension before lexical association? evidence from event-related potentials. *Language and Cognitive Processes*, *27*(5), 698-733.

Bryan, J., Luszcz, M. A., & Crawford, J. R. (1997). Verbal knowledge and speed of information processing as mediators of age differences in verbal fluency performance among older adults. *Psychology and Aging*, *12*(3), 473-478.

Camblin, C. C., Gordon, P. C., & Swaab, T. Y. (2007). The interplay of discourse congruence and lexical association during sentence processing: Evidence from ERPs and eye-tracking. *Journal of Memory and Language*, *56*(1), 103-128.

Caramazza, A. (1997). How many levels of processing are there in lexical access? *Cognitive Neuropsychology*, *14*(1), 177-208.

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*(5), 1098-1110.

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

Chen, E., Gibson, E., & Wolf, F. (2005). Online syntactic storage costs in sentence comprehension. *Journal of Memory and Language*, *52*(1), 144-169.

Chiarello, C., Liu, S., & Faust, M. (2001). Bihemispheric sensitivity to sentence anomaly. *Neuropsychologia*, *39*(13), 1451-1463.

Chwilla, D. J., & Kolk, H. H. (2002). Three-step priming in lexical decision. *Memory Cognition*, *30*(2), 217-225.

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

Clark, H. H. (1996). *Using language.* Cambridge, MA: Cambridge University Press.

Clark, H. H., & Krych, M. A. (2004). Speaking while monitoring addressees for understanding. *Journal of Memory and Language*, *50*(1), 62-81.

Cooper, R. M. (1974). Control of eye fixation by meaning of spoken language - New methodology for real-time investigation of speech perception, memory, and language processing. *Cognitive Psychology*, *6*(1), 84-107.

Costa, A., Caramazza, A., & Sebastian-Galles, N. (2000). The cognate facilitation effect: Implications for models of lexical access. *Journal of Experimental Psychology: Learning Memory Cognition*, *26*(5), 1283-1296.

Coulson, S., Federmeier, K. D., Van Petten, C., & Kutas, M. (2005). Right hemisphere sensitivity to word- and sentence-level context: Evidence from event-related brain potentials. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*(1), 129-147.

Cumming, G. (2012). *Understanding the new statistics: Effect sizes, confidence intervals, and meta-analysis.* New York: Routledge.

Cumming, G. (2014). The new statistics: Why and how. *Psychological Science*, *25*(1), 7-29.

De Deyne, S., Navarro, D., & Storms, G. (2013). Associative strength and semantic activation in the mental lexicon: Evidence from continued word associations. In *Proceedings of the Annual Conference of the Cognitive Science Society (CogScie 2013, Berlin, Germany).* Cognitive Science Society.

De Deyne, S., Navarro, D. J., & Storms, G. (2012). Better explanations of lexical and semantic cognition using networks derived from continued rather than single-word associations. *Behavior Research Methods*, 1-19.

de Groot, F., Huettig, F., & Olivers, C. N. L. (2015). When meaning matters: The temporal dynamics of semantic influences on visual attention. *Journal of Experimental Psychology: Human Perception and Performance*, *Advance online publication.*

de Groot, F., Koelewijn, T., Huettig, F., & Olivers, C. N. (in press). An extended stimulus set in Dutch and English containing words and pictures matched for visual and semantic similarity. *Journal of Cognitive Psychology*.

Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review*, *93*(3), 283-321.

Dell, G. S., & Chang, F. (2014). The p-chain: Relating sentence production and its disorders to comprehension and acquisition. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *369: 20120394*(1634).

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.

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

Duñabeitia, J. A., Aviles, A., Afonso, O., Scheepers, C., & Carreiras, M. (2009). Qualitative differences in the representation of abstract versus concrete words: Evidence from the visual-world paradigm. *Cognition*, *110*(2), 284-92.

Dunn, L. M., & Dunn, D. (1997). *Peabody picture vocabulary test (3rd edition)*. Circle Pines: American Guidance Service.

Engel, A. K., Fries, P., & Singer, W. (2001). Dynamic predictions: Oscillations and synchrony in top-down processing. *Nature Review Neuroscience*, *2*(10), 704-716.

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175-191.

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. (2002). Picture the difference: Electrophysiological investigations of picture processing in the two cerebral hemispheres. *Neuropsychologia*, *40*(7), 730-747.

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*(3), 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*(2), 133-146.

Ferreira, F., & Swets, B. (2002). How incremental is language production? Evidence from the production of utterances requiring the computation of arithmetic sums. *Journal of Memory and Language*, *46*(1), 57-84.

Ferretti, T. R., McRae, K., & Hatherell, A. (2001). Integrating verbs, situation schemas, and thematic role concepts. *Journal of Memory and Language*, *44*(4), 516-547.

Fidler, F., & Loftus, G. R. (2009). Why figures with error bars should replace p-values. *Zeitschrift für Psychologie/Journal of Psychology*, *217*(1), 27-37.

Fitzpatrick, S., Gilbert, S., & Serpell, L. (2013). Systematic review: Are overweight and obese individuals impaired on behavioural tasks of executive functioning? *Neuropsychological Review*, *23*(2), 138-156.

Foucart, A., Martin, C. D., Moreno, E. M., & Costa, A. (2014). Can bilinguals see it coming? Word anticipation in L2 sentence reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(5), 1461-1469.

Frisson, S., Rayner, K., & Pickering, M. J. (2005). Effects of contextual predictability and transitional probability on eye movements during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*(5), 862-877.

Friston, K. (2010). The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience*, *11*(2), 127-138.

Frith, C. D. (2007). The social brain? *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *362*(1480), 671-678.

Gagnepain, P., Henson, R. N., & Davis, M. H. (2012). Temporal predictive codes for spoken words in auditory cortex. *Current Biology*, *22*(7), 615-621.

Gambi, C., Cop, U., & Pickering, M. J. (2015). How do speakers coordinate? Evidence for prediction in a joint word-replacement task. *Cortex*, *68*, 111-128.

Garrett, M. F. (1980). Levels of processing in speech production. In B. Butterworth (Ed.), *Language production* (p. 177-220). New York: Academic Press.

Gilbert, D. T., & Wilson, T. D. (2007). Prospection: Experiencing the future. *Science*, *317*(5843), 1351-4.

