# Mispronunciation detection with pupillometry:

## A window to study lexical representations

Katalin Tamási, Cristina McKean, Adamantios Gafos, & Barbara Höhle

International Doctorate Experimental Approaches to Language and the Brain (IDEALAB),
Universities of Potsdam, Newcastle, Groningen, Trento and Macquarie University (Sydney)

# How detailed are early words?

Infants:

- Excellent discriminators of phonetic detail

  Jusczyk & Aslin (1995), Werker & Tees (1984)
- Ability to form phonetic categories

  Hochmann & Papeo (2014)
- Performance in discrimination > word learning

  Stager & Werker (1997)

→ What details are stored in the early mental lexicon?

Fikkert (2010), Pater, Stager & Werker (2004)

# Assessing lexical development

I. Production studies

- Spontaneous speech: consonant harmonies, assimilations

  Ferguson & Farwell (1975)

- Elicitation tasks: strong lexical effects

  Storkel (2002)

- Metalinguistic tasks: inability to manipulate phonemes

  Treiman & Baron (1981; 1983)

Limitation:

- Motor immaturity or genuine representational deficit?

  McLeod, Doorn, & Reed, (2001)

# Assessing lexical development

II. Perception studies: e.g., mispronunciation detection

- 14 month-olds
  - PoA change (e.g., *bin* – *din*)

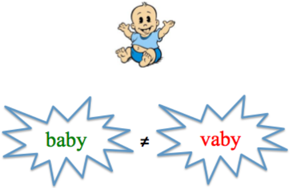
Jusczyk & Aslin (1995); Yoshida, Fennell, Swingley, & Werker (2009)

- 19 month-olds
  - Voicing change (e.g., *dog* – *tog*)
  - MoA change (e.g., *bird* – *vird*)
  - Height and backness change (e.g., *bed-bid, brush-brash*)

Swingley & Aslin, (2000; 2002); White & Morgan (2008); Mani, Coleman, & Plunkett (2008)

# Assessing lexical development

Detecting mispronunciation:
→ Early words contain sub-phonemic information



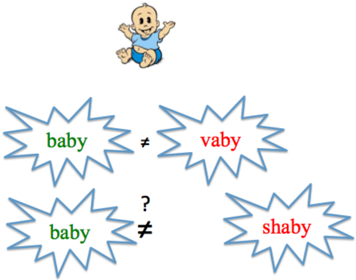
# Assessing lexical development

Detecting degrees of mispronunciation

White & Morgan (2008), Mani, Coleman, & Plunkett (2008)

$\rightarrow$ Sensitivity to featural distance

$\rightarrow$ Lexical representations: contain featural information



# Pupillometry

Tobii T1750 eye tracker:
Detecting changes in pupil dilation

- A tool for mispronunciation detection

  Fritzsche & Höhle, (2015); Hochmann & Papeo (2014)

- Proxy of cognitive effort
  (surprising / unexpected / incongruous stimuli)

  Kahnemann, (1973); Karatekin (2007)

- Prediction: sensitivity to the degree of featural distance

# Participants

48 children (5 excluded due to insufficient data)

- Mean age: 30 months (*SD 0.57*)
- Monolingual German background
- Familiarity with experimental words:
  - 82.1 % (*SD 14.6*)
- Vocabulary size (max. 600 words):
  - 410 words (*SD 112*)

# Stimuli

20 words chosen from the German CDI

Szagun, Stumper, & Schramm (2009)

- Part of productive vocabulary of children at 30 months
- CVC and CVCV items, diverse featural makeup
- Word frequency, positional biphone probability, & neighborhood density info collected from Clearpond

Marian, Bartolotti, Chabal, & Shook (2012)

- Produced & recorded by a native German speaker

# Manipulation

- Four-way: number of feature changes (0, 1, 2, 3)
- Counterbalanced for feature types (PoA, MoA, V)

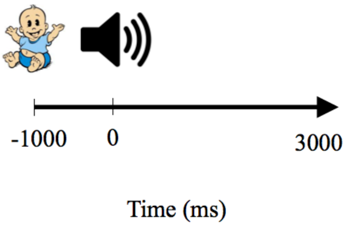
| Correct | Δ1F | Δ2F | Δ3F |
|---|---|---|---|
| /k/amm (*comb*) | /p/ | /f/ | /v/ |
| /z/onne (*sun*) | /d/ | /f/ | /p/ |

