The Effects of Semantic and Syntactic Prediction on Reading Aloud

Elisa Gavard, Johannes C Ziegler

The effects of semantic and syntactic prediction on reading aloud

Elisa Gavard[1] and Johannes C. Ziegler[1]

[1] Laboratoire de Psychologie Cognitive (UMR 7290), Aix-Marseille Univ, CNRS, Marseille, France

Author Note

Elisa Gavard https://orcid.org/0000-0001-5144-7767

Johannes C. Ziegler https://orcid.org/0000-0002-2061-5729

Correspondence should be addressed to Elisa Gavard or Johannes Ziegler, Laboratoire de Psychologie Cognitive, Aix-Marseille University, 3, place Victor Hugo, 13003 Marseille. France. Email: elisa.Gavard@univ-amu.fr or johannes.Ziegler@univ-amu.fr

The information needed to reproduce all the reported results and methodology is available at https://osf.io/geyhs/?view_only=caf59bfdacae4f309833caecdd5b55ad

Abstract

Semantic and syntactic prediction effects were investigated in a word naming task using semantic or syntactic contexts that varied between three and six words. Participants were asked to read the contexts silently and name a target word, which was indicated by a color change. Semantic contexts were composed of lists of semantically associated words without any syntactic information. Syntactic contexts were composed of semantically neutral sentences, in which the grammatical category but not the lexical identity of the final word was highly predictable. When the presentation time of the context words was long (1200 ms), both semantically and syntactically related contexts facilitated reading aloud latencies of target words and syntactically related contexts produced larger priming effects than semantically related contexts in two out of three analyses. When presentation time was short (200 ms), however, syntactic context effects disappeared, while semantic context effects remained significant. Across the three experiments, longer contexts produced faster response latencies, but longer contexts did not produce larger priming effects. Results are discussed in the context of the extant literature on semantic and syntactic priming and more recent evidence suggesting that syntactic information constrains single word recognition.

Keywords: semantic prediction, syntactic prediction, reading aloud, word recognition

The effects of semantic and syntactic prediction on reading aloud

The idea of the *predictive brain* has become a key concept in neuroscience (Friston & Kiebel, 2009), cognitive science and psychology (Hohwy, 2013; Lupyan & Clark, 2015) over the past decade. In the context of this theory, the brain could be viewed as a “prediction engine” rather than a passive receiver of data (Clark, 2013). That is, it makes predictions about future events and seeks to minimize errors in these predictions, which is the basis for implicit and adaptive learning (Grossberg, 2012). The key idea is that the brain proactively anticipates what will happen next (Bar, 2009; den Ouden et al., 2012; Friston, 2010; Rao & Ballard, 1999) and this is seen as having two advantages in information processing. First, if an outcome conforms to expectations, the processing of that event might be facilitated because some of the features of an upcoming event might have been pre-activated (priming). Second, if an outcome does not conform to expectations, then the brain must update its model of that domain, which provides an opportunity for adaptive learning and the formation of rich and accurate representations.

This idea about the brain being a prediction engine has been rather influential in psycholinguistics, where prediction is seen as one of the cognitive mechanisms involved in language comprehension, language production and reading (Chang et al., 2006; Federmeier & Kutas, 1999; Pickering & Garrod, 2013; Staub, 2015). For instance, it has been argued that effective language processing in children and adults results from efficient use of prediction in both word and speech perception (Bar, 2011; Drake & Corley, 2015; Mani & Huettig, 2012; Pickering & Gambi, 2018). In word production, speakers constantly predict their own utterances for the purpose of speech planification and speech monitoring. Similarly, listeners predict future utterances as they listen (Johnson et al., 2013).

The influence of word predictability on reading has been extensively investigated in research on eye movements. Ehrlich and Rayner (1981) were the first to show that fixations were shorter and less frequent to highly predictable words than to less predictable words. Moreover, when a word can be predicted from the preceding context, the eyes are relatively likely to skip over the word rather than directly fixate it. These effects have been replicated in dozens of eye movement experiments (for review, see Staub, 2015). Typically, predictability is operationalized in terms of a word’s cloze probability (Taylor, 1953), that is the proportion of subjects who provide a given word in an offline sentence completion task. For example, in the sentence “It was a windy day, the boy went outside to fly his …”, “kite” has a higher cloze probability than “airplane” although both words perfectly fit the context. Indeed, predictability along with word frequency is one of the two main variables that affect eye movements in current computational models of eye movements in reading, such as E-Z Reader (Reichle et al., 1998) and SWIFT (Engbert et al., 2005). Word predictability also plays an important role in language processing models, in which processing difficulty is a function of a word's probability in its context (Hale, 2001; Levy, 2008).

The influence of word predictability on reading and on-line language comprehension has also been studied using event-related brain potentials (for reviews, see DeLong et al., 2005; Kutas & Federmeier, 2011). The N400, which is the brain's electrophysiological response to any potentially meaningful item, is especially large to nouns that do not meaningfully fit with their preceding contexts ("the pizza is on the shoe"; Kutas & Hillyard, 1980). Even when there is no semantic mismatch and both target words fit the context ("airplane" versus "kite" in the above example), the N400 is strongly inversely related (~-.90) to the most expected target (Kutas & Hillyard, 1984).

Note that most of the above research focused on semantic rather than syntactic predictions or a mix of both (i.e., cloze probability). In the sentence "it was a windy day, so the boy went outside to fly his __", people not only make a semantic prediction (the target word must be a flying object) but also a syntactic prediction (the target word must be a noun). Indeed, any time a semantic prediction is made, a syntactic prediction is generated as well (Ferreira & Qiu, 2021). However, this does not work the other way around. That is, people can make a syntactic prediction without being able to make a semantic prediction. For example, if the sentence starts with "Those_", people would be able to predict a plural noun phrase, but it would be difficult to make a semantic prediction about the upcoming word (see Ferreira & Qiu, 2021).

Both semantic and syntactic prediction effects have extensively been investigated using the priming paradigm, in which a single word prime influences the reading of a subsequent, related target word (Becker, 1980; den Heyer et al., 1985; Forster, 1998). The seminal finding goes back to Meyer & Schvaneveldt (1971) who showed that the recognition of a target word (e.g., "nurse") was facilitated if was preceded by a highly associated priming word (e.g., "doctor"), rather than an unrelated prime (e.g., "chair"). Although there has been some discussion as to whether semantic priming effects require associative relations between primes and targets (Fischler, 1977; Lupker, 1984) and whether semantic priming can be obtained with short or subliminal primes (Draine & Greenwald, 1998; Fischler & Goodman, 1978; Hohenstein et al., 2010), it seems to be the case that pure semantic priming effects can be observed in both lexical decision and naming even at relatively short stimulus-onset asynchronies (SOAs) (Perea & Gotor, 1997) suggesting that at least one part of the effect reflects the automatic "spreading of activation" to words that are semantically related to the primes (Collins & Loftus, 1975).

