# Language Learning: A Data-Driven Approach

Day 4: Exploring online processing using Peekbank

Peekbank

**Michael C. Frank**

**(with slides from Virginia Marchman, Claire Bergey, and Martin Zettersten)**

LOT Winter School

# Outline

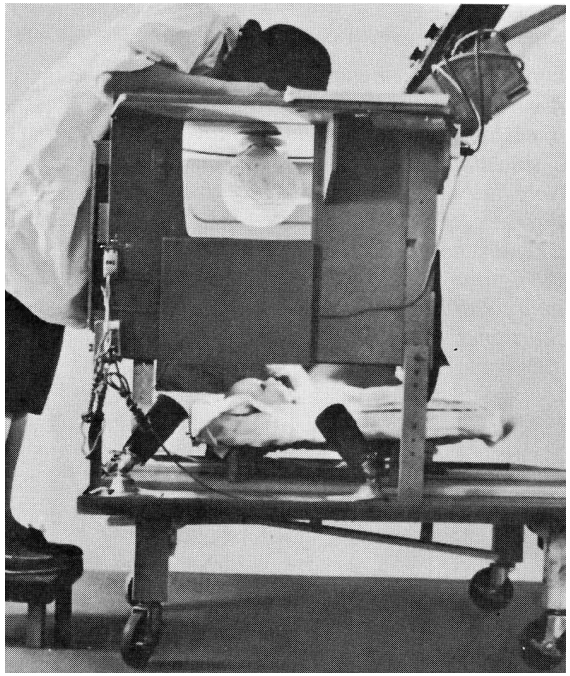
1. **Introducing the looking-while-listening method**
2. Peekbank, a new resource
3. Explorations of LWL methods using Peekbank
4. Insights into online word recognition using Peekbank

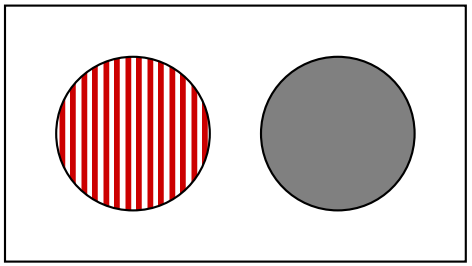
# Using eye movements to study language understanding

*"There is one perceptual activity that can be objectively observed and measured - **the movement of the eyes**. Here is a very promising technique for studying perception!"*

