Predictive processing suppresses form-related words with overlapping onsets

Katja I. Haeuser (khaeuser@coli.uni-saarland.de)
Department of Psychology, Saarland University,
66123 Saarbrucken, Germany

Arielle Borovsky (aborovsky@purdue.edu)
College of Health and Human Sciences, Purdue University,
Lafayette, Ind., USA

Abstract

Do language users predict word forms as readily as they predict semantic features? Previous studies are conflicting, possibly because they did not differentiate between two types of word form relationship: Head and rhyme relationships, sharing onset or offset features with predictable words. Here, we investigated prediction of form and meaning by means of a priming lexical decision task. People read constraining sentences that disconfirmed their expectations, and indicated, at sentence offset, whether a letter string was a word. Targets were predictable but not presented nouns, semantically related nouns, as well as head- and rhyme-related nouns. Unrelated control nouns were also presented. Results showed facilitation for predictable and semantically related words, with no difference between the two. While no effects emerged for rhymes, head-related words showed slowing, indicating suppression of lexical neighbors following prediction of word forms. Our findings align with word recognition models and prediction-by-production models of predictive processing.

Keywords: language processing, prediction, reading, sentence comprehension

Introduction

Language users leverage linguistic context to generate predictions about upcoming words (Huettig, 2015; Kuperberg & Jaeger, 2016; Ryskin & Nieuwland, 2023). But what are the contents of the predictions that language users generate? Do people predict form-related features of words (e.g. a word's orthography) as readily as they predict meaning-related features? We explore these questions by using a novel sentence-reading lexical decision task, designed to measure prediction of form and meaning in a sample of native speakers of German. Our results show that prediction of form does occur – albeit in a very different manner than was previously reported. We begin by reviewing the previous literature.

Prediction of meaning

It is well attested that comprehenders predict semantic features of upcoming words. For example, in a visual-world paradigm that contains multiple semantically matching referents (Kamide et al., 2003), listeners use sentence context predictively (e.g., "The girl will now ride the ..." vs "The man will now ride the ...") to quickly fixate on the most likely referent (e.g., a carousel or a motorbike). Crucially, the prediction of semantic features of words seems to have "spill-over" effects on words that are unpredictable but share semantic features. Using ERPs, Federmeier and Kutas (1999) showed that when people are reading constraining mini-stories that render a particular word highly predictable (e.g., "palms"), facilitated processing also occurs for unpredictable but semantically related targets (e.g., "pines"). This finding suggests that comprehenders can probabilistically pre-activate a range of related but unpredictable meanings based on semantic features relevant to a sentence context.

Prediction of form

Prediction of word form is debated in the literature. Even though most researchers agree that comprehenders can, in principle, predict orthographic or morpho-syntactic features of upcoming words (e.g., DeLong et al., 2005; Fleur et al., 2020; Haeuser et al., 2022; Van Berkum et al., 2005, Wicha et al., 2003), it is less clear how readily such form prediction occurs and under which conditions. One line of research has suggested that language users predict word forms less quickly and readily than semantic features. For example, using ERPs, Ito and colleagues (2016) investigated N400 modulation for predictable words, semantically related words and form-related words that were presented in moderately and highly constraining sentences. They additionally compared a long and a short word presentation rate (500ms vs. 700ms SOA). The results indicated that comprehenders pre-activated form-related features later than semantic ones, whereas prediction of form occurred only in highly constraining sentences. Similarly, other studies have documented individual differences in word form prediction. For example, older adults, children, and foreign language learnersare less likely to predict word forms (e.g., DeLong et al., 2012; Gambi et al., 2018; Martin et al., 2013), or may show form prediction effects that are qualitatively different from younger adults (e.g., Gambi et al., 2021; Haeuser et al., 2022; Lew-Williams et al., 2010). These results suggest that form prediction may quite variable, and may not occur in all learning and communicated contexts.