Gollan, T. H., Slattery, T. J., Goldenberg, D., van Assche, E., Duyck, W., & Rayner, K. (2011). Frequency drives lexical access in reading but not in speaking: The frequency-lag hypothesis. *Journal of Experimental Psychology: General*, *140*(2), 186-209.

Gratton, G., Coles, M. G. H., & Donchin, E. (1983). A new method for off-line removal of ocular artifact. *Electroencephalography and Clinical Neurophysiology*, *55*, 468-484.

Greenhouse, S. W., & Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrika*, *24*(2), 95-112.

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*(3), 313-338.

Griffin, Z. M., & Bock, K. (2000). What the eyes say about speaking. *Psychological Science*, *11*(4), 274-279.

Hagoort, P., Hald, L., Bastiaansen, M., & Petersson, K. M. (2004). Integration of word meaning and world knowledge in language comprehension. *Science*, *304*(5669), 438-441.

Hald, L. A., Steenbeek-Planting, E. G., & Hagoort, P. (2007). The interaction of discourse context and world knowledge in online sentence comprehension. evidence from the n400. *Brain Research*, *1146*(0), 210-218.

Heldner, M., & Edlund, J. (2010). Pauses, gaps and overlaps in conversations. *Journal of Phonetics*, *38*(4), 555-568.

Henry, J. D., & Crawford, J. R. (2004). A meta-analytic review of verbal fluency performance in patients with traumatic brain injury. *Neuropsychology*, *18*(4), 621-628.

Hintz, F., & Meyer, A. S. (2015). Prediction and production of simple mathematical equations: Evidence from anticipatory eye movements. *PLoS One*, *10*(7).

Huettig, F. (2015). Four central questions about prediction in language processing. *Brain Research*, *Advance online publication*.

Huettig, F., & Altmann, G. T. M. (2005). Word meaning and the control of eye fixation: semantic competitor effects and the visual world paradigm. *Cognition*, *96*(1), B23-32.

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., & Brouwer, S. (2015). Delayed anticipatory spoken language processing in adults with dyslexia. *Dyslexia*, *21*(2), 97-122.

Huettig, F., & Janse, E. (2015). Individual differences in working memory and processing speed predict anticipatory spoken language processing in the visual world. *Language, Cognition and Neuroscience, Advance online publication.*

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

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*(2), 151-171.

Hutchison, K. A. (2003). Is semantic priming due to association strength or feature overlap? A microanalytic review. *Psychonomic Bulletin & Review*, *10*(4), 785-813.

Hutchison, K. A., Balota, D. A., Cortese, M. J., & Watson, J. M. (2008). Predicting semantic priming at the item level. *Quarterly Journal of Experimental Psychology*, *61*(7), 1036-1066.

Iordanescu, L., Grabowecky, M., & Suzuki, S. (2011). Object-based auditory facilitation of visual search for pictures and words with frequent and rare targets. *Acta Psychologica*, *137*(2), 252-259.

Iordanescu, L., Guzman-Martinez, E., Grabowecky, M., & Suzuki, S. (2008). Characteristic sounds facilitate visual search. *Psychonomic Bulletin & Review*, *15*(3), 548-554.

James, A., & Watson, D. (2013). Language experience and prediction in auditory sentence comprehension. In *Proceedings of the 19th Architectures and Mechanisms for Language Processing Conference (AMLaP 2013, Marseille, France).*

Kahnemann, D. (2011). *Thinking, fast and slow.* New York: Farrar, Straus and Giroux.

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. (2008). Anticipatory processes in sentence processing. *Language and Linguistics Compass*, *2*(4), 647-670.

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*(1), 37-55.

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.

Keuleers, E., Brysbaert, M., & New, B. (2010). SUBTLEX-NL: A new measure for Dutch word frequency based on film subtitles. *Behavior Research Methods*, *42*(3), 643-650.

Kintsch, W., & Mross, E. F. (1985). Context effects in word identification. *Journal of Memory and Language*, *24*(3), 336-349.

Kleiman, G. M. (1980). Sentence frame contexts and lexical decisions: Sentence-acceptability and word-relatedness effects. *Memory & Cognition*, *8*(4), 336-344.

Klem, G. H., Lüders, H. O., Jasper, H. H., & Elger, C. (1999). The ten-twenty electrode system of the international federation of clinical neurophysiology. *Electroencephalography and Clinical Neurophysiology*, *Supplement 1999*(52), 3-6.

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*(1), 95-127.

Korvorst, M., Roelofs, A., & Levelt, W. J. (2007). Telling time from analog and digital clocks: A multiple-route account. *Experimental Psychology*, *54*(3), 187-191.

Kretzschmar, F., Schlesewsky, M., & Staub, A. (2015). Dissociating word frequency and predictability effects in reading: Evidence from coregistration of eye movements and EEG. *Journal of Experimental Psychology: Learning, Memory, and Cognition, Advance online publication*.

Kubose, T. T., Bock, K., Dell, G. S., Garnsey, S. M., Kramer, A. F., & Mayhugh, J. (2006). The effects of speech production and speech comprehension on simulated driving performance. *Applied Cognitive Psychology*, *20*(1), 43-63.

Kuchinsky, S. E., Bock, K., & Irwin, D. E. (2011). Reversing the hands of time: Changing the mapping from seeing to saying. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*(3), 748-756.

Kukona, A., Fang, S. Y., Aicher, K. A., Chen, H., & Magnuson, J. S. (2011). The time course of anticipatory constraint integration. *Cognition*, *119*(1), 23-42.

Kuperberg, G. R. (2007). Neural mechanisms of language comprehension: Challenges to syntax. *Brain Research*, *1146*(1), 23-49.

Kuperberg, G. R., Caplan, D., Sitnikova, T., Eddy, M., & Holcomb, P. J. (2006). Neural correlates of processing syntactic, semantic, and thematic relationships in sentences. *Language and Cognitive Processes*, *21*(5), 489-530.

Kuperberg, G. R., Paczynski, M., & Ditman, T. (2010). Establishing causal coherence across sentences: An ERP study. *Journal of Cognitive Neuroscience*, *23*(5), 1230-1246.

Kutas, M., DeLong, K., & Smith, N. (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 predict the future* (p. 190-203). New York, USA: 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.