Much less research has been conducted on syntactic priming. In the seminal study by Goodman et al. (1981), the authors used single words as primes (e.g., articles and pronouns) that unequivocally predicted the syntactic class of the target (i.e., noun and verb, respectively). They found that lexical decision latencies to targets were significantly shorter when they were preceded by a syntactically appropriate word (e.g., my oven) compared to a syntactically inappropriate word (e.g., he oven). Seidenberg et al. (1984) replicated their study in both a lexical decision and naming task. They found significant syntactic priming effects only in lexical decision but not in naming. Finally, Sereno (1991) investigated orthographic, associative, and syntactic priming effects in both a lexical decision and naming using fast masked priming with a prime duration of only 60 ms. She found that primes that were orthographically similar or identical to targets facilitated both lexical decision and naming. However, associatively or syntactically related primes facilitated only lexical decision but not naming. She argued that fast-acting semantic and syntactic priming affects a post-lexical decision stage that would be operative in lexical decision but not in naming. Together then, it seems that different processes underlie semantic and syntactic priming.

In the present study, we decided to move away from the classic priming paradigm and use longer contexts to build up stronger semantic and syntactic predictions. We also used a task that allowed us to dissociate semantic from syntactic contexts while investigating responses to exactly the same target words. Finally, we focused on naming rather than lexical decision because naming is not subject to post-lexical decision strategies (e.g., Seidenberg et al., 1984). In our task, a succession of words (between 4 and 7) was presented on the screen and participants were asked to read the words silently until a color change indicated a target word that had to be read aloud as quickly as possible. To assess the effects of semantic prediction, we compared a context of semantically related nouns (cat - dog – rabbit - mouse) with a context of unrelated nouns (table - green - flower - mouse). To assess the effects of syntactic prediction for the same target words, we compared syntactically correct sentences (she - likes - this - mouse) with syntactically incorrect scrambled sentences (this - likes - she - mouse). Predictive strength of the context was manipulated by using contexts that varied between three and six words. Experiments 1 and 2 used a rather long presentation time of the context words (1200 ms) to give more time for semantic and syntactic predictions to build up; Experiment 3 used a shorter presentation time (200 ms) to test semantic and syntactic prediction in more ecological reading conditions.

We obtained the semantic relatedness between our context words and between the context and target words using a common vector space model (word2vec, Mikolov et al., 2013, 2013) trained on a large French database (WikiFr). To compute syntactic prediction strength, we calculated conditional probabilities for grammatical categories of targets words as a function of preceding contexts using the Universal Dependencies corpus (de Marneffe et al., 2021). This allowed us to create sentences with syntactically highly predictive target words, for which the semantic relatedness was extremely weak.

We expected related contexts to facilitate reading aloud compared to unrelated or scrambled context. Given that highly predictive semantic contexts made it possible to pre-activate the precise target word, whereas highly predictive syntactic contexts could only pre-activate the grammatical class of the target word, we expected stronger context effects for semantic as opposed to syntactic contexts. We also predicted that the size of the priming effects should vary as a function of context length and SOA with greater priming effects for longer contexts and longer SOAs.

Experiment 1

Method

Participants. We recruited 20 participants from Aix-Marseille University (Marseille, France), all French native speakers (13 women), reported normal or corrected-to-normal vision, no learning disorders (dyslexia) and no neurological or psychiatric disorders. They were between 19 and 23 years old (M = 21.8; SD = 1.5) and signed informed-consent forms prior to participation.

Design and Stimuli.

Semantic prediction. We first selected semantic contexts, for which we could find a relatively large number of semantically related words (e.g., flowers, animals, transportation etc.). We then calculated the semantic relatedness between all context words using a distributional vector space model of semantic associations trained on a very large number of words (Frwiki, 11GB, 914,601,321 tokens, DISCO, Kolb, 2008). This allowed us to select from these semantic contexts 80 target words that were the most predictive in a given context (average cosine value = .75). Note that vector space models compute semantic similarities rather than co-occurrences or associative relationships (for a discussion, see Günther et al., 2019). For the same target words, we created 80 unrelated contexts matched in word length and frequency that had no semantic relatedness according to DISCO (average cosine value = .12). Context length varied systematically with 20 target words per context length (3, 4, 5, 6). Note that in terms of cosine similarities (semantic prediction), shorter context were as predictive as longer contexts ($F(1,3) = 1.32$, $p = .27$). An example is given in Table 1.

Table 1

Example of target words presented in a semantic or syntactic context, in a related or unrelated condition for one subject (length of 4 words).

	Condition	
Context	Related	Unrelated
Semantic	cat – dog – rabbit – mouse	table – green – flower – car
Syntactic	they – want – a – car	this – likes – she – mouse

Syntactic Prediction. For the syntactic condition, we constructed 80 sentences using the same target words as in the semantic condition. Importantly, the syntactic contexts were constructed such that the target word was highly predictive on a syntactic basis but not at all on a semantic basis (verified using the same vector space model as above). Syntactic predictability was calculated using the Universal Dependencies corpus (de Marneffe et al., 2021). For each target word and sentence length, we calculated conditional probabilities of predicting the grammatical class of the target word taken account the preceding context (P(target/context) > .60). Eighty syntactically nonpredictive contexts were created by using the same target words but scrambling the words of the context such that syntactic prediction was close to zero (P(target/context) < .01). As in the semantic condition, context length varied systematically with 20 target words per context length (3, 4, 5, 6). Note that in terms of probabilities (syntactic prediction), shorter context were as predictive as longer contexts ($F(1,3) = .94$, $p = .43$).

Experimental Design. As illustrated in Table 1, we had three factors that were fully crossed: condition (semantic vs. syntactic), context (related vs. unrelated) and context length (3, 4, 5, 6 words). To limit target repetition, we created two stimulus lists (for two groups of participants) such that a related target in the semantic condition would be an unrelated target in the syntactic condition and vice versa. This was done such that a given participant would not encounter the same target word twice in the same condition. Altogether, the experiment consisted of two lists of 160 trials composed of 80-word sequences in the semantic condition (40 related and 40 unrelated) and 80 sentences in the syntactic condition (40 related and 40 unrelated). Semantic and syntactic conditions were blocked, and their order of presentation was counterbalanced across subjects. In each condition, the word sequences or sentences were presented in random order to each participant.

Procedure. All participants agreed to take part in the study by filling out a consent form after the experimenter had explained the course of the study. The experiment was programmed with E-Prime 2.0. Stimuli were presented on a computer screen. Each trial began with the presentation of a fixation cross in the center of the screen for 1000 ms. Context words in both conditions were presented one by one in the center of the screen for 1200 ms with a 100 ms inter-stimuli-interval (ISI). The target word remained on the screen until the participant read it out loud. The next trial started after an inter-trial-interval (ITI) of 350 ms. The words were displayed in the center of the screen, in lowercase 30-point Arial font. Participants received instructions verbally by the experimenter and visually on the screen.

