Is reanalysis selective when regressions are manually controlled?

Dario Paape & Shravan Vasishth (University of Potsdam)

To what extent is rereading in syntactically complex sentences under conscious control? It has been argued by, e.g., [1] that in cases of syntactic misanalysis, the problematic part of the sentence is selectively reread ("selective reanalysis"). However, it is not clear to what extent readers consciously decide to reread earlier material. It is possible that selective reanalysis is, to some extent, consciously triggered when the reader notices that misanalysis has occurred. Conscious awareness of garden-pathing may occur in extremely difficult examples (*The horse raced past the barn fell*) but may also stochastically arise in milder cases.

Due to the SARS-CoV-2 pandemic, it is currently difficult to conduct lab-based eye-tracking studies. We thus switched to a web-based, modified version of self-paced reading that allows rereading ("bidirectional SPR"): readers press the right arrow key to move forward through the sentence and the left arrow key to move backward ("manual regression"). Readers can also return to the beginning of the sentence or move directly to the comprehension question with specific keys. Sentence presentation is masked and non-cumulative. While this method has obvious drawbacks, such as the inability to skip words, and is quite dissimilar to eye tracking, it also has advantages: oculomotor noise is eliminated as a source of fixation errors and regressions require conscious deliberation.

We designed an experiment in German with two conditions (early versus late disambiguation) and two types on ambiguity: the NP/S coordination ambiguity [2] and the German SVO/OVS ambiguity [3] (see examples on page 2). We recruited 100 participants through Prolific, each of whom read 32 critical sentences plus 52 fillers. In addition to a baseline monetary compensation, participants earned bonus compensation by reading both quickly and accurately. Points were awarded based on the time taken to complete the study and the percentage of correct answers to comprehension questions. Detailed comprehension questions were asked after each trial.

Overall, rereading was quite common in our study: at least one manual regression occurred in about 51% of trials for critical sentences. We analyzed first-pass reading times and rereading times by region, analogously to eye-tracking studies. First-pass reading times showed a divergence between coordination and SVO/OVS sentences in the early disambiguation region (see Figure 1, left), in that coordination sentences showed an ambiguity slowdown (95% CrI: [6 ms, 65 ms]) while SVO/OVS sentences showed a disambiguation slowdown (CrI: [7 ms, 69 ms]). Both sentence types showed a garden-path effect in the late disambiguation region, both in first-pass reading times (CrI: [-2 ms, 35 ms]) and in rereading times (CrI: [12 ms, 127 ms]).

We used multidimensional scaling and model-based clustering to identify clusters of scanpaths across both sentence types [4]. Scanpaths clustered along two dimensions: amount of rereading and location of regressions (see Figure 1, right). Across conditions, participants who scored high on the speed/accuracy measure showed more rereading. Garden-pathing also led to higher values in this dimension. Correspondingly, garden-pathing decreased the probability of scanpath membership in cluster 2 (CrI: [-9%, 0%]), in which there are almost no regressions (see Figure 2), and increased the probability of membership in cluster 3 (CrI: [0%, 6%]), in which large portions of the sentence are reread. Membership in other, smaller clusters, such as cluster 5 with short regressions to the disambiguating region, showed no indication of a difference between conditions.

Our results suggest that manually-controlled rereading as a response to syntactic misanalysis is overall relatively unselective, though there was some increased focus on the disambiguating region. To the extent that these findings are transferable to eye tracking, it may be that conscious disruption of the reading flow results in a global rereading strategy, while more subtle disruptions result in more locally selective strategies.

Coordination, early disamb.		*The make-up artists powdered the singers . . .*		
. . . und	der Schauspieler (NOM)	mit blauen Augen	wurde	parfümiert, . . .
. . . *and*	*the actor*	*with blue eyes*	*was*	*perfumed . . .*
Coordination, late disamb. (garden path)		*The make-up artists powdered the singers . . .*		
. . . und	die Schauspielerin (NOM/ACC)	mit blauen Augen	wurde	parfümiert . . .
SVO/OVS, early disamb.				
. . .	den Forscher (ACC)	im Dschungel	stachen	die Moskitos, . . .
. . .	*the researcher.SG*	*in.the jungle*	*bit.PL*	*the mosquitoes . . .*
SVO/OVS, late disamb. (garden path)				
. . .	die Forscherin (NOM/ACC)	im Dschungel	stachen	die Moskitos, . . .

