Modeling Cue-Based Retrieval and Prediction — One Morpheme at a Time

Elise Oltrogge (Oltrogge@em.uni-frankfurt.de)
Goethe University Frankfurt
60629 Frankfurt am Main, Germany

Umesh Patil (umesh.patil@gmail.com)
t2k GmbH, 01277 Dresden, Germany

Abstract

This work proposes an extension to the cue-based retrieval theory of sentence processing: memory retrieval and prediction processes during sentence comprehension take place at the morpheme level instead of at the word level. We illustrate this proposal by extending an existing cue-based retrieval model from word-level to morpheme-level processing, and show that our model better captures the interactions between memory retrieval and predictive processing. Specifically, we extend the model reported in Patil and Lago (2021), which accounted for the interaction between retrieval and prediction during the processing of German possessive pronouns, but failed to generalize to structures involving determiners (Oltrogge, Veríssimo, Patil, & Lago, accepted). Our results show that modeling at the morpheme level captures retrieval-prediction interactions more precisely. The model successfully predicts the pattern of prediction onsets across German possessive pronouns and determiners. The proposed morpheme-by-morpheme model is further supported by psycholinguistic evidence suggesting that humans naturally decompose words into their constituent morphemes.

Keywords: cue-based retrieval modeling; prediction; possessive pronoun; determiner

Introduction

Sentence processing relies on at least two key mechanisms: memory retrieval and predictive processing. Memory retrieval is involved in resolving retrospective linguistic dependencies. For instance, in the sentence *Martin is looking for his friend* the possessive pronoun *his* triggers a memory search to identify its gender-matching antecedent, *Martin*. Predictive processing, on the other hand, is essential for forming prospective dependencies. For example, in the French translation *Martin cherche sa copine*, the possessive pronoun *sa* is gender-marked to match the upcoming noun, *copine*. Upon encountering *sa* comprehenders use this feminine cue to predict that the following noun will also be feminine.

Typically, memory retrieval and prediction are studied separately. However, recent studies suggest that memory retrieval and predictive processing can interact (Campanelli, Van Dyke, & Marton, 2018; Chow, Momma, Smith, Lau, & Phillips, 2016; Schoknecht, Roehm, Schlesewsky, & Bornkessel-Schlesewsky, 2022; Stone et al., 2021). This highlights the need for a model that can account for both processes. While memory retrieval is often modeled within frameworks such as cue-based retrieval theory (Lewis & Vasishth, 2005), these frameworks typically focus only on retrospective dependencies. Patil & Lago (2021) extended this framework to incorporate predictive processing, positing that predictions also rely on memory. Like retrieved antecedents, predictions must be stored in memory until they are either confirmed or disconfirmed. Their model focused on the processing of German possessive pronouns which present a compelling test case for investigating the interaction of these mechanisms, because processing them requires both retrieving an antecedent and generating a gender-based predictions.

German possessive pronouns

German possessive pronouns like *ihre* encode two agreement dependencies: the stem (*ihr*$-$) refers to a preceding feminine possessor, and the suffix ($-e$) signals an upcoming feminine possessee noun. Consequently, processing German possessive pronouns requires resolving an antecedent dependency and predicting an upcoming noun. Crucially, the retrieval-prediction interaction has also been found for German pronoun resolution. In a visual-world study, Stone et al. (2021) investigated whether the time course of noun prediction could experience interference from the preceding antecedent-retrieval—a process that should be grammatically irrelevant to the prediction process. Participants listened to instructions that contained a possessive pronoun (such as in 1) while seeing two gender mismatching objects on a screen (e.g., a target such as bottle$_{FEM}$ and a competitor such as button$_{MASC}$). Participants were told at the beginning of the experiment that these objects were belonging to one of two fictional characters, Martin or Sarah. Stone et al. (2021) analyzed participants' prediction onsets, defined as the point in time when participants began to look significantly more at the target object than at the competitor object. Their analysis revealed two key findings: first, participants predicted the target object before hearing its name, suggesting they used the pronoun's suffix to generate predictions; second, a match advantage such that predictions were faster when the gender features in the possessive matched (as in 1a) compared to when they mismatched (as in 1b).

