Gamma Function Modeling of Visual World Eye-Tracking Data

Emily Atkinson, Akira Omaki, & Colin Wilson

Contact: atkinson@cogsci.jhu.edu

Johns Hopkins University

AMLaP XX Architectures and Mechanisms for Language Processing

Introduction

Goal: To model fixation patterns and individual differences in the Visual World using a simple function with links to cognitive processing: time-delayed gamma functions sensitive to parser state (a function drawn from models of BOLD responses).

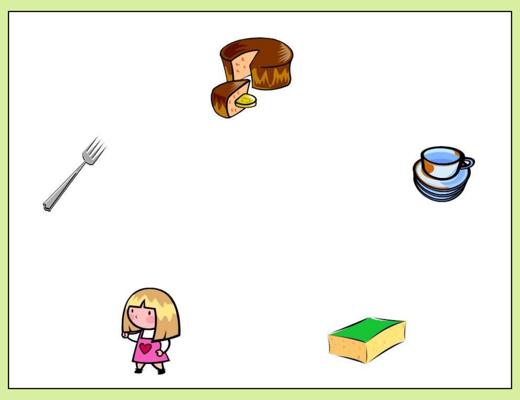
Previous VWP models: TRACE [1, 2], SRNs [3], dynamical systems [4], HMMs [5], and Growth Curve Analysis [6], against which we compare our model.

Proposal: Model time-evolving fixation patterns for multiple objects using a time-delayed unimodal continuous function (e.g., gamma).

1. Processing of each syntactic region of interest 'activates' one object
2. Each region described by a gamma distribution; probabilities derived by normalization
3. Log likelihood is the (log) probability of the model as applied to individual trial data

Visual World Paradigm Data (Atkinson et al., 2013)

Display: The subject & 2 events, each with an associated object & instrument



Critical items: Temporarily ambiguous wh-questions following a story

(1) Can you tell me what Emily was eating the cake with ___?

Predicted Eye Movements

Anticipatory fixations on the relevant object during the verb region

- Presence of a verb predicts compatible object [7]; presence of a *wh*-filler predicts earliest possible integration location (active gap-filling [8-11])

Anticipatory fixations on the relevant instrument during the object / preposition region

Fixation on associated picture as referent is processed

Gamma Model

Syntactic regions of interest are used to predict fixations. Fixations during each region are modeled by region and by time bin (30ms) with the following gamma function (inspired by models of hemodynamic brain responses [12]):

$$p_i(t) = \begin{cases} a\dfrac{[t-(T_0+T_L)]^{n-1}}{\lambda^n(n-1)!}e^{\frac{-t-(T_0+T_L)}{\lambda}} & \text{for } t > (T_0+T_L) \\ 0 & \text{for } t \leq (T_0+T_L) \end{cases}$$

- $\boldsymbol{a}$ = amplitude; $\boldsymbol{n}$ = shape of gamma; $\boldsymbol{\lambda}$ = rate of gamma
- $\boldsymbol{T_0}$ = onset of syntactic region, constants determined by stimuli
- $\boldsymbol{T_L}$ = lag between word onset & behavioral response, set to 180ms
  - represents time to plan & execute saccades [13]

Distributions that activate the same item are summed.