La Heij, W. (2005). Selection processes in monolingual and bilingual lexical access. In J. Kroll & A. M. B. de Groot (Eds.), *Handbook of Bilingualism: Psycholinguistic Approaches* (p. 289-307). New York: University Press.

LaBerge, D. (1995). *Attentional Processing: The Brain's Art of Mindfulness.* Cambridge, MA: Harvard University Press.

Landauer, T. K., & Dumais, S. T. (1997). A solution to plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review*, *104*(2), 211-240.

Laszlo, S., & 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*(3), 326-338.

Laszlo, S., Stites, M., & Federmeier, K. D. (2012). Won't get fooled again: An event-related potential study of task and repetition effects on the semantic processing of items without semantics. *Language and Cognitive Processes*, *27*(2), 257-274.

Lau, E. F., Holcomb, P. J., & Kuperberg, G. R. (2013). Dissociating N400 effects of prediction from association in single-word contexts. *Journal of Cognitive Neuroscience*, *25*(3), 484-502.

Ledoux, K., Camblin, C. C., Swaab, T. Y., & Gordon, P. C. (2006). Reading words in discourse: The modulation of lexical priming effects by message-level context. *Behavioral and Cognitive Neuroscience Reviews*, *5*(3), 107-127.

Levelt, W. J. M. (1999). Models of word production. *Trends in Cognitive Sciences*, *3*(6), 223-232.

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

Levinson, S. C. (2013). Cross-cultural universals and communication structures. In M. A. Arbib (Ed.), *Language, music, and the brain: A mysterious relationship* (p. 67-80). Cambridge, MA: MIT Press.

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

Loftus, G. R., & Masson, M. E. J. (1994). Using confidence intervals in within-subject designs. *Psychonomic Bulletin & Review*, *1*(4), 476-490.

Luo, L., Luk, G., & Bialystok, E. (2010). Effect of language proficiency and executive control on verbal fluency performance in bilinguals. *Cognition*, *114*(1), 29-41.

Lupyan, G., & Clark, A. (in press). Words and the world: Predictive coding and the language-perception-cognition interface. *Current Directions in Psychology*.

MacDonald, M. C. (2013). How language production shapes language form and comprehension. *Frontiers in Psychology*, *4*.

Macmillan, N., & Creelman, C. (1991). *Detection theory: A user's guide*. New York: Cambridge University Press.

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.

Mani, N., & Huettig, F. (2013). Towards a complete multiple-mechanism account of predictive language processing. *Behavioral and Brain Sciences*, *36*(4), 365-366.

Mani, N., & Huettig, F. (2014). Word reading skill predicts anticipation of upcoming spoken language input: A study of children developing proficiency in reading. *Journal of Experimental Child Psychology*, *126*, 264-279.

Masson, M. E. J., & Loftus, G. R. (2003). Using confidence intervals for graphically-based data interpretation. *Canadian Journal of Experimental Psychology/Revue Canadienne De Psychologie Experimentale*, *57*(3), 203-220.

Matsuki, K., Chow, T., Hare, M., Elman, J. L., Scheepers, C., & McRae, K. (2011). Event-based plausibility immediately influences on-line language comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*(4), 913-934.

McCauley, S., & Christiansen, M. H. (2011). Learning simple statistics for language comprehension and production: The CAPPUCCINO model. In *Proceedings of the 33rd Annual Conference of the Cognitive Science Society* (p. 1619-1624).

McDonald, S., & Shillcock, R. (2003). Low-level predictive inference in reading: The influence of transitional probabilities on eye movements. *Vision Research*, *43*(16), 1735-1751.

McQueen, J. M., & Huettig, F. (2014). Interference of spoken word recognition through phonological priming from visual objects and printed words. *Attention, Perception, & Psychophysics*, *76*(1), 190-200.

McRae, K., Ferretti, T. R., & Amyote, L. (1997). Thematic roles as verb-specific concepts. *Language and Cognitive Processes*, *12*(2-3), 137-176.

McRae, K., Hare, M., Elman, J. L., & Ferretti, T. (2005). A basis for generating expectancies for verbs from nouns. *Memory & Cognition*, *33*(7), 1174-1184.

McRae, K., & Matsuki, K. (2009). People use their knowledge of common events to understand language, and do so as quickly as possible. *Language and Linguistics Compass*, *3*(6), 1417-1429.

Metusalem, R., Kutas, M., Urbach, T., Hare, M., McRae, K., & Elman, J. (2012). Generalized event knowledge activation during online sentence comprehension. *Journal of Memory and Language*, *66*(4), 545-567.

Meyer, A. S., Sleiderink, A. M., & Levelt, W. J. M. (1998). Viewing and naming objects: Eye movements during noun phrase production. *Cognition*, *66*(2), B25-B33.

Meyer, A. S., van der Meulen, F., & Brooks, A. (2004). Eye movements during speech planning: Talking about present and remembered objects. *Visual Cognition*, *11*(5), 553-576.

Mishra, R. K., Singh, N., Pandey, A., & Huettig, F. (2012). Spoken language-mediated anticipatory eye-movements are modulated by reading ability - Evidence from Indian low and high literates. *Journal of Eye Movement Research*, *5*(1), 1-10.

Mitchell, D. C. (1994). An evaluation of subject-paced reading tasks and other methods for investigating immediate processes in reading. In D. E. Kieras & M. A. Just (Eds.), *New methods in reading comprehension research* (p. 69-90). Hillsdale: Erlabum.

Mitchell, D. C. (2004). On-line methods in language processing: Introduction and historical review. In M. Carreiras & C. Clifton Jr. (Eds.), *The on-line study of sentence comprehension: Eyetracking, ERP and beyond* (p. 15-33). New York: Psychology Press.

Moss, H. E., McCormick, S., & Tyler, L. K. (1997). The time course of activation of semantic information during spoken word recognition. *Language and Cognitive Processes*, *12*(5-6), 695-732.

Neely, J. H. (1991). Semantic priming effects in visual word recognition: A selective review of current findings and theories. In D. Besner & G. Humphreys (Eds.), *Basic processes in reading: Visual word recognition* (p. 264-336). Hillsdale (New Jersey, US): Erlbaum.

