Computational models of retrieval processes in sentence comprehension in aphasia

Paula Lisson, Dorothea Pregla, Dario Paape, Bruno Nicenboim*, Nicole Stadie, Frank Burchert, Shravan Vasishth. University of Potsdam. *Tilburg University.
paula.lisson@uni-potsdam.de

Individuals with aphasia (IWA) are known to experience difficulties in the comprehension of non-canonical sentences (Caramazza & Zurif, 1976), especially when thematic roles can be reversed, such as in (1b).

(1) a. **Subject Relative (SR):** The girl who chased the mother hugged the boy.
b. **Object Relative (OR):** The girl who the mother chased hugged the boy.

There are several theories of language processing deficits in IWA that could explain these difficulties: delayed lexical access (Ferrill et al., 2012), slowed syntax (Burkhardt et al., 2008), resource reduction (Caplan, 2012), and intermittent deficiencies (Caplan et al., 2015), among others. Some of these theories have been computationally implemented using the Lewis & Vasishth (2005, henceforth LV05) cue-based retrieval model of sentence processing (Maetzig et al., 2018, Patil et al., 2016). However, there is another cue-based retrieval model that could explain the performance of IWA: the direct-access model (McElree, 2000, henceforth DA). These two models hold different assumptions: in LV05 the latency and retrieval probability of an item depend on its activation. In contrast, DA assumes that items' retrieval probability can differ, whereas retrieval times are constant. We implement LV05 and DA, and we analyze how their parameters can be linked to the above-mentioned theories in aphasia.

Methods. We selected a subset of the dataset in Caplan et al. (2015) that contains self-paced listening times and picture selection accuracies for subject and object relative clauses. We model aggregated listening times at the verb of the subordinate clause and the second noun phrase regions, and picture-selection accuracies from 33 IWA and 46 control participants. We follow Nicenboim & Vasishth (2018) and implement both models in the Bayesian framework: LV05 as a lognormal race of accumulators, and the DA as a mixture model. **The activation-based model:** In LV05, the element in memory whose activation is higher gets retrieved. In the race of accumulators, the rate of accumulation is equated to activation in LV05: the accumulator that wins the race corresponds to the item retrieved from memory. **The direct-access model:** The correct item is retrieved from memory with probability θ (see Equation 5). If the initial retrieval is incorrect, bakctracking is performed with probability P_b, and this leads to the retrieval of the correct item. If the initial retrieval is incorrect and there is no backtracking, a misretrieval is predicted. We implemented the models in Stan (Carpenter et al., 2017) and performed 10-fold cross-validation (Vehtari et al., 2015) to quantitatively compare their predictive performance.

Results. The activation-based model shows a better quantitative fit. However, the difference in predictive performance is not decisive (see Equation 7). The results of the activation-based model are compatible with the intermittent deficiencies theory, which claims that IWA have intermittent breakdowns in the parsing system: IWA are estimated to have a noisier rate of accumulation. In addition, the means of the accumulators are higher for IWAs relative to controls. This would be in line with the lexical access deficit or with slowed syntax. The results of the direct-access model are compatible with the resource reduction hypothesis and with intermittent deficiencies, since the model estimates that IWA do not backtrack as much as controls. This could indicate that the mechanism of backtracking is disrupted in IWA, or that it is intermittently disrupted. However, IWA are estimated to have a higher probability of initial correct retrieval in object relative clauses relative to controls, which is an unrealistic estimate (see Figure 2). Finally, IWA are also estimated to have higher noise in the retrieval mechanism. This is could be in line with intermittent deficiencies. Overall, both models can account for language processing deficits in aphasia, but some parameters of DA show unrealistic estimates for IWA.