# Procedure

Block structure

- four versions
- 5 x 4 blocks = 20 trials
- semi-randomized order
- between-block attention getters

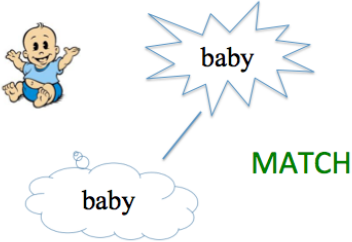
Trial structure




Afterwards: parental questionnaire

# Predictions

Correct trials



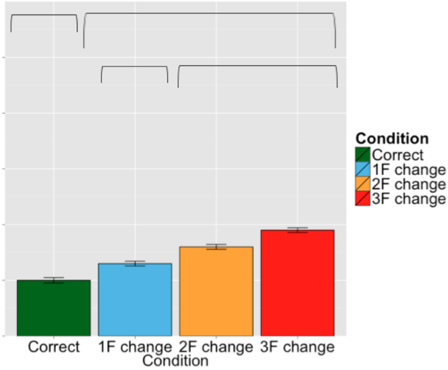

Incorrect ($\Delta$1F, $\Delta$2F, $\Delta$3F) trials


Semantic integration: more complex with mispronounced words
$\rightarrow$ more cognitive effort
$\rightarrow$ larger degree of pupil dilation

# Predictions

Pupil dilation reflects:

- Effect of mispronunciation →
- Effect of featural distance →



# Exploratory analysis

Fixed effects:

- Featural distance
- Lexical factors: familiarity, word frequency, positional biphone probability, neighborhood density

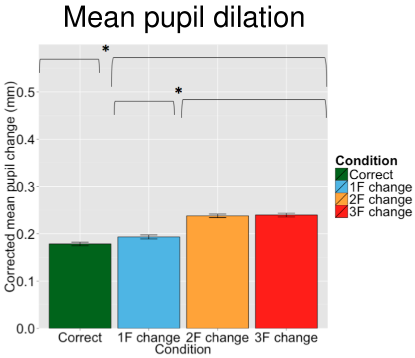
Random effects:

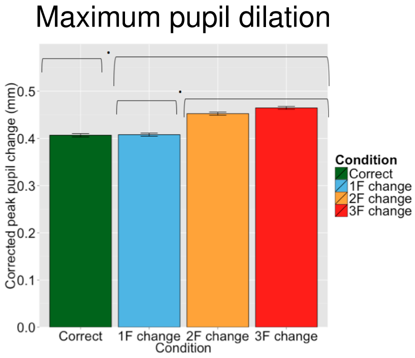
- Participants (*N = 43*) (featural distance in random slope)
- Items (*N = 20*)

Outcome measures:

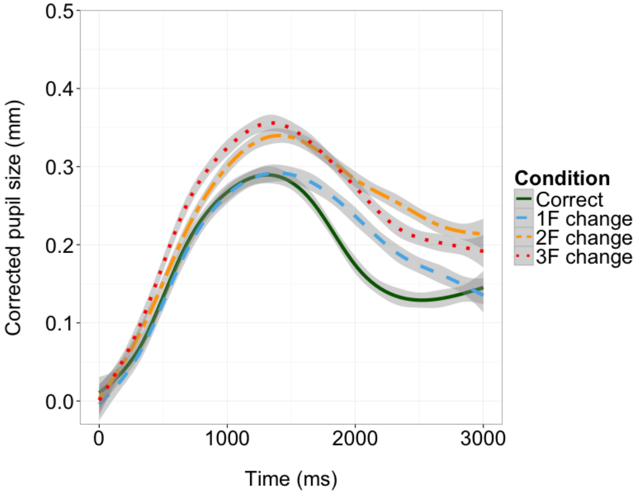
- Mean pupil dilation
- Maximum pupil dilation

# Featural distance



# Featural distance



# Discussion

1. Pupillometry registers differential response to
   - Mispronunciation
   - Featural distance

   → Viable method in child language research
2. Detecting mispronunciations
   - Lexical representations contain sub-phonemic information
3. Detecting degrees of mispronunciations
   - Suggests sensitivity to featural distance

→ Lexical representations contain featural information

# Future research

- Abstractness in lexical representations
  - Discounting acoustic / perceptual similarity
  - Effect of type & direction of feature change
  - Possible interactions between features
- Extending the paradigm to...
  - other languages
  - bilinguals
  - adults
- Methodological considerations: dependent measure?
- Impact of lexical factors

# Acknowledgements

Babylab personnel:
Tom Fritzsche
Katja Schneller
Carolin Jaekel
Steffi Meister
Maike Riegel

And, of course:
The participants and their families

This research was supported by grant Erasmus Mundus EMJD 520101-1-2011-1-DE-ERA.