Nelson, D., McEvoy, C., & Schreiber, T. (2004). The University of South Florida free association, rhyme, and word fragment norms. *Behavior Research Methods*, *36*(3), 402-407.

Nieuwland, M. S., & van Berkum, J. J. A. (2005). Testing the limits of the semantic illusion phenomenon: ERPs reveal temporary semantic change deafness in discourse comprehension. *Cognitive Brain Research*, *24*(3), 691-701.

Norris, D. (2006). The Bayesian reader: explaining word recognition as an optimal Bayesian decision process. *Psychological Review*, *113*(2), 327-57.

Noyes, J. M., & Garland, K. J. (2008). Computer- vs. paper-based tasks: Are they equivalent? *Ergonomics*, *51*(9), 1352-1375.

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*(0), 166-177.

Otten, M., & van Berkum, J. J. A. (2008). Discourse-based word anticipation during language processing: Prediction or priming? *Discourse Processes*, *45*(6), 464-496.

Perea, M., & Gotor, A. (1997). Associative and semantic priming effects occur at very short stimulus-onset asynchronies in lexical decision and naming. *Cognition*, *62*(2), 223-240.

Pickering, M. J., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences*, *27*(2), 169-189.

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. (2013). An integrated theory of language production and comprehension. *Behavioral and Brain Sciences*, *36*(4), 329-347.

Pickering, M. J., & Garrod, S. (2014). Self-, other-, and joint monitoring using forward models. *Frontiers in Human Neuroscience*, *8*.

Prinz, W. (2006). What re-enactment earns us. *Cortex*, *42*(4), 515-517.

Rapp, B., Buchwald, A., & Goldrick, M. (2013). Integrating accounts of speech production: The devil is in the representational details. *Language, Cognition and Neuroscience*, *29*(1), 24-27.

Raven, J., Raven, J. C., & Court, J. H. (1998). *Raven manual section 4: Advanced Progressive Matrices.* Oxford, UK: Oxford Psychologists Press.

Rayner, K., Ashby, J., Pollatsek, A., & Reichle, E. D. (2004). The effects of frequency and predictability on eye fixations in reading: Implications for the 'E-Z Reader Model'. *Journal of Experimental Psychology: Human Perception and Performance*, *30*(4), 720-732.

R-Core-Team. (2012). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. *Open access available at: http://cran.rproject.org*.

Roelofs, A. (1997). The WEAVER model of word-form encoding in speech production. *Cognition*, *64*(3), 249-284.

Rommers, J., Meyer, A. S., & Huettig, F. (2015). Verbal and nonverbal predictors of language-mediated anticipatory eye movements. *Attention, Perception, & Psychophysics*, *77*(3), 720-730.

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.

Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language*, *50*, 696-735.

Saslow, M. G. (1967). Latency for saccadic eye movement. *Journal of the Optical Society of America*, *57*(8), 1030.

Schacter, D. L., Addis, D. R., & Buckner, R. L. (2007). Remembering the past to imagine the future: The prospective brain. *Nature Reviews Neuroscience*, *8*(9), 657-661.

Scheepers, C., & Sturt, P. (2014). Bidirectional syntactic priming across cognitive domains: From arithmetic to language and back. *Quarterly Journal of Experimental Psychology*, *67*(8), 1643-1654.

Scheepers, C., Sturt, P., Martin, C. J., Myachykov, A., Teevan, K., & Viskupova, I. (2011). Structural priming across cognitive domains: From simple arithmetic to relative-clause attachment. *Psychological Science*, *22*(10), 1319-1326.

Schiller, N. O., Horemans, I., Ganushchak, L., & Koester, D. (2009). Event-related brain potentials during the monitoring of speech errors. *NeuroImage*, *44*(2), 520-530.

Schlichting, L. (2005). *Peabody Picture Vocabulary Test Dutch-III-NL.* Amsterdam, NL: Harcourt Assessment BV.

Schröger, E. (2007). Mismatch negativity: A microphone into auditory memory. *Journal of Psychophysiology*, *21*(3-4), 138-146.

Schwanenflugel, P. J., & Shoben, E. J. (1985). The influence of sentence constraint on the scope of facilitation for upcoming words. *Journal of Memory and Language*, *24*(2), 232-252.

Shao, Z., Janse, E., Visser, K., & Meyer, A. S. (2014). What do verbal fluency tasks measure? predictors of verbal fluency performance in older adults. *Frontiers in Psychology*, *5*(772).

Sjerps, M. J., & Meyer, A. S. (2015). Variation in dual-task performance reveals late initiation of speech planning in turn-taking. *Cognition*, *136*, 304-324.

Smith, A. C., Monaghan, P., & Huettig, F. (2013). An amodal shared resource model of language-mediated visual attention. *Frontiers in Psychology*, *4*(528).

Smith, N., & Levy, R. (2013). The effect of word predictability on reading time is logarithmic. *Cognition*, *128*(3), 302-3019.

Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures:

Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory*, *6*(2), 174-215.

Spalek, K., Hoshino, N., Wu, Y. J., Damian, M., & Thierry, G. (2014). Speaking two languages at once: Unconscious native word form access in second language production. *Cognition*, *133*(1), 226-231.

Staub, A. (2011). The effect of lexical predictability on distributions of eye fixation durations. *Psychonomic Bulletin & Review*, *18*(2), 371-376.

Staub, A., & Benatar, A. (2013). Individual differences in fixation duration distributions in reading. *Psychonomic Bulletin & Review*, *20*(6), 1304-1311.

Staub, A., & Clifton Jr, C. (2006). Syntactic prediction in language comprehension: Evidence from either... or. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*(2), 425-436.

Szewczyk, J. M., & Schriefers, H. (2013). Prediction in language comprehension beyond specific words: An ERP study on sentence comprehension in Polish. *Journal of Memory and Language*, *68*(4), 297-314.

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*(5217), 1632-1634.

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*(1), 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.

van Berkum, J. J. A., Hagoort, P., & Brown, C. (1999). Semantic integration in sentences and discourse: Evidence from the N400. *Journal of Cognitive Neuroscience*, *11*(6), 657-671.

van der Elst, W. I. M., van Boxtel, M. P. J., van Breukelen, G. J. P., & Jolles, J. (2006). Normative data for the animal, profession and letter m naming verbal fluency tests for dutch speaking participants and the effects of age, education, and sex. *Journal of the International Neuropsychological Society*, *12*(1), 80-89.

van Heuven, W. J., & Dijkstra, T. (2010). Language comprehension in the bilingual brain: fMRI and ERP support for psycholinguistic models.

