A Joint Modeling Approach for Reaction Time and Accuracy in Psycholinguistic Experiments

Tom Loeys, Yves Rosseel & Kristof Baten
Department of Data Analysis - Ghent University
Tom.Loeys@UGent.be

BSS 2010, Spa

Psycholinguistic Experiment

Homograph recognition modulated by categorial overlap?

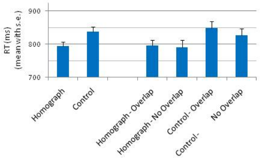
Condition	Target sentence (homograph / control word)
Overlap	She looked up and there seemed to be an ANGEL / ALIEN.
No Overlap	He told me he thinks this news is very BIG / SAD.

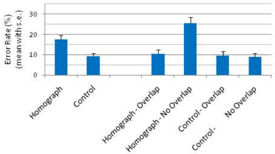
96 experimental trials:

- 16 homographs and 16 controls
- 16 filler sentences and 48 non-words

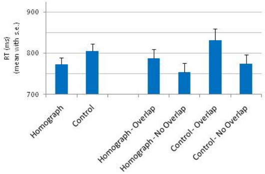
- L2-lexical decision task
- homographs and their controls in final position of low-constraint English sentences
- presented through serial visual presentation
- Study 1:
  32 highly proficient Dutch-English bilinguals
- Study 2:
  31 intermediate proficient Dutch-English bilinguals

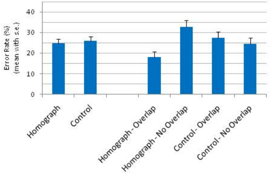
HIGH PROFICIENCY GROUP





INTERMEDIATE PROFICIENCY GROUP





Notation

- Subjects $i = 1, \ldots, N$ and items by $j = 1, \ldots, K$
- Two sources of information:

  (1) the reaction time $T_i = (T_{i1}, \ldots, T_{iK})$

  (2) the response accuracy $Y_i = (Y_{i1}, \ldots, Y_{iK})$

- X_{1i} a subject specific characteristic
- X_{2j} an item specific characteristic

Analysis of reaction times

Common approach: F1- and F2-statistics

F1: ANOVA on the mean per participant per item condition
$\Rightarrow$ ignores variability due to items

F2: ANOVA on the mean values per item

Are both significant?

A better model for the reaction time

A linear mixed model with crossed random effects:

$$T_{ij} = \alpha_0 + \alpha_1 X_{1i} + \alpha_2 X_{2j} + \theta_{1i} + \tau_{1j} + \epsilon_{ij}$$

θ_{1i}: speed (large values=slower responders)

τ_{1j}: time intensity (large values=larger reaction times)

with $\epsilon_{ij} \sim N(0, \sigma^2)$

Further improvements?

(i) participant and item variability

(ii) increasing variance with means

(iii) a non-zero minimum

$\Rightarrow$ shifted 3-parameter Weibull distribution for T_{ij}

$$f(t \mid \psi, \lambda, \gamma) = \lambda\gamma(t-\psi)^{\gamma-1} \exp\left[-\lambda(t-\psi)^{\gamma}\right], \quad t \geq \psi$$

with participant specific shifts ψ_i and shapes γ_i
with participant and item specific rate parameter λ_{ij}

$$\log \lambda_{ij} = -\tilde{\alpha}_0 - \tilde{\alpha}_1 x_{1i} - \tilde{\alpha}_2 x_{2j} - \tilde{\theta}_{1i} - \tilde{\tau}_{1j}$$

A model for the accuracy

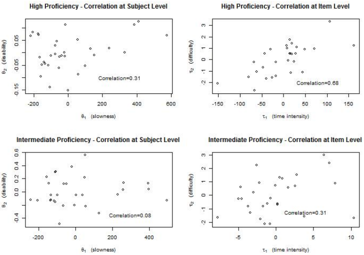
A mixed effects logistic regression model
for the probability of incorrect response:

$$\mathrm{logit}(P(Y_{ij} = 1)) = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2j} + \theta_{2i} + \tau_{2j}$$

θ_{2i}: participant's ability

τ_{2j}: item's difficulty

Inspecting the correlations from the separate models...



based on normal distribution for RT

Joint modeling approach

$$\begin{cases} T_{ij} = \alpha_0 + \alpha_1 X_{1i} + \alpha_2 X_{2j} + \theta_{1i} + \tau_{1j} + \epsilon_{ij} \\ \text{logit}(P(Y_{ij} = 1)) = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2j} + \theta_{2i} + \tau_{2j} \end{cases}$$

- subjects parameters θ_{1i} and θ_{2i} follow a bivariate normal distribution with variance-covariance

$$\Sigma_S = \begin{pmatrix} \sigma^2_{\theta 1} & \rho_\theta \sigma_{\theta 1} \sigma_{\theta 2} \\ \rho_\theta \sigma_{\theta 1} \sigma_{\theta 2} & \sigma^2_{\theta 2} \end{pmatrix}$$

$\Rightarrow \rho_\theta$: correlation between ‘speed’ and ‘ability’