(1) a. **MATCH condition**
Klicke auf ihre blaue Flasche!
Click on **her**$_{feminine}$ **blue**$_{feminine}$ **bottle**$_{feminine}$

b. **MISMATCH condition**
Klicke auf seine blaue Flasche!
Click on **his**$_{masculine}$ **blue**$_{feminine}$ **bottle**$_{feminine}$

In D. Barner, N.R. Bramley, A. Ruggeri and C.M. Walker (Eds.), *Proceedings of the 47th Annual Conference of the Cognitive Science Society*

Predictions as memory retrievals – past modeling work

Patil & Lago (2021) created a model that accounted for the effects observed by Stone et al. (2021) by conceptualizing predictive processing as a form of memory retrieval within a cue-based retrieval framework. This framework, the cue-based retrieval theory (Lewis & Vasishth, 2005), models memory retrievals in sentence processing based on the principles of ACT-R (Adaptive Control of Thought-Rational, Anderson et al. (2004)). It conceptualizes sentence processing as a series of activation-based skilled memory retrievals. Declarative memory stores lexical knowledge and partial sentence representations in the form of chunks, while procedural memory encodes psycholinguistic processes as production rules. These rules facilitate the retrieval and integration of information, enabling the incremental construction of sentence meaning.

In Patil & Lago's (2021) model, predicting the possessee noun based on a possessive pronoun uses the same cue-based mechanism as retrieving the pronoun's antecedent. Since both processes rely on the same architecture, they can interact, leading to similarity-based interference. The model explains the match advantage as follows: upon encountering a possessive pronoun, an antecedent retrieval is initiated, guided by cues like gender (e.g. $feminine$) and animacy ($animate$). The chunk for *Sarah* matches both cues, while the *bottle* partially matches $feminine$, receiving some spreading activation. In the subsequent prediction step, the *bottle* chunk gains further activation. Since it fully matches the expected features (being $feminine$ and $animate$) and has already been boosted, prediction is facilitated in the match condition. In the mismatch condition, the competitor object (e.g., the *button*) receives spreading activation during antecedent retrieval due to its match with one of the retrieval cues (*masculine*). As a result, predictions for the target object are only triggered during the picture prediction step.

Patil & Lago (2021) could successfully model the fixation patterns observed in Stone et al. (2021). However, this model was fitted post hoc to the empirical data. To address this limitation, recent work (Oltrogge et al., accepted) extended the model to test its ability to account for a new configuration: a determiner condition (as in 1c). To ensure comparability across conditions, all possessive pronouns and determiners were aligned to begin at the same time point, and word length was controlled by using indefinite forms instead of definites. The determiner configuration was intended to provide a baseline measure of prediction speed in the absence of antecedent retrieval. Therefore, the antecedent retrieval step was absent. Despite these adjustments, the extended model failed to account for the empirical data collected subsequently.

(1) c. **DETERMINER condition**
Klicke auf eine blaue Flasche!
Click on **a blue**$_{feminine}$ **bottle**$_{feminine}$

The model predicted a two-way split, with predictions being equally speeded in both the match and the determiner conditions compared to the mismatch condition (match = determiner < mismatch). This result stemmed from the absence of an antecedent retrieval step in the determiner condition, allowing the model to proceed directly to the picture prediction step. As a result, target fixations in the determiner condition increased at the same time as the antecedent retrieval boosted the activation of the target object in the match condition, increasing target fixations. However, the empirical data showed a three-way split or a staircase pattern, with the determiner condition falling between the match and mismatch conditions in prediction speed (match < determiner < mismatch). This discrepancy highlights the limitations of the extended model in capturing the nuanced time course of predictive processing in this new configuration.