Participants were instructed to read all context words silently and read aloud the target word as quickly as possible. The presence of a target word was indicated by a color change. Participants could not fully anticipate when exactly a target word appeared because the contexts varied in length. The experimenter registered the accuracy of the pronunciations using a Chronos® device. Responses were recorded using a microphone and response latencies (RTs) were calculated as the differences between the onset of the target word presentation and the onset of the naming response. The onset of the naming response was automatically detected by the Eprime software and the onset tags were verified manually for each response. A session typically lasted for about 20 minutes.

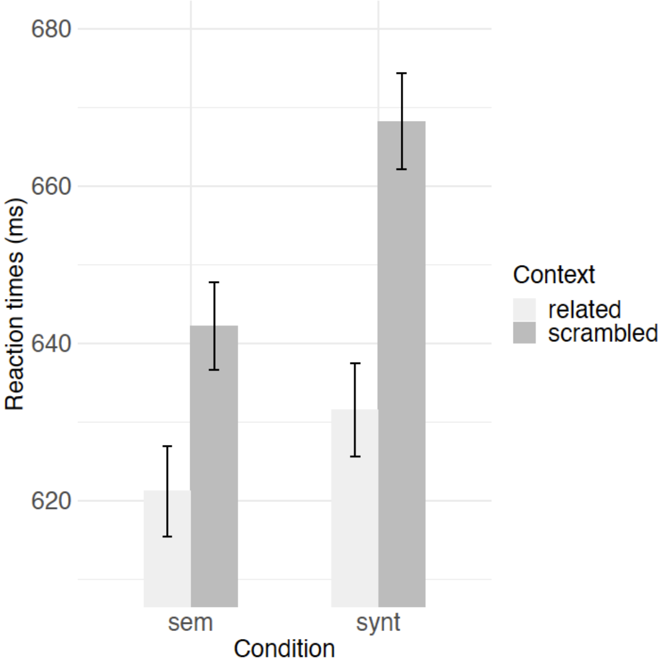
Analyses. We used R (Version 4.0.2) and *lme4* (Bates et al., 2015) to perform a linear mixed-effects (LME) analysis with participants and items as random effects and condition, context, and length as fixed effects. Out of all 3200 observations (RTs), 339 were removed because of a dysfunction of the microphone. We then removed extreme values (i.e., RTs below 200 ms or above 4000 ms) and we further considered as outliers data points that were above or below 2.5 standard deviations from each individual participant's mean RT. Outliers were replaced by the cutoff RT corresponding to each participant's ± 2.5 standard deviation (see Hoaglin et al., 1986). We report unstandardized regression coefficients (b), standard errors (*SEs*), and t-values. Fixed effects were considered significant if $|t|$ was greater than 1.96 (Baayen, 2008). Fixed effects, random effects, and random slopes were only included if they significantly improved the model's fit in a forward stepwise model selection procedure.

Results

Participants did not make any naming errors. Thus, accuracy was not further analyzed. Because the RT distribution was slightly skewed, we log-transformed the RTs. This resulted in a perfectly normal RT distribution (Skewness: -.14, Kurtosis: -.21) (Blanca et al., 2013). For the LME analysis, the final model included condition, context, length, and all interactions between these variables. As random effects, we included by-participant intercepts, random slopes for condition and random slopes for context by participants. The results showed a significant effect of context (b = 0.02, *SE* = 0.006, *t* = 3.14) but no effect of condition. Although the syntactic priming effect appeared larger than the semantic priming effect, the interaction between the effects of context and condition failed to reach significance (b = 0.01, *SE* = 0.01, *t* = 1.14). Contrast analysis confirmed that participants were faster at reading the target word following related than unrelated contexts in both the semantic (b = -0.02, *SE* = 0.003, *t* = -4.58) and syntactic condition (b = -0.02, *SE* = 0.004, *t* = -6.31) (see Figure 1). There was also an effect of context length (b = -0.02, *SE* = 0.007, *t* = -3.3), which reflected the fact that participants were faster as the context was getting longer, but this effect did not interact with the context (related versus scrambled).

Figure 1

Mean reaction times (ms) as a function of context (related vs scrambled) and condition (semantic vs syntactic) in Experiment 1. Errors bars indicate within-participant standard errors.



Discussion

In line with our hypotheses, we found a rather large and significant prediction effect for both conditions. Semantically and syntactically related contexts, as compared to unrelated contexts, facilitated reading aloud of single target words. Although the priming effect appeared stronger in the syntactic than the semantic condition, the interaction between the effects of context and condition was not significant. Furthermore, both types of contexts (semantic and syntactic) produced an increase in response latencies as a function of context length (the longer the context, the faster the responses). However, this effect occurred whether the context was related or scrambled, which suggests that the effect was due to general response preparation (i.e., the longer participants wait, the faster they get in responding to the target word).

Experiment 2

Experiment 2 was designed as a replication of Experiment 1 with two main differences. First, we used the same target word in each of the four conditions (semantically related and unrelated, syntactically related and unrelated) for each participant. Second, we introduced a prediction error condition in the semantically and syntactically related condition. That is, on 20% of the trials, a semantically or syntactically related context was followed by a mismatching (incongruent) target word. All other aspects of the design and procedure were identical to those of Experiment 1.

Method

Participants. We recruited 20 participants (17 women) at Aix-Marseille University (Marseille, France). All were native speakers of French and reported normal or corrected-to-normal vision, no learning disorders (dyslexia) and no neurological or psychiatric disorders. They were between 18 and 21 years old (M = 19.4; SD = 0.8) and signed informed consent forms prior to participation.

Design and Stimuli. Because each target word was seen in all contexts by all participants, we selected 40 target words that were yoked with related and unrelated semantic and syntactic contexts (see Table 2 for examples). In addition, we added 8 sequences in each related condition where the target word violated the semantic or syntactic prediction (incongruous condition). Thus, in the semantic condition, we had 88 sequences composed of 40 semantically related targets, 40 unrelated targets, and 8 incongruous targets. The target words were identical in each condition, only the semantic context changed. The unrelated condition was made up of the same context words as the related condition, but they were randomly scrambled, controlling for word length and frequency. Analyses with the vector space model (DISCO) confirmed that related context had an average cosine value off .75, whereas unrelated context had an average cosine value off .12. Context length was evenly distributed with 22 targets per context length (3, 4, 5, 6). Note that in terms of cosine similarities (semantic prediction), shorter context were also as predictive as longer contexts than in the Experiment 1.

Similarly, in the syntactic condition, we had 88 sentences made up of 40 syntactically correct, 40 incorrect (scrambled), and 8 incongruent sentences (see Table 2 for examples). The same target words were the same in each condition. As for the semantic condition, the unrelated context was composed of the same sentences as in the related context, but context words were scrambled such that conditional syntactic prediction probability (P(target/context)) was lower than .01. Context length was evenly distributed with 22 targets per context length (3, 4, 5, 6). Note that in terms of probabilities (syntactic prediction), shorter context were also as predictive as longer contexts than in the Experiment 1. The semantic and syntactic conditions were blocked, and the order was counterbalanced across participants. The trials were presented in a different random order for each participant.

