Sentence processing depends on rapidly encoding and retrieving grammatical structure. However, properties of short-term memory remain controversial. Classical models posit specialized buffers with limited capacity [1–3], but recent models prioritize dynamic retrieval and interference without such buffers [4–6]. Processing of filler-gap dependencies provides insight into these mechanisms. To understand the phrase the book that Mary wrote __, the filler the book must be associated with the gap after wrote [7,8]. Some results propose that the filler is maintained in a privileged memory buffer [3,9]. For instance, [9] found differential effects of length on syntactic and semantic features in filler-gap dependency processing. Comprehenders sensed syntactic errors across dependency lengths, but sensitivity to implausibility decreased over long distances. Such data indicate that syntactic features are maintained over long distances, but semantic features are too costly to maintain over time and must be retrieved at the gap site, suggesting that different linguistic information may have different status in memory.

In a self-paced reading study, we found a similar asymmetry in phonological features. Short dependencies show no effect of phonological similarity between filler and verb, but we find that phonological overlap between filler and verb eases processing for long dependencies. This is surprising; phonological cues are not features for retrieval [11], and phonological activation dissipates rapidly in typical processing [12]. We propose that phonological features are not stored in working memory, but may be reactivated along with the syntactic/semantic features of the filler in a later retrieval event, supporting the model described in [9–10].

Experiment. Our experiment sought to replicate the length effect on plausibility in [9] and to determine whether length similarly affected phonological information. We expected fillers to facilitate processing of verbs with overlapping onsets if phonological features were active as a kind of phonological priming effect [11]. We manipulated Filler (Control/Implausible/Phonological Overlap) and Length (Long/Short). In the Control condition, the filler was a plausible argument of the verb but shared no phonological onset segments with the verb onset. In the Implausible condition, the filler was an implausible argument. In the Phonological Overlap condition, the first 3-5 phonological segments of the filler were the same as that of the verb. In the Long condition, there was a 4-word PP between the filler and verb. Forty-two participants read 36 sets of six items and 24 complexity-matched fillers at their own pace using a self-paced reading paradigm.

We conducted a linear mixed effects model on log residual reading times, following [13], at the critical verb (cited) and spillover regions (disliked; how). We report only on the critical verb and second spillover region. We removed two items due to unintended phonological overlap between control filler and verb. At the critical verb, there was a main effect of Length ($\beta = 0.015 \pm 0.0068, p = 0.02$) and an interaction effect between Length and Filler ($\beta = 0.024 \pm 0.0097, p = 0.01; \beta = 0.017 \pm 0.0098, p = 0.08$). Pairwise comparisons revealed increased reading times for Implausible fillers in this region for Short conditions only ($\beta = 0.063 \pm 0.0026, p = 0.02$). This suggests an immediate sensitivity to short implausible filler-gap dependencies, but not long implausible filler-gap dependencies. In the second spillover region, there was a main effect of Filler ($\beta = 0.040 \pm 0.0091, p < 0.01$), and a marginal interaction effect between Filler and Length ($\beta = 0.018 \pm 0.0093, p = 0.06$). Pairwise comparisons revealed a marginal effect of Phonologically Overlapping fillers, resulting in facilitated processing for Long dependencies only ($\beta = 0.017 \pm 0.0098, p = 0.07$). Implausible fillers were again longer in the Short conditions only ($\beta = 0.074 \pm 0.0023, p < 0.01; \beta = 0.068 \pm 0.023, p < 0.01$). This suggests that phonological overlap facilitates processing, but only if the filler is no longer actively maintained in working memory and must be retrieved, thereby reactivating the phonological features.
Figure 1. Materials for Experiment; critical region; spillover region.

Figure 2. Timeline schematic for differential treatment of lexical information in memory. Solid lines represent readily accessible lexical information over time. Dashed lines represent decayed but retrievable information.

Figure 3. Word-by-word log mean reading times per condition in Experiment.