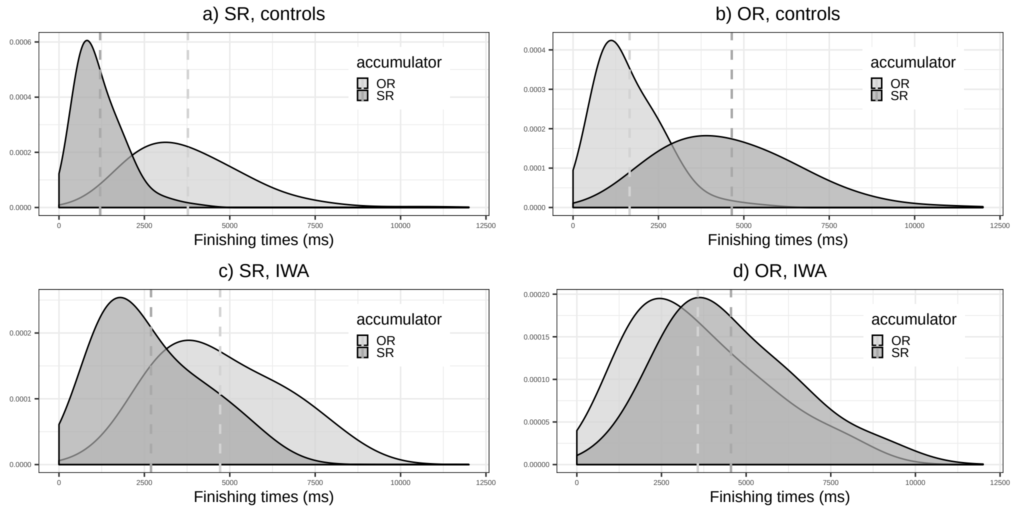

Figure 1: Activation-based model: Distribution of finishing times (FT) for each accumulator split by group and condition (see Equation 1 for the definition of *finishing times*). Panels (a) and (b) indicate that the mean FT for control participants are lower. Panels (b) and (d) indicate that the mean FT for the object relatives are higher relative to the subject relatives. The great overlap between the distributions in panel (d) suggests that the model estimates IWA to have more difficulties in object relatives.

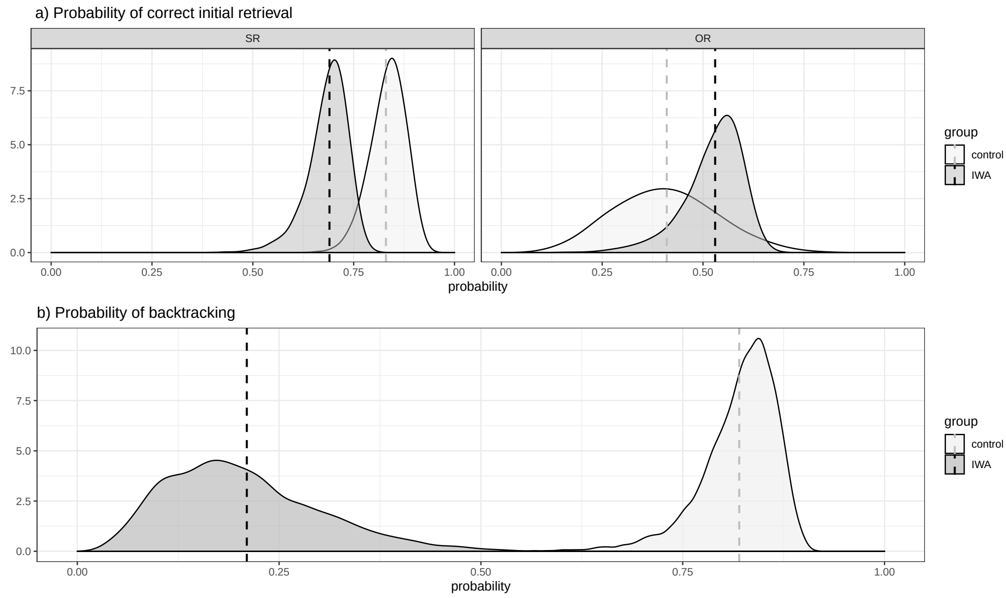

Figure 2: DA model: Panel (a) depicts the estimated probability of initial correct retrieval for both groups across conditions. Controls have a higher initial correct probability in subject relatives. However, IWAs are estimated to have a higher probability of correct initial retrieval in object relatives. This is surprising because IWAs are known to experience difficulties in object relatives. Panel (b) shows that controls are estimated to perform backtracking much more often than IWA.