Morpheme-by-morpheme model

We propose a model that processes sentences on a morpheme-by-morpheme level rather than on a word-by-word level, as proposed by the classic cue-based retrieval theory (Lewis & Vasishth, 2005). This increased granularity of input processing is well-supported by robust psycholinguistic evidence showing that humans naturally decompose words into their constituent morphemes during language processing (Gafni, Yablonski, & Ben-Shachar, 2019; Rastle, Davis, & New, 2004; Taft & Nguyen-Hoan, 2010). Visual masked priming experiments, in particular, have demonstrated that morphological decomposition occurs very early and facilitates faster recognition of morphologically complex words, with a morphologically related prime boosting processing speed (Diependaele, Sandra, & Grainger, 2009; Marslen-Wilson, Bozic, & Randall, 2008). Neural evidence from MEG studies (Solomyak & Marantz, 2010) supports the early decomposition of morphologically complex words. ERP findings (Regel, Opitz, Müller, & Friederici, 2019; Leminen & Clahsen, 2014; Smolka & Eulitz, 2015; Marslen-Wilson & Tyler, 2007) and fMRI data (Marslen-Wilson & Tyler, 2007) further demonstrate that even inflected words undergo morphological decomposition. Although these studies do not specifically focus on inflected determiners or pronouns, the general pattern of decomposition observed in inflected words likely applies to these forms as well.

The granular processing approach is also seen in other domains of cognitive science, such as visual processing. In visual processing, information is first decomposed by small receptive fields into basic features such as lines. These simple features are then combined step by step at higher levels, forming larger patterns and eventually more complex shapes such as objects. This kind of hierarchical processing helps the brain efficiently interpret complex visual information by decomposing it (for an overview of hierarchical models, see Ricci and Serre (2022).

Importantly, the current data also suggest that such a morphemic decomposition occurs: participants showed early divergence between the target and competitor fixations even before the pronoun suffix was heard. Specifically, in the match condition, a fixation preference for the target emerged at around 400 ms, whereas in the mismatch condition, such a preference was observed for the competitor. These early and opposite divergence patterns, which are consistent with findings from Stone et al. (2021), indicate that participants reacted to the gender information in the pronoun stem (e.g. *ihr−* or *sein−*) prior to hearing the full pronoun. Adopting a morpheme-by-morpheme processing model should better capture this behavior by simulating antecedent retrieval during stem processing, with retrieval interference driving the divergence between target and competitor curves. We evaluate whether this morpheme-based model generates a more accurate fit to the observed three-way split pattern, providing further insights into the underlying cognitive processes.

Model assumptions

The model presented in this paper refines and extends the assumptions proposed by Patil & Lago (2021), which are as follows:

(1) The smallest unit of processing linguistic input is the morpheme, rather than a whole word. Words are processed incrementally, morpheme-by-morpheme, and then combined to form a whole word.

(2) At each morpheme, the model attempts to predict the target object based on the linguistic input available up to that point. This reflects the participant's task of clicking on the target object, and assumes that participants incrementally use every piece of information to perform the task.

(3) All objects on screen are stored as referents in declarative memory which includes, for instance, *bottle*, *button* but also *Martin* and *Sarah*. Thus, these memory representations can be accessed during sentence processing.

(4) Predictions are implemented as memory retrievals.

(5) The probability of fixating on an object depends on its activation in memory: the object with higher activation is assumed to be the one that is fixated.

The key assumption of the morpheme-by-morpheme model is Assumption (1). It introduces morphemes as the smallest unit of sentence processing, refining the traditional approach of modeling sentence processing as a word-by-word process (e.g. cue-based retrieval theory (Lewis & Vasishth, 2005), Dependency Locality Theory (Gibson, 2000), self-organized parsing (Smith, Franck, & Tabor, 2018), expectation-based parsing (Levy, 2008)). All morphemes are treated equally, with no effect of morpheme-level frequency. Assumption (2) adapts the incremental prediction mechanism from the Patil & Lago (2021) framework to align with the morpheme-by-morpheme approach. This is task-dependent: in tasks without a selection component, predictions for the target would likely be less frequent. Assumptions (3–5) align with Patil & Lago (2021), reflected in the processing steps across the three conditions.