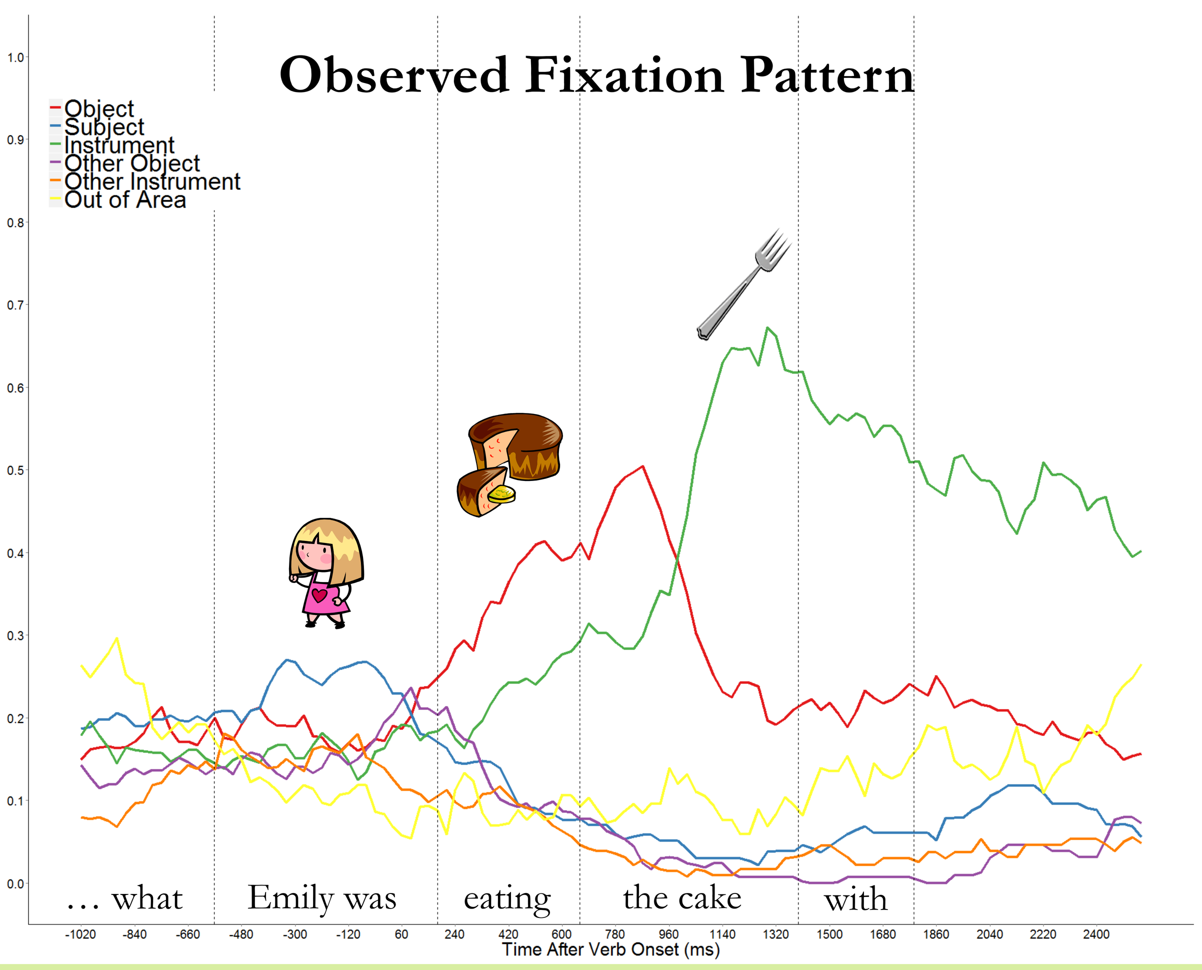
Normalization

By-region values were converted to probabilities by bin using the Luce choice rule with a temperature parameter [14].