Table 2

Example of target words presented in a semantic or syntactic context, in a related, unrelated, or incongruent condition for one subject (length of 4 words).

	Condition		
Context	Related	Unrelated	Incongruent
Semantic	cat – dog – rabbit – mouse	table – green – flower – mouse	blue – red – yellow – mouse
Syntactic	she – likes – this – mouse	this – likes – she – mouse	The – dog – comes – mouse

Procedure. The procedure was identical to Experiment 1.

Analyses. Analyses were identical to Experiment 1. Out of all 3520 observations (RTs), 243 were removed because of a dysfunction of the microphone. Extreme values were removed (i.e., RTs below 200 ms or above 4000 ms) and outliers were replaced by the cutoff RT corresponding to each participant's ± 2.5 standard deviation (see Hoaglin et al., 1986). We report unstandardized regression coefficients (b), standard errors (*SEs*), and t-values. Fixed effects were considered significant if |t| was greater than 1.96 (Baayen, 2008). Fixed effects, random effects, and random slopes were only included if they significantly improved the model's fit in a forward stepwise model selection procedure.

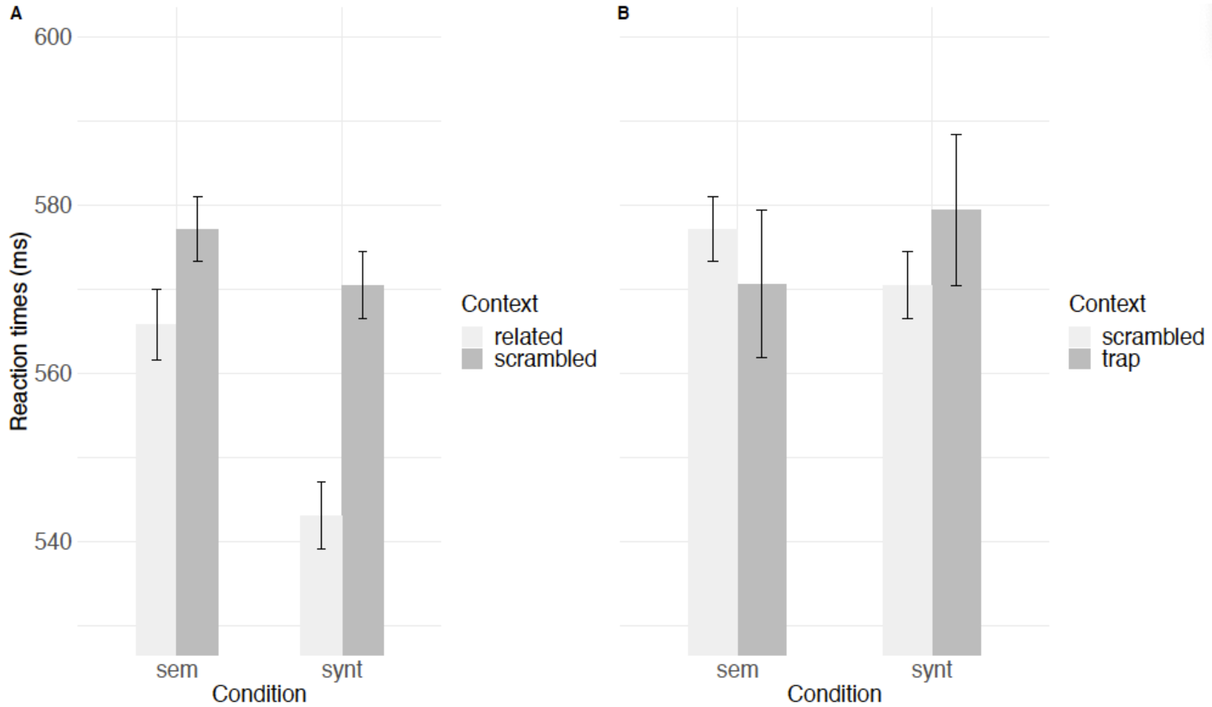
Results

Participants did not make any naming errors. Thus, accuracy was not further analyzed. Because the RT distribution was slightly skewed, we log-transformed the RTs. This resulted in an acceptable normal RT distribution (Skewness: -.24, Kurtosis: 1.05) (Blanca et al., 2013). We separated the analyses of the standard context by condition effect (as in Experiment 1) from the analysis of prediction errors. We then provide a combined analysis of Experiment 1 and 2.

Standard Analysis. This analysis was identical to that of Experiment 1. The final model included condition, context (related vs scrambled), length, and all interactions. As random effects, we included by-participant intercepts, random slopes for condition and random slopes for context by participants. The results showed a significant effect of context (b = 0.02, *SE* = 0.005, *t* = 3.79), no effect of condition, but a significant interaction between the effects of condition and context (b = 0.02, *SE* = 0.01, *t* = 2.01), which reflected the fact that the context effect was stronger for the syntactic than for the semantic condition. Contrast analyses confirmed that participants were faster at reading the target word following related than scrambled contexts in both the semantic (b = -0.01, *SE* = 0.005, *t* = -2.69) and syntactic condition (b = -0.02, *SE* = 0.007, *t* = -3.23) (see Figure 2). There was also an effect of context length (b = -0.03, *SE* = 0.009, *t* = -3.32), which reflected that participant were faster when the context was longer, but this effect did not interact with the context (related versus scrambled).

Figure 2

Mean reaction times (ms) in Experiment 2. Panel A presents the classic analysis with context (related vs scrambled) and condition (semantic vs syntactic) as factors. Panel B presents the prediction error analysis with context (scrambled vs incongruent) and condition (semantic vs syntactic) as factors. Errors bars indicate within-participant standard errors.



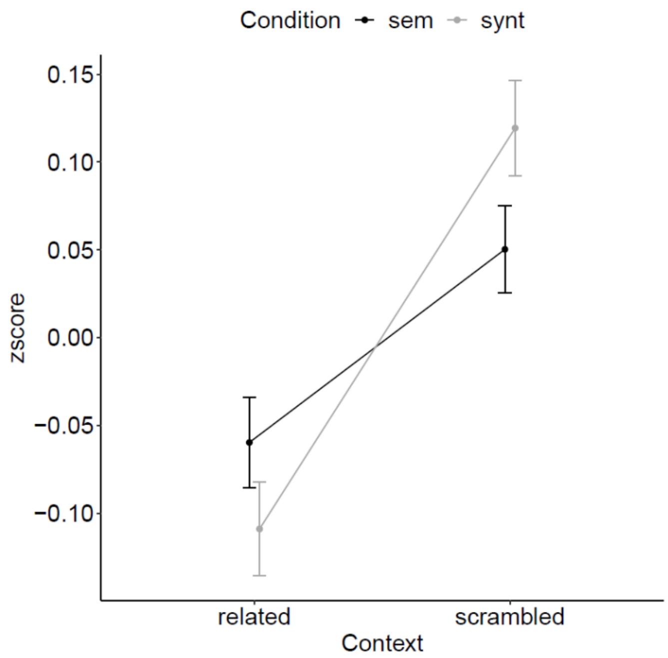
Analysis of Prediction errors. To analyze the effects of prediction errors, the context factor now opposed incongruent (prediction error) with scrambled contexts. The final model included condition, context (scrambled vs incongruent), length, and all interactions. As random effects, we had by-participant intercepts, random slopes for condition and random slopes for context by participants. The results showed no effect of condition or context and no interaction between the fixed effects. There was an effect of length (b = -0.03, *SE* = 0.01, *t* = -2.95) which reflected that participant were faster when the context was longer, but this effect did not interact with the context (scrambled or incongruent).