Brain Research Reviews, *64*(1), 104-122.

van Heuven, W. J., Mandera, P., Keuleers, E., & Brysbaert, M. (2014). Subtlex-UK: A new and improved word frequency database for British English. *Quarterly Journal of Experimental Psychology*, *67*, 1176-1190.

Van Petten, C. (1993). A comparison of lexical and sentence-level context effects in event-related potentials. *Language and Cognitive Processes*, *8*(4), 485-531.

Van Petten, C. (2014). Examining the N400 semantic context effect item-by-item: Relationship to corpus-based measures of word co-occurrence. *International Journal of Psychophysiology*, *94*(3), 407-419.

Van Petten, C., & Luka, B. (2012). Prediction during language comprehension: Benefits, costs, and erp components. *International Journal of Psychophysiology*, *83*, 176-190.

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*(2), 63-72.

Whitford, V., & Titone, D. (2013). The effects of reading comprehension and launch site on frequency-predictability interactions during paragraph reading. *The Quarterly Journal of Experimental Psychology*, *67*(6), 1151-1165.

Wicha, N. Y. 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. 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.

Willems, R. M., Frank, S. L., Nijhof, A. D., Hagoort, P., & van den Bosch, A. (2015). Prediction during natural language comprehension. *Cerebral Cortex, Advance online publication*.

Wilson, M., & Knoblich, G. (2005). The case for motor involvement in perceiving conspecifics. *Psychological Bulletin*, *131*(3), 460-473.

Wlotko, E. W., & Federmeier, K. D. (2013). Two sides of meaning: The scalp-recorded N400 reflects distinct contributions from the cerebral hemispheres. *Frontiers in Psychology*, *4*, 181.

Wolpert, D. M., Doya, K., & Kawato, M. (2003). A unifying computational framework for motor control and social interaction. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *358*(1431), 593-602.

Yee, E., Overton, E., & Thompson-Schill, S. (2009). Looking for meaning: Eye movements are sensitive to overlapping semantic features, not association. *Psychonomic Bulletin & Review*, *16*(5), 869-874.

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

Ziefle, M. (1998). Effects of display resolution on visual performance. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *40*(4), 554-568.

Nederlandse samenvatting

Tijdens het begrijpen van taal maken lezers en luisteraars vaak voorspellingen over informatie die nog moet komen. Er worden verschillende functies beschreven van het anticiperen op taalinput. Over het algemeen wordt aangenomen dat het het begripsproces versnelt. In het dagelijks leven zijn mensen gewend taal te begrijpen in veel verschillende situationele contexten, waaronder bijvoorbeeld het begrijpen van geschreven en gesproken taal, dialogen en taal die refereert naar de visuele wereld. Deze dissertatie gaat over *hoe* de taalbegrijper informatie voorspelt in verschillende situaties van taalverwerking.

Eerder onderzoek heeft geleid tot een aantal verschillende theorieën over de mechanismes die ten grondslag liggen aan voorspelling. Sommige wetenschappers suggereren dat mensen tijdens het verwerken van zinnen mentale modellen opzetten over de gebeurtenissen die beschreven worden en hun algemene kennis over gebeurtenissen gebruiken om te voorspellen wat een logisch vervolg zou kunnen zijn. Andere taalonderzoekers beweren dat voorspelling tijdens het begrijpen van taal gebaseerd is op processen die ook betrokken zijn bij taalproductie. Meer specifiek gezegd wordt het voorspellen van een woord vaak beschreven als verwant aan het innerlijk produceren van een woord. Gedrag als dit kan herkend worden wanneer we naar de spraak van iemand anders luisteren en zijn zinnen (voor hem) afmaken. Een derde mogelijkheid is dat taalbegrijpers gebruik maken van frequenties van woordcombinaties die vaak samen voorkomen, waardoor het herkennen van het ene woord automatisch leidt tot het herkennen of activeren van een ander woord. Tenslotte is er in eerder onderzoek ook beschreven dat de waarschijnlijkheid van voorspelling gemediëerd wordt door de mate van geletterdheid van de begrijper. In hoofdstuk 2 tot en met 6 heb ik de betrokkenheid getest van de drie mechanismes en de mediërende factor op verschillende situaties van taalverwerking.

Taal wordt vaak gebruikt als referentie naar de onmiddellijk waarneembare visuele wereld. Eerder onderzoek liet zien dat informatie uit gesproken werkwoorden gebruikt kan worden om onze visuele aandacht ergens naartoe te leiden. Dit betekent dat bij het horen van een zin als “De man pelt de banaan” tijdens het zien van een visuele weergave met vier objecten waarvan één pelbaar was (banaan), participanten al naar de banaan keken bij het horen van “pelt”, nog voordat de banaan genoemd werd. In Hoofdstuk 2 heb ik de bijdrage onderzocht die associaties, productiegebaseerde mechanismen en geletterdheid leveren aan dergelijk werkwoordgemediëerd anticiperend kijkgedrag met behulp van drie eye-trackingexperimenten. In Experiment 1 en 2 zagen participanten de visuele weergave al voordat de gesproken zin werd afgespeeld. In Experiment 3 zagen ze de visuele weergave pas 500 ms voordat ze het targetwoord (“banaan”) hoorden. Het bleek dat beide mechanismen en de mediërende factor bijdragen aan voorspelling. Een belangrijke bevinding is dat productiegebaseerde mechanismen alleen bij bleken te dragen aan voorspelling wanneer luisteraars genoeg tijd hadden om vooraf de afbeeldingen te bestuderen, maar niet wanneer de afbeeldingen pas verschenen kort voordat het targetwoord gehoord werd.