References. Burkhardt et al., *Journal of Neurolinguistics* (2008). Caplan, in *Perspectives on aggramatism* (2012). Caplan et al., *Cognitive Neuropsychology* (2015). Carpenter et al., *JSS* (2017). Ferrill et al., *AJSLP* (2012). Lewis and Vasishth, *Cog Sci* (2005), Maetzig et al., *Topics in Cog Sci* (2018), McElree, *JPR* (2000). Nicenboim Vasishth, *JML* (2018). Patil et al., *Cog Sci* (2016). Vehtari et al., *Statistics and Computing* (2016).

Activation-based model. The race of accumulators is generated by sampling values from two lognormal distributions (1). These sampled values are the *finishing times* (FT) of each accumulator. For each observation, the accumulator with the faster FT stands for the winning interpretation (SR or OR), and its value is the estimated listening time (LT). The hierarchical structure is shown in (2),(3) and (4), where *u* and *w* are the by-subject and by-item adjustments.

$$\begin{aligned} \text{SR accumulator: } FT_{SR_i} &\sim lognormal(\mu_{SR}, \sigma) \\ \text{OR accumulator: } FT_{OR_i} &\sim lognormal(\mu_{OR}, \sigma) \\ LT_i &= min(FT_{SR_i}, FT_{OR_i}) \end{aligned} \tag{1}$$

$$\begin{aligned} \mu_{SR} = \alpha_1 + u_{\alpha_1} + w_{\alpha_1} + \\ (\beta_1 + w_{\beta_1}) \cdot group + (\beta_3 + u_{\beta_3}) \cdot rc_{type} + \\ \beta_5 \cdot group \cdot rc_{type} \end{aligned} \tag{2}$$

$$\begin{aligned} \mu_{OR} = \alpha_2 + u_{\alpha_2} + w_{\alpha_2} + \\ (\beta_2 + w_{\beta_2}) \cdot group + (\beta_4 + u_{\beta_4}) \cdot rc_{type} + \\ \beta_6 \cdot group \cdot rc_{type} \end{aligned} \tag{3}$$

$$\sigma = \sigma_0 + \beta_7 \cdot group \tag{4}$$

Direct-access model. The distribution of listening times for correct responses are a mixture of directly accessed retrievals, and initial failed retrievals followed by backtracking. The distribution of incorrect responses corresponds to initial failed retrievals without backtracking. The estimated listening times are sampled from two lognormal distributions, as shown in (5). The hierarchical adjustments embedded in θ (probability of initial correct retrieval), P_b (probability of backtracking), μ (mean listening time), δ (time needed for backtracking) and σ (noise) are shown in (6).

$$LT \sim \begin{cases} lognormal(\mu, \sigma), & \text{initial retrieval succeeds, probability } \theta \\ lognormal(\mu + \delta, \sigma), & \text{initial retrieval fails, backtracking, probability } (1-\theta) \cdot P_b \\ lognormal(\mu, \sigma), & \text{initial retrieval fails, no backtracking, probability } (1-\theta) \cdot (1-P_b) \end{cases} \tag{5}$$

$$\begin{aligned} \mu &= \mu_0 + u_{\mu 0} + w_{\mu 0} + \beta_1 \cdot group \\ \theta &= \alpha + u_\alpha + w_\alpha + (\beta_2 + u_{\beta_2}) \cdot rc_{type} + \\ &(\beta_3 + w_{\beta_3}) \cdot group + \beta_4 \cdot group \cdot rc_{type} \\ P_b &= \gamma + u_\gamma + \beta_5 \cdot group \\ \delta &= \delta_0 + \beta_6 \cdot group \\ \sigma &= \sigma_0 + \beta_7 \cdot group \end{aligned} \tag{6}$$

Model comparisons. 10-fold cross-validation was run for both models. The data was split in 10 balanced subsets, the models were run for each one of the subsets, and the predictive performance with respect to the other remaining 9 subsets was computed. This yields the *expected log point-wise predictive density* measure ($\widehat{elpd}$). The model with the smaller $\widehat{elpd}$ provides a better fit for the data. The $\widehat{elpd}$ values yielded a difference of 115 (SE = 69), suggesting that the activation-based model furnishes a better fit to the data, but the large SE indicates that this difference is not decisive.

$$elpd_{act} = -12515 \ (SE = 49) \quad elpd_{DA} = -12630 \ (SE = 52) \tag{7}$$