# Questions

Any questions regarding...?

- background
- method
- analysis
- adult performance

What about lexical effects such as...?

- familiarity
- word frequency
- positional biphone probability
- neighborhood density

# White & Morgan (2008)

19 month-olds are differentially sensitive to the number of feature changes in the onset

- Preferential looking paradigm: target + distractor image
  - Novel approach: distractor is an unfamiliar object, more likely to be a possible match with the mispronounced label
  - Auditory stimuli: *Where is the X*? *Find the X*!
  - Dependent measure: looking time at the target object

# Critical manipulation
## White & Morgan (2008)

- Onset features: PoA, MoA, Voicing
- Number of features changed: 1, 2, 3
  - Δ 1F: PoA ({keys} → {teys})
  - Δ 2F: PoA + Voicing ({keys} → {deys})
  - Δ 3F: PoA + Voicing + MoA ({keys} → {zeys})

# Results
## White & Morgan (2008)

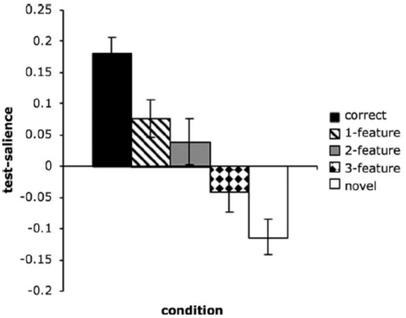



Fig. 2. Proportional looking times and standard errors, Experiment 1. Condition is represented on the *x*-axis. The *y*-axis represents the difference between proportion looking at the familiar object in the test phase and proportion looking at the familiar object in the salience phase.

# Possible limitations

1. Preferential looking paradigm
   - Indirect measure
   - Potential confound with distractor (even with unknown label)
2. Stimuli set
   - Predominance of labials, especially of {b}
   - Unbalanced for type of feature change

# Pupillometry

Tobii T1750 eye tracker

- Detecting changes in pupil dilation

Why use pupillometry?

1. Easy to administer
2. Inexpensive and easy to learn
3. Simpler design

# Exploratory analysis

Transformation, exclusion criteria

- Linear interpolation of blinks (no longer than 400 ms)
- Averaging left and right pupil values
- Successful trials = more than 50% pupil data
- 43/48 children: more than 50% successful trials

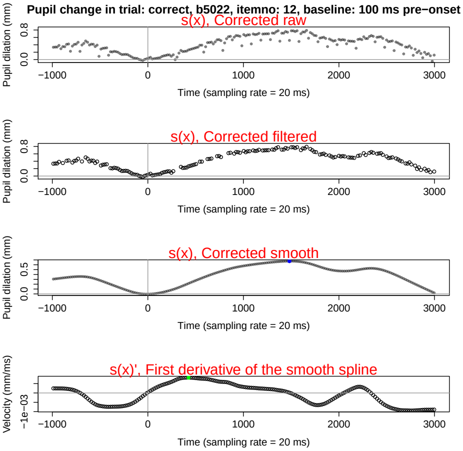
# Exploratory analysis

Potential outcome measures (per trial):

- Mean pupil dilation (mm) (baseline corr.: 100 ms pre-onset)
- Peak dilation of smooth spline (mm)
- Latency to peak dilation
- Peak velocity of smooth spline (mm/ms)
- Latency to peak velocity
- Wavelet basis function

