Long filler-gap dependencies can make a main verb analysis less tempting

Brian Dillon (University of Massachusetts, Amherst)
brian@linguist.umass.edu

Race models of syntactic processing claim that the human sentence processor adopts the first available syntactic analysis [e.g. 1-4]. Though details vary, these models hold that multiple parses compete for selection at points of (incremental) ambiguity, and the parse that becomes available most quickly is selected as the ultimate analysis. However, there is little direct evidence for this claim. We provide a novel test of this claim by investigating sentences like (1).

(1) [**SHORT**]: *I'm not sure which treasure the pirate **noticed** at the beach found.*
[**LONG**]: *I'm not sure which treasure the gnarly old captain said that the pirate **noticed** at the beach found.*

In (1), ***noticed*** is temporarily ambiguous between a main verb (MV) reading and a reduced relative clause (RRC) reading. Since comprehenders immediately associate unassigned fillers to potential gap sites [5], an incremental MV parse of the sentence up to ***noticed*** involves the retrieval and integration of the filler which treasure. An RRC analysis (e.g. *the pirate who was noticed…)*, however, can be computed without retrieving the filler. For this reason, the time it takes to compute a MV parse in these examples depends on the time it takes to retrieve and integrate the filler at the ambiguous verb ***noticed***, while this is not true for the RRC parse.

Importantly, increasing the distance between a filler and its gap has been shown to increase the time required to fully process a filler-gap analysis [5,6]. In (1), this puts a MV analysis at a relative disadvantage in the LONG condition, because the parser may need more time to retrieve and successfully integrate a more distant filler phrase [5,6]. This should make the MV analysis slower to compute in the LONG condition, making an RRC analysis of ***noticed*** in this condition relatively more competitive. Thus if syntactic ambiguity is resolved by adopting the analysis that is computed most quickly, a garden-path effect at the disambiguating verb *found* should be smaller in the LONG condition than in the SHORT condition, because comprehenders will be more likely to select the (ultimately correct) RRC analysis initially.

We test this prediction in three pre-registered experiments. Experiments 1a (original) + b (confirmatory replication) test this using a speeded acceptability judgment design (both N_{SUBJ} = 32; N_{ITEM} = 28; native American English speakers recruited on Prolific Academic). We measured speeded acceptability judgments to AMBIGuous sentences like (1), as well as UNAMBIGuous counterparts created by inserting *who was* before the ambiguous verb ***noticed***. Results are shown in Table 1. We found that ambiguity impacted acceptability judgments to a much greater degree in SHORT conditions than in LONG conditions; this pattern was replicated in the confirmatory replication. This is initial evidence that comprehenders are more strongly garden-pathed in the SHORT conditions than in the LONG conditions, confirming our prediction. In addition, we saw a main effect of length, due to the increased complexity of LONG conditions.

In Experiment 2, we tested the same conditions in a self-paced reading experiment (N_{SUBJ} = 60; N_{ITEM} = 24). To avoid wrap-up effects, we added spillover regions to the stimuli from E1a/b. We found a significant slowdown at the point of disambiguation. This garden path penalty was significantly reduced for LONG conditions at the spillover +2 region. At the ambiguous region ***noticed***, the LONG filler-gap dependencies were associated with longer RTs than the SHORT dependencies, which might reflect increased difficulty in completing the filler-gap dependency [5], or the difficulty of accessing the reduced relative clause analysis [6].

In sum, we found that longer filler-gap dependencies led to smaller garden path effects when an ambiguous string is disambiguated to the dispreferred RRC structure. We interpret this result as evidence that an MV analysis of an ambiguous verb can be made relatively costly to compute by making the filler it relies on more distant, because this increases the time needed to successfully integrate the filler at its gap site. We take this as support for race-based models of syntactic processing, and evidence that at least part of the robust MV advantage in ambiguity resolution is rooted in how quickly/easily the parser can compute the main verb analysis [1-4].

