**A rational model of word skipping in reading: ideal integration of visual and linguistic information**

Yunyan Duan (Northwestern University) & Klinton Bicknell (Northwestern University; Duolingo)
yduan@u.northwestern.edu

Readers combine visual and linguistic information to recognize words and comprehend sentences. While many studies have examined effects of these two information sources in reading, exactly *how* they are combined is relatively unexplored. If readers combine them optimally, the results should be a complex function of the particular word, linguistic context, and the visual information. For example, visual information about only the beginning of some words is enough for identification, e.g., seeing that the initial letters `xyl` of the word `xylophone`[1]. Similarly, in certain contexts, a reader only needs to see a few of the initial letters of a word to be confident in its identification, such as in 'The children went outside to pl…'. However, whether readers do in fact combine visual and linguistic information optimally in this way remains unclear.

One promising way to study this process is with word skipping. It is generally hypothesized that when a reader intentionally skips a word, i.e., moves their eyes past it without fixating it, it is because the reader was highly confident in its identity. Crucially, this (implicit) decision about whether to skip the word is made when the reader is fixating a prior word, and thus when the reader only has high quality visual information about some of the word's initial letters (if any) but not yet high quality visual information about the whole word. The amount of visual information the reader has at this time is a function of the *launch site,* the distance from that fixation position to the beginning of the word. As such, studying skipping decisions made for a variety of words in a variety of contexts with a variety of launch sites provides an ideal testbed for studying the complex interactions of context and visual information in word identification.

**Method.** To test whether readers display signatures of optimal integration across these contexts, we build a computational implementation of an ideal-integration model predicting identification confidence for each skipping decision. We show that these model predictions explain significant variance in human skipping rates when added to a strong baseline model.

**Baseline model**. We analyzed first-pass skipping in the Dundee corpus with a generalized additive mixed-effects regression model (GAMM) predicting words' skipping from variables shown to influence skipping, including word length, launch site, word frequency, surprisal (5-gram), and contextual constraint measured by entropy. We also controlled for previous word's properties such as word length and frequency (see Table 1). Crucially, this generalized additive model allowed for arbitrary non-linear effects of each of these variables, providing a strong baseline.

**Simulation**. We implemented a rational model (Figure 1) with Bayesian belief updating, using linguistic knowledge (frequency and 5-grams) as prior and visual input as likelihood [2], with the visual information sampling process computationally simplified as random walk in a multidimensional Gaussian distribution. For each potential word to be skipped, the model receives noisy visual information about it conditional on launch site and full information about the context. It computes a posterior distribution on the word, and we then extract the entropy of this distribution (postH) and add it to our baseline. The visual information in this model has two parameters: overall visual input quality and the width of acuity function. We used ten/six sets of parameter pairs for the models with frequency/5-gram priors; these parameters were chosen to be reasonable values, respecting the trade-off between width of the acuity function and its overall quality.

**Results.** Results showed that in every combination of free parameters postH had a significant effect over and above the baseline model in predicting skipping (Table 2). This result suggests that the rational model captured complex interactions between visual input and linguistic knowledge that human readers utilized in making skipping decisions.

**Conclusion.** Overall, these results paint a picture of word identification in reading – at least in word skipping decisions – as resulting from optimal integration of linguistic information and the particular visual information obtained.

Table 1. GAMM results of baseline model.

| | $\chi^2$ | $p$-value |
|---|---|---|
| word length | 6026.25 | < 2e-16 *** |
| launch site | 9123.73 | < 2e-16 *** |
| frequency | 527.94 | < 2e-16 *** |
| 5-gram surprisal | 38.40 | 1.01e-06 *** |
| context entropy | 71.16 | 8.28e-11 *** |
| word length × frequency | 89.06 | 7.73e-16 *** |
| launch × frequency | 36.09 | 2.85e-05 *** |
| launch × surprisal | 29.39 | 1.13e-04 *** |
| launch × entropy | 66.82 | 2.24e-11 *** |
| (previous word's properties) | … | … |

Figure 1. Bayesian model of skipping.

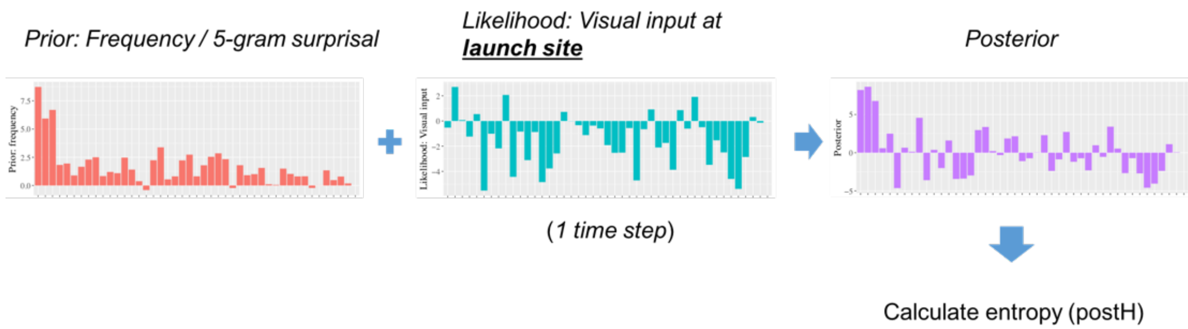


Table 2. Model improvement after adding entropy of a rational model's posterior distribution. $\sigma$ denotes the deviation of a visual acuity function (normal distribution); a large $\sigma$ value means flat acuity function and good preview quality from a far launch site. $\Lambda$ denotes the overall visual input quality; a large $\Lambda$ value means good visual acuity and good preview quality from a far launch site.

| | Prior: Frequency | | Prior: 5-gram surprisal | |
|---|---|---|---|---|
| $(\sigma, \Lambda)$ | $\chi^2$ | $p$-value | $\chi^2$ | $p$-value |
| (0.5, 5) | 40.14 | 1.67e-06 *** | | |
| (0.5, 15) | 62.15 | 4.11e-10 *** | 20.45 | 0.004 ** |
| (1, 15) | 37.02 | 6.79e-07 *** | | |
| (1, 30) | 38.27 | 3.22e-07 *** | 16.42 | 0.002 ** |
| (1, 40) | 30.68 | 4.85e-06 *** | 19.63 | 4.76e-04 *** |
| (2, 40) | 12.90 | 0.002 ** | 18.16 | 4.68e-04 *** |
| (4, 4) | 100.25 | < 2e-16 *** | | |
| (4, 30) | 36.78 | 2.07e-05 *** | 23.03 | 2.01e-04 *** |
| (5, 0.2) | 48.58 | 4.48e-09 *** | | |
| (5, 3) | 91.12 | 9.92e-16 *** | 19.25 | 6.17e-04 *** |


References

[1] Hyönä, J., Niemi, P., & Underwood, G. (1989). Reading long words embedded in sentences: Informativeness of word halves affects eye movements. *Journal of Experimental Psychology: Human Perception and Performance*, 15(1), 142.

[2] Bicknell, K., & Levy, R. (2010, July). A rational model of eye movement control in reading. In *Proceedings of the 48th annual meeting of the Association for Computational Linguistics* (pp. 1168-1178). Association for Computational Linguistics.