The interpretation of focus in contrastive stress sentences

Repeated measures ANOVA

vs.

Mixed-effect models

How do we focus?

- By intonation (stress)

- By certain words: focus particles
  - *even*
  - ***only***

- Two theoretical accounts:
  1. Reference Set Computation (Reinhart, 2004)
  2. (bidirectional) Optimality Theory (Hendriks, 2010)

Narrow focus vs. wide focus

- De prinses heeft alleen een T-shirt aan de COWBOY gegeven
- (The princess has only given a T-shirt to the COWBOY.)

- Narrow focus reading
  1. *The princess gave a T-shirt to the cowboy*
  2. *The only person who got a T-shirt is the cowboy*

- Wide focus reading
  1. *The princess gave a T-shirt to the cowboy*
  2. *The only thing the princess did was giving a t-shirt to the cowboy*

Narrow focus vs. wide focus

- De prinses heeft alleen een T-SHIRT aan de cowboy gegeven
- (The princess has only given a T-SHIRT to the cowboy.)

- <u>Narrow focus reading</u>
  1. *The princess gave a T-shirt to the cowboy*
  2. *The only thing the princess has given to the cowboy, is a T-shirt.*

Research questions

How do children assign focus in sentences with the Dutch focus particle *alleen*?

Method

- Participants
  - 35 Dutch children
  - Age 8;0-10;11 (m = 9;2)
  - 4 participants excluded due to high trackloss

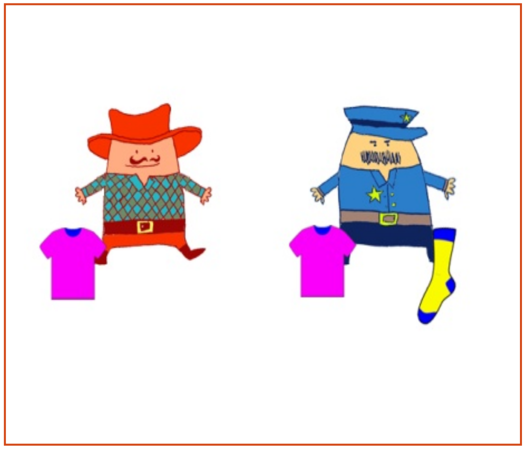
- Materials & design
  - Picture-verification task
  - 2 practice trials, 36 test trials
  - 6 different verbs
  - 6 character combinations

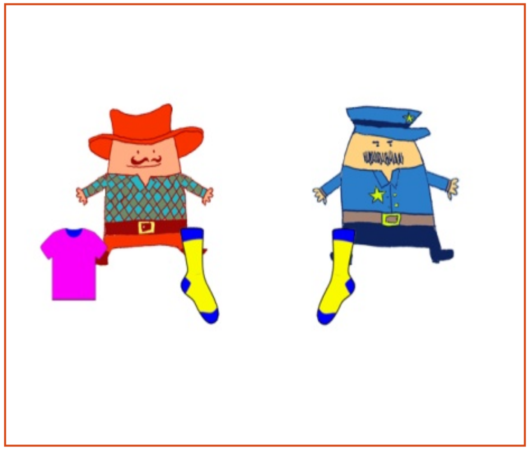
Method

- 2x2 within-subjects design
    1. **STRESS** (default vs. marked)
    2. **PICTURE** (1-item vs. 2-item)

