Limited lexical prediction among content experts:
Evidence from an ERP study of language and memory

Melissa Troyer (melissa.troyer@unlv.edu), Psychological and Brain Sciences, UNLV

The functional role of lexical prediction in language comprehension is not well specified, with many assuming we predict to ease comprehension, and others suggesting (non-mutually-exclusively) that prediction is part of a learning mechanism [1]. These functional attributions have differing implications for when individuals will engage in prediction, and why. If prediction's main role is to ease comprehension, then domain experts—those with a great deal of relevant knowledge about a sentence's contents—might routinely engage in (and benefit from) prediction. By contrast, if prediction is primarily used to learn new information from (and about) language, then experts might be less inclined to predict specific words.

Methods. To examine the neural dynamics of prediction, as well as related downstream behavior, we recorded EEG as participants (current N = 37[1]) read sentence pairs. The first 90 were control items about everyday topics containing critical words that were predictable, unpredictable but plausible, or anomalous; the next 90 were about a specific fictional domain, the narrative world of Harry Potter (HP), with the same types of critical words (Table 1). After every 30 words, participants completed a memory task including previously seen critical words, never-seen words, and, importantly "lures"—predictable but never-seen words. We were especially interested in whether high-knowledge individuals would false-alarm to the HP "lures," implying they had been predicted (and made active) during the experiment [2].

Memory: For control items, we replicated previous findings [2] showing "lingering expectations": participants lured more frequently to *predictable* never-seen words compared to other "new" words, suggesting these words had been predicted and thus were more available in memory (Figure 1a). Luring effects were *not* seen in the HP blocks (neither across all participants nor within high-knowledge individuals; Figure 1b). However, domain knowledge increased memory performance overall (both hits and correct rejections). Additionally, high-knowledge individuals (alone) showed effects of condition on HP "hits" (expected > plausible > anomalous; Figure 2).

ERPs: As expected, we replicated several existing findings [3-4]: (1) for **both control and HP items**, N400s (larger in amplitude when more new information needs to be made available) were graded in amplitude (predictable < unpredictable-plausible < anomalous), and (2) late frontal positivities (LFPs; associated with revising predictions when disconfirmed) were observed only to unpredictable-plausible words; (3) **for HP items only**, N400 effects were driven by those with the greatest domain knowledge. Interestingly, our findings also show (for the first time) that for these "experts", late (posterior) positive components (LPCs; indicators of deep, endogenously captured attention known to promote memory processes) were largest to *predictable* words; for non-experts, LPCs were largest to *unpredictable-plausible* (i.e., "everyday" English) words, implicating a rapid effect of domain knowledge on subjective, explicit evaluation processes indexed by LPCs (Figure 3).

Subsequent memory effects on ERPs: In HP items, for high-knowledge, but not low-knowledge, individuals, we observed smaller N400s to critical words for subsequent hits < misses, suggesting that contextual facilitation during reading was positively related to subsequent memory (Figure 4). Among *plausible-unpredictable* words, high-knowledge (but not low-knowledge) individuals showed larger LPCs for (a) hits > misses (Figure 5) and (b) when they subsequently correctly rejected the associated (predictable) lure (compared to false alarms; Figure 6). Thus, endogenous capture of attention during processing of the unexpected, but plausible words was linked to higher accuracy on the memory test later on.

Summary. Unsurprisingly, individuals made immediate use of their domain knowledge to facilitate word processing; however, memory results suggest these processes may rely more on readily available *meaning* (based on rich mental models) rather than specific word-form prediction. Domain knowledge also changed how individuals processed and remembered unpredictable words (i.e., those not consistent with existing knowledge). These findings align with proposals that lexical prediction is part of a learning mechanism and suggest that this learning mechanism may be engaged less when relevant existing knowledge is high.

Supplementary figures and tables

Table 1: Representative materials.

	Context	Expected	Plausible	Anomalous	Ending
Control	We had been watching the blue jay for days. The bird laid her eggs in the	nest	shed	inch	up high.
	Eva put the whiskey bottle up high. She wanted it out of the baby's	reach	stroller	performer	as a precaution.
HP	Harry was practicing his Patronus. Eventually he conjured a	stag	spirit	statute	out of thin air.
	Peter Pettigrew often changes shape. This morning he turned into a	rat	zombie	salary	at breakfast.