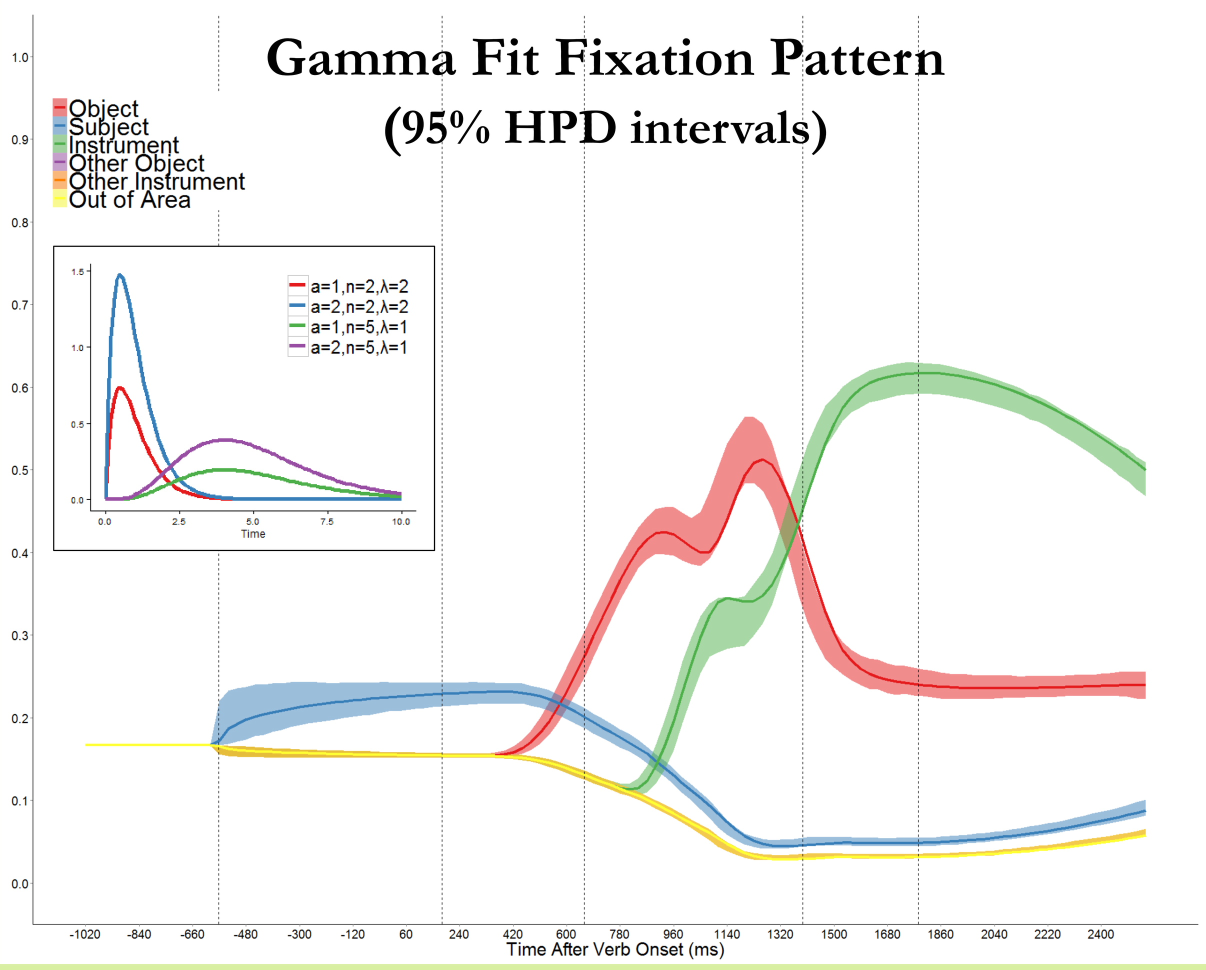
Log Likelihood

Probabilities applied to *individual trial data* according to the multinomial distribution. Parameters were sampled 1000 times with rStan [15].