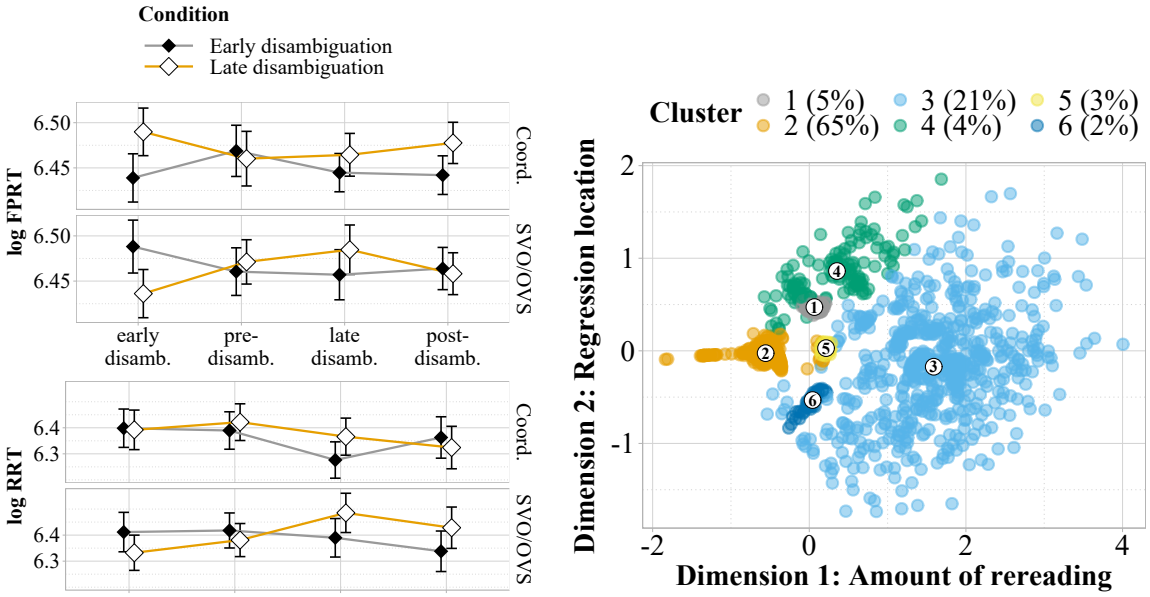

Figure 1. Left: First-pass reading times and rereading times by region of interest, residualized against region length. Error bars show 95% confidence intervals. Right: Location of clusters in scanpath space. "Regression location" refers to the region of interest in the sentence where regressions occur.

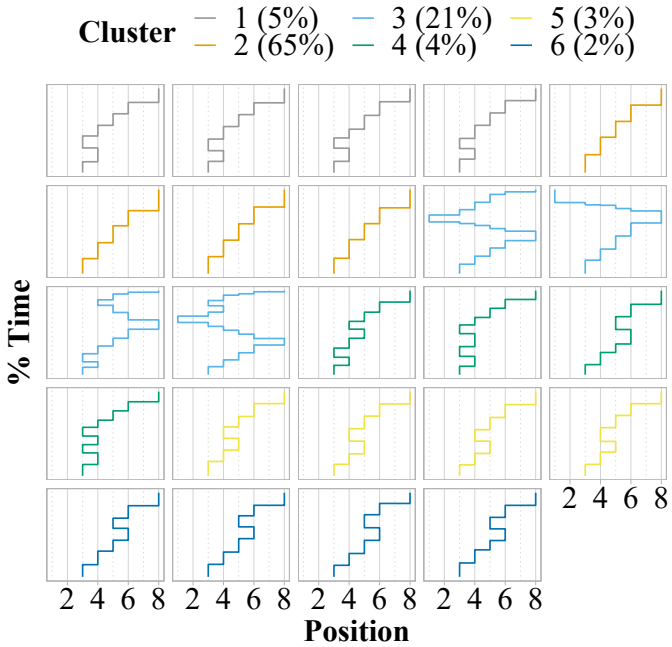

Figure 2. Typical scanpaths belonging to each of the clusters. Scanpaths were analyzed starting at the first visit to the early disambiguation region (3). Only data from regions starting with the early disambiguation region (3) and ending with the post-disambiguation region (6), as well as visits to the start (1) and end (8) regions, were included in the analysis.

References. [1] **Frazier & Rayner** (1982), Cognitive Psychol. [2] **Frazier** (1987), Nat Lang Linguist Th. [3] **Hemforth** (1993), Kognitives Parsing: Repräsentation und Verarbeitung sprachlichen Wissens. [4] **von der Malsburg & Vasishth** (2011), J Mem Lang.