Match/mismatch condition The model of processing possessive pronouns included the following key steps (as illustrated in Figure 1a). First, the stem initiates lexical processing, antecedent retrieval and picture prediction. During lexical processing, the model retrieves the stem morpheme and its features from declarative memory, simulating the auditory input of the morpheme. The antecedent retrieval step performs the pronoun resolution. It carries out a search in memory, as a cue-based retrieval process, for the possessor using the retrieval cues: gender (either *masculine* for the stem *sein-* or $feminine$ for the stem *ihr−*), animacy (*animate*) and number (*singular*). In the subsequent picture prediction stage, a search for the upcoming noun is initiated, again as a cue-based retrieval process. The retrieval is carried out using the cues: gender ($neuter$), animacy ($inanimate$), and number ($singular$). The gender cue is $neuter$, reflecting the grammatical property of the morpheme, assuming it forms a complete word, such as in *sein Auto* (his car$_{neuter}$) or *ihr Auto*(her car$_{neuter}$). The picture prediction step based on the stem is a reflection of Assumption (2) above. Next, the suffix initiates lexical processing and picture prediction. To simulate the auditory input of the morpheme, the model retrieves the suffix morpheme and its features from declarative memory during lexical processing. During picture prediction, the search in memory for the upcoming noun is driven by a gender cue (either *masculine* for the suffix *-en* or $feminine$ for the suffix *-e*), an animacy cue (*inanimate*) and number cue (*singular*). Finally, the stem and suffix are combined to form a possessive pronoun.

Determiner condition For the model of processing indefinite determiners, the antecedent retrieval step is absent which reduces the number of processing steps (see Figure 1b). The stem handles lexical processing and picture prediction using memory retrieval cues: gender (*neuter*), animacy (*inanimate*), and number (*singular*). The gender cue is derived from interpreting the stem morpheme as a full determiner such as in *ein Auto* (a car$_{neuter}$). Subsequently, the suffix performs the same steps, using the same cues during picture prediction as in possessive pronouns. Finally, the stem and suffix combine to form the determiner. Thus, the model posited fewer processing steps for the determiner condition compared to the match and mismatch conditions. This reduction in steps led to an earlier processing of the suffix in the determiner condition, as illustrated in Figure 1. Note that the process of attaching the suffix to the stem to form the whole word was modeled in an under-specified manner using the explicit morphemes as retrieval cues to access the full word from memory. A more comprehensive approach would be to implement morpheme buffers in ACT-R for maintaining currently processed morphemes, which would allow for the attachment of the two morphemes by using them as retrieval cues and then combining them to form the word. This approach would be similar to the lexical buffer and lexical-retrieval buffer implemented in Lewis & Vasishth (2005). However, since the process of attaching morphemes was the same across all conditions and

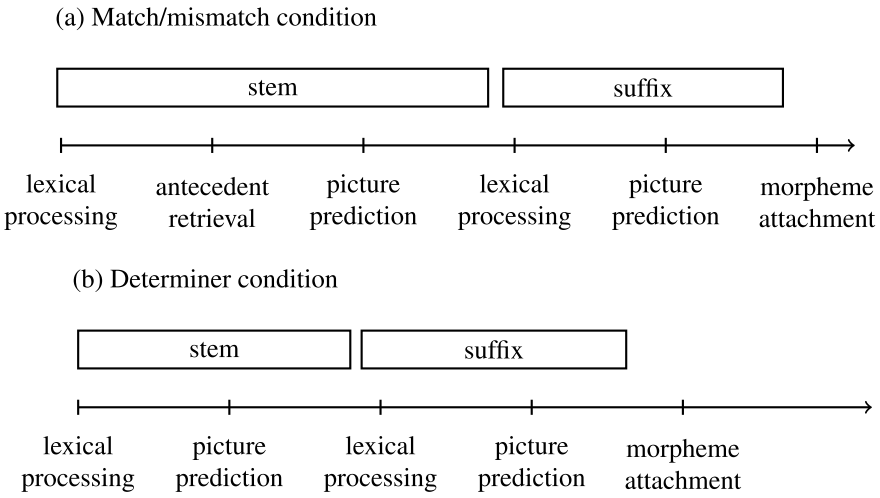

Figure 1: Schematic illustration of the model steps in the match and mismatch conditions vs. the determiner condition.

occurred after the second picture prediction process, it did not influence the picture prediction process. Therefore, the key difference between the models was the absence of the antecedent retrieval step in the determiner condition, which captured the main difference between the experimental design in Oltrogge et al. (accepted) with respect to the Patil & Lago (2021) model.

