No scope for planning – latent mixture modelling of sentence onsets

Before a speaker can start speaking a sentence they must pre-plan some minimal linguistic unit (e.g. Martin et al., 2014). Whether or not this minimal unit is syntactically organised is highly controversial (Bock & Ferreira, 2014). The argument that pre-planning must involve syntax hinges, partially, on the observation that, when the target sentence starts with a conjoined NP (*The A and the B*), the onset is delayed compared to matched sentences starting in simple NPs (*The A*) (for recent reproductions see Hardy et al., 2018; Roeser et al., 2018).

An alternative view is that planning beyond the first noun, and thus the syntax of the first NP, is not obligated by the linguistic encoder: Incremental theories of sentence planning hold that the minimum obligated planning unit might, in fact, encompass just the first lexical word. Under this account extended onset latencies may occur for sentences starting with a conjoined NP but this will be sporadically across trials and for reasons other than obligatory planning scope (Allum & Wheeldon, 2007). We contrasted these two views directly in a reanalysis of onset latencies from 3 experiments (subjects = 96, items = 144), published in XXX et al. (20xx) in which participants described arrays of moving images starting with either a conjoined NP or a simple NP.

We fitted a series of Bayesian models in Stan (Carpenter et al., 2017), incrementally increasing model complexity. These are summarised in Tab. 1. This approach allows a direct test of the predictions from the two contrasting theories. All models include random intercepts for subjects and items (Bates et al., 2015). The onset latency is assumed to come from a log-normal distributions (Baayen, 2008). Comparison of linear mixed models (LMM) M0 and M1 represents the standard approach to evaluating the effects of initial NP complexity as a test of the hypothesis that sentence planning requires a phrasal scope, and thus syntax.

Model M2 tests the alternative hypothesis that there is no consistent effect of initial NP complexity but that on a subset of trials in the conjoined NP condition participants opted to plan beyond the initial noun (i.e. planning syntax is more likely for conjoined NPs but not obligated by the production system). Under this account latencies come from a mixture of log-normal distributions. In M2 the parameters of each of these distributions – the mean and the variance – were fixed across condition, but the mixing proportion p – the proportion of trials associated with each distribution – was allowed to vary for each NP type. This captures the probability of latencies coming from a particular mixture component K (Vasishth et al., 2017). This mixture model therefore allowed us to test whether extreme values were more likely for conjoined NPs even if the location of the normal distributions remain constant across NP types. Model fit was compared by taking the difference of expected log pointwise predictive density ($\Delta\widehat{elpd}$) (Vehtari et al., 2017); see Tab. 2.

M1 revealed a slowdown of $\hat{\beta}$ = 95ms (95% HPDI [77.2, 114.4]) for conjoined NPs and better fit than M0 confirming the standard analysis. By contrast, the mixture models revealed a higher predictive performance; M2 showed better fit than M1. The other models not presented here revealed a poorer predictive performance than M2. Results of M2 can be found in Fig. 1 showing a combination of three mixture components. Conjoined NP types show a larger proportion of long values – component K_3 – than simple NPs; see Fig. 2.

We therefore found a greater tendency for longer onset latencies for conjoined NPs. However, these longer latencies remained in the minority. The majority of observations were associated with the same underlying distribution as was found for simple NPs. This is not consistent with obligatory phrasal scope (i.e. scope of the entire initial verb-argument NP), but rather that advance planning is highly incremental, scoping, at most, over the sentence-initial noun. We suggest that the frequently replicated slowdown for conjoined NPs might be better explained by non-syntactic processing demands that act only for some participants on some trials.

Tab. 1: Model specifications; models ordered by ascending complexity.

- M0 LMM without slope parameter (null model)
- M1 LMM with NP effect (standard analysis)
- M2 Mixture model with three mixture components; mixing proportions included for each NP type

Tab. 2: Model comparisons. A positive $\Delta\widehat{elpd}$ (*SE* = standard error) signifies improved predictive performance of a more complex model.

Models	$\Delta\widehat{elpd}$	*SE*
M2 vs M1	1103.5	102.6
M1 vs M0	53.5	10.8

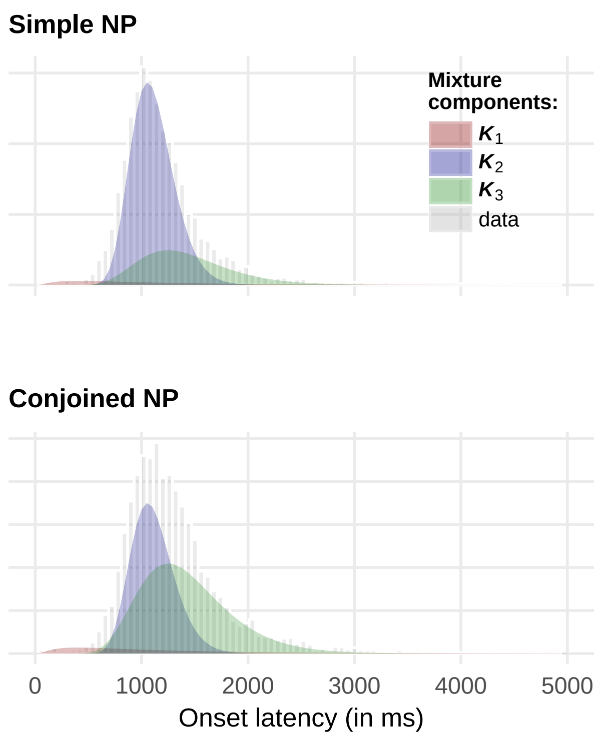

Fig. 1: Posterior distributions of mixture components inferred from M2.

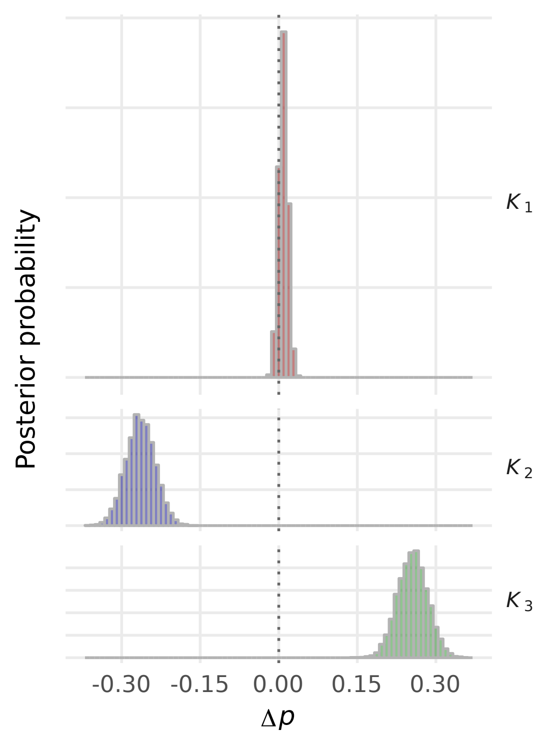

Fig. 2: Posterior probability of NP type difference of mixing proportions Δp = conjoined NP - simple NP by *K*.