# Exploratory analysis

## Pupil dilation over time in a representative trial



# Statistical model, mean pupil dilation

| | Coefficients (SD) |
|---|---|
| (Intercept) | **0.24**(0.02)*** |
| cond1_vs_234 *(Effect of mispronunciation)* | **0.04**(0.02)* |
| cond2_vs_34 *(Effect of featural distance)* | **0.05**(0.02)* |
| c.PTAF *(Neighborhood density)* | −0.02(0.01) |
| c.PBPP *(Positional biphone frequency)* | 0.01(0.01) |
| c.LOGFREQ *(Logged frequency)* | −0.02(0.01) |
| c.PTAF:cond1_vs_234 | **0.02**(0.00)*** |
| c.PTAF:cond2_vs_34 | **0.01**(0.00)*** |
| c.PBPP:cond1_vs_234 | −**0.03**(0.00)*** |
| c.PBPP:cond2_vs_34 | −**0.04**(0.00)*** |
| c.LOGFREQ:cond1_vs_234 | **0.04**(0.00)*** |
| c.LOGFREQ:cond2_vs_34 | −**0.07**(0.00)*** |

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

# Statistical model, maximum pupil dilation

| | Coefficients (SD) |
|---|---|
| (Intercept) | **0.42**(0.03)*** |
| cond1_vs_234 *(Effect of mispronunciation)* | **0.03**(0.02)· |
| cond2_vs_34 *(Effect of featural distance)* | **0.05**(0.03)· |
| c.PTAF *(Neighborhood density)* | −0.02(0.02) |
| c.PBPP *(Positional biphone frequency)* | 0.01(0.01) |
| c.LOGFREQ *(Logged frequency)* | −0.01(0.02) |
| c.PTAF:cond1_vs_234 | −0.00(0.00) |
| c.PTAF:cond2_vs_34 | **0.02**(0.00)*** |
| c.PBPP:cond1_vs_234 | −**0.04**(0.00)*** |
| c.PBPP:cond2_vs_34 | −**0.03**(0.00)*** |
| c.LOGFREQ:cond1_vs_234 | **0.04**(0.00)*** |
| c.LOGFREQ:cond2_vs_34 | −**0.07**(0.00)*** |

***$p < 0.001$, **$p < 0.01$, *$p < 0.05$

# Lexical effects hypotheses, children

More cognitive effort required
(as indicated by larger pupil dilation):

- Unknown words
- Low-frequency words

Goodman, Dale, & Li (2008)

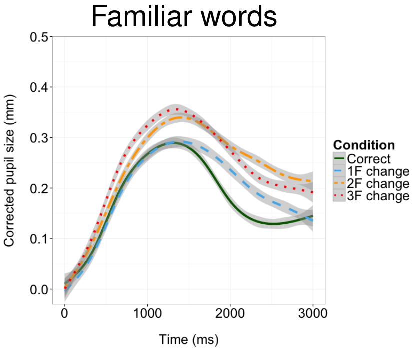
- Words with higher positional biphone probability

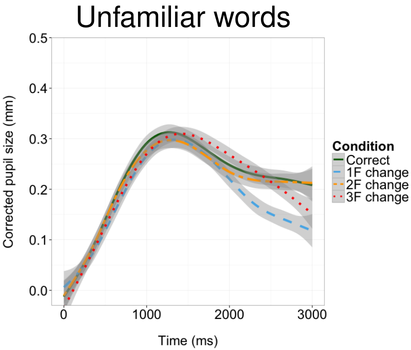
Hoover, Storkel, & Hogan (2010)

- Words in sparser lexical neighborhoods

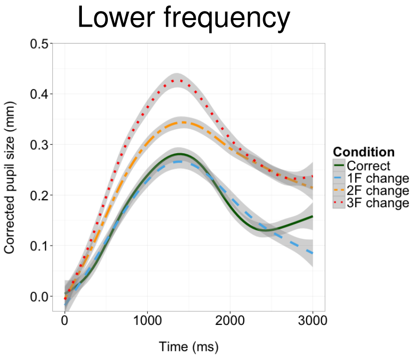
Hollich, Jusczyk, & Luce (2002)