We used the same values for all ACT-R parameters as in Patil & Lago (2021) except for the *creation-time* parameter for the chunks *Martin* and *Sarah*. We set the value to -0.1. This parameter stores the time-point at which a chunk is created with respect to the beginning of the model run (which is $time = 0$). This sets the time for the creation of two chunks prior to the model's start time, resulting in a lower base-level activation compared to other chunks that are created after the start time. This adjustment is intended to reflect the fact that the pictures of *Martin* and *Sarah* are presented in the instructions *before* the experimental trial begins.

Model Predictions

We generated the predictions of the model by running 10,000 simulations for each condition. We obtained activation values for the target and the competitor object before each production fired and after each memory retrieval event. This yielded 9 activation values for each object for the match and mismatch condition and 7 for the determiner condition. To reflect Assumption (5), the object with the higher activation at each time-point is assumed to be fixated.

Results

Figure 2 shows the model predictions for prediction onsets–the time-point when the target is fixated significantly more often than the competitor–across three conditions. To analyze the prediction onset in the model predictions, we followed a procedure similar to that used in Patil & Lago (2021). Their approach involved conducting binomial t-tests on the target vs. competitor fixation predictions at each processing step (with the null hypothesis of fixation probability on the target being 0.5 and the alternative hypothesis of it being greater than 0.5). The prediction onset was identified as the first step showing a significant number of target fixations, which needed to be followed by at least four consecutive significant steps. Since our model sampled fewer activation values than those used in Patil & Lago (2021) (10 and 13 in P&L vs. 7 and 9 in ours), we adopted a more lenient criterion: we defined prediction onset as the first step with a significant t-test result, followed by at least one consecutive significant step.

The model successfully captures two key effects observed in the empirical data. First, in all three conditions, the model predicts predictive fixations on the target: higher fixations on the target compared to the competitor while processing the possessive pronoun or indefinite determiner (see Table 1). This captures the predictive effect seen in the empirical data, where participants show preferential fixations on the target. Second, the model predicts the staircase pattern observed in the empirical data for prediction onsets (match < determiner < mismatch). It predicts the earliest onsets in the match condition, later onsets in the determiner condition compared to the match condition, and the most delayed onsets in the mismatch condition (see Fig. 1 middle panel & right panel).

The model accounts for the predictive fixation patterns through similarity-based interference in both the match and mismatch conditions. During antecedent retrieval based on the possessive's stem, the gender cue boosts activation of gender-matching objects in memory. In the match condition, the retrieval cue (e.g., $feminine$ for the stem *ihr*) boosts the activation of the target object ($\text{bottle}_{feminine}$), resulting in more fixations on the target during stem processing. In contrast, in the mismatch condition, the retrieval cue (e.g., $masculine$ for the stem *sein-*) boosts the activation of the competitor object ($\text{button}_{masculine}$), leading to more fixations on the competitor during stem processing. The model predicts above-chance target fixations only when picture prediction

Table 1: Predicted fixations on the target and competitor objects by the morpheme-by-morpheme model across conditions and model steps. A '✓' indicates a fixation preference, '×' no preference and 'NA' a missing step. Fixations were sampled based on the current activation of the object chunk before the production rule fired. For antecedent retrieval and picture prediction, activation was sampled before and after chunk retrieval with the first instance of the model step representing the fixation before and the second after retrieval.

	Model Step	MATCH		DETERMINER		MISMATCH	
		Target	Competitor	Target	Competitor	Target	Competitor
Stem	Lexical Processing	×	×	×	×	×	×
	Antecedent Retrieval (before)	×	×	NA	NA	×	×
	Antecedent Retrieval (after)	✓	×	NA	NA	×	✓
	Picture Prediction (before)	✓	×	×	×	×	✓
	Picture Prediction (after)	✓	×	×	×	×	✓
Suffix	Lexical Processing	✓	×	×	×	×	✓
	Picture Prediction (before)	✓	×	×	×	×	✓
	Picture Prediction (after)	✓	×	✓	×	✓	×
	Attachment of morphemes	✓	×	✓	×	✓	×

based on the suffix occurs, as this retrieval cue (e.g., $feminine$ for the suffix *-e*) boosts the activation of the gender-matching target object (bottle$_{feminine}$).