Experiment 1 and 2 combined. In the combined LME analyses, the final model included experiment (experiment 1 vs experiment 2), condition, context, length, and all interactions. As random effects, we included by-participant intercepts, random slopes for condition and random slopes for context by participants. Because the participants were faster in the second experiment (*M* = 565; *SD* = 110) than in the first (*M* = 641; *SD* = 158), we z-scored the raw data for each subject to eliminate absolute speed differences between Experiment 1 and 2 (see Faust et al., 1999).The results showed a significant effect of context (b = 0.17, *SE* = 0.04, *t* = 4.79), no effect of condition but a significant interaction between the effects of condition and context (b = 0.22, *SE* = 0.07, *t* = 2.92). This interaction indicates that the context effect was stronger for the syntactic condition (effect size = .22) than for the semantic condition effect size = .13). Contrast analyses confirmed that participants were faster to read the target word for related than for scrambled contexts in the semantic condition (b = -

0.13, *SE* = 0.03, *t* = -4.8) and in the syntactic condition (b = -0.22, *SE* = 0.03, *t* = -8.4). There was also an effect of length (b = -0.26, *SE* = 0.08, *t* = -3.29) which reflected that participant were faster when the context is longer, but this effect did not interact with the context (related versus scrambled). The results of the combined analysis are shown in Figure 3.

Figure 3

Combined analysis of Experiments 1 and 2 (z-score-transformed data) with context (related vs scrambled) and condition (semantic vs syntactic) as factors. Error bars show within-participant standard errors.



Discussion

As in Experiment 1, we replicated both semantic and syntactic prediction effects on reading aloud latencies using the exact same target words across all conditions in a within-subject design. Unlike Experiment 1, we observed a significantly stronger context effect for the syntactic than the semantic condition. The combined analysis confirmed this result suggesting that the failure to find a significant context by condition interaction in Experiment 1, despite the apparent trend for larger context effects in the syntactic condition, was due to a lack of power. In contrast to our expectations, however, trials that induced prediction errors did not show additional costs compared to scrambled contexts. As in Experiment 1, we found context length effects for both conditions but again this effect occurred whether the context was related or unrelated, which suggest that the effects was due to response preparation.

Experiment 3

In the previous experiments, the presentation time for the context words was quite long (1200 ms). Such a long presentation time was chosen to make sure there was enough time to build strong expectations about upcoming target words. Thus, the question arose as to whether the prediction effects would persist if context words were presented for much shorter times (i.e., 200 ms). Indeed, Perea and Gotor (1997) found that reliable semantic priming effects with rather short prime durations (SOA = 67 ms) and Snell and Grainger (2017) found strong syntactic effects on word identification (i.e., better partial report for words embedded in syntactically correct sequences) when 4-word sentences were presented for only 200 ms (see also Massol et al., 2021; Wen et al., 2021). We therefore replicated Experiment 2 while reducing the presentation times of the context words to 200 ms.

Method

Participants. We recruited 21 participants (18 women) at Aix-Marseille University (Marseille, France). All were native speakers of French and reported normal or corrected-to-normal vision, no learning disorders (dyslexia) and no neurological or psychiatric disorders. They were between 18 and 26 years old (*M* = 20,7; *SD* = 2,1) and signed informed consent forms prior to participation.

Design and Stimuli. Design and stimuli were identical to Experiment 2, each target word was seen in all contexts by all participants, but we have removed the incongruent condition. Participants saw 160 target words: 80 in the semantic condition (40 in a related context and 40 in an unrelated context) and 80 in the syntactic condition (same pattern).

Procedure. The procedure was identical to Experiment 1 and 2 except for the timing of the context words, that is, context words in both conditions were presented one by one in the center of the screen for 200 ms with a 100 ms ISI. The experiment lasted about 6 minutes.

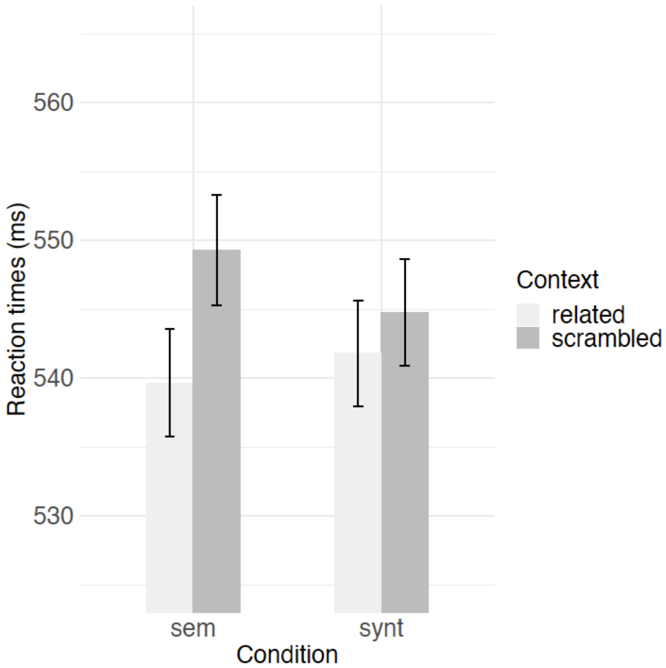
Analyses. Analyses were identical to Experiment 1 and 2. Out of the 3360 observations (RTs), 89 were removed because of a dysfunction of the microphone. Extreme values were removed (i.e., RTs below 200 ms or above 4000 ms) and outliers were replaced by the cutoff RT corresponding to each participant's ± 2.5 standard deviation (see Hoaglin et al., 1986). We report unstandardized regression coefficients (b), standard errors (*SEs*), and t-values. Fixed effects were considered significant if |t| was greater than 1.96 (Baayen, 2008). Fixed effects, random effects, and random slopes were only included if they significantly improved the model's fit in a forward stepwise model selection procedure.