Method

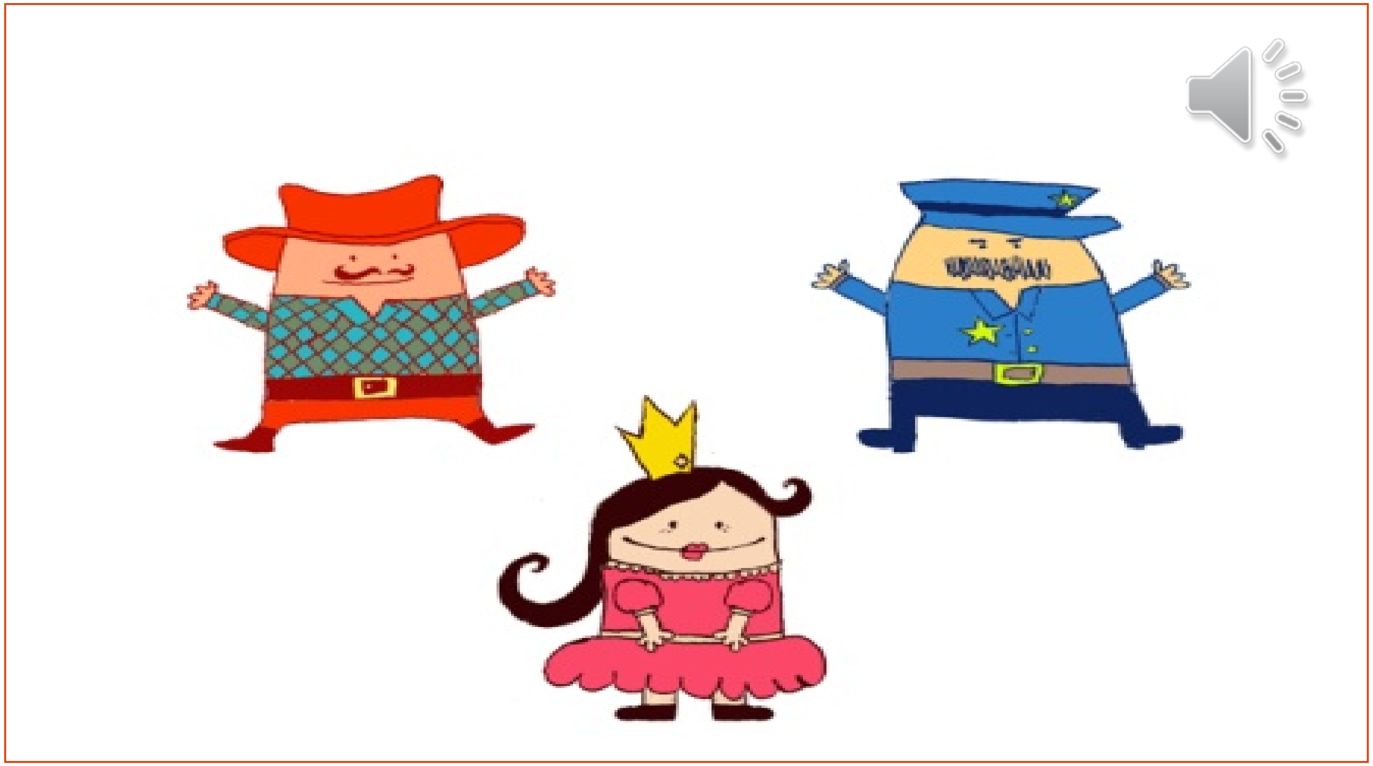
2. **PICTURE** (1-item vs. 2-item)