In the determiner condition, the model does not perform antecedent retrieval. When processing the determiner's stem, the model moves directly to picture prediction. Since this prediction process is driven by the gender retrieval cue $neuter$, it does not significantly boost the activation of either the target or the competitor and fixations remain at chance levels for both. However, when the suffix is processed, the model predicts an increase in target fixations because the retrieval cue used for the picture prediction (e.g., $feminine$ for the suffix *-e*) boosts the activation of the gender-matching target object (bottle$_{feminine}$).

The staircase pattern in prediction onsets in the model prediction emerges as a combined effect of similarity-based retrieval interference and the timing of the picture prediction step across three conditions. The prediction onset in the match condition occurs earlier during antecedent retrieval based on the stem, whereas in the determiner and mismatch conditions it is triggered by the picture prediction based on the suffix. Since the determiner condition involves fewer processing steps, the model processes the suffix and initiates the picture prediction earlier than in the mismatch condition.

Discussion

This study addresses the interaction between memory retrieval and predictive processing in sentence comprehension, focusing on German possessive pronouns. These pronouns require both antecedent retrieval and prediction of an upcoming noun based on gender-marked morphemes. The study builds on the computational framework proposed by Patil & Lago (2021), which modeled predictions as a form of memory retrieval within a cue-based retrieval theory. This framework posits that both antecedent retrieval and noun prediction rely on shared memory mechanisms, allowing for interactions that influence the timing of predictions. The primary objective of this study was to refine and test the Patil & Lago (2021) model by implementing morpheme-by-morpheme processing as the fundamental unit of analysis instead of word-by-word processing. This shift is motivated by psycholinguistic evidence suggesting that humans naturally decompose words into their constituent morphemes during language processing. By processing possessive pronouns at the morpheme level, the current model seeks to better capture the nuances of predictive processing and its interaction with memory retrieval.

The morpheme-by-morpheme model reveals several key findings. First, consistent with Patil & Lago (2021), it accurately replicates the speed-up in prediction onsets observed in matching configurations compared to mismatching ones, as reported by Stone et al. (2021). Second, in contrast to the Patil & Lago model, when extended to the determiner condition, the model successfully accounts for the staircase pattern observed in prediction onsets in Oltrogge et al. (accepted) data, highlighting that a more granular approach can better capture the nuanced timing of these effects. Furthermore, the morpheme-based processing strategy is supported by the early divergence between target and competitor fixations, which occurs before the pronoun suffix is heard–a pattern observed in both Stone et al. (2021) and Oltrogge et al. (accepted).

However, the model struggles to capture certain finer details of the staircase pattern. The data show that the difference in prediction onset between conditions is more pronounced: the difference between the mismatch and determiner conditions is considerably larger than the difference between the determiner and match conditions. The model, however, predicts a smaller difference between the determiner and mismatch conditions, with a more pronounced difference between the determiner and match conditions.

