Expectations and prediction in sentence comprehension: German particle verbs as a test case

Kate Stone, Shravan Vasishth, Titus von der Malsburg (University of Potsdam)
stone@uni-potsdam.de

The expectation-based account of sentence processing posits that incremental accumulation of contextual constraint strengthens expectations for the syntactic and lexical properties of downstream words, making them easier to process when they are encountered.[1,2] However, does strong expectation entail commitment to a specific lexical item even before it is encountered? While there is some support for this idea,[3-6] the evidence is by no means conclusive.[7,8] Here we hypothesise that, while expectations may always be generated during comprehension, a constraining-enough context may trigger a phase transition during which fluid expectations crystallise into a specific lexical prediction. German particle verbs represent an ideal test case for this hypothesis.[9,10] In a sentence like "Das Fest ging früh los" (*The celebration started early PARTICLE*), the finite base verb "ging" strongly predicts the particle "los" which resides in clause-final position. However, other base verbs would be compatible with a range of particles and may therefore not allow a specific lexical prediction. In two experiments, we manipulated base verbs to create expectations for either a small or a large set of possible downstream particles (see 1). We predicted that a small set of expected downstream particles would encourage the prediction of a specific particle and hence speed up processing at the particle site. In contrast, a large set of possible particles should discourage prediction due to high uncertainty, resulting in increased processing times at the particle.

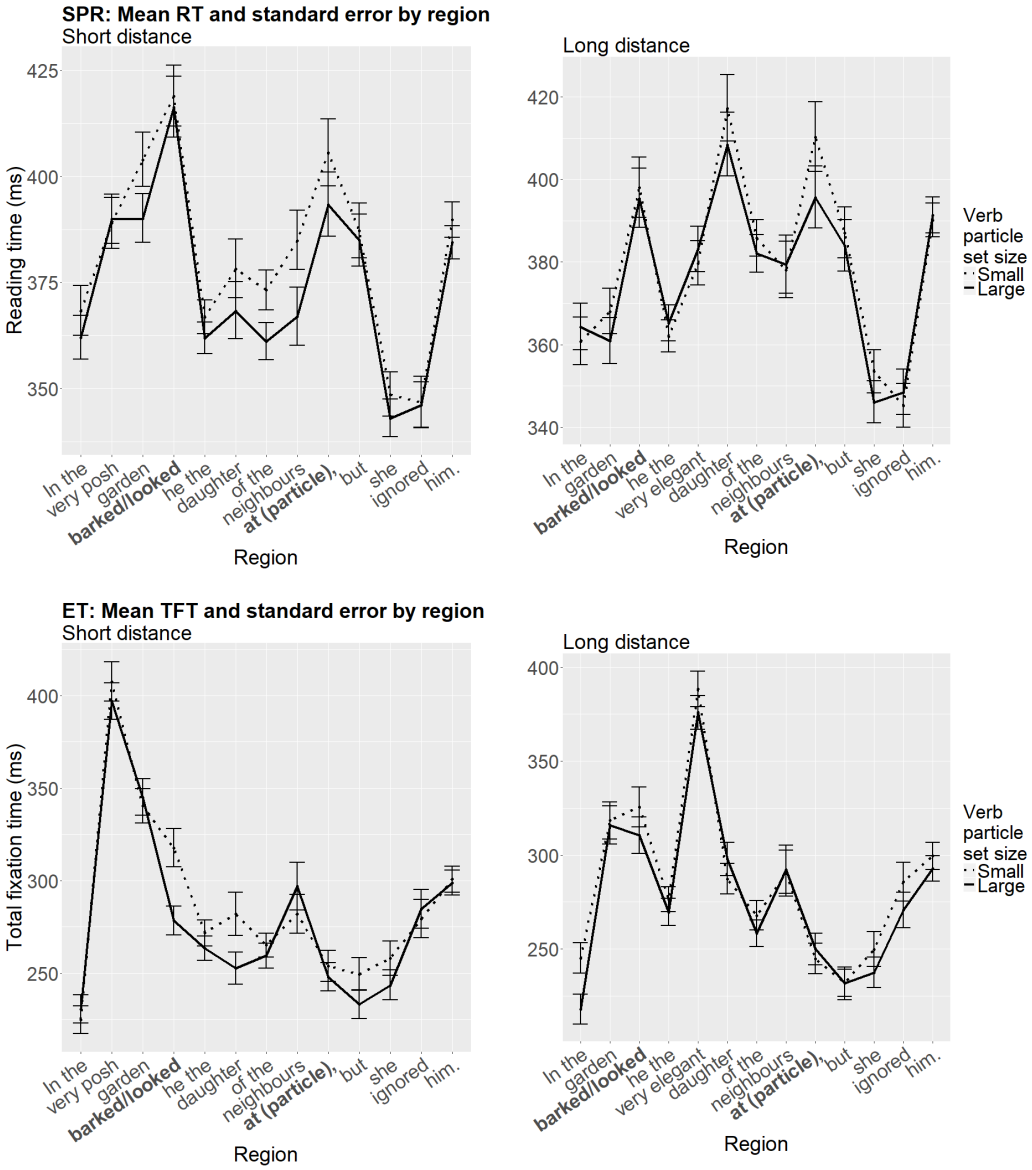
Design: We employed a 2×2 design with factors particle set size and particle-verb distance. In the small set size condition, verbs licensed 6 or fewer particles (1a, 1b) whereas verbs in the large set condition licensed 10 or more particles (1c, 1d). Distance was manipulated by shifting material either into the region between verb and particle (1a, 1c) or into the region before the base verb, effectively strengthening contextual constraint at the verb (1b, 1d). Sentences were designed such that the target particle had the highest cloze probability of all possible particles (a norming study confirmed this for 92% of the items). 24 items and 72 fillers were presented to 120 German native speakers. In the interest of cross-methodological validity, half read the material in a self-paced reading task (SPR), and half in a standard eye-tracking task (ET).