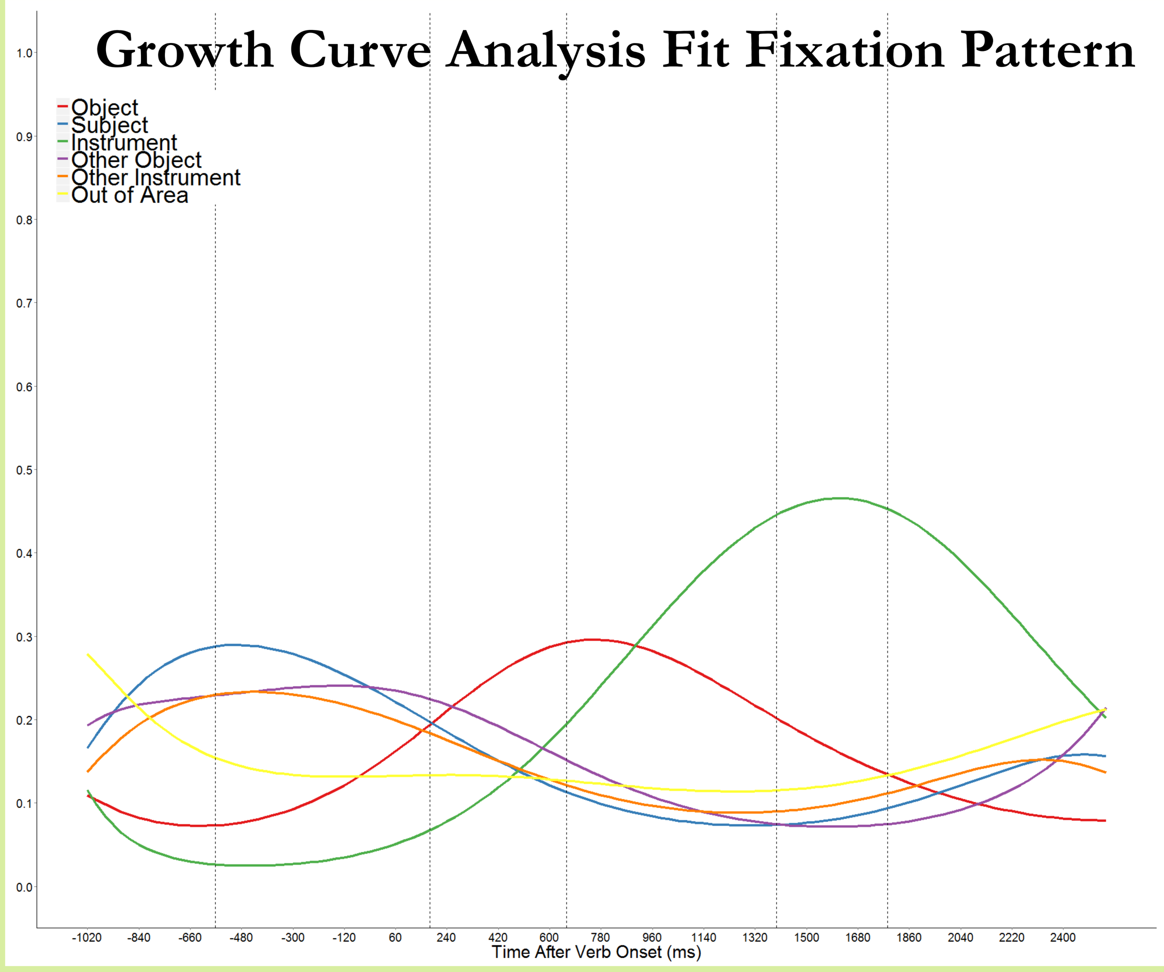
Results



Observed Fixation Pattern



Gamma Fit Fixation Pattern (95% HPD intervals)



Growth Curve Analysis Fit Fixation Pattern

Method	# of Parameters	Log Likelihood
Gamma Model: *1000 samples*	13	-19556.6 ± 2.6 (mean ± SD)
Growth Curve Analysis: *quartic polynomials*	29	-21494.2 (optimum)
Hierarchical Gamma Model: *1000 samples*	13/participant + 15 hyper-parameters	-17870.5 ± 17.8 (mean ± SD)

Hierarchical Model: Individual Differences

Gamma parameters & temperature vary by participants (N = 27, 5 trials each)

