The CoFI Reader: A Continuous Flow of Information approach to modeling reading

Bruno Nicenboim[1]

[1] Cognitive Science & AI, Tilburg University, the Netherlands

b.nicenboim@tilburguniversity.edu

Cognitive processing is classically depicted as a sequence of discrete, non-overlapping stages, each concluding before the next begins [1, 2]. Although cascading alternatives were proposed decades ago [3], most mechanistic psycholinguistic theories and computational models rely on discrete stages (explicitly or implicitly).

I introduce the *CoFI reader*, a novel cognitive model of (self-paced) reading grounded in a continuous-flow-of-information framework [4]. The CoFI reader assumes that visual, lexical, and syntactic/semantic information are processed simultaneously across hierarchical layers. Information at each layer continuously accumulates from the moment a word appears, and partial outputs from each stage immediately become available to subsequent stages. Specifically, visual processing (v) begins shortly after word onset, rapidly providing spatial information to a low-level lexical stage (f), which then continuously transmits partial outputs onward to the high-level syntactic/semantic stage (g). The CoFI reader incorporates a stochastic timer (h) governing when readers press a button to reveal the next word. Information from the higher-level processing stage continuously feeds into this timer, dynamically influencing the speed at which it reaches a threshold that triggers the motor response. Processing difficulty at any stage, due to the characteristics of the current or preceding words, inhibits this timer and directly prolongs reading times. Additionally, because each high-level processing stage integrates partial information from both the current and preceding words, processing difficulty cascades forward, producing spillover effects across multiple subsequent words (see Figure 1).

The model addresses two critical issues that complicate interpreting word-level reading times as straightforward indices of processing difficulty, previously recognized in cognitive psychology [5]. (i) Observed latencies at individual words (whether fixation durations or self-paced reaction times) are too short to accommodate all serial processes that should occur if stages were strictly non-overlapping [6]. (ii) Spillover effects extending beyond the current word are pervasive in both eye-tracking and self-paced reading paradigms. CoFI explains both phenomena without relying on parafoveal preview mechanisms, unlike oculomotor control models such as E-Z Reader or SWIFT [7, 8].

The framework is deliberately flexible: any variable hypothesized to influence processing can be mapped to the accumulation rate of the relevant layer. In the implementation presented here, word frequency modulates the low-level rate, and contextual surprisal modulates the high-level rate; both rates follow hierarchical log-normal distributions, while the timer threshold follows a weakly stationary AR(1) process.

The model was fitted in a fully hierarchical Bayesian framework (using Stan) to a subset of a self-paced reading corpus [9]. CoFI Reader captures the full sequence of word-level reading times to a reasonable extent, despite somewhat overestimating their variability (Figure 3), as well as first- and second-order spillover effects (Figure 2).

Together, these results demonstrate that a mechanistic model with a continuous-flow architecture can jointly explain rapid reading times and spillover in the absence of parafoveal information, offering a principled computational alternative to stage-based models and a coherent link between cognitive processing and behavior.
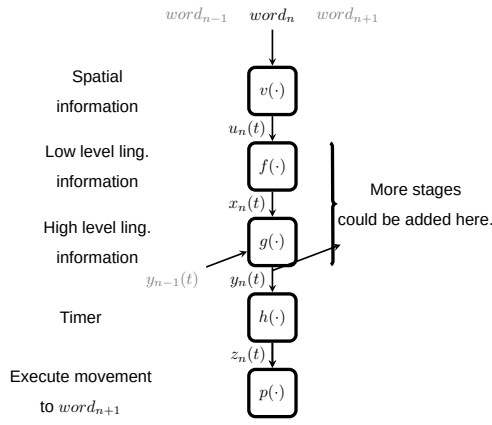
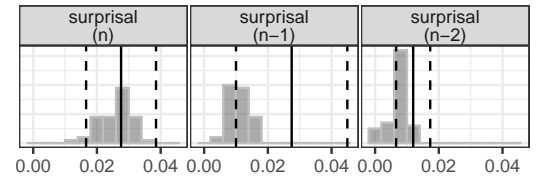Figure 1: Schematic of the CoFI model.

Figure 2: Distribution of (posterior predicted) simulated effects of surprisal and two spillovers. The empirical effects (with their SE) are represented by black lines.
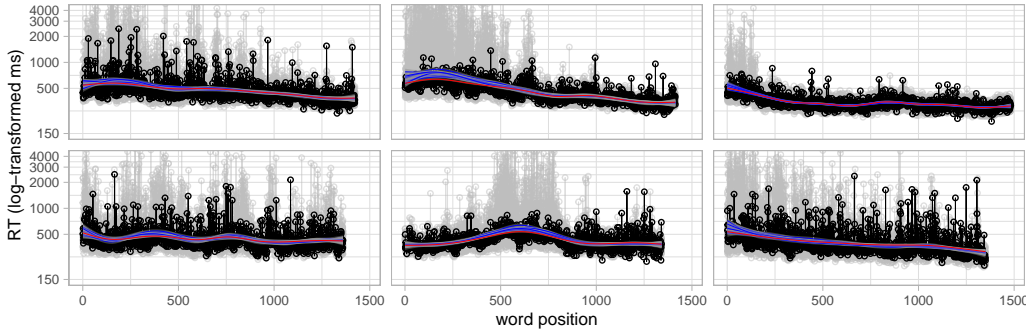

Figure 3: Each panel shows the observed RTs for a subject (black points, with smoothed conditional means in red) together with the model's posterior predictive distributions based on the same data (grey points, with smoothed conditional means in blue).

References

[1] Donders, F. C. (1868–1969). On the speed of mental processes. *Acta Psychologica*, *30*, 412–431.

[2] Sternberg, S. (1969). The discovery of processing stages: Extensions of Donders' method. *Acta psychologica*, *30*, 276–315.

[3] Mcclelland, J. (1979). On the time relations of mental processes: An examination of systems of processes in cascade. *Psychological Review*, *86*, 287–330.

[4] Townsend, J. T., & Wenger, M. J. (2021). A beginning quantitative taxonomy of cognitive activation systems and application to continuous flow processes. *Attention, Perception, & Psychophysics*, *83*(2), 748–762.

[5] Kliegl, R., Nuthmann, A., & Engbert, R. (2006). Tracking the mind during reading: The influence of past, present, and future words on fixation durations. *Journal of experimental psychology: General*, *135*(1), 12.

[6] Reichle, E. D., & Reingold, E. (2013). Neurophysiological constraints on the eye-mind link. *Frontiers in Human Neuroscience*, *7*.

[7] Engbert, R., Nuthmann, A., Richter, E. M., & Kliegl, R. (2005). SWIFT: A Dynamical Model of Saccade Generation During Reading. *Psychological Review*, *112*(4), 777–813.

[8] Reichle, E. D., Warren, T., & McConnell, K. (2009). Using E-Z Reader to model the effects of higher level language processing on eye movements during reading. *Psychonomic Bulletin & Review*, *16*(1), 1–21.

[9] Frank, S. L., Fernandez Monsalve, I., Thompson, R. L., & Vigliocco, G. (2013). Reading time data for evaluating broad-coverage models of English sentence processing. *Behavior Research Methods*, *45*(4), 1182–1190.