(1) [**SHORT-AMBIG**]: *I'm not sure which treasure the pirate **noticed** at the beach found.*
(2) [**SHORT-UNAMBIG**]: *I'm not sure which treasure the pirate who was **noticed** at the beach found.*
(3) [**LONG-AMBIG**]: *I'm not sure which treasure the gnarly old captain said that the pirate **noticed** at the beach found.*
(4) [**LONG-UNAMBIG**]: *I'm not sure which treasure the gnarly old captain said that the pirate who was **noticed** at the beach found.*

Condition	% Yes (*1a/1b*)	RT (*1a/1b*)
SHORT-AMBIG	.38 / .44	933 (61) / 875 (60)
SHORT-UNAMBIG	.70 / .80	717 (30) / 695 (32)
LONG-AMBIG	.41 / .45	859 (43) / 830 (47)
LONG-UNAMBIG	.53 / .63	791 (46) / 762 (41)
Unacceptable fillers	.26 / .31	

Table 1: Percent yes/RT to correct responses for all four conditions, plus bad fillers, in Experiments 1a/b. Maximal logistic LMERs revealed a significant interaction of ambiguity and length on accuracy in both experiments ($t = -2.56$; $t = -3.27$, respectively). No consistent, reliable effects on RT were observed across experiments, but in both experiments the ambiguity penalty on RTs was more than twice as large in SHORT conditions as in LONG conditions.

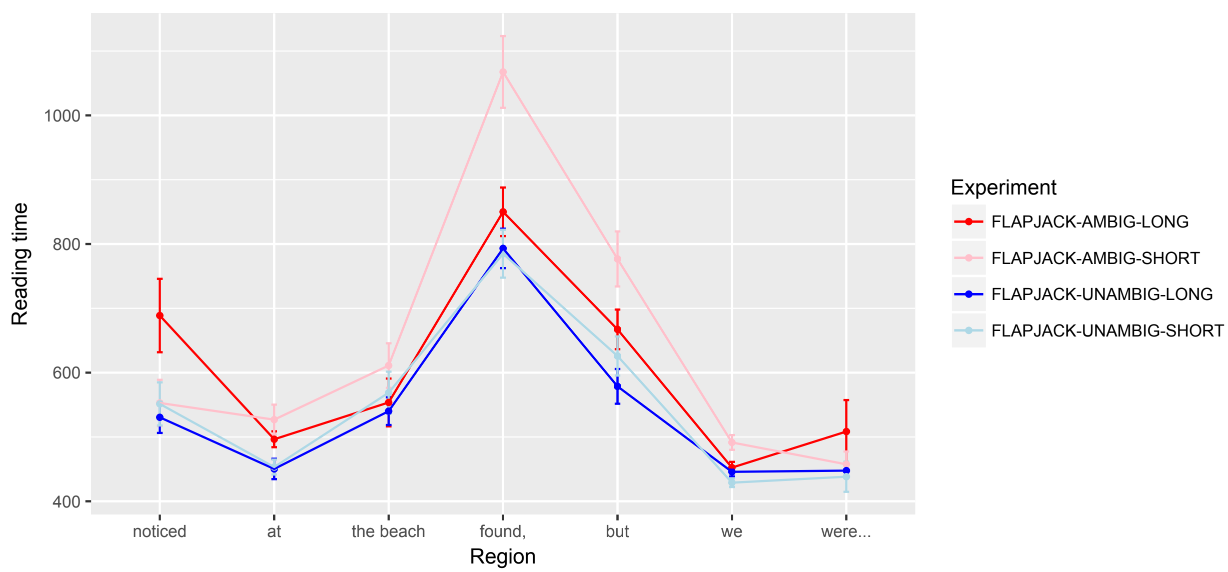

Figure 1: Mean reading times (ms) by condition by region. Error bars represent standard error by participants, corrected for between-participant variance. Maximal LMERs on log RTs revealed a significant effect of ambiguity at the critical region and spillover regions, qualified by an interaction of length and ambiguity ($t = -3.25$) at the 2nd spillover region, and a marginal interaction in log RTs at the critical region ($t = -1.89$). At the ambiguous verb *noticed,* an interaction with the opposite sign was observed ($t = 2.16$).

[1] Frazier, L. (1979). *PhD Thesis.* [2] Traxler, M., Pickering M., & Clifton, C. (1998). *Journal of Memory and Language*. [3] van Gompel, R., Pickering, M., & Traxler, M. (2000). *Reading as a perceptual process.* [4] Logacev, P. & Vasishth, S. (2016). *Cognitive Science.* [5] Phillips, C., azanina, N., & Abada, S. *Cognitive Brain Research.* [6] Wagers, M. & Phillips, C. (2014). *QJEP.* [6] Groder, D., Gibson, E. & Tunstall, S. (2002). *Journal of Memory and Language.*