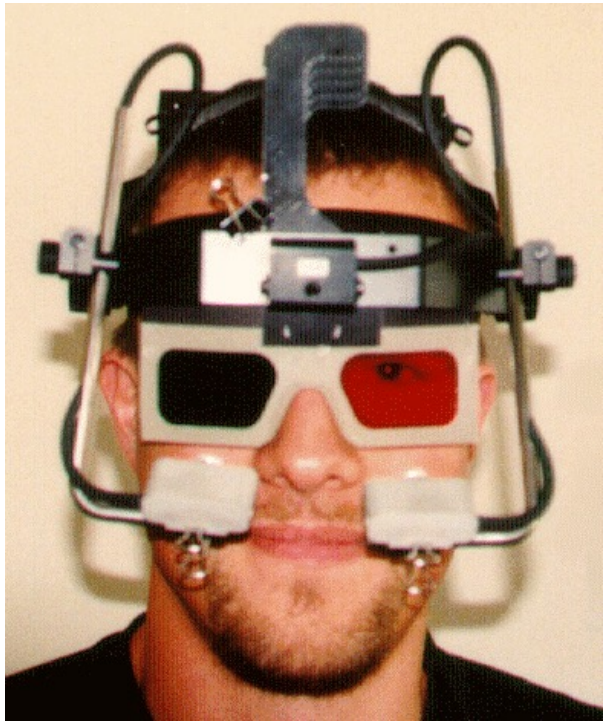
*Eleanor Gibson (1969)*

# Fantz' visual preference procedure (1960)





# Head-mounted eye-tracker



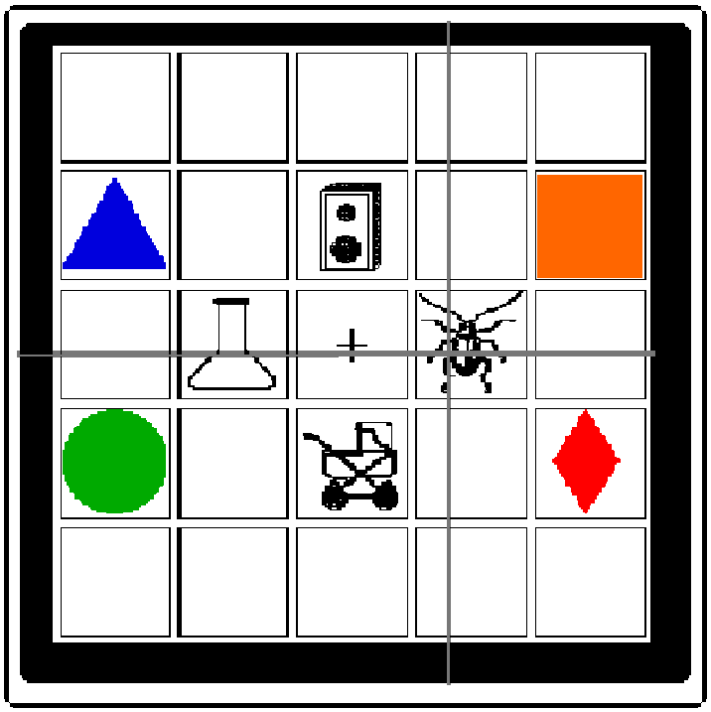
Do rhymes compete with words in lexical access?