Results: Box-Cox transformed reading times were analysed using Bayesian linear mixed models with maximal random effects structures. Contrary to our predictions, the SPR data showed that reading times at the particle were slower in sentences with small-set verbs than in sentences with large-set verbs ($-12ms$, $Pr(\hat{\beta} < 0) = 0.98$, credible interval: $[-24, 0]$). There was no evidence for a main effect of set size in the ET task, however, there was evidence for an interaction of set size with verb-particle distance: Reading times for small-set verbs were slower, but only when distance was short ($25ms$, $Pr(\hat{\beta} > 0) = 0.98$, credible interval: $[1, 49]$).

Conclusions: These results disconfirm the prediction that a small particle set size (i.e. strong constraint) would facilitate processing; instead the SPR data show the opposite effect. Fig. 1 suggests an explanation for this puzzle: Starting from the base verb, reading times in SPR and ET were substantially slowed when set size was small and when the verb was preceded by material that helped narrow down the set of plausible particles (condition 1a). We speculate that only in this situation was contextual constraint strong enough to give rise to a specific lexical prediction. If true, the observed slowdown may reflect the added cost of forming a prediction and the deeper semantic analysis afforded once a concrete prediction is made. A follow-up experiment testing these ideas is currently underway.

(1) a. Im sehr vornehmen Garten **bellte** er die Tochter der Nachbarn **an**, . . .
b. Im Garten **bellte** er die sehr vornehme Tochter der Nachbarn **an**, . . .
c. Im sehr vornehmen Garten **schaute** er die Tochter der Nachbarn **an**, . . .
d. Im Garten **schaute** er die sehr vornehme Tochter der Nachbarn **an**, . . .
*In the very posh garden **barked/looked** he the very elegant daughter of the neighbours **at**, ...*

Figure 1: Mean reading times and standard errors by region in SPR and ET experiments.



References: [1] Hale (2001) *Proceedings NAACL* [2] Levy (2008) *Cognition* [3] De Long et al. (2005) *Nat. Neurosci* [4] Wicha et al. (2004) *J Cog Neurosci* [5] Van Berkum et al. (2005) *J Exp Psych* [6] Husain et al. (2014) *PloS One* [7] Nieuwland et al. (2017) doi:10.1101/111807 [8] Safavi et al. (2016) *Front Psych* [9] Piai et al. (2013) *Brain & Lang* [10] Müller (2002) *Complex Predicates*