Results

Participants did not make any naming errors except for one participant who uttered the word “isocèle” as “isolé” in all four conditions. Those trials were removed. Otherwise, accuracy was not further analyzed. Because the RT distribution was slightly skewed, we log-transformed the RTs. This resulted in an acceptable normal RT distribution (Skewness: .7, Kurtosis: 3.2) (Blanca et al., 2013). The final model included condition, context (related vs scrambled), length, and all interactions. As random effects, we included by-participant intercepts, random slopes for condition and random slopes for context by participants. The results showed a small effect but significant of context (b = 0.005, *SE* = 0.003, *t* = 1.92), but no effect of condition, nor an interaction between these two effects. Contrast analyses confirmed that priming effects were significant in the semantic condition (b = -0.008, *SE* = 0.004, *t* = -2.31) but not in the syntactic condition (b = -0.002, *SE* = 0.003, *t* = -0.47) (see Figure 4). There was also an effect of length (b = -0.03, *SE* = 0.007, *t* = -5.2) which reflected that participant were faster when the context is longer, but this effect did not interact with the context (related or scrambled).

Figure 4

Mean reaction times (ms) in Experiment 3 as a function of context (related vs scrambled) and condition (semantic vs syntactic). Errors bars show within-participant standard errors.



Discussion

The results showed that the syntactic context effect disappeared with shorter SOAs while the semantic context effect remained significant. The effect size of the latter effect was smaller in Experiment 3 than the two previous experiments, which suggests that strong semantic and syntactic predictions need time to build up. As in Experiment 1 and 2, we found context length effects for both conditions but again this effect occurred whether the context was related or unrelated.

General discussion

Semantic priming effects have been previously reported in lexical decision and naming (see Neely, 1991 for a review). However, syntactic priming effects have only been obtained in lexical decision but not naming (Carello et al., 1988; Seidenberg et al., 1984; Sereno, 1991) except for a study by Farrar (1998) who obtained syntactic priming effects for high frequency words in naming when presenting primes and targets together ("her cars" versus "she cars"). The present study showed that syntactic priming effects can be found in reading aloud but only when context words were presented for a long time (1200 ms). Under these presentation conditions, syntactic priming was even bigger than semantic priming. However, syntactic priming effects disappeared when the context words were presented more swiftly (200 ms), while semantic priming effects remained significant, although somewhat smaller in size.

This result suggests that syntactic constraints on reading aloud require time to build up. The finding is consistent with our initial assumption that the syntactic context does not allow to pre-activate a specific target word but only its grammatical class, whereas semantic contexts allow to pre-activate specific target words. It seems that short SOAs tap this automatic pre-activation phase of individual target words. Syntactic constraints cannot pre-activate individual target words but they can considerably reduce the search space. For example, syntactic constraints make certain continuations strictly impossible ("He wants to → table"), whereas semantic constraints make them only highly implausible ("He pets his → table"). However, our data suggest that this process (i.e., reduction of search space) needs time to affect reading aloud responses to individual words.

Our results are somewhat different from those of Snell and Grainger (2017) who found strong syntactic effects (i.e., faster partial report word identification accuracy for words that were embedded in syntactically correct sentences as opposed to scrambled agrammatical sequences) when 4-word sequences were presented for only 200 ms in a rapid parallel visual presentation (RPVP). However, there is a notable difference between their and our experiments. In Snell and Grainger, the 4-word sequences were presented all at once for only 200 ms, they were backward-masked, and participants had to make a partial report of one of the words. Thus, we are clearly in an "iconic memory" situation (Sperling, 1960), where the contents of brief alphanumeric displays are initially held in a high-capacity fast-decay visual-information store ("iconic memory"). In Sperling's (1960) famous partial report experiments, it has been shown that partial-report performance decreases until it finally reaches the level of whole report at cue delays of about 500-800 ms. According to Coltheart (1983), "the ability to report material from brief visual displays is seen as depending upon parallel (and perhaps unlimited) transfer from iconic memory to a post-categorical memory mode, followed by a limited (and perhaps serial) transfer to an output stage" (p. 283). It seems plausible that in Snell and Grainger's experiments syntactic information provides important constraints about which items are transferred from the rapidly decaying iconic memory into a longer lasting storage or output stage. This is very different from our experiment, in which context words were presented one by one for 200 ms followed by the target word. Clearly, the serial presentation mode of our experiment does not put any constraints on iconic memory. Despite these differences, both experiments show that syntactic constraints can affect word identification at various stages and time scales in conditions that minimized any potential influence of between-word semantic relatedness or predictability (i.e., cloze probability measures).

Coming back to our experiments, we would argue that semantically related contexts make it possible to make specific, discrete predictions of the upcoming target words. Given that we find semantic priming effects in a (non-semantic) naming task and at short SOAs, we believe that the pre-activation of a potential upcoming word facilitates letter and word identification rather later stages of lexical access. This is also in line with the observation that, in eye movements, predictability influences very early eye movement measures, including word skipping, which is the earliest possible measure (Staub, 2015). Syntactically predictive contexts, however, cannot pre-activate a specific word. They can only broadly activate words that are likely to appear in the input or exclude or inhibit categories of words that are impossible to occur, thus considerably reducing the lexical search space. This is in line with the observation that predictability effects are present even at the low end of the cloze probability range. According to Staub (2015), "it seems unlikely that readers would specifically predict a word that has relatively low cloze probability" (p. 322).

Note that these explanations would assume that the locus of semantic and syntactic context effects is different, and it would be interesting to investigate to what extent they can be separated in such a tightly controlled paradigm. Indeed, previous research suggested that semantic and syntactic priming mechanisms can be dissociated. Kutas and Hillyard (1980) initially observed a N400 in response to a semantically incongruent word placed at the end of a sentence and in syntactically unrelated word pairs (Bentin et al., 1985). In contrast, syntactic violations seem to generate P600 components (Friederici, 2003; Osterhout & Holcomb, 1992). Finally, Dapretto and Bookheimer (1999) indicated that part of Broca's area and the inferior part of the left inferior frontal gyrus are involved in the processing of syntactic information and semantic aspects of a sentence, respectively. However, the claim that the locus of semantics and syntax can be dissociated at the neural level is still a matter of considerable debate (Siegelman et al., 2019). The design of our study is ideal to examine a dissociation at the electrophysiological or neural level because the last word is identical in all conditions, which allows one to investigate the locus and time course of syntactic versus semantic prediction in a more controlled fashion.

Context length had a strong effect on latencies (the longer the context, the faster the response latencies) but context length did not interact with context (related versus unrelated). That is, we did not find that context length increased the size of the semantic or syntactic priming effects. This result suggests that context length only affected response preparation. We initially predicted that longer contexts should make semantic and syntactic predictions stronger thus producing bigger priming effects. It therefore seems that a three-word context is sufficient to generate strong semantic and/or syntactic expectations. Note that in terms of conditional probabilities (syntactic prediction) and cosine similarities (semantic prediction), shorter contexts were as predictive as longer contexts (see Methods).

Altogether our results are consistent with the concept of a "predictive brain" (Clark, 2013). An aim for future research is to test these prediction effects at various time scales in a larger sample while adding measures of reading skill and statistical learning to further investigate whether semantic and syntactic prediction skills predict reading skills and whether genuinely predictive or anticipatory processes (e.g., statistical learning) predict linguistic prediction skills (Arciuli, 2018; Hung et al., 2018; Isbilen et al., 2022).