In contrast, another line of research has shown time-equivalent prediction of form and meaning (DeLong et al., 2019, 2021; also see Laszlo & Federmeier, 2009; Kim & Lai, 2012), even in conditions when language was presented

In L. K. Samuelson, S. L. Frank, M. Toneva, A. Mackey, & E. Hazeltine (Eds.), *Proceedings of the 46th Annual Conference of the Cognitive Science Society.*

very rapidly (e.g., DeLong et al., 2021). For example, in an ERP study that conceptually replicated the one by Ito and colleagues (2016), DeLong and colleagues (2019) demonstrated time-equivalent prediction of form and meaning. In another study (DeLong et al., 2021), the same authors even found very rapid form prediction effects when using a very short word presentation rate (250ms SOA), intended to mimic real-time language comprehension.

In part, it may be possible to attribute these conflicting findings to methodological differences – among them, for example, differences in the selection of the N400 time window, or differences with respect to sentence length and critical word position (for discussion, see Delong et al., 2019). However, another important (and previously not considered) aspect could be that many studies did not sufficiently differentiate between two well-known kinds of word-form relationship: conditions when target and orthographic neighbor share onset similarity (e.g., "cat"-"car", HEAD relationship) and conditions when they share offset similarity (e.g., "cat"-"mat", RHYME relationship).

Crucially, previous research has suggested that word onsets are more critical to language processing than offsets. For example, head relationships affect processing during the earliest stages and throughout a word's acoustic presentation, whereas rhyme relationships affect processing later and less consistently (e.g., Allopenna et al. 1998; Magnuson et al., 2007). In line with this, prediction-by-production models (which assume that comprehenders engage their production system to generate predictions) postulate earlier effects for word onsets than offsets: Since orthographic assembly of onset features likely precedes the one for offset features, the rhyme of the predictable word would become available later during processing than the head.

It is important to point out here that most previous studies on predictive processing have assumed that form prediction will *facilitate* processing of form-related neighbors – for example, expecting "ice" will facilitate processing "dice". However, orthographic priming studies have shown *inhibitory*, not facilitatory effects for form-related words, in particular for head-related neighbors (e.g., Frisson et al., 2014). For example, when a head-related word is used as a prime (e.g., "cat"), that word will become a strong competitor during the recognition of the target word (e.g., "car"), which results in slowed target word recognition (e.g., Grainger & Ferrand, 1994). Similar results have emerged for spoken word recognition where words with many form neighbors are recognized more slowly, not quickly (e.g., Magnuson et al., 2007).

The present study

In this study, we investigated the prediction of form and meaning by means of a combined sentence-reading lexical-decision task. Participants read strongly constraining German sentences that ended in an unpredictable but plausible word (self-paced word-by-word reading[1]). The goal of presenting an unpredictable continuation was to demonstrate that the results of the subsequent priming task could only be driven by predicting the specific word form – which was never specifically presented. At sentence offset, participants decided as quickly as possible whether a visually presented letter string was an existing German word or not (i.e., lexical decision task; LDT). Targets were predictable words, semantically-related words, as well as head- and rhyme-related words. Unrelated control words were also presented.

For predictable and semantically-related words, we expected to find facilitated processing compared to unrelated controls, with a larger facilitation effect for predictable than semantically-related targets, in line with previous research reporting evidence for graded prediction of meaning. For form-related targets, in particular for head-related words, we expected to find slowed processing (compared to unrelated controls), in line with prediction-by-production accounts and evidence from priming and spoken word recognition.

Method

Participants

Ninety-eight Prolific workers (62m, 35f, 1nb) between the ages of 18 and 40 (*Mean* = 28 years, *SD* = 6) completed the experiment. All participants were right-handed native speakers of German, had normal or corrected-to-normal vision, and reported no neuropsychological disorders.