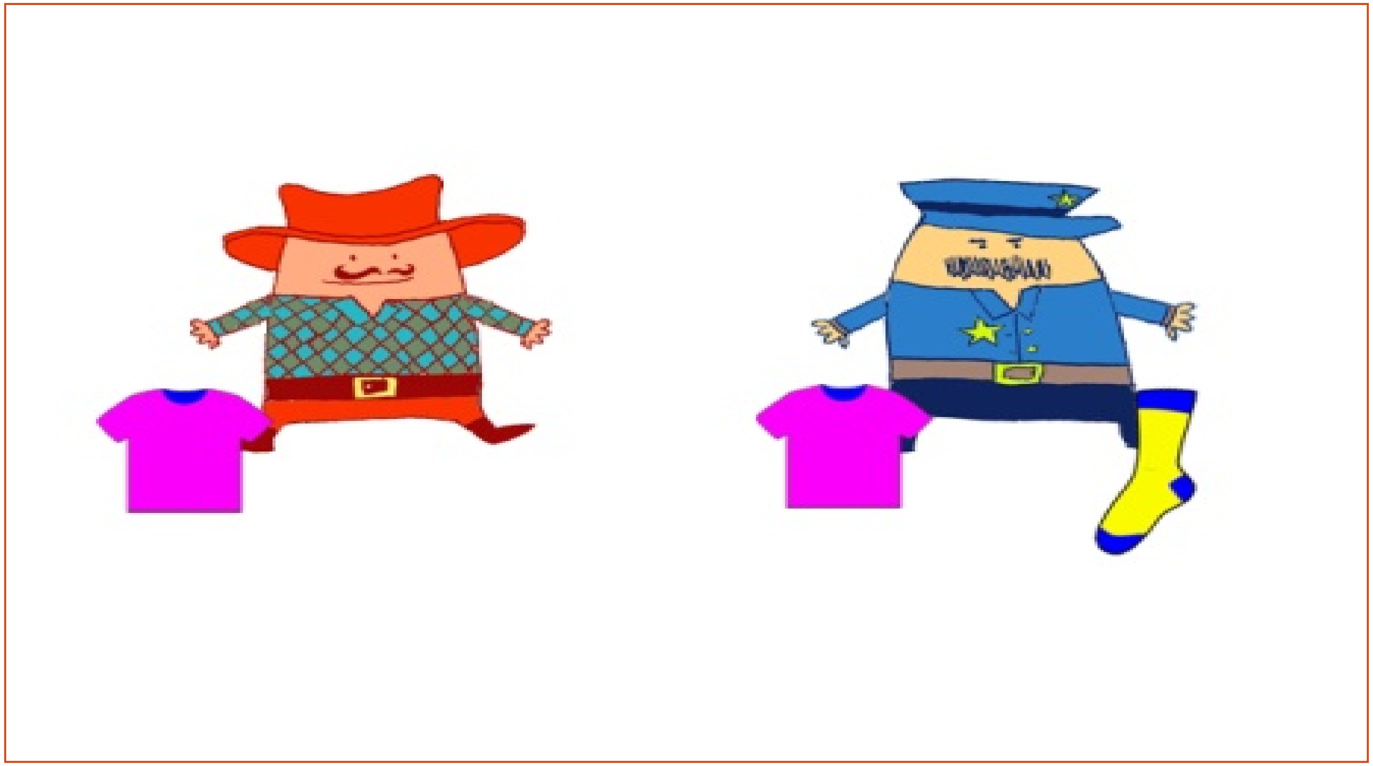


STRESS	MATCH	MISMATCH
IO (default)	IO_2	IO_1
DO (marked)	DO_1	DO_2

Introduction stimulus + sentence

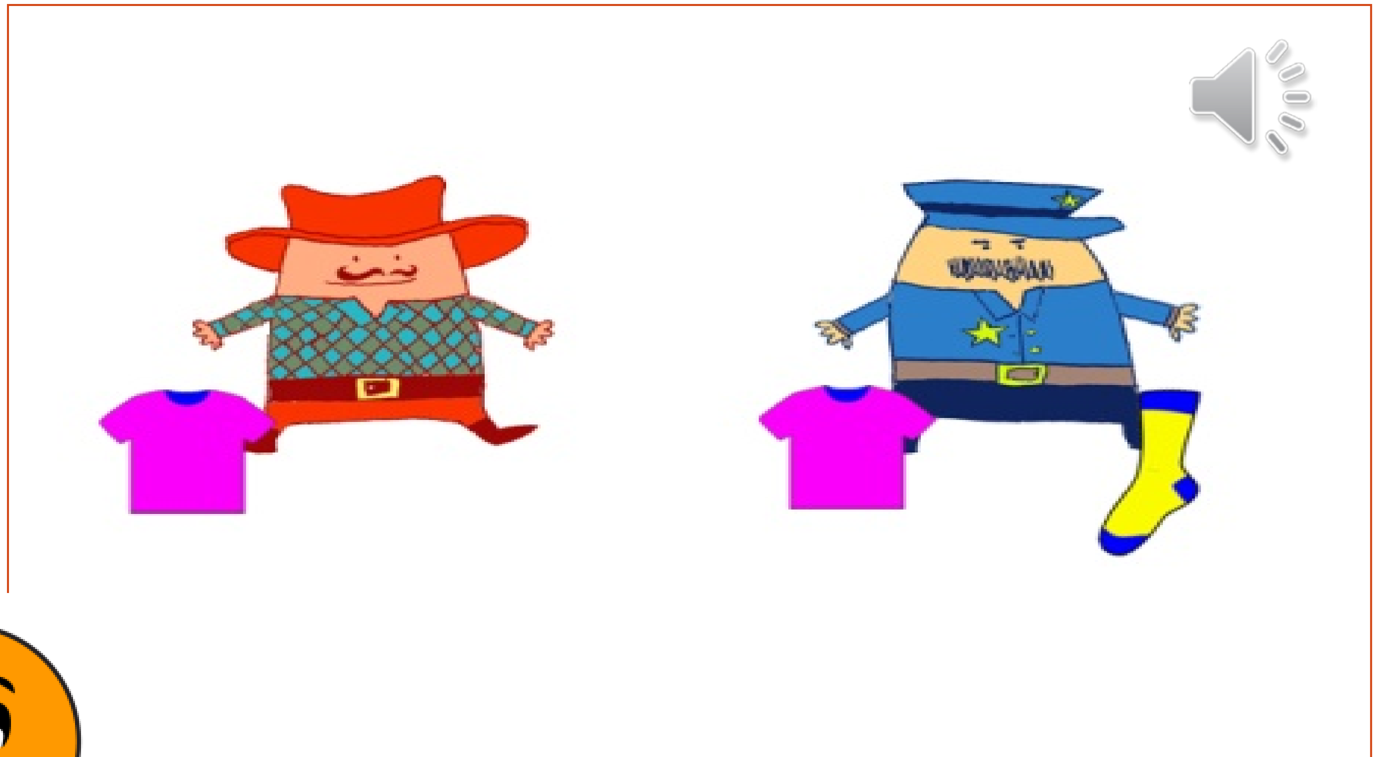


Preview



Fixation cross

Test stimulus + sentence



Test stimulus + sentence

Original analysis: RM ANOVA

- Repeated measures ANOVA for Accuracy
- 2x2 within-subjects design
- No significant interaction between STRESS and PICTURE
- Strong effect for PICTURE (p <.001, F = 79.868, η_p^2 = .701)
    - Children distinguish on the basis of the situation

Original analysis: RM ANOVA

```
> m1.aov<-aov(Correct~stress*match+Error(subj/(stress*match)),data=lezing.by.subj)
> summary(m1.aov)

Error: subj
          Df Sum Sq Mean Sq F value Pr(>F)
Residuals 30  58.98   1.966

Error: subj:stress
          Df Sum Sq Mean Sq F value   Pr(>F)
stress     1  627.7   627.7    91.2 1.32e-10 ***
Residuals 30  206.5     6.9
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Error: subj:match
          Df Sum Sq Mean Sq F value Pr(>F)
match      1   4.27   4.266   3.461 0.0727 .
Residuals 30  36.98   1.233
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Error: subj:stress:match
             Df Sum Sq Mean Sq F value Pr(>F)
stress:match  1  0.395  0.3952   1.203  0.281
Residuals    30  9.855  0.3285
```

Original analysis: RM ANOVA

```
> ezANOVA(data=lezing.by.subj,dv=Correct,wid=.(subj),within=.(stress,match))
$ANOVA
       Effect DFn DFd         F            p p<.05        ges
2      stress   1  30 91.198547 1.322851e-10     * 0.667767588
3       match   1  30  3.460532 7.267921e-02       0.013475304
4 stress:match   1  30  1.202946 2.814599e-01       0.001263636
```

Why use mixed-effects?

- The Language-as-Fixed-Effect Fallacy
  - Implicit generalizations for subjects and/or items

- Generalization not shown in statistical analyses
  - $F_1 \rightarrow$ what would happen with a new sample of subjects?
  - $F_2 \rightarrow$ what would happen with a new sample of sentences?