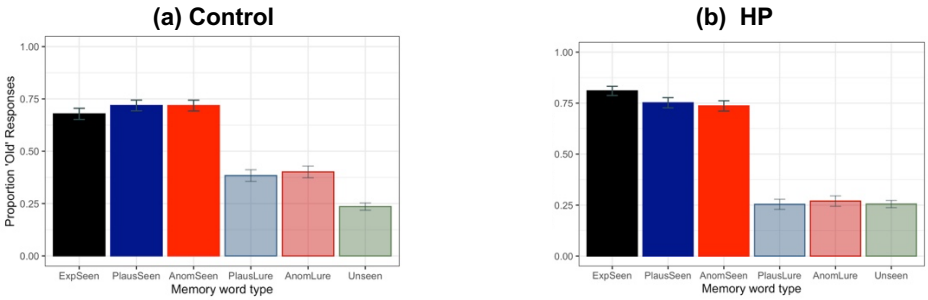


Figure 1. Proportion 'old' responses by sentence type (Control, HP) and memory word type.

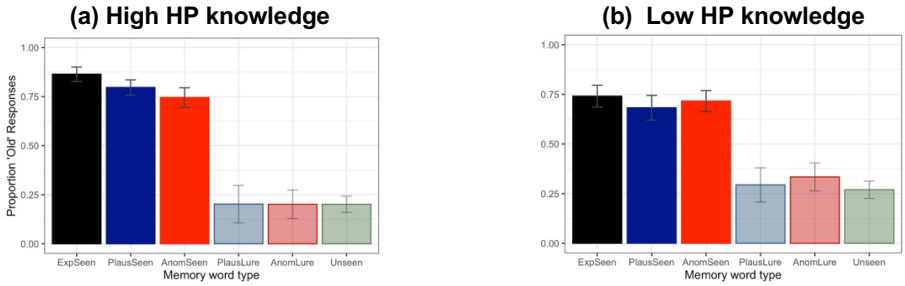
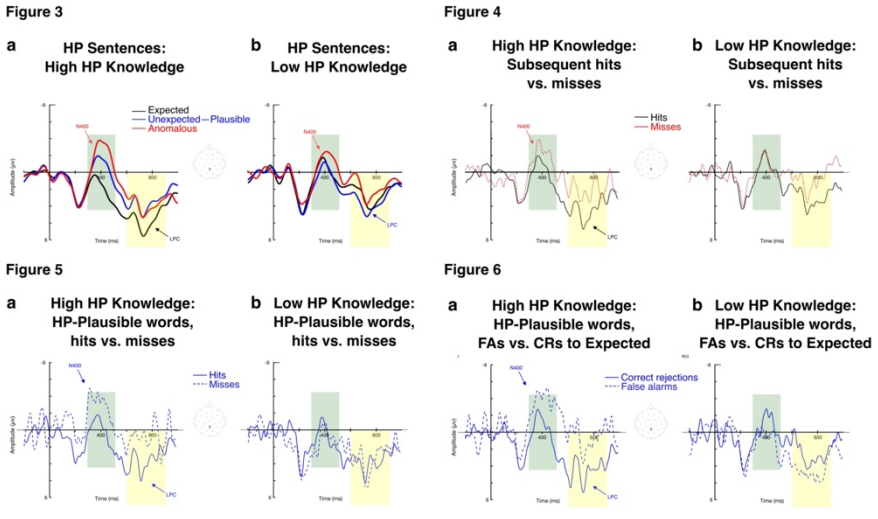


Figure 2. Proportion 'old' responses for HP words split by knowledge subgroup and memory word type.



ERPs from a representative midline parietal electrode. **Figure 3.** Domain knowledge-based differences in N400 (shaded in green) and LPC (yellow) effects to critical words in HP sentences. **Figure 4.** Knowledge-based subsequent memory effects for hits vs. misses in HP sentences, across all conditions. **Figure 5.** Knowledge-based subsequent memory effects for hits vs. misses to plausible continuations in HP sentences. **Figure 6.** Knowledge-based subsequent luring effects for plausible continuations in HP sentences when they were followed by false alarms vs. correct rejections to the corresponding expected words.

References

[1] Ryskin & Nieuwland, 2023. *TICS*. [2] Rommers & Federmeier, 2018. *NeuroImage*. [3] Troyer & Kutas, 2018. *LCN*. [4] Troyer & Federmeier, 2026. *Frontiers in Lang. Sci*.

[1] We have currently analyzed ERPs from 37 participants. The behavioral memory data come from a larger set of 81 participants, many of whose ERP data has not yet been analyzed.