Materials

Each experimental trial consisted of a sentence, which was followed by a LDT target noun. Experimental sentences were forty German sentence frames (e.g., *Gerald trinkt seinen Tee gerne aus der* ... [Gerald likes drinking his tea from the …]) which constrained expectations towards a particular noun (e.g., *Tasse* [cup]) but instead were completed with an unpredictable noun (e.g., *Schüssel* [bowl]) and three following words (e.g., *vom Markt nebenan* [from the nearby market]). Cloze probability ratings, obtained from forty native speakers of German who did not participate in the main experiment, confirmed that the sentence frames were highly constraining: On average, predictable nouns had a cloze probability of 89% (range: 63% - 100%), whereas unpredictable nouns had near-zero cloze probabilities. Unpredictable nouns were selected to be as plausible given the sentence context as possible, but no formal plausibility rating was conducted. To assess potential semantic overlap between the unpredictable and predictable nouns, we computed cosine similarities between the two. Cosine similarities were derived from a small German language model, trained with word2vec on the German Wikipedia and German news articles (see

[1] We chose self-paced reading over fixed word-by-word presentation to create more natural processing conditions.

https://devmount.github.io/GermanWordEmbeddings/). The resulting cosine similarities varied between .16 (*heater-clock*) and .89 (*tomato-onion*), M = .49. Finally, the experiment also contained fifty predictable filler sentences (taken from the Potsdam sentence corpus; Kliegl et al., 2006), which were inserted to make sure that participants continued to generate predictions during the experiment.

LDT targets consisted of 200 German nouns, which were presented over five conditions (forty items per condition): predictable words (i.e., the predictable nouns from the experimental sentences, e.g., *Tasse* [cup], PRED), nouns that were semantically related to the predictable word (e.g., *Kaffee* [coffee], SEM[2]; verified by means of German association norms; Melinger & Weber, 2006), form-related targets that shared the first three letters (i.e. the head) with the predictable word (e.g., *Tasche* [bag], HEAD), and form related-targets that shared a rhyme relationship with the predictable word (e.g., *Kasse* [check-out], RHYME). The fifth condition consisted of unrelated control nouns that did not have a semantic or form relationship with the predictable word (e.g., *Lehrer* [teacher], UR). There were also 50 non-words (created through the pseudoword generation program *Wuggy*; Keuleers & Brysbaert, 2010), which matched the experimental words with respect to their number of letters and syllables.

HEAD and RHYME targets were matched with UR with respect to word length and frequency (p's > .20). PRED and SEM targets were matched to UR with respect to word length (p's > .20), but differed from UR with respect to frequency (both p's < .01). Our statistical analysis controlled for these differences. Table 1 reports lexical characteristics of lexical decision targets.

Table 1: Lexical characteristics of words used in the LDT.

	PRED	SEM	HEAD	RHYME	UR
Frequency	2.65	2.93	2.19	2.03	2.25
Length	5.40	5.32	5.58	5.38	5.6

Note. Frequency values reflect log per-million values from the German movie subtitle corpus (Brysbaert et al., 2011). Length values reflect number of characters.

Procedure

One experimental trial consisted of a sentence, followed by a lexical-decision target noun, presented in one of the five experimental conditions. Non-words always followed after filler sentences. Participants read the sentences in a self-paced manner using a word-by-word reading paradigm. Pushing the space bar with their right hand revealed the next word in the sentence. Immediately at sentence offset, a fixation cross appeared for 700ms. The target noun then appeared in the center of the screen until the participant made a response about its lexicality by pressing the S-key with their left index-finger for non-words, and the K-key with their right index finger for words. In 40% of all trials, a yes-no comprehension question appeared after the lexical decision target, inserted to make sure that participants read the sentences for comprehension. Participants were instructed to read the sentences as quickly as possible, and respond to the lexical decision targets as quickly and accurately as possible. In total, there were five experimental lists with 90 trials each; each participant only saw one version of a single item.

Results

Performance on the LDT and the comprehension questions indicated that participants were attentive during the experiment. In the LDT, the average hit rate for words was high (97%, range = 88% - 100%), and the false alarm rate to nonwords was low (3%, range: 0% - 30%). Comprehension questions accuracy was near ceiling (Mean = 90%, range = 85% - 100%).

Prior to statistical analysis, we screened subjects for potential speed-accuracy trade-offs in the LDT by computing inverse efficiency (IE) scores (Townsend & Ashby, 1978). IE scores were computed by dividing each subject's reaction times by their accuracy rate. Using the interquartile range method, we identified one subject whose IE score exceeded the upper bound (Q3+1.5*IQR). That subject was excluded from further analysis.