- Some coefficients as a fixed instead of a random effect
  - E.g. item (verb, characters)

- Adding more sources of error

New analysis: mixed-effect modeling

- $y_i = a_{j[i]} + bx_i + e_i$
  - $aj = \mu_a + \epsilon j$

$\Downarrow$

- $y_i = \mu_a + \epsilon_{j[i]} + bx_i + e_i$

- Random variation due to each subject → every subject has its own intercept

- We assume that every subject has a different baseline level for correctly determining focus in contrastive stress sentences.

New analysis: mixed-effect modeling

- $y_i = a_{j[i]} + bx_{j[i]} + e_i$
  - $a_j = \mu_a + \epsilon_j$
  - $bj = \mu_b + \epsilon_j$

$\Downarrow$

- $y_i = \mu_a + \epsilon_{j[i]} + \mu_b + \epsilon_{j[i]} + e_i$

- Random variation due to each sentence (verb/characters) → every subject has its own slope for each coefficient

- We assume that every subject react to the (different parts of the) experimental manipulation in a different way.

Why mixed-effect logistic GLM?

- Mixed-effect → multiple sources of variation; some of the coefficients can be random instead of fixed

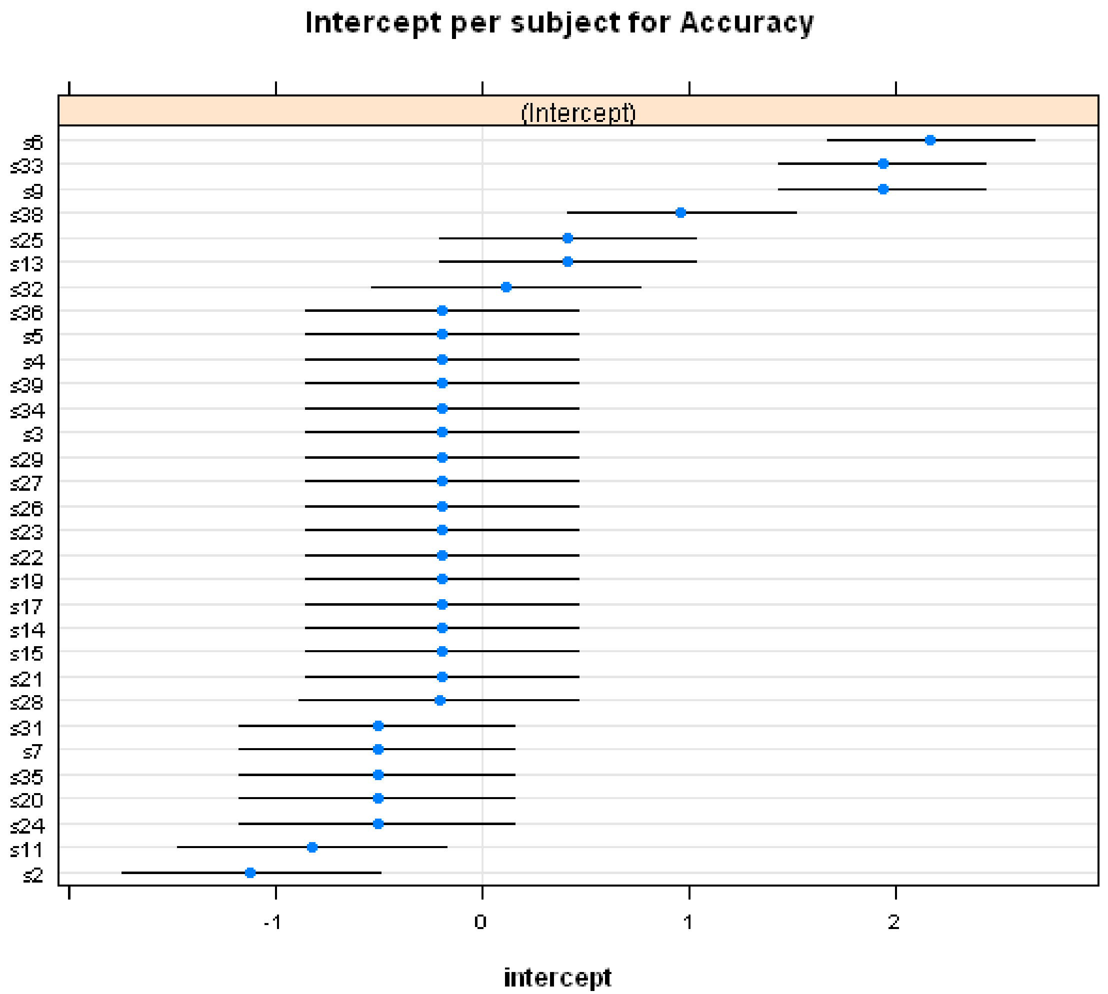
- Logistic regression → binomially distributed error
  - So no normally distributed error with zero mean