BEAKER - target word

BEETLE - cohort competitor

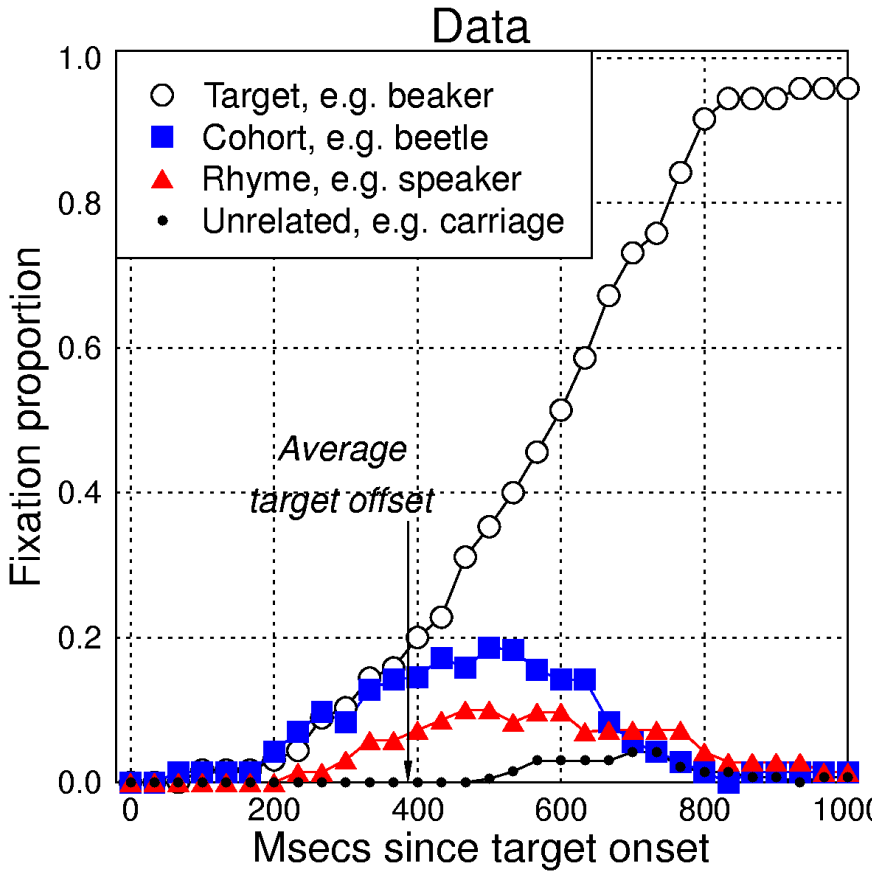
SPEAKER - rhyme

# STIMULI (Allopenna, Magnuson & Tanenhaus, 1998)



*“Pick up the beaker”*

# Allopenna et al. Results

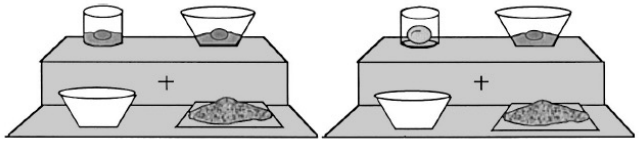


# Why use eye movements to measure language understanding in infants?

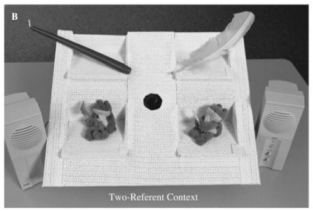
- Even very young infants have lots of experience moving their eyes; by 3 months it's a highly practiced response
- The "task demands" are low, a crucial consideration in research with infants
- Eye movements are rapid ballistic responses tightly linked to attention  (Haith: 200-300 msec for infants, as for adults)
- Eye movements are time-locked to speech, providing a measure of the time course of spoken language processing

# Different names for different procedures...

## Visual World Paradigm



Chambers, Tanenhaus, & Magnuson (2004)



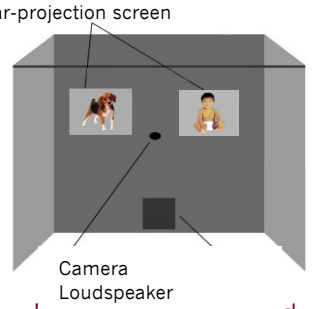

Snedeker & Trueswell (2004)

## Intermodal Preferential Looking

Naigles lab - U. Conn.