Reaction time outliers among the 97 remaining subjects were identified based on visual inspection of the data. Accordingly, data points faster than 250ms and slower than 4000ms were excluded, a procedure that affected less than 1% of all data points.

Table 2 shows by-condition average accuracy rates and correct reaction times in the remaining subjects. Figure 1 shows a plot of the LDT reaction time data.

Table 2: Means (and standard deviations) of accuracy and correct reaction times in the LDT.

	PRED	SEM	HEAD	RHYME	UR
Accuracy	.99	.99	.93	.95	.95
	(.06)	(.08)	(.25)	(.23)	(.21)
Correct RT	814	797	925	884	877
	(308)	(263)	(358)	(305)	(343)

To statistically analyze the reaction time data, we ran linear mixed effects models (LMER) as implemented in the lm4 library (Bates et al., 2013; version 1.1-31) in R (R Development Core Team, 2016; version 4.1.3).

The outcome variable were trial-by-trial correct reaction times, log-transformed to avoid skewness. The critical predictor was condition (five levels: PRED, SEM, HEAD, RHYME, UR), dummy-coded, with the unrelated control condition (UR) set as the reference/baseline category. Control predictors consisted of trial number (to capture effects of customization with the experiment), word frequency (log-per million values from the SUBTLEX-DE data base; Brysbaert et al., 2011), and word length. All

[2] Though semantically related, SEM targets were unpredictable given the sentence context, with a mean cloze probability of .06 (Range: 0 - .07).

predictor variables were scaled (centered and standardized around their means) to ease comparison of beta-coefficients.

Initially the model was fit using by-subject and by-item random intercepts, as well as by-subject and by-item random slopes for condition (i.e., a fully maximal model; see Barr et al., 2013). Due to convergence issues, we reduced the model using the least-variance approach. The final model converged with by-subject and by-item random intercepts (see Table 3, for model output). The formal specification of the LMER model used was log(RT) ~ condition + scale(trial) + scale(frequency) + scale(length) + (1|subject) + (1|item).

As expected, reaction times for PRED and SEM targets were facilitated, compared to the unrelated control condition (UR, both p's < .01; see Figure 1), though facilitation was similar for PRED and SEM (due to similar beta-coefficients "b's"). Reaction times to RHYME targets did not statistically differ from those to UR (p = .58). Notably, reaction times to HEAD targets resulted in slowing when compared to the unrelated control condition UR (p = .01).

Table 3: Output of LMER model estimating the effects of condition and control predictors on log-transformed correct LDT times.

	b	SE	*t*	*p*
(Intercept)	***6.71***	***0.02***	***333.33***	***< .001***
PRED	**-0.04**	**0.01**	**-2.67**	**< .01**
SEM	**-0.04**	**0.01**	**-2.57**	**< .01**
HEAD	**0.03**	**0.01**	**2.47**	**.01**
RHYME	0.01	0.01	0.55	.58
Scaled Trial	**-0.03**	**0.01**	**-6.70**	**< .001**
Scaled Length	0.005	0.01	1.07	.28
Scaled Frequency	**-0.05**	**0.01**	**-9.16**	**< .001**

Note. Significant results are shown in bold. The intercept reflects fitted log-RT for UR. For PRED, SEM, HEAD and RHYME, negative beta coefficients show facilitation compared to UR; positive coefficients reflect slowing.

To examine possibility that the presented unpredictable nouns primed predictable nouns by means of low-level semantic association (e.g. *bowl* priming *cup*), we ran a follow-up model. This new model specified the same variables as the main RT model, but also included the scaled cosine similarity between unpredictable and predictable nouns as a fixed effect (both as a main effect and in interaction with condition). We conjectured that if the presented unpredictable nouns associatively primed the predictable ones irrespective of any prior prediction, then our condition effects should interact with cosine similarities, such that condition effects should be larger for items in which there was greater semantic overlap between unpredictable and predictable nouns. The model identified no interactions with cosine similarity for any of the specified contrasts (all p's > .10); the main effect of cosine similarity was also not significant (p > .50).