Acknowledgments

The authors declare that there is no conflict of interest. Both authors approved the final version of the manuscript. This research has been supported by the Institute of Convergence ILCB (ANR-16-CONV-0002) and the Excellence Initiative of Aix-Marseille University A*MIDEX (ANR-11-IDEX-0001-02). The work that led to this publication has also been supported by AMPIRIC, a center of excellence on research in education and teacher training.

References

Arciuli, J. (2018). Reading as Statistical Learning. *Language, Speech, and Hearing Services in Schools, 49*(3S), 634-643. https://doi.org/10.1044/2018_LSHSS-STLT1-17-0135

Baayen, R. H. (2008). *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge University Press.

Bar, M. (2009). Predictions : A universal principle in the operation of the human brain. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1521), 1181-1182. https://doi.org/10.1098/rstb.2008.0321

Bar, M. (2011). *Predictions in the Brain: Using our past to generate a future*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195395518.001.0001

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software, 67*(1). https://doi.org/10.18637/jss.v067.i01

Becker, C. A. (1980). Semantic context effects in visual word recognition : An analysis of semantic strategies. *Memory & Cognition*, *8*(6), 493-512. https://doi.org/10.3758/BF03213769

Bentin, S., McCARTHY, G., & Wood, C. C. (1985). Event-related potentials, lexical decision and semantic priming. *Electroencephalography and Clinical Neurophysiology*, *60*(4), 343-355. https://doi.org/10.1016/0013-4694(85)90008-2

Blanca, M. J., Arnau, J., López-Montiel, D., Bono, R., & Bendayan, R. (2013). Skewness and Kurtosis in Real Data Samples. *Methodology, 9*(2), 78-84. https://doi.org/10.1027/1614-2241/a000057

Carello, C., Lukatela, G., & Turvey, M. T. (1988). Rapid naming is affected by association but not by syntax. *Memory & Cognition, 16*(3), 187-195. https://doi.org/10.3758/BF03197751

Chang, F., Dell, G. S., & Bock, K. (2006). Becoming syntactic. *Psychological Review, 113*(2), 234-272. https://doi.org/10.1037/0033-295X.113.2.234

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences, 36*(3), 181-204. https://doi.org/10.1017/S0140525X12000477

Collins, A. M., & Loftus, E. F. (1975). A Spreading-Activation Theory of Semantic Processing. *Psychological Review, 82*(6), 407. https://doi.org/10.1037/0033-295X.82.6.407

Coltheart, M. (1983). Iconic memory. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences, 302*(1110), 283-294. https://doi.org/10.1098/rstb.1983.0055

Dapretto, M., & Bookheimer, S. Y. (1999). Form and Content : Dissociating Syntax and Semantics in Sentence Comprehension. *Neuron*, *24*(2), 427-432. https://doi.org/10.1016/S0896-6273(00)80855-7

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. https://doi.org/10.1038/nn1504

de Marneffe, M.-C., Manning, C. D., Nivre, J., & Zeman, D. (2021). Universal Dependencies. *Computational Linguistics*, 1-54. https://doi.org/10.1162/coli_a_00402

den Heyer, K., Briand, K., & Smith, L. (1985). Automatic and strategic effects in semantic priming : An examination of Becker's verification model. *Memory & Cognition, 13*(3), 228-232. https://doi.org/10.3758/BF03197685

den Ouden, H. E. M., Kok, P., & de Lange, F. P. (2012). How Prediction Errors Shape Perception, Attention, and Motivation. *Frontiers in Psychology, 3,* 548. https://doi.org/10.3389/fpsyg.2012.00548

Draine, S. C., & Greenwald, A. G. (1998). Replicable Unconscious Semantic Priming. *Journal of Experimental Psychology: General*, *127*(3), 286. https://doi.org/10.1037/0096-3445.127.3.286

Drake, E., & Corley, M. (2015). Effects in production of word pre-activation during listening : Are listener-generated predictions specified at a speech-sound level? *Memory & Cognition*, *43*(1), 111-120. https://doi.org/10.3758/s13421-014-0451-9

Ehrlich, S. F., & Rayner, K. (1981). Contextual effects on word perception and eye movements during reading. *Journal of Verbal Learning and Verbal Behavior*, *20*(6), 641-655. https://doi.org/10.1016/S0022-5371(81)90220-6

Engbert, R., Nuthmann, A., Richter, E. M., & Kliegl, R. (2005). SWIFT: a dynamical model of saccade generation during reading. *Psychological review*, *112*(4), 777. https://doi.org/10.1037/0033-295X.112.4.777

Farrar, W. T. F. (1998). Investigating Single-Word Syntactic Primes in Naming Tasks : A Recurrent NetworkApproach. *Journal of Experimental Psychology: Human Perception and Performance*, *24*(2), 648. https://doi.org/10.1037/0096-1523.24.2.648

Faust, M. E., Balota, D. A., & Spieler, D. H. (1999). Individual Differences in Information-Processing Rate and Amount : Implications for Group Differences in Response Latency. *Psychological Bulletin*, *125*(6), 777-799. https://doi.org/10.1037/0033-2909.125.6.777

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. https://doi.org/10.1006/jmla.1999.2660

Ferreira, F., & Qiu, Z. (2021). Predicting syntactic structure. *Brain Research*, *1770*, 147632. https://doi.org/10.1016/j.brainres.2021.147632

Fischler, I. (1977). Semantic facilitation without association in a lexical decision task. *Memory & Cognition*, *5*(3), 335-339. https://doi.org/10.3758/BF03197580

Fischler, I., & Goodman, G. O. (1978). Latency of Associative Activation in Memory. *Journal of Experimental Psychology: Human Perception and Performance*, *4*(3), 455. https://doi.org/10.1037/0096-1523.4.3.455

Forster, K. I. (1998). The Pros and Cons of Masked Priming. *Journal of Psycholinguistic Research*, *27*(2), 203-233. https://doi.org/10.1023/A:1023202116609

Friederici, A. D. (2003). The Role of Left Inferior Frontal and Superior Temporal Cortex in Sentence Comprehension : Localizing Syntactic and Semantic Processes. *Cerebral Cortex*, *13*(2), 170-177. https://doi.org/10.1093/cercor/13.2.170

Friston, K. (2010). The free-energy principle : A unified brain theory? *Nature Reviews Neuroscience*, *11*(2), 127-138. https://doi.org/10.1038/nrn2787

Friston, K., & Kiebel, S. (2009). Predictive coding under the free-energy principle. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1521), 1211-1221. https://doi.org/10.1098/rstb.2008.0300

Goodman, G. O., McClelland, J. L., & Gibbs., R. W. (1981). The role of syntactic context in word recognition. *Memory & Cognition*, *9*(6), 580-586. https://doi.org/10.3758/BF03202352

Grossberg, S. T. (2012). *Studies of mind and brain: Neural principles of learning, perception, development, cognition, and motor control* (Vol. 70). Springer Science & Business Media.