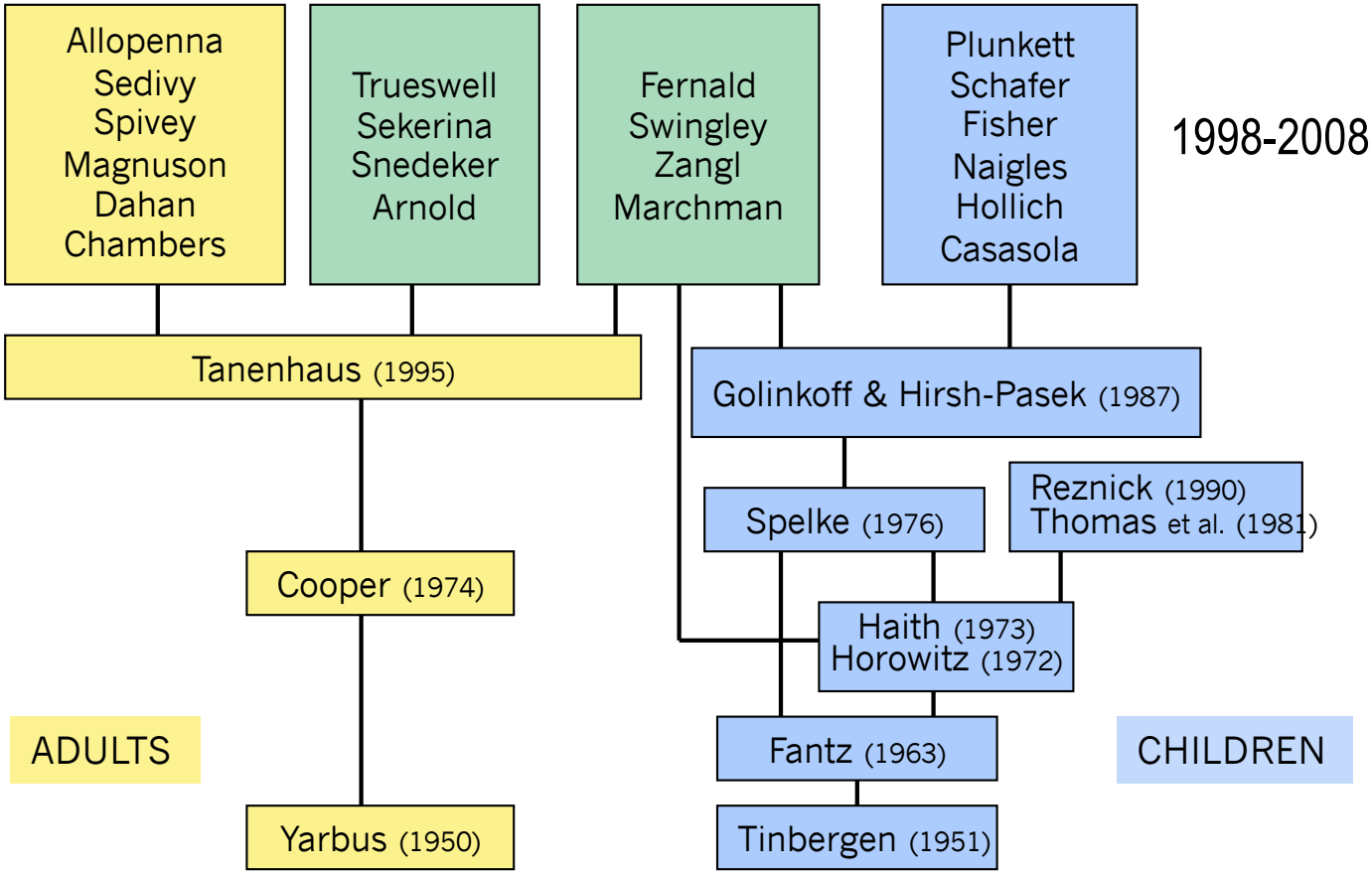
Hollich lab - Purdue

## Looking-While-Listening

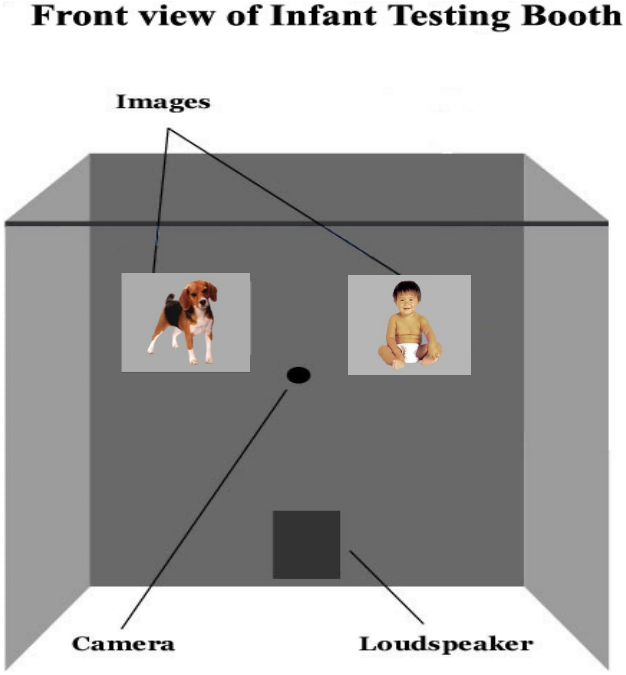



Fernald lab - Stanford

# A family tree of eye-tracking methods



# The "looking-while-listening" procedure




"Where's the baby?"

