Sentence-recall involves de- and encoding of syntax

Jens Roeser, Mark Andrews, Mark Torrance, Rodilene Gittoes (Nottingham Trent University)
jens.roeser@ntu.ac.uk

What we remember about an utterance plays a central role for our understanding of how language comprehension relates to production. Some representation of an utterance is necessary to correctly respond to questions. For example, Christianson et al. (2001) found that people are more likely to answer *yes* to the question *Did the man hunt the deer?* when omitting the comma after "hunted" in *While the man hunted the deer run into the woods*. This finding was taken as evidence that syntax decoding happens to some extent off-line (e.g. Logačev & Vasishth, 2016). As linguistic representations rapidly decay from memory (Christiansen & Chater, 2016), the production system might be tasked to encode syntax from a general conceptual representation of the previously read sentence. We contrasted these two views.

In two experiments participants read sentences and then typed each sentence from memory. Experiment 1 tested whether ambiguous sentences show increased typing-onset latencies. Participants (*N*=64) were presented with 24 items taken from Christianson et al. (2001) including ambiguous and unambiguous sentences (example above). Bayesian mixed effects models showed that onset latencies were 53 msecs (95% HPDI[10, 96]) longer for ambiguous sentences.

Experiment 2 tested whether latencies increase because of difficulty related to syntax de- or encoding. Participants (*N*=80) were presented with 24 simplified items taken from Van Gompel et al. (2001); involving global ambiguities (1), and temporary ambiguous sentences with high (2) or low attaching PPs (3).

1. The caretaker cleaned the pail with the brush. (global ambiguity)
2. The caretaker cleaned the suit with the brush. (high attachment)
3. The caretaker cleaned the pail with the holes. (low attachment)

We compare two possible explanations for recall difficulty: first, structures were parsed incorrectly and required reanalysis (syntax decoding); thus longer latencies are expected for the subset of temporary ambiguous trials that was parsed incorrectly; global ambiguities allow both parses. Second, a more complex syntactic arrangement has to be generated from a conceptual representation (syntax encoding); thus, longer latencies are predicted for high-attachment sentences and the subset of global ambiguous trials that was encoded as high attachment.

To compare these two views we implemented a series of Bayesian models in Stan (Carpenter et al., 2017). We modelled typing-onset latencies as coming from a mixture of two log-normal distributions with means β and δ where δ is the result of some onset inhibition. The probability of onsets of type δ was captured by the mixing proportion θ. Crucially, we tested how θ varied by attachment type. Models were compared to a linear mixed effects model with attachment type as predictor and an intercept-only model. Model fit was assessed using expected log predictive density ($\widehat{elpd}$) which penalises models with a larger number of parameters (Vehtari et al., 2017).

Highest predictive performance was found for the model with mixing proportions for each attachment type; see Tab. 1. In other words, onset latencies were found to be a combination of two distributions with different mixing proportions for each attachment type. Every attachment type showed a mixing proportion around $\hat{\theta}$=.5. Parameter estimates showed a higher proportion of long latencies for high attachment compared to other attachment types; see Fig. 1.

These results are unexpected as we predicted mixture for either temporary or global ambiguities. Instead we found mixtures for all attachment types. These may reflect parsing errors for temporary ambiguities and increased encoding demands for high attachment interpretations. Therefore, sentence recall may involve both decoding and encoding of syntactic representations.

Tab. 1: Model comparisons. The difference in expected log predictive density ($\Delta\widehat{elpd}$) and standard errors (SE) are shown in comparison to the model with the highest predictive performance (top row).

Model	$\Delta\widehat{elpd}$	SE
MM: mixture proportion by-attachment type	0	0
MM: mixture proportion for temporary ambiguities	-15.6	9.1
LMM: attachment type as predictor	-29.6	11.1
MM: mixture proportion for global ambiguities	-29.9	9.9
LMM: intercept-only (null model)	-32.6	10.7

Note. MM = mixture model; LMM = linear mixed effects model; all models were fitted with random intercepts for subjects and items (Bates et al., 2015) and distributed as log-normals (e.g. Baayen, 2008).

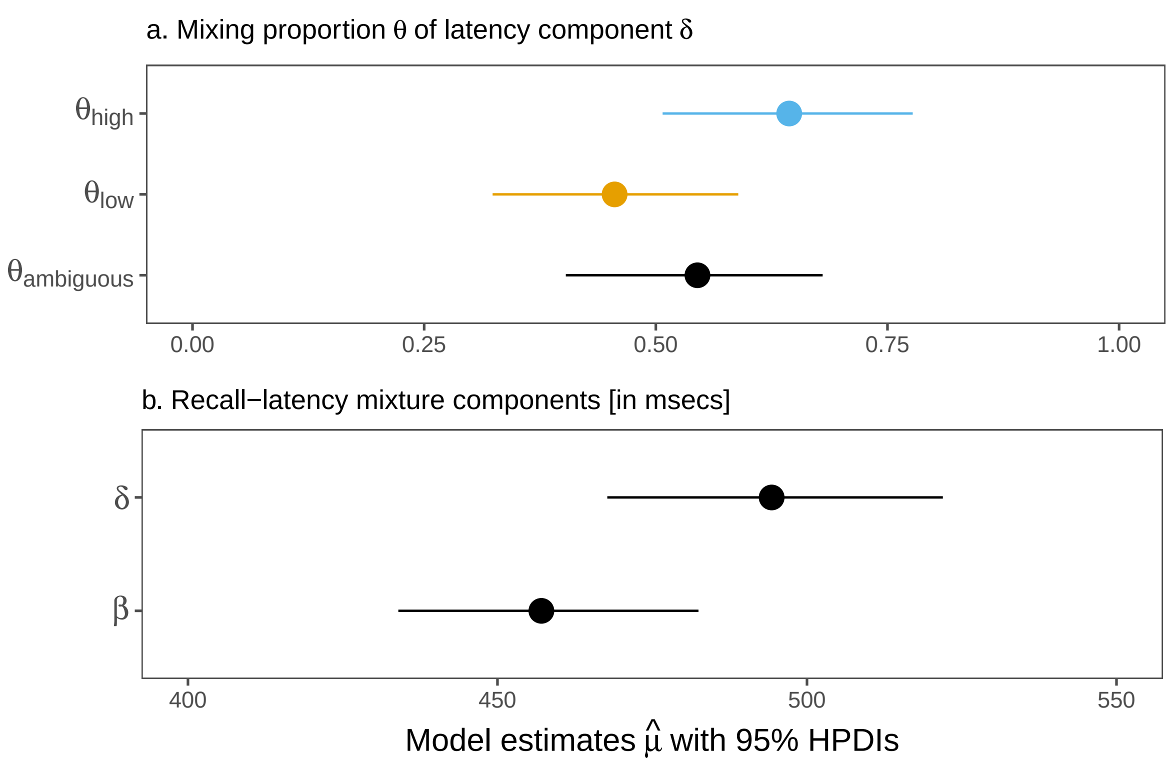

Fig. 1: Parameter estimates of the best fitting model with mixing proportions for each attachment type. Panel (a) shows the probability of observing long values θ by attachment type. Panel (b) shows the range of the latency distribution β and long latencies δ.