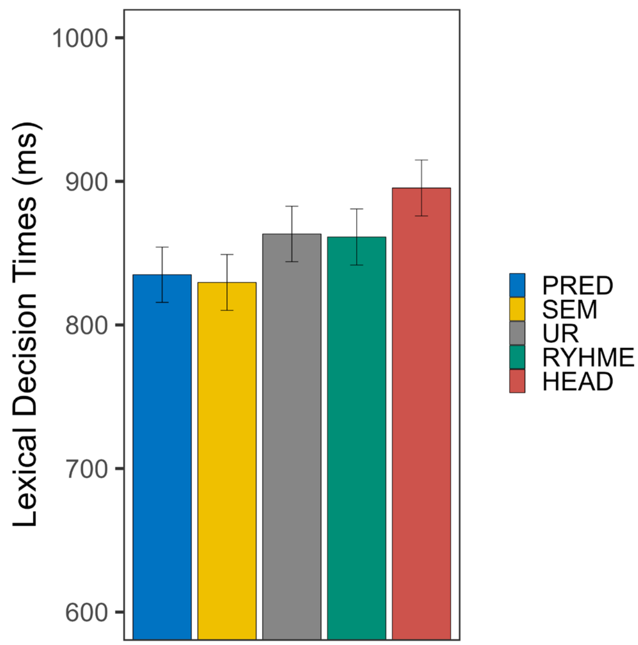

Figure 1: Bar graph (95% SE) showing correct lexical decision times across five experimental conditions. PRED=Predictable noun; SEM=Semantically related to predictable noun; UR=Unrelated control; RHYME=Offset-related noun; HEAD=Onset-related noun.

Discussion

Previous research has shown that language users leverage the linguistic context to predict meaning-related aspects of upcoming words, but it is much less clear whether comprehenders as readily predict form-related aspects of words, for example a word's orthography.

We hypothesized that the conflicting findings pertaining to form prediction may result from conflating two types of form relationship that are known to have diverging effects on language processing. Therefore, we examined whether two prominent types of form-related words, head- and rhyme-related words, might yield different effects on prediction.

As expected, we found that participants predicted meaning and form during sentence reading, and we showed that these effects were unlikely to be driven by associative semantic priming in which the presented unpredictable nouns may have implicitly primed the predictable ones, irrespective of any prediction. Importantly, and uniquely, we find that the lexical activation mechanisms resulting from prediction manifested differently for meaning and form conditions.

Predicting Meaning

In line with previous studies, our results show that readers predict semantic features of upcoming words. Specifically, lexical decision times for PRED and SEM targets were facilitated, compared to unrelated controls. This pattern aligns with previous research demonstrating that pre-activation of the semantic features of predictable words may also activate unpredictable but semantically related words (Kim & Lai, 2012; Federmeier & Kutas, 1999), for example trough spreading activation.

However, with respect to the magnitude of the facilitation gained by these two conditions, our results differ from previous investigations. Specifically, the facilitation gained by PRED and SEM targets was comparable. This finding differs from ERP studies which have observed a graded pattern, where PRED targets' processing was more facilitated than SEM targets'**.**

It is likely that this discrepancy stems from methodological differences between ERP and LDT. While the N400 ERP component is known to index very fine-grained semantic differences between words, such differences may not be easily picked up in lexical decision latencies.

It is worthwhile pointing out that we found facilitation for PRED and SEM targets even though the previous sentence had ended in an unpredictable word that disconfirmed participants' expectations. This aligns with previous studies showing that predictable words remain accessible in memory (e.g., Haeuser & Kray, 2022; Hubbard et al., 2019). Hence, having predictions disconfirmed does not necessarily result in their immediate suppression (e.g., Ness & Meltzer-Asscher, 2018).

Predicting Form

We found a critical difference between the processing of rhyme-related and head-related words. While rhyme-related words were recognized no differently than unrelated controls, head-related words were recognized more *slowly* than unrelated controls. This finding indicates that participants experienced interference, and slowing of recognition, for head-related words from the predicted (but not presented) target word in the sentence.