**Reaction Time (RT)**

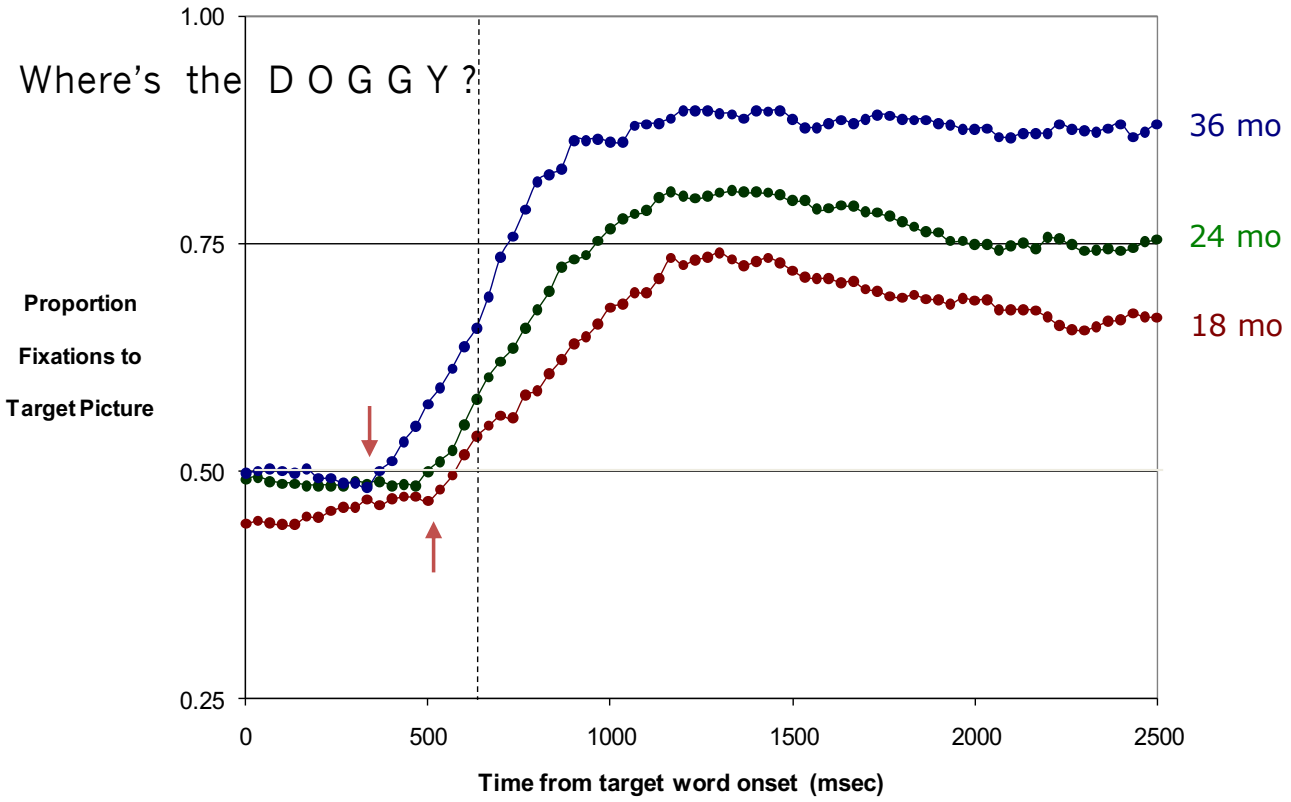
**Accuracy**

Fernald et al. (2008)