$a_{ij} \sim \mathcal{N}(\mu_{a_i}, \sigma_a)$ for participant j & region i

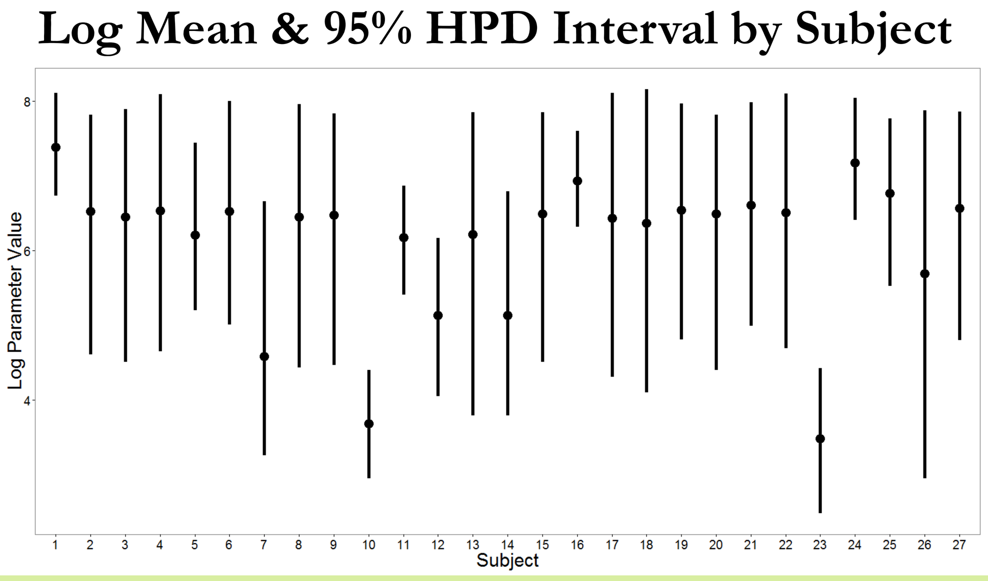
$\mu_{a_i} \sim \mathcal{N}(\mu_A, \sigma_A)$ for region i

$\sigma_a \sim G(s_A, r_A)$ for all regions

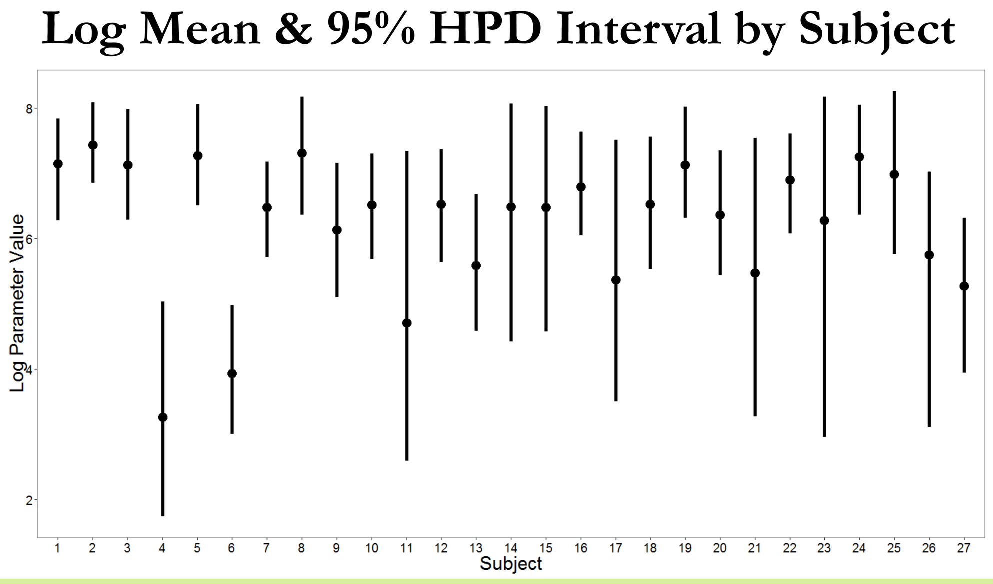
Similar design for other gamma parameters