- Generalized Linear Models → allow other than normal error distributions

Random intercept(s) models

```
> m1<-glmer(Accuracy~stress*match+(1|subj),data=lezing,family=binomial)
> summary(m1)
Generalized linear mixed model fit by maximum likelihood ['glmerMod']
 Family: binomial ( logit )
Formula: Accuracy ~ stress * match + (1 | subj)
   Data: lezing

      AIC       BIC    logLik  deviance
 521.5188  544.5723 -255.7594  511.5188

Random effects:
 Groups Name        Variance Std.Dev.
 subj   (Intercept) 0.8556   0.925
Number of obs: 743, groups: subj, 31

Fixed effects:
                      Estimate Std. Error z value Pr(>|z|)
(Intercept)             3.2792     0.4029   8.138 4.01e-16 ***
stressIO               -5.0590     0.4288 -11.797  < 2e-16 ***
matchMismatch          -1.1545     0.4292  -2.690  0.00715 **
stressIO:matchMismatch  0.7703     0.5333   1.444  0.14860
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Correlation of Fixed Effects:
            (Intr) strsIO mtchMs
stressIO    -0.789
matchMsmtch -0.770  0.724
strssIO:mtM  0.617 -0.772 -0.805
```

Random intercept(s) models



Random intercept

```
> m3<-glmer(Accuracy~stress*match+(1|subj)+(1|verb),data=lezing,family=binomial)
> summary(m3)
Generalized linear mixed model fit by maximum likelihood ['glmerMod']
 Family: binomial ( logit )
Formula: Accuracy ~ stress * match + (1 | subj) + (1 | verb)
   Data: lezing

      AIC       BIC    logLik  deviance
 523.5187  551.1829 -255.7594  511.5187

Random effects:
 Groups Name        Variance  Std.Dev.
 subj   (Intercept) 8.556e-01 9.250e-01
 verb   (Intercept) 5.704e-10 2.388e-05
Number of obs: 743, groups: subj, 31; verb, 6

Fixed effects:
                       Estimate Std. Error z value Pr(>|z|)
(Intercept)              3.2792     0.4029   8.138 4.02e-16 ***
stressIO                -5.0590     0.4288 -11.797  < 2e-16 ***
matchMismatch           -1.1545     0.4292  -2.690  0.00715 **
stressIO:matchMismatch   0.7703     0.5333   1.444  0.14860
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Correlation of Fixed Effects:
            (Intr) strsIO mtchMs
stressIO    -0.789
matchMsmtch -0.770  0.724
strssIO:mtM  0.617 -0.772 -0.805
```

Random intercept

```
> anova(m1,m3)
Data: lezing
Models:
m1: Accuracy ~ stress * match + (1 | subj)
m3: Accuracy ~ stress * match + (1 | subj) + (1 | verb)
   Df    AIC    BIC  logLik deviance Chisq Chi Df Pr(>Chisq)
m1  5 521.52 544.57 -255.76   511.52
m3  6 523.52 551.18 -255.76   511.52     0      1     0.9951
```