Moreover, the model operates in time frame that does

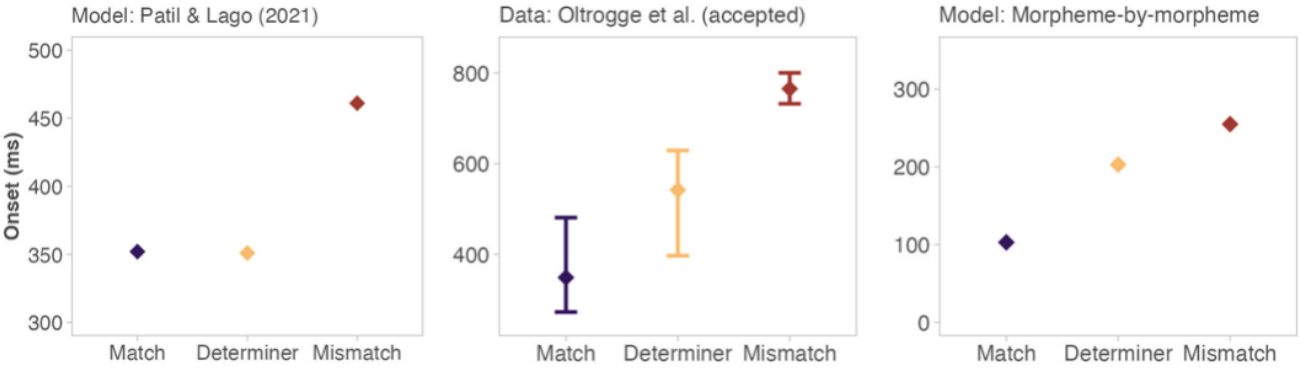

Figure 2: Left panel: Plot shows the median prediction onsets by Condition across 10,000 simulations. This data stem from Oltrogge et al. (accepted) who extended the Patil & Lago (2021) model by the determiner condition. Middle panel: Plot shows the median prediction onsets by Condition and their respective 95% highest-density intervals in the empirical data by Oltrogge et al. (accepted). Right panel: Plot shows the median prediction onsets by Condition across 10,000 simulations in the new morpheme-by-morpheme model.

not reflect real-time processing. For example, as shown in Figure 2, the median onset in the match condition appears around 380 ms in the empirical data, but much earlier (around 100 ms) in the model. The model's shorter onset times are likely a result of assumed durations individual processes in the model which are constrained by the duration for production rule firings and retrieval processes in ACT-R. This discrepancy in temporal resolution could be addressed in future work by fitting a free parameter to map model time onto empirical time—for instance, by adjusting ACT-R's latency factor to lengthen retrieval times or by introducing an external scaling factor.

An additional limitation is that we are not yet able to capture the variability around the prediction onset in the model data. This is due to the fact that we could not apply the GAMM-based onset detection method by Veríssimo and Lago (under review) to the model predictions that has been used to analyze the empirical data in Oltrogge et al. (accepted). The method requires grouping data by participants and items, which is not possible with the ACT-R model's independent simulations. Additionally, the model's discontinuous sampling process, compared to the continuous sampling in the behavioral data, further complicates the application of this method to the model data.

Another limitation of the model is its inability to capture the empirical pattern of incremental fixations. Instead of the stable or rising target fixations observed during picture prediction from the pronoun stem, the model predicts a decline—likely due to the weak activation from the *neuter* cue at that point. It also anticipates a sharp decrease in fixations after the suffix, which does not match empirical trends. This may reflect how morpheme attachment is implemented. Introducing a morpheme buffer, similar to the lexical buffer in Lewis & Vasishth (2005), could offer a more accurate account of incremental processing.

Further exploration is needed regarding the model's assumption that prediction relies on memory retrieval. While this aligns with the empirical design, where upcoming referents are limited and known, it contrasts with surprisal-based approaches that account for a broader range of continuations. Future work should explore whether retrieval-based prediction generalizes to less constrained contexts. Finally, since the model was fitted post hoc to the dataset of Oltrogge et al. (accepted), its generalizability remains uncertain. Future research should therefore empirically test the model in novel configurations.

Conclusion

By adopting a morpheme-by-morpheme approach to sentence processing, this study demonstrates that refining a psycholinguistic model to a more granular level–morphemes instead of words–enhances its ability to capture psycholinguistic effects more accurately. This refinement better accounts for morphosyntactic complexity, where cues used for memory retrieval are linked to the stem and cues used for prediction to the suffix. These findings not only have implications for cue-based models, but also help refine our understanding of when exactly prediction and retrieval take place by providing a working model of the two processes unfolding during sentence comprehension. Additionally, the observed limitations highlight the challenges of modeling real-time language comprehension, emphasizing the need for further refinements to capture incremental response patterns.

References

Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., & Qin, Y. (2004). An integrated theory of the mind. *Psychological review*, *111*(4), 1036.

Campanelli, L., Van Dyke, J. A., & Marton, K. (2018). The modulatory effect of expectations on memory retrieval during sentence comprehension. *Cognitive Science*, *42*(5), 789–815.

Chow, W.-Y., Momma, S., Smith, C., Lau, E., & Phillips, C. (2016). Prediction as memory retrieval: Timing and mechanisms. *Language, Cognition and Neuroscience*, *31*(5), 617–627. doi: 10.1080/23273798.2016.1151331

Diependaele, K., Sandra, D., & Grainger, J. (2009). Semantic transparency and masked morphological priming: The case of prefixed words. *Memory & Cognition*, *37*(6), 895–908. doi: 10.3758/MC.37.6.895

Gafni, C., Yablonski, M., & Ben-Shachar, M. (2019). Morphological sensitivity generalizes across modalities. *The Mental Lexicon*, *14*(1), 37–67. doi: 10.1075/ml.00013.gaf

Gibson, E. (2000). The dependency locality theory: A distance-based theory of linguistic complexity. *Image, language, brain/MIT Press*.