$temp_j \sim \mathcal{N}(\mu_T, \sigma_T)$ for participant j

Amplitude: Subject
Log Mean & 95% HPD Interval by Subject



Amplitude: Gap Prediction at Verb
Log Mean & 95% HPD Interval by Subject



- Differences in **amplitude** in the **verb region** indicate varying degrees of active-gap filling (i.e., prediction of an object gap) *OR* action (i.e., fixation) based on that prediction.
- Differences in **amplitude** in the **subject region** indicate amount of interest in the subject as it is named.
- Comparison: subjects do not necessarily have consistent amplitudes across regions (R^2=0.3, p>0.1)

Interpretation & Discussion

Gamma model connects processing of syntactic regions to fixations

- Assumption that fixation distributions are well-approximated by a particular function
- Data reduced to only 13 parameters that capture qualitative patterns (~16,000 total binned data points)
- Greater likelihood than growth curves with less parameters

Relation of gamma distribution to parser?

- Inherent to parser: Gamma distribution describes time evolution of word/object activation during parsing
- Inherent to action (eye movement): Aggregate action on the parser output (i.e., where & when to look) results in gamma-shaped activations

Remaining Issues & Further Work

- Syntactic regions are currently hard-coded into the model. We would like to incorporate an explicit parsing model from which the fixation distributions observed in VWP arise.
- Gamma functions best fit unimodal fixation patterns, but data from other conditions in the same study are multimodal/cyclic.
- Allow time lag parameter (T_L) to vary across participants to model differences in saccade execution
- Apply to additional populations (e.g., children [8]) & study designs to test generalizability of method.

References

[1] Allopenna, P. D., Magnuson, J. S., & Tanenhaus, M. K. (1998). Tracking the time course of spoken word recognition using eye movements. *JML*, *38*, 419-439.
[2] Spivey, M. J. (2007). *The Continuity of Mind*. New York: Oxford Press.
[3] Mayberry, M., Crocker, M., & Knoeferle, P. (2005). A Connectionist Model of sentence comprehension in visual worlds. *CogSci Proceedings*, 1437-1442.
[4] Kukona, A. & Tabor, W. (2011). Impulse processing: A dynamical systems model of incremental eye movements in the Visual World Paradigm. *Cognitive Science*, *35*, 1009-1051.
[5] Dziemianko, M. (2013). Modeling eye movements and visual attention in synchronous visual and linguistic processing. Doctoral dissertation, University of Edinburgh.
[6] Mirman, D., Dixon, J.A., Magnuson, J.S. (2008). Statistical and computational models of the visual world paradigm: Growth curves and individual differences. *JML*, *59*, 475-494.
[7] Altmann, G. T.M. & Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition*, *73*, 247-264.
[8] Stowe, L.A. (1986). Parsing WH-constructions: Evidence for on-line gap location. *Language and cognitive processes*, *1*(3), 227-245.
[9] Traxler, M.J. & Pickering, M.J. (1996). Plausibility and the processing of unbounded dependencies: An eye-tracking study. *JML*, *35*, 454-475.
[10] Atkinson, E., Simeon, K., & Omaki, A. (2013). The time course of filler-gap dependency processing in the developing parser. Poster presented at CUNY. Colombia, SC.
[11] Sussman, R.S. & Sedivy, J.C. (2003). The time-course of processing syntactic dependencies: Evidence from eye movements. *LCP*, *18*(2), 143-163.
[12] Ashby, F.G., Waldschmidt, J.G. (2008). Fitting computational models to fMRI data. *Behavioral Research Methods*, *40*(3), 713-721.
[13] Matin, E., Shao, K.C., & Boff, K.R. (1993). Saccadic overhead: Information-processing time with and without saccades. *Perception & Psychophysics*, *53*(4), 372-380.
[14] Luce, R.D. (1959). *Individual Choice Behavior*. New York: Wiley.
[15] Stan Development Team (2014). Rstan: the R interface to Stan, Version 2.4. URL http://mc-stan.org/rstan.html.