Random intercept

```
> m4<-glmer(Accuracy~stress*match+(1|subj)+(1|char),data=lezing,family=binomial)
> summary(m4)
Generalized linear mixed model fit by maximum likelihood ['glmerMod']
 Family: binomial ( logit )
Formula: Accuracy ~ stress * match + (1 | subj) + (1 | char)
   Data: lezing

      AIC       BIC    logLik  deviance
 523.5187  551.1829 -255.7594  511.5187

Random effects:
 Groups Name        Variance  Std.Dev.
 subj   (Intercept) 8.556e-01 9.250e-01
 char   (Intercept) 8.477e-10 2.912e-05
Number of obs: 743, groups: subj, 31; char, 4

Fixed effects:
                        Estimate Std. Error z value Pr(>|z|)
(Intercept)               3.2792     0.4029   8.138 4.01e-16 ***
stressIO                 -5.0590     0.4288 -11.797  < 2e-16 ***
matchMismatch            -1.1545     0.4292  -2.690  0.00715 **
stressIO:matchMismatch    0.7704     0.5333   1.445  0.14860
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Correlation of Fixed Effects:
            (Intr) strsIO mtchMs
stressIO    -0.789
matchMsmtch -0.770  0.724
strssIO:mtM  0.617 -0.772 -0.805
```

Random intercept

```
> anova(m1,m4)
Data: lezing
Models:
m1: Accuracy ~ stress * match + (1 | subj)
m4: Accuracy ~ stress * match + (1 | subj) + (1 | char)
   Df    AIC    BIC  logLik deviance Chisq Chi Df Pr(>Chisq)
m1  5 521.52 544.57 -255.76   511.52
m4  6 523.52 551.18 -255.76   511.52     0      1     0.9951
```

Random intercept & slope

```
> m5<-glmer(Accuracy~stress*match+(1|subj)+
(1+verb|subj),data=lezing,family=binomial)
Warning message:
In function (fn, par, lower = rep.int(-Inf, n), upper = rep.int(Inf,  :
  failure to converge in 10000 evaluations
> summary(m5)
Generalized linear mixed model fit by maximum likelihood ['glmerMod']
 Family: binomial ( logit )
Formula: Accuracy ~ stress * match + (1 | subj) + (1 + verb | subj)
   Data: lezing

      AIC       BIC    logLik  deviance
 556.2257  676.1038 -252.1129  504.2257

Random effects:
 Groups Name           Variance  Std.Dev. Corr
 subj   (Intercept)    0.5502392 0.74178
 subj.1 (Intercept)    1.3112086 1.14508
        verbKopen      1.0808734 1.03965  -1.00
        verbPakken     1.4106298 1.18770  -1.00  1.00
        verbUitkiezen  0.5287392 0.72714  -1.00  1.00  1.00
        verbUitlenen   0.0237898 0.15424  -1.00  1.00  0.99  1.00
        verbVerkopen   0.0003678 0.01918  -0.50  0.50  0.49  0.52  0.57
Number of obs: 743, groups: subj, 31

Fixed effects:
                       Estimate Std. Error z value Pr(>|z|)
(Intercept)              3.4403     0.4148   8.295  < 2e-16 ***
stressIO                -5.2256     0.4430 -11.796  < 2e-16 ***
matchMismatch           -1.1815     0.4409  -2.680  0.00736 **
stressIO:matchMismatch   0.7805     0.5493   1.421  0.15534
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Correlation of Fixed Effects:
            (Intr) strsIO mtchMs
stressIO    -0.802
matchMsmtch -0.774  0.725
strssIO:mtM  0.620 -0.771 -0.803
```