# What are some things we know so far?

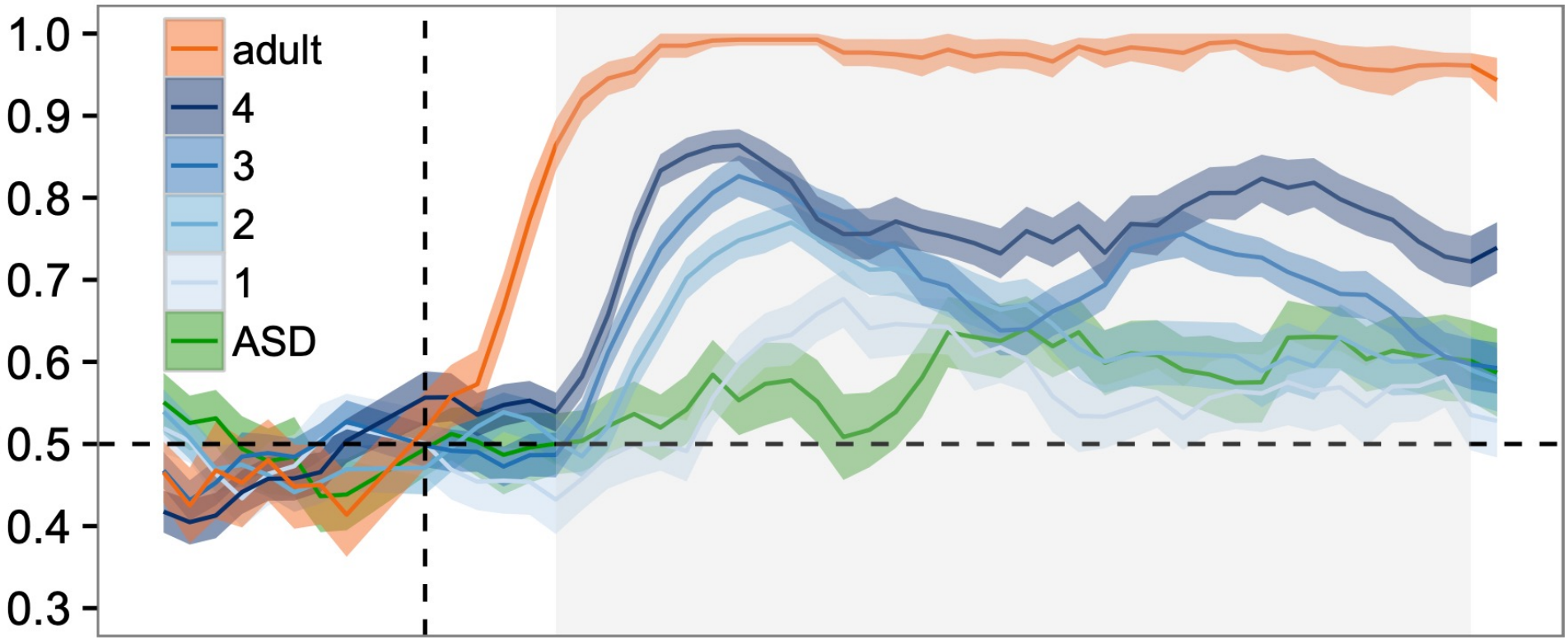
- Children, like adults, make sense of the language they hear *incrementally,* as it unfolds over time
  - attend to word initial information
  - identification within 300 msec of the word
- Infants get faster and more accurate over the 2nd and 3rd years of life
  - Age-related changes from 15- to 24-months of age
- Efficiency of speech processing is related to offline measures of language (e.g., vocabulary size and growth, standardized tests)

# Children become faster and more accurate



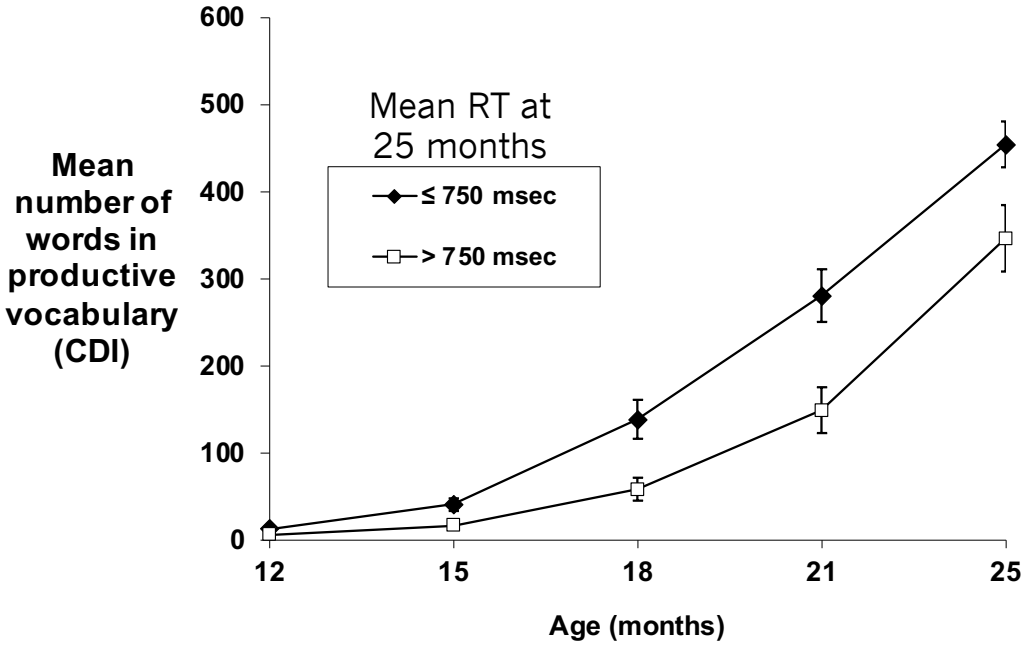

Zangl & Fernald (2007)

# This process continues in childhood




Yurovsky et al. (unpublished)