Leminen, A., & Clahsen, H. (2014). Brain potentials to inflected adjectives: Beyond storage and decomposition. *Brain Research*, *1543*, 223–234.

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

Lewis, R. L., & Vasishth, S. (2005). An activation-based model of sentence processing as skilled memory retrieval. *Cognitive Science*.

Marslen-Wilson, W. D., Bozic, M., & Randall, B. (2008). Early decomposition in visual word recognition: Dissociating morphology, form, and meaning. *Language and Cognitive Processes*, *23*(3), 394–421. doi: 10.1080/01690960701872516

Marslen-Wilson, W. D., & Tyler, L. K. (2007). Morphology, language and the brain: the decompositional substrate for language comprehension. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *362*(1481), 823–836. doi: 10.1098/rstb.2007.2101

Oltrogge, E., Veríssimo, J., Patil, U., & Lago, S. (accepted). *Memory retrieval and prediction interact in sentence comprehension: An experimental evaluation of a cue-based retrieval model.* Retrieved from `https://osf.io/7vz3m` (Preprint)

Patil, U., & Lago, S. (2021). Prediction advantage as retrieval interference: An act-r model of processing possessive pronouns. In T. C. Stewart (Ed.), *Proceedings of the 19th international conference on cognitive modelling* (pp. 213–219). Applied Cognitive Science Lab, Penn State. Retrieved from `http://acs.ist.psu.edu/papers/ICCM2021Procee-dings.pdf`

Rastle, K., Davis, M. H., & New, B. (2004). The broth in my brother's brothel: Morpho-orthographic segmentation in visual word recognition. *Psychonomic Bulletin & Review*, *11*, 1090–1098. doi: 10.3758/BF03196706

Regel, S., Opitz, A., Müller, G., & Friederici, A. D. (2019). Processing inflectional morphology: Erp evidence for decomposition of complex words according to the affix structure. *Cortex*, *116*, 143–153.

Ricci, M., & Serre, T. (2022). Hierarchical models of the visual system. In *Encyclopedia of computational neuroscience* (pp. 1533–1546). Springer. doi: 10.1007/978-3-030-12189-9_479

Schoknecht, P., Roehm, D., Schlesewsky, M., & Bornkessel-Schlesewsky, I. (2022). The interaction of predictive processing and similarity-based retrieval interference: An erp study. *Language, Cognition and Neuroscience*, *37*(7), 883–901. doi: 10.1080/23273798.2022.2069826

Smith, G., Franck, J., & Tabor, W. (2018). A self-organizing approach to subject–verb number agreement. *Cognitive Science*, *42*, 1043–1074. doi: 10.1111/cogs.12571

Smolka, E., & Eulitz, C. (2015). Electrophysiological evidence for stem access in regular and irregular german participles. In *Poster presented at the eight international morphological processing conference (moproc), potsdam, germany.*

Solomyak, O., & Marantz, A. (2010). Evidence for early morphological decomposition in visual word recognition. *Journal of Cognitive Neuroscience*, *22*(9), 2042–2057.

Stone, K., Veríssimo, J., Schad, D. J., Oltrogge, E., Vasishth, S., & Lago, S. (2021). The interaction of grammatically distinct agreement dependencies in predictive processing. *Language, Cognition and Neuroscience*, *36*(9), 1159–1179. doi: 10.1080/23273798.2021.1925133

Taft, M., & Nguyen-Hoan, M. (2010). A sticky stick? the locus of morphological representation in the lexicon. *Language and Cognitive Processes*, *25*(2), 277–296. doi: 10.1080/01690960903151546

Veríssimo, J., & Lago, S. (under review). *A novel method for detecting onsets of experimental effects in visual world eye-tracking.* Retrieved from `https://osf.io/preprints/psyarxiv/yk4xb_v3` (Preprint)

Appendix

An ACT-R trace with a single simulation of each condition is available on OSF: `https://osf.io/7epu2`.