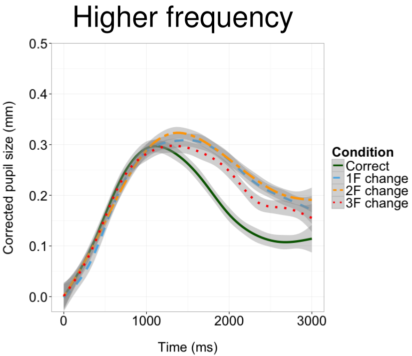
# Familiarity



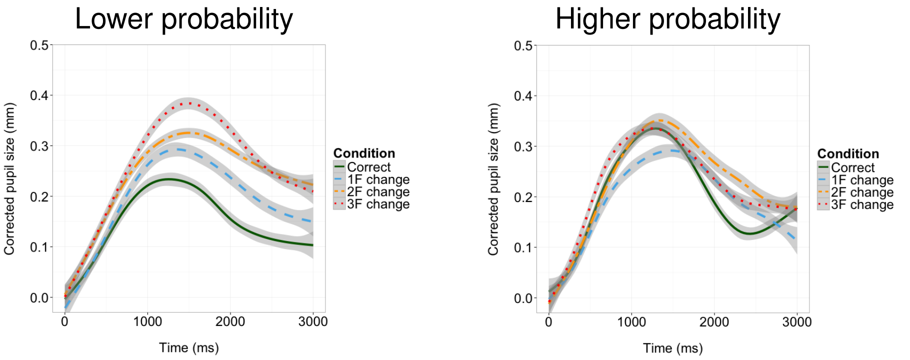
# Word frequency (only familiar words)



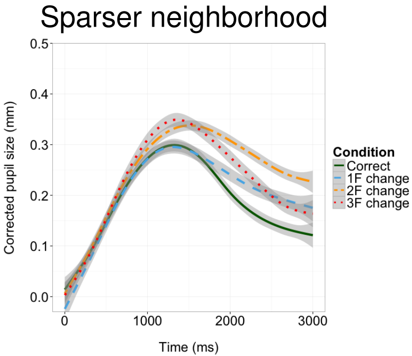

Adults

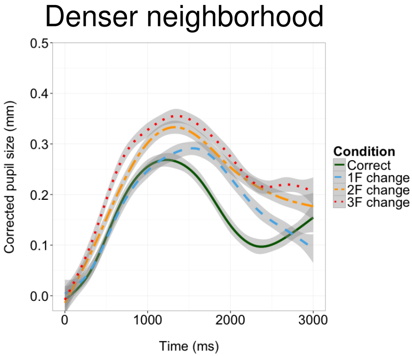
# Positional biphone probability (only familiar words)




▸ Adults

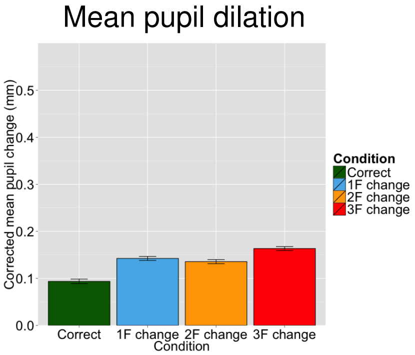
# Neighborhood density (only familiar words)