Günther, F., Rinaldi, L., & Marelli, M. (2019). Vector-Space Models of Semantic Representation From a Cognitive Perspective : A Discussion of Common Misconceptions. *Perspectives on Psychological Science*, *14*(6), 1006-1033. https://doi.org/10.1177/1745691619861372.

Hale, J. 2001. A probabilistic Earley parser as a psycholinguistic model. *Proceedings of NAACL 2*. 159–166.

Hoaglin, D. C., Iglewicz, B., & Tukey, J. W. (1986). Performance of Some Resistant Rules for Outlier Labeling. *Journal of the American Statistical Association*, *81*(396), 991-999. https://doi.org/10.1080/01621459.1986.10478363

Hohenstein, S., Laubrock, J., & Kliegl, R. (2010). Semantic preview benefit in eye movements during reading : A parafoveal fast-priming study. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*(5), 1150-1170. https://doi.org/10.1037/a0020233

Hohwy, J. (2013). *The predictive mind*. Oxford University Press.

Hung, Y.-H., Frost, S. J., & Pugh, K. R. (2018). Domain Generality and Specificity of Statistical Learning and its Relation with Reading Ability. In T. Lachmann & T. Weis (Éds.), *Reading and Dyslexia* (Vol. 16, p. 33-55). Springer International Publishing. https://doi.org/10.1007/978-3-319-90805-2_2

Isbilen, E. S., McCauley, S. M., & Christiansen, M. H. (2022). Individual differences in artificial and natural language statistical learning. *Cognition*, *225*, 105123. https://doi.org/10.1016/j.cognition.2022.105123

Johnson, M. A., Turk-Browne, N. B., & Goldberg, A. E. (2013). Prediction plays a key role in language development as well as processing. *Behavioral and Brain Sciences*, *36*(4), 360-361. https://doi.org/10.1017/S0140525X12002609

Kolb, P. (2008). DISCO: A Multilingual Database of Distributionally Similar Words. *Proceedings of KONVENS-2008, 156*, 8.

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*(1), 621-647. https://doi.org/10.1146/annurev.psych.093008.131123

Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. Science (American Association for the Advancement of Science), 207(4427), 203-205. https://doi.org/10.1126/science.7350657

Kutas, M., & Hillyard, S. A. (1984). Brain potentials during reading reflect word expectancy and semantic association. *Nature*, *307*(5947), 161-163. https://doi.org/10.1038/307161a0

Levy, R. (2008). Expectation-based syntactic comprehension. *Cognition*, *106*(3), 1126-1177. https://doi.org/10.1016/j.cognition.2007.05.006

Lupker, S. J. (1984). Semantic priming without association : A second look. *Journal of Verbal Learning and Verbal Behavior*, *23*(6), 709-733. https://doi.org/10.1016/S0022-5371(84)90434-1

Lupyan, G., & Clark, A. (2015). Words and the World : Predictive Coding and the Language-Perception-Cognition Interface. *Current Directions in Psychological Science*, *24*(4), 279-284. https://doi.org/10.1177/0963721415570732

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. https://doi.org/10.1037/a0029284

Massol, S., Mirault, J. & Grainger, J. (2021). The contribution of semantics to the sentence superiority effect. *Scientific Reports*, *11*, 1-6. 20148. https://doi.org/10.1038/s41598021-99565-6

Meyer, D. E., & Schvaneveldt, R. W. (1971). Facilitation in recognizing pairs of words : Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology*, *90*(2), 227-234. https://doi.org/10.1037/h0031564

Mikolov, T., Le, Q. V., & Sutskever, I. (2013). *Exploiting Similarities among Languages for Machine Translation* (arXiv:1309.4168). arXiv. http://arxiv.org/abs/1309.4168

Neely, J. H. (1991). Semantic priming effects in visual word recognition: A selective review of current findings and theories. Dans Besner, D., & Humphreys, G.W. (Eds.), *Basic processes in reading: Visual ord REcognition (1st ed.)* (p. 272-344). Routledge.

Osterhout, L., & Holcomb, P. J. (1992). Event-related brain potentials elicited by syntactic anomaly. *Journal of Memory and Language*, *31*(6), 785-806. https://doi.org/10.1016/0749-596X(92)90039-Z

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. https://doi.org/10.1016/S0010-0277(96)00782-2

Pickering, M. J., & Gambi, C. (2018). Predicting while comprehending language : A theory and review. *Psychological Bulletin*, *144*(10), 1002-1044. https://doi.org/10.1037/bul0000158

Pickering, M. J., & Garrod, S. (2013). An integrated theory of language production and comprehension. *Behavioral and Brain Sciences*, *36*(4), 329-347. https://doi.org/10.1017/S0140525X12001495

Rao, R. P. N., & Ballard, D. H. (1999). Predictive coding in the visual cortex : A functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience*, *2*(1), 79-87. https://doi.org/10.1038/4580

Reichle, E. D., Pollatsek, A., Fisher, D. L., & Rayner, K. (1998). Toward a model of eye movement control in reading. *Psychological review*, *105*(1), 125. https://doi.org/10.1037/0033-295X.105.1.125

Seidenberg, M. S., Waters, G. S., Sanders, M., & Langer, P. (1984). Pre- and postlexical loci of contextual effects on word recognition. *Memory & Cognition*, *12*(4), 315-328. https://doi.org/10.3758/BF03198291

Sereno, J. A. (1991). Graphemic, Associative, and Syntactic Priming Effects at a Brief Stimulus Onset Asynchrony in Lexical Decision and Naming. *Journal of experimental psychology. Learning, memory, and cognition, 17*(3), 459-477. https://doi.org/10.1037/0278-7393.17.3.459

Siegelman, M., Blank, I. A., Mineroff, Z., & Fedorenko, E. (2019). An Attempt to Conceptually Replicate the Dissociation between Syntax and Semantics during Sentence Comprehension. *Neuroscience, 413*, 219-229. https://doi.org/10.1016/j.neuroscience.2019.06.003

Snell, J., & Grainger, J. (2017). The sentence superiority effect revisited. *Cognition, 168*, 217-221. https://doi.org/10.1016/j.cognition.2017.07.003

Sperling, G. (1960). The information available in brief visual presentations. *Psychological monographs: General and applied, 74*(11), 1. https://doi.org/10.1037/h0093759

Staub, A. (2015). The effect of lexical predictability on eye movements in reading: Critical review and theoretical interpretation. *Language and Linguistics Compass, 9*(8), 311 327. https://doi.org/10.1111/lnc3.12151.

Taylor,W. L. (1953). “Cloze procedure”: a new tool for measuring readability. *Journalism Quarterly, 30*(4), 415–433. https://doi.org/10.1177/107769905303000401

Wen, Y., Mirault, J. & Grainger, J. (2021). Fast syntax in the brain: Eletrophysiological evidence from the rapid parallel visual presentation paradigm (RPVP). *Journal of Experimental Psychology: Learning, Memory, and Cognition, 47*(1), 99-112. https://doi.org/10.1037/xlm0000811