It is important to point out that the slowed processing we find for heads is not in line with the usually facilitatory effects associated with form prediction that previous studies on predictive processing have often reported. For example, several previous ERP studies have shown facilitated processing for unpredictable words and pseudo-words that share global orthographic features with predictable nouns (e.g., "The student is going to the library to borrow a hook", "She measured the flour so she could bake a ceke"; DeLong et al., 2019; Ito et al., 2016; Kim & Lai, 2012). Hence, compared to earlier studies that quantified form prediction as *facilitation* for words with global similarity, our results here demonstrate *slowing* for words with onset similarity.

Notably though, our findings match with a large body of literature suggesting that word onsets are more crucial for lexical processing than word offsets (e.g., Allopenna et al., 1998; Frisson et al., 2014; Magnuson et al., 2007). More broadly, our findings align with models of word processing that postulate competition between words and its onset neighbors. For example, the word recognition model by McClelland and Rumelhart (1981) argues that words are recognized by suppressing the activation of closely related competing candidate words. When a word is predicted, like in our study, onset neighbors compete for activation due to their lexical similarity, which ultimately results in activation suppression and slowed processing for the onset neighbor. In other words, when reading constraining sentences, participants predict the orthographic form of predictable words, which, in turn, inhibits the activation of closely related competitors.

Even though such competitor effects may ultimately inhibit both onset and offset related neighbors (i.e. both head and rhymes), inhibitory effects may be more prominent for heads during early stages of processing when participants have already accrued activation of orthographic forms but did not have enough time to fully assemble these forms before reading the unpredictable sentence-final noun. Hence, at sentence offset, activation was measurable for word onsets but not offsets.

This interpretation of our findings is also supported by prediction-by-production models (e.g., Pickering & Garrod, 2007; Martin et al., 2018), which assume that predictions are generated by means of covert simulation in the language production system. Word form assembly is likely one of the final stages of this simulation process that involves a linear alignment of selected phonemes. Naturally then, word onsets are assembled before word offsets.

Conclusion

Taken together, our results suggest that language users activate both meaning- and form-related features when reading constraining sentences. While semantically related words show facilitation, possibly resulting from spreading activation, orthographic neighbors that share the same onset become inhibited. Ongoing work in our lab tracks the time course with which onset and offset features of predictable words become activated during predictive processing.

Acknowledgments

We would like to thank the anonymous reviewers for their valuable comments. This work was funded by the *Deutsche Forschungsgesellschaft* DFG (German Research Foundation) – Project-ID 232722074 – SFB 1102, Project A5.

References

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.