In Hoofdstuk 3 heb ik verder onderzoek gedaan naar de hypothese dat contextuele visuele informatie invloed heeft op werkwoordgemedieerd anticiperend kijkgedrag. De oogbewegingen van participanten werden geregistreerd terwijl ze luisterden naar voorspelbare zinnen met dezelfde structuur als de voorspelbare zinnen in Hoofdstuk 2 (bijv. “De man pelt een banaan”). Tijdens het luisteren keken participanten naar verschillende typen situaties. Het targetobject (banaan) was óf aanwezig, óf het werd vervangen door een object met dezelfde visuele vorm als het targetobject (kano), óf het werd vervangen door een object dat semantisch gerelateerd was aan het concept dat door de target opgeroepen werd (aap). Elke trial werd gepresenteerd in zowel de vroege vertoning, waar participanten de afbeeldingen zagen voordat ze het werkwoord hoorden, als in de late vertoning, waar participanten de afbeeldingen pas na het werkwoord te zien kregen, vlak voor de aanvang van het targetwoord. Het bleek dat participanten anticipeerden op de targetobjecten in beide condities. Echter, robuust bewijs voor effecten van visuele vorm en semantische competitie werd alleen gevonden in de late vertoning en niet in de vroege vertoning. De afwezigheid van kijkgedrag naar de objecten met dezelfde visuele vorm en semantisch gerelateerde objecten voordat het targetwoord gehoord was, suggereert dat de visuele context restricties oplegt aan de dynamiek van ancitiperende taalverwerking. Dat wil zeggen: wanneer participanten de visuele vertoning eerder te zien kregen, inspecteerden ze deze en zeer waarschijnlijk werd hierdoor kennis over de afgebeelde objecten geactiveerd, waaronder ook bepaalde acties waarin de objecten vaak gebruikt worden. Bij trials waar het targetobject aanwezig was, ontstond er een match tussen de kennis die geactiveerd werd door de visuele input en de kennis die geactiveerd werd bij het horen van het werkwoord, wat leidde tot oogbewegingen naar het targetobject. Echter, bij trials waar de target afwezig was ontstond geen match; daarom werden er geen oogbewegingen waargenomen naar visueel- en semantisch gerelateerde objecten. De resultaten beschreven in Hoofdstuk 3 bevestigen de notie dat taalgemedieerde anticiperende oogbewegingen substantieel beïnvloed worden door contextuele visuele informatie.

In Hoofdstuk 4 werd onderzocht wat associaties en kennis over gebeurtenissen bijdragen aan voorspellingen tijdens het begrijpen van contexten. Het EEG-signaal van participanten werd opgenomen terwijl ze targetwoorden lazen, voorafgegaan door het lezen van associatief gerelateerde woorden in een coherente context die een gebeurtenis beschrijft (Experiment 1) of in zinnen die geen coherente context vormden (Experiment 2). Vergelijkbaar met eerder onderzoek bleek in Experiment 1 dat contextueel onverwachte targetwoorden die geassocieerd werden met de beschreven gebeurtenis een N400 veroorzaakten met een gereduceerde amplitude vergeleken met contextueel onverwacht targetwoorden die ongerelateerd waren aan de gebeurtenis. Naar alle waarschijnlijkheid reflecteert dit de gecombineerde bijdrage van gebeurtenissenkennis en associaties aan voorspellingen. In Experiment 2, waar de invloed van gebeurtenissenkennis geminimaliseerd was, werd een vergelijkbaar maar gereduceerd N400-effect gevonden. Het voorspellingseffect in Experiment 2 geeft dus waarschijnlijk voornamelijk de bijdrage van associatieve priming weer.

In Hoofdstuk 5 werd een nieuwe versie van het *visual world paradigm* gebruikt om te testen hoe vergelijkbaar voorspellings- en productieprocessen zijn, een vereiste voor productiegebaseerde benaderingen van voorspelling. De participanten luisterden naar rekensommen terwijl ze keken naar een wijzerplaat met daarop de getallen 1 tot en met 12. Ze hoorden om en om een som inclusief antwoord ("drie plus acht is elf") of hoorden enkel de som ("drie plus acht is") en moesten het antwoord zelf produceren. De participanten werden aangemoedigd om naar de relevante getallen te kijken gedurende de trials. Ze fixeerden op het juiste getal voordat de opgenomen spreker het noemde, en ook voordat ze het zelf noemden. De oogbewegingen gerelateerd aan het anticiperen op de spraak van iemand anders en het voorbereiden op het zelf spreken waren over het algemeen erg vergelijkbaar. De tijdsduur tot de fixatie was echter korter bij productietrials dan bij begripstrials, wat suggereert dat de processen betrokken bij productie en voorspelling – hoewel vergelijkbaar – verschillen met betrekking tot hoe alert of betrokken mensen zijn wanneer ze van plan zijn te reageren of wanneer ze enkel luisteren naar iemand anders. De resultaten van Hoofdstuk 5 benadrukken dat er rekening gehouden moet worden met motivatieverschillen tussen de twee processen in theorieën die productiegebaseerde mechanismen veronderstellen.

Gerelateerd aan verschillen in motivatie tussen voorspellings- en productieprocessen werd in Hoofdstuk 6 onderzocht hoe verschillende taken invloed uitoefenen op de mate waarin voorspelling een rol speelt in het begrijpen van taal. Ik heb onderzocht of het aanmoedigen van voorspelling in een taalproductietaak ook voorspelling in begrip aanmoedigt. De participanten luisterden naar het eerste deel van een zin en gaven het laatste woord door een afgebeeld object te benoemen (Experiment 1a), of ze lazen dezelfde zinnen in een self-paced leestaak (Experiment 1b). Het laatste woord in de zin was voorspelbaar of niet-voorspelbaar. In Experiment 1a werd de reactietijd tot het benoemen van het object gemeten; in Experiment 1b werden de leestijden gemeten. De resultaten lieten zien dat voorspelbare trials sneller benoemd werden dan niet-voorspelbare trials (Experiment 1a). Echter, een dergelijk effect werd niet gevonden wanneer de targetwoorden gelezen werden (Experiment 1b). Bij een combinatie van trials met lezen en trials met benoemen (Experiment 2) vond ik wel een faciliterend effect in beide taken. Deze resultaten suggereren dat het aanmoedigen van voorspelling ten gunste van taalproductie ook leidt tot een toenemende waarschijnlijkheid van voorspelling in taalbegrip.

