The art of the state: Mixed effect regression modeling in the Visual World

Monitoring eye movements in the "visual world" has become a paradigm of choice for examining real-time spoken language processing. The most informative data, proportions of fixations to objects over time present challenges for standard analyses (ANOVA), which assume continuous data and cannot naturally model time. Advances in regression modeling, e.g., mixed (empirical) logit models and curve fitting [1,2], provide powerful new tools for visual world data, while avoiding some of the problematic assumptions of ANOVA. We provide a solution to a common challenge in analyzing eye-tracking (and comparable time series) data that persists in these new models (contrary to [1]) and is at least as likely to lead to spurious results as inappropriate use of ANOVA.

At the onset of a region of interest, fixations on the target picture will often differ between conditions (e.g., because of differences in anticipatory eye movements). Analyses of rate effects (differences in speed to converge on the target) must take these baseline differences into account. Using mixed effect models, Barr proposes using an intercept term.

We show that this approach is flawed: intercept terms do *not* account for baseline differences and can lead to spurious positive results and mask real effects. We show that contingent analyses, where trials with prior looks to the target referent are excluded from the analysis are more appropriate.

Consider a two-condition experiment. Participants are looking at the target 40% of the time before the target is named in A compared to 10% in B. Assume A and B do *not* differ: processing the target name provides evidence that would cause a saccade to the target 50% of the time in each condition. However, the participant is in one of two states: in the off-target state, but not in the on-target state, s/he can make a saccade to the target. So in A, looks to the target would rise from 40% to 70% [=40% + (100% - 40%) * 50%] compared to 10% to 55% in B, *in*correctly suggesting a larger rate effect for B (even more so in logit space). The correct analyses must therefore evaluate looks with respect to the participant's state. A contingent analysis on trials where the participant is not already looking at the target does just that, thereby avoiding the spurious rate effect.

Indeed, using real data that should not contain a rate effect, a non-contingent analysis (Barr's approach) returns a *spurious* rate effect. The contingent analysis correctly returns no significant rate effect (both analyses used weighted mixed effect regression against empirical logit, as in [1]; collinearity is not an issue in either model). Other effect sizes remain equally strong.

The major challenge to contingent analyses is a non-arbitrary criterion for excluding data from the analysis. We discuss several solutions to this challenge, including one that integrates prior looks to the target ("the baseline") into the mixed effect model. We conclude with a brief discussion of a multi-state multilevel multinomial model currently under development that incorporates state-modeling into the analysis of eye-tracking and other time series data.

[1] Barr, D.J. (in press). Analyzing 'visual world' eyetracking data using multilevel logistic regression. *Journal of Memory and Language*

[2] Mirman, D., Dixon, J.A. & Magnuson, J.S. (in press). Statistical and computational models of the visual world paradigm: Growth curves and individual differences. *Journal of Memory and Language*