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*, 255-278.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*, 1-48.
Brysbaert, M., Buchmeier, M., Conrad, M., Jacobs, A. M., Bölte, J., & Böhl, A. (2011). The word frequency effect: A review of recent developments and implications for the choice of frequency estimates in German. *Experimental Psychology*, *58*, 412–424.
DeLong, K. A., Chan, W. H., & Kutas, M. (2019). Similar time courses for word form and meaning preactivation during sentence comprehension. *Psychophysiology*, *56*(4), e13312.
DeLong, K. A., Chan, W. H., & Kutas, M. (2021). Testing limits: ERP evidence for word form preactivation during speeded sentence reading. *Psychophysiology*, *58*(2), e13720.
DeLong, K. A., Groppe, D. M., Urbach, T. P., & Kutas, M. (2012). Thinking ahead or not? Natural aging and anticipation during reading. *Brain and Language*, *121*(3), 226-239.
DeLong, K. A., Urbach, T. P., & Kutas, M. (2005). Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. *Nature Neuroscience*, *8*(8), 1117-1121.
Federmeier, K. D., & 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.
Ferreira, F., & Chantavarin, S. (2018). Integration and prediction in language processing: A synthesis of old and new. *Current Directions in Psychological Science*, *27*(6), 443-448.
Fleur, D. S., Flecken, M., Rommers, J., & Nieuwland, M. S. (2020). Definitely saw it coming? The dual nature of the pre-nominal prediction effect. *Cognition*, *204*, 104335.
Frisson, S., Bélanger, N. N., & Rayner, K. (2014). Phonological and orthographic overlap effects in fast and masked priming. *Quarterly Journal of Experimental Psychology*, *67*(9), 1742-1767.
Gambi, C., Gorrie, F., Pickering, M. J., & Rabagliati, H. (2018). The development of linguistic prediction: Predictions of sound and meaning in 2-to 5-year-olds. *Journal of Experimental Child Psychology*, *173*, 351-370.
Gambi, C., Pickering, M. J., & Rabagliati, H. (2021). Prediction error boosts retention of novel words in adults but not in children. *Cognition*, *211*, 104650.
Haeuser, K. I., & Kray, J. (2022). Uninvited and unwanted: False memories for words predicted but not seen. In J. Culbertson, A. Perfors, H. Rabagliati & V. Ramenzoni (*Eds.*), *Proceedings of the 44th Annual Meeting of the Cognitive Science Society* (pp. 2401-2408).
Haeuser, K. I., Kray, J., & Borovsky, A. (2022). Hedging Bets in Linguistic Prediction: Younger and Older Adults Vary in the Breadth of Predictive Processing. *Collabra: Psychology*, *8*(1), 36945.
Huettig, F. (2015). Four central questions about prediction in language processing. *Brain Research*, *1626*, 118-135.
Ito, A., Corley, M., Pickering, M. J., Martin, A. E., & Nieuwland, M. S. (2016). Predicting form and meaning: Evidence from brain potentials. *Journal of Memory and Language*, *86*, 157-171.
Kamide, Y., Altmann, G. T., & Haywood, S. L. (2003). The time-course of prediction in incremental sentence processing: Evidence from anticipatory eye movements. *Journal of Memory and Language*, *49*(1), 133-156.
Keuleers, E., & Brysbaert, M. (2010). Wuggy: A multilingual pseudoword generator. *Behavior Research Methods*, *42*, 627-633.
Kim, A., & Lai, V. (2012). Rapid interactions between lexical semantic and word form analysis during word recognition in context: Evidence from ERPs. *Journal of Cognitive Neuroscience*, *24*(5), 1104-1112.
Kliegl, R., Nuthmann, A., & Engbert, R. (2006). Tracking the mind during reading: The influence of past, present, and future words on fixation durations. *Journal of Experimental Psychology: General*, *135*, 12–35.
Kuperberg, G. R., & Jaeger, T. F. (2016). What do we mean by prediction in language comprehension? *Language, Cognition and Neuroscience*, *31*, 32-59.
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.
Lew-Williams, C., & Fernald, A. (2010). Real-time processing of gender-marked articles by native and non-native Spanish speakers. *Journal of Memory and Language*, *63*(4), 447-464.
Martin, C. D., Thierry, G., Kuipers, J. R., Boutonnet, B., Foucart, A., & Costa, A. (2013). Bilinguals reading in their second language do not predict upcoming words as native readers do. *Journal of Memory and Language*, *69*(4), 574-588.
Magnuson, J. S., Dixon, J. A., Tanenhaus, M. K., & Aslin, R. N. (2007). The dynamics of lexical competition during spoken word recognition. *Cognitive Science*, *31*(1), 133-156.
Melinger, A. & Weber, A. (2006). Database of Noun Associations for German. Retrieved at http://www.coli.uni-saarland.de/projects/nag/.
McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: Part I. An account of basic findings. *Psychological Review, 88*(5), 375–407.
Ness, T., & Meltzer-Asscher, A. (2018). Lexical inhibition due to failed prediction: Behavioral evidence and ERP correlates. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *44*, 1269–1285.
Pickering, M. J., & Garrod, S. (2007). Do people use language production to make predictions during comprehension? *Trends in Cognitive Sciences*, *11*(3), 105-110.

R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Ryskin, R., & Nieuwland, M. S. (2023). Prediction during language comprehension: What is next? *Trends in Cognitive Sciences*.

Townsend, J.T., & Ashby, F.G. (1978). Methods of modeling capacity in simple processing systems. In J. Castellan & F. Restle (Eds.), *Cognitive theory. Vol. 3.* (pp. 200-239). Erlbaum.

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

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