Samengevat laten de beschreven resultaten zien dat wanneer volwassen lezers en luisteraars voorspellingen maken over taal, verschillende mechanismen en mediërende factoren parallel worden gebruikt. Ook passen mensen het gebruik van deze mechanismen aan als reactie op situationele en taakgerelateerde behoeften. Taalgebruikers kunnen begripsstrategieën dynamisch aanpassen aan de situationele en taakgerelateerde context om op die manier een zo efficiënt mogelijke vorm van begrip te bereiken.

Acknowledgements

When I started my PhD in 2012, I could not predict what lay ahead of me. Ironically, I could not predict that I would be working on 'prediction' as I initially started on a different topic. Despite, and in several cases because of unpredicted events along the way, I am looking back on a great time at a great research institute with great people who have contributed to making this a unique phase of my life. Unfortunately, I cannot mention all persons here, but I am grateful to everybody.

First of all, I would like to thank my doctoral supervisors Falk Huettig and Antje Meyer who provided the stimulating work environment in which this thesis was written. Falk, thanks for trusting in me, for the countless scientific things that I have learned from you and for all the humor. Your sharp interpretation of experimental results and your creative approach to conducting research have shaped my way of thinking. I have been enjoying working with you since my internship at the MPI in 2010. Antje, thank you for heading such a great department, for the excellent support throughout my PhD, for always demanding clarity and precision in thinking and writing, and for providing freedom and possibilities.

I am indebted to Prof. Peter Hagoort, Prof. Pia Knoeferle, and Dr. Roel Willems who took the time to read and evaluate this thesis.

Many thanks to the members and associates of the Psychology of Language department—those who are still around and those working in other places now: Agnieszka, Alastair, Alma, Amie, Annelies, Caitlin, Chu, Conny, Dan, Ernesto, Esther, Evelyn, Hans Rutger, Jana, Jannie, Johanne, Loan, Markus, Maartje, Matthias, Neil, Nina, Renske, Shiri, Suzanne, Svetlana, Will, Xuefei, and Zeshu—thanks for all the fun moments at work, at conferences, and in general for making my time at the MPI so enjoyable. I would also like to thank the research assistants of the Psychology of Language department for their help in preparing and running my experiments, in particular: Ilse, Laura, Marloes G., Marloes W., Sophie, and Stephanie. Chapter 5 in this thesis benefitted from the input of Elisabeth Lehmann and Benedikt Langenbach who worked as interns on the project—thanks for your contributions.

Here, I would also like to mention Katharina von Kriegstein who hired me as a student assistant at the MPI in Leipzig in 2009 and guided me on my first steps into the world of cognitive science. I will always remember the two wonderful years that I have worked in her research group.

Thank you to the technical group of the MPI, especially Reiner Dirksmeyer, Alex Dukers, Ronald Fischer, and Johan Weustink who ensured that my computer and experiments run smoothly. Apart from putting up with my rudimentary command of Dutch, Albert Russel always did a tremendous job in implementing my online experiments. Many thanks, Albert, for being a great help and showing a friendly face. Thanks also to the MPI's library team, Karin and Meggie, and the coordinators of the IMPRS, Dirkje and Els.

Thank you, Antje (Strauß) and Joost for your input on the EEG chapter. Thanks to Tijn, for translating the English summary of my dissertation into Dutch.

After a week in the office, I really enjoyed playing football on Friday evenings. Many thanks to the MPI football team, especially to Guilherme for booking the pitch.

To my paranymphs, Suzanne, Alastair, and Paul (in spirit), I am extremely happy to have you stand next to me during my defense. Suzanne, thank you very much for always listening to my stories and laughing about my jokes. I realized at a certain point that a lot of politeness must have been involved in the latter—I still appreciate it though. I also appreciate that your office door was always open for chats and quick (or sometimes lengthy) scientific discussions. Thank you for being such a loyal friend and colleague! Alastair, my dear office mate, it was on a Friday after a match of football at the Sports Center when we first talked. I had not yet started at the MPI but thought for myself that I would probably learn a lot about football, the UK and science from this guy. Ok, the order of topics was primed at the time, but I was right: I learned a lot from you and I am extremely grateful for your unconditional support, even in times when you were super busy yourself. Thank you for introducing me to golf, for great conferences in Riva del Garda, Budapest, and Edinburgh and for the nice atmosphere in our office! I hope that we will be able to continue this relationship, albeit by enjoying the exquisite cuisine of the Refter once per week. Paul, was mit einer harmlosen Begegnung beim MPI Fußball begann, endete in einer mich stets bereichernden Freundschaft. Ob Wissenschaft, das gesellschaftliche Gesamtgefüge oder einfach nur platte Wortwitze—die Unterhaltungen mit dir bereiten mir grundsätzlich gute Laune und haben mich so manches Mal vom Ernst des ‚Day-Jobs' abgelenkt. Vielen Dank für dein immer offenes Ohr, für deine guten Ratschläge in allen Lebenssituationen, für das Organisieren meines Junggesellenabschiedes, und für alle anderen witzigen Momente in Nimwegen, Kleve, Leipzig und am Bodensee.

Es bleiben die Menschen, ohne die nichts geht und ohne die nichts gegangen wäre. Mama, Papa, und Kati, obwohl ich für meine Doktorarbeit im wahrsten Sinne des Wortes ans andere Ende des Landes gezogen bin, habt ihr mir das Gefühl gegeben, nicht weit von der Heimat zu sein. Ich danke euch für eure Unterstützung während meines Studiums und der Promotion, für euer Vertrauen und für die Freiheit, das zu machen, was mir Spaß macht.

Liebe Cornelia, mein kleiner Julius, 2015 war wahrhaftig ein aufregendes Jahr für uns drei. Es war ein Jahr, das mich in vielerlei Hinsicht sehr stolz gemacht hat. Stolz darauf, meine Promotion abgeschlossen zu haben, aber mindestens genauso stolz, der Mann einer so tollen Frau und der Vater eines so goldigen Sohnes zu sein. Vielen Dank für eure Unterstützung und für eure Liebe.

Curriculum Vitae

Florian Hintz was born in 1987 in Bernau, Germany. In 2010, he obtained his Bachelor's degree and in 2012 his Master's degree in Linguistics, both from the University of Leipzig, Germany. In the fall of 2011, he spent one semester studying Linguistics and Cognitive Neuroscience at the Radboud University. In 2012, Florian started his PhD project at the Max Planck Institute for Psycholinguistics in the Psychology of Language Department. He is currently a postdoctoral researcher at the Center for Language Studies at the Radboud University.