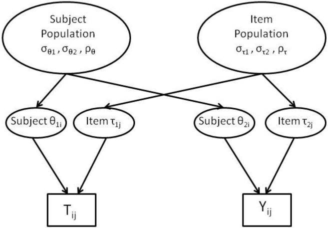
- item parameters τ_{1i} and τ_{2i} follow a bivariate normal distribution with variance-covariance

$$\Sigma_I = \begin{pmatrix} \sigma^2_{\tau 1} & \rho_\tau \sigma_{\tau 1} \sigma_{\tau 2} \\ \rho_\tau \sigma_{\tau 1} \sigma_{\tau 2} & \sigma^2_{\tau 2} \end{pmatrix}$$

$\Rightarrow \rho_\tau$ correlation between 'intensity' and 'difficulty'

- similarly in case of shifted Weibull for reaction time distribution

$$\begin{cases} \log \lambda_{ij} = -\tilde{\alpha}_0 - \tilde{\alpha}_1 x_{1i} - \tilde{\alpha}_2 x_{2j} - \tilde{\theta}_{1i} - \tilde{\tau}_{1j} \\ \text{logit}(P(Y_{ij} = 1)) = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2j} + \theta_{2i} + \tau_{2j} \end{cases}$$


Subject
Population
σθ1 , σθ2 , ρθ
Item
Population
στ1 , στ2 , ρτ
Subject θ1i
Item τ1j
Subject θ2i
Item τ2j
Tij
Yij

Estimation: a Bayesian approach

- independent normal distributions with zero mean and large variances for the fixed effect parameters
- the inverse Wishart distribution for the covariance matrix of a multivariate normal distributions,

$$\Sigma_S \sim \mathrm{Inverse-Wishart}(\Sigma_{S0}^{-1}, \kappa_{S0}) \text{ and } \Sigma_I \sim \mathrm{Inverse-Wishart}(\Sigma_{I0}^{-1}, \kappa_{I0})$$

$\Rightarrow$ provides info about the scale of random effects

$$\begin{aligned} \kappa &= 2 : \text{ least informative} \\ \Sigma^{-1} &= \omega \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \text{ informative through } \omega. \end{aligned}$$

- $\Gamma(\eta_1, \eta_2)$ for measurement precision $1/\sigma$

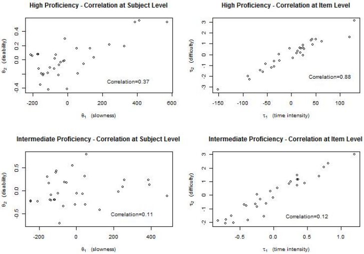
- in case of shifted Weibull distribution:

  γ_i: $\Gamma(\zeta_1, \zeta_2)$-distribution

  ψ_i: flat uniform distribution

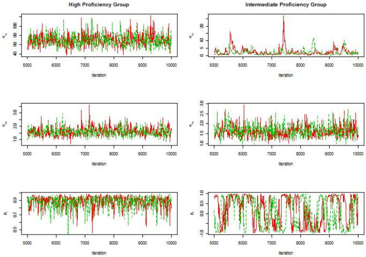
- implemented in WINBUGS
  (requires ‘zeros trick’ based on Poisson for shifted Weibull)

Joint modeling approach: estimated correlations

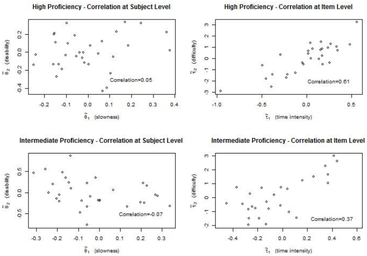


based on normal distribution for RT

Convergence of covariance parameters at the item level



- 2 chains of length 10000
- burn-in period of length 5000



based on shifted Weibull distribution for RT

Questions?

- correlations more biased towards 0 for separate modeling?
- maximum likelihood vs. Bayesian framework?
- gain precision for fixed effects when joint modeling?
- criterium to assess need for joint modeling?

Findings simulation study

Compare: SEP-B (separate modeling - Bayesian),
SEP-M (separate modeling - maximum likelihood)
JOINT (joint modeling - Bayesian framework)

- little overall difference between SEP-B and SEP-M
- SEP-B and SEP-M tend to give more biased estimates towards zero for correlations ρ_θ and ρ_τ compared to JOINT
- increasing the number of subjects or increasing the item variability on reaction time and accuracy resulted in less biased estimates of the item correlation ρ_τ, but with JOINT still outperforming SEP-B and SEP-M

- similar findings for the subject correlation are observed when the number of items or the subject variability are increased
- while some gain in efficiency for the fixed effect parameters could be expected from JOINT we did not see such effects
- DIC favored the JOINT above SEP-B with a difference of at least 3 in 97.6%, 25.6% and 4.4% of the cases when correlations were high, low and zero, respectively
  (DIC: mimics AIC: looks at the posterior expectation of the deviance function)

Future research

- ‘censoring’ totally ignored so far
- correlation in other psycholinguistic experiments
- diffusion versus hierarchical models