Random intercept & slope

```
> m6<-glmer(Accuracy~stress*match+(1|subj)+
(1+char|subj),data=lezing,family=binomial)
> summary(m6)
Generalized linear mixed model fit by maximum likelihood ['glmerMod']
 Family: binomial ( logit )
Formula: Accuracy ~ stress * match + (1 | subj) + (1 + char | subj)
   Data: lezing

      AIC       BIC    logLik  deviance
 538.9592  608.1196 -254.4796  508.9592

Random effects:
 Groups Name        Variance Std.Dev. Corr
 subj   (Intercept) 0.003644 0.06037
 subj.1 (Intercept) 1.407459 1.18636
        charFPR     1.210417 1.10019  -0.71
        charPCA     0.167882 0.40973  -0.76  1.00
        charVPC     0.123835 0.35190  -0.64  1.00  0.99
Number of obs: 743, groups: subj, 31

Fixed effects:
                       Estimate Std. Error z value Pr(>|z|)
(Intercept)              3.2619     0.3985   8.185 2.72e-16 ***
stressIO                -5.1072     0.4351 -11.737  < 2e-16 ***
matchMismatch           -1.0848     0.4326  -2.508   0.0121 *
stressIO:matchMismatch   0.7079     0.5423   1.305   0.1918
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Correlation of Fixed Effects:
            (Intr) strsIO mtchMs
stressIO    -0.774
matchMsmtch -0.769  0.714
strssIO:mtM  0.606 -0.772 -0.801
```

Random intercept & slope

```
> anova(m1,m5)
Data: lezing
Models:
m1: Accuracy ~ stress * match + (1 | subj)
m5: Accuracy ~ stress * match + (1 | subj) + (1 + verb | subj)
   Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
m1  5 521.52 544.57 -255.76   511.52
m5 26 556.23 676.10 -252.11   504.23 7.2931     21     0.9975
```

```
> anova(m1,m6)
Data: lezing
Models:
m1: Accuracy ~ stress * match + (1 | subj)
m6: Accuracy ~ stress * match + (1 | subj) + (1 + char | subj)
   Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
m1  5 521.52 544.57 -255.76   511.52
m6 15 538.96 608.12 -254.48   508.96 2.5596     10       0.99
```

Conclusion

- Children interpret focus differently from adults
- Small effect of match-mismatch
- Controlled experiment → items have little influence on variance
- Further research: same experiment with autistic adults

Discussion

- ! Significant effect for stress in replicated original analysis AND new analysis
- Exactly the same AIC for models with random intercept for verb and characters
- Random intercept per subject → many exactly equal intercepts

References

- Clark, H. H. (1973). The language-as-fixed-effect fallacy: A critique of language statistics in psychological research. *Journal of verbal learning and verbal behavior*, *12*(4), 335-359.
- Gelman, A., & Hill, J. (2007). *Data analysis using regression and multilevel/hierarchical models*. Cambridge University Press.
- Hendriks, P. (2010). *Conflicts in Interpretation*. London, UK: Equinox Publishing.
- Reinhart, T. (2004). The processing of cost of reference set computation: Acquisition of stress shift and focus. *Language Acquisition* 12;2, 109-155.

- *Pictures + audio files all come from the experiment performed by Bart Hollebrandse and Petra Hendriks*

F1, F2

- F1: treatment effect against treatment by subject interaction
  - $F_1(p\text{-}1, (p\text{-}1)(r\text{-}1)) = MS_t / MS_{t\ x\ s}$
- F2: treatment effect against Words-within-Treatments effect
  - $F_2(p\text{-}1, p(q\text{-}1)) = MS_t / MS_{WwT}$

TABLE 1

SOURCES OF VARIANCE AND EXPECTED MEAN SQUARES FOR MIXED HIERARCHICAL THREE FACTOR DESIGN WITH ONE FIXED EFFECT AND TWO RANDOM EFFECTS

Label	Sources of variance	Degrees of freedom	Expected value of mean square
T	Treatments (p)	$p - 1$	$\sigma_e^2 + \sigma_{ws}^2 + q\sigma_{ts}^2 + r\sigma_w^2 + qr\sigma_t^2$
WwT	Words (q) within Treatments	$p(q - 1)$	$\sigma_e^2 + \sigma_{ws}^2 + r\sigma_w^2$
S	Subjects (r)	$r - 1$	$\sigma_e^2 + \sigma_{ws}^2 + pq\sigma_s^2$
T × S	Treatments × Subjects	$(p - 1)(r - 1)$	$\sigma_e^2 + \sigma_{ws}^2 + q\sigma_{ts}^2$
S × WwT	Subjects × Words within Treatments	$p(q - 1)(r - 1)$	$\sigma_e^2 + \sigma_{ws}^2$