Publications

Hintz, F., Meyer, A.S., & Huettig, F. (in preparation). Predictors of verb-mediated anticipatory eye movements in the visual world.

Hintz, F., Meyer, A.S., & Huettig, F. (in preparation). The relative contribution of event knowledge and word associations to prediction during discourse reading.

Hintz, F., Meyer, A.S., & Huettig, F. (submitted). I see what you're going to say: Listeners' use of contextual visual information during predictive language processing restricts visual and semantic competition.

Hintz, F., Meyer, A.S., & Huettig, F. (submitted). Encouraging prediction during production facilitates subsequent comprehension: Evidence from interleaved object naming in sentence context and sentence reading.

Hintz, F. & Huettig, F. (2015). The complexity of the visual environment modulates language-mediated eye gaze. In R.K. Mishra, N. Srinivasan, & F. Huettig (Eds.), *Attention and Vision in Language Processing* (pp. 39-56). New Delhi: Springer.

Hintz, F., & Meyer, A.S. (2015). Prediction and production of simple mathematical equations: Evidence from anticipatory eye movements. *PLoS One, 10*(7).

Roswandowitz, C., Mathias, S.R., Hintz, F., Kreitewolf, J., Schelinkski, S., & von Kriegstein, K. (2014). Two cases of selective voice-recognition impairments. *Current Biology, 24*(19), 2348-2353.

Díaz, B., Hintz, F., Kiebel, S.J., & von Kriegstein, K. (2012). Dysfunction of the medial geniculate body during speech perception in developmental dyslexia. *Proceedings of the National Academy of Sciences of the United States of America* (PNAS), *109*(34), 13841-13846.

MPI Series in Psycholinguistics

1. The electrophysiology of speaking: Investigations on the time course of semantic, syntactic, and phonological processing. *Miranda I. 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 Petra 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 J.Janssen*

12. Valence and transitivity in Saliba: An Oceanic language of Papua New Guinea. *Anna Margetts*

13. From speech to words. *Arie H. 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 Christine Weber*

17. Moving eyes and naming objects. *Femke Frederike 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 Helena de Jong*

21. Fixed expressions and the production of idioms. *Simone Annegret 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 cross-linguistic experimental studies. *Fermín Moscoso del Prado Martín*

24. Contextual influences on spoken-word processing: An electrophysiological approach. *Danielle van den Brink*

25. Perceptual relevance of prevoicing in Dutch. *Petra Martine 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 Henriëtte Wilhelmina Meeuwissen*

28. Morphology in auditory lexical processing: Sensitivity to fine phonetic detail and insensitivity to suffix reduction. *Rachèl Jenny Judith Karin 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. *Marijtje Elizabeth Debora 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 Elisabeth Broersma*

35. Retrieving semantic and syntactic word properties: ERP studies on the time course in language comprehension. *Oliver Müller*

36. Lexically-guided perceptual learning in speech processing. *Frank Eisner*
37. Sensitivity to detailed acoustic information in word recognition. *Keren Batya 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 non-constrative variation. *Christiane Dietrich*
41. Cognitive cladistics and the relativity of spatial cognition. *Daniel Haun*
42. The acquisition of auditory categories. *Martijn Bastiaan Goudbeek*
43. Affix reduction in spoken Dutch: Probabilistic effects in production and perception. *Mark Plumaekers*
44. Continuous-speech segmentation at the beginning of language acquisition: Electrophysiological evidence. *Valesca Madalla Kooijman*
45. Space and iconicity in German sign language (DGS). *Pamela M. 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 Eva 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 Ursula 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 Payton 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 Mary Magteld Junge*

68. Adjusting to different speakers: Extrinsic normalization in vowel perception. *Matthias Johannes Sjerps*

69. Structuring language: Contributions to the neurocognition of syntax. *Katrien Rachel 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 speech 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*

82. Mental representation and processing of reduced words in casual speech. *Iris Hanique*

83. The many ways listeners adapt to reductions in casual speech. *Katja Pöllmann*

84. Contrasting opposite polarity in Germanic and Romance languages: Verum Focus and affirmative particles in native speakers and advanced L2 learners. *Giuseppina Turco*

85. Morphological processing in younger and older people: Evidence for flexible dual-route access. *Jana Reifegerste*

86. Semantic and syntactic constraints on the production of subject-verb agreement. *Alma Veenstra*

87. The acquisition of morphophonological alternations across languages. *Helen Buckler*

88. The evolutionary dynamics of motion event encoding. *Annemarie Verkerk*

89. Rediscovering a forgotten language. *Jiyoun Choi*

90. The road to native listening: Language-general perception, language-specific input. *Sho Tsuji*

91. Infants'understanding of communication as participants and observers. *Gudmundur Bjarki Thorgrímsson*

92. Information structure in Avatime. *Saskia van Putten*

93. Switch reference in Whitesands. *Jeremy Hammond*
94. Machine learning for gesture recognition from videos. *Binyam Gebrekidan Gebre*
95. Acquisition of spatial language by signing and speaking children: A comparison of Turkish sign language (TID) and Turkish. *Beyza Sumer*
96. An ear for pitch: On the effects of experience and aptitude in processing pitch in language and music. *Salomi Savvatia Asaridou*
97. lncrementality and Flexibility in Sentence Production. *Maartje van de Velde*
98. Social learning dynamics in chimpanzees: Reflections on (nonhuman) animal culture. *Edwin van Leeuwen*
99. The request system in Italian interaction. *Giovanni Rossi*
100. Timing turns in conversation: A temporal preparation account. *Lilla Magyari*
101. Assessing birth language memory in young adoptees. *Wencui Zhou*
102. A social and neurobiological approach to pointing in speech and gesture. *David Peeters*
103. Investigating the genetic basis of reading and language skills. *Alessandro Gialluisi*
104. Conversation electrified: The electrophysiology of spoken speech act recognition. *Rósa Signý Gisladottir*
105. Modelling multimodal language processing. *Alastair Charles Smith*
106. Predicting language in different contexts: The nature and limits of mechanisms in anticipatory language processing. *Florian Hintz*