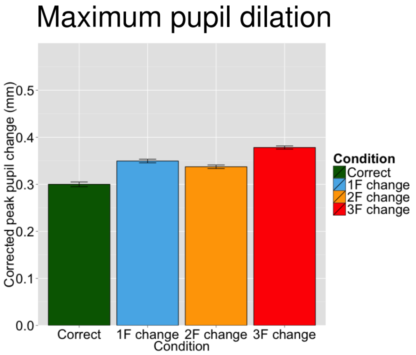


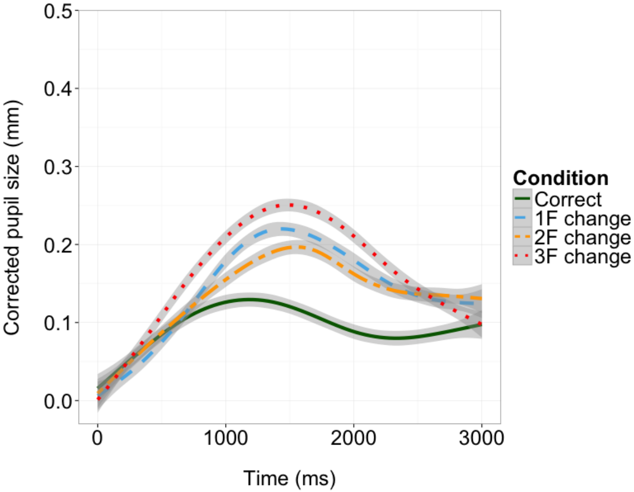

Adults

# Featural distance, adults

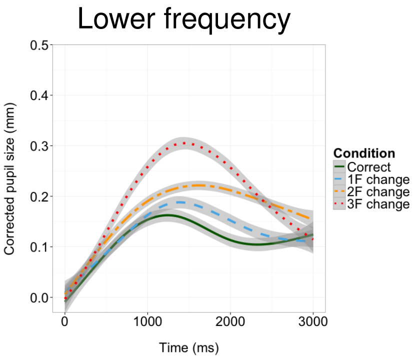


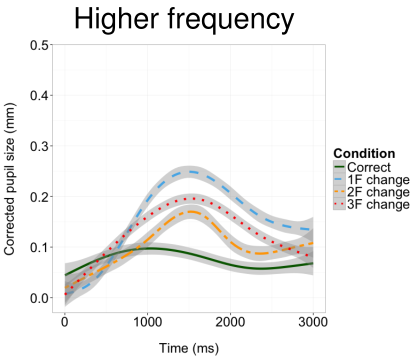



# Featural distance, adults



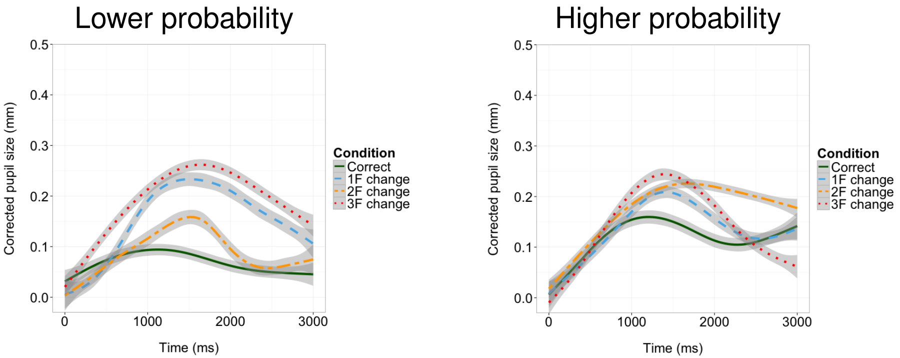
# Word frequency, adults



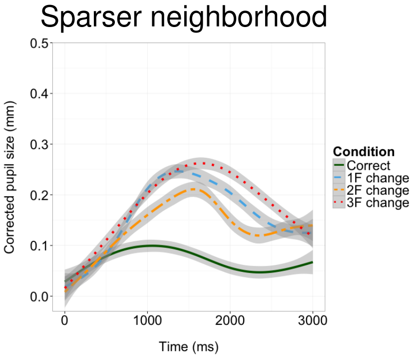

Children

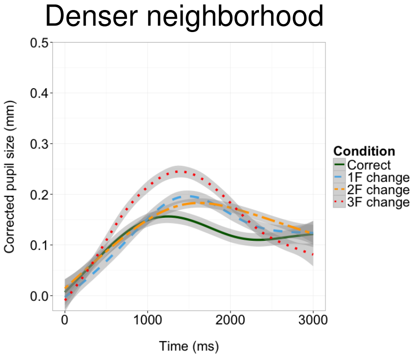
# Positional biphone probability, adults




▸ Children

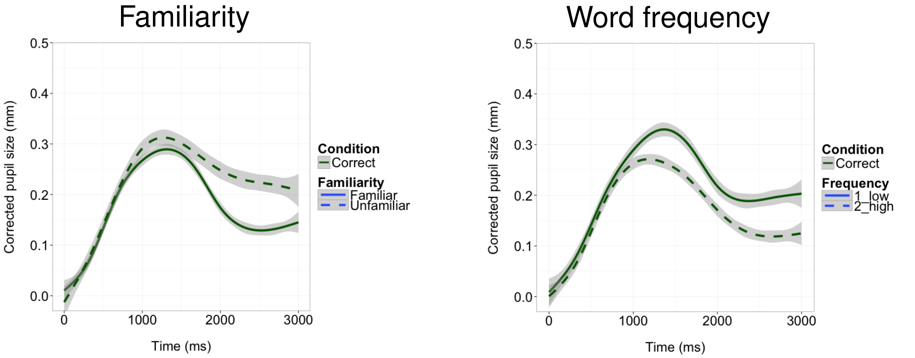
# Neighborhood density, adults




Children

# Lexical effects (only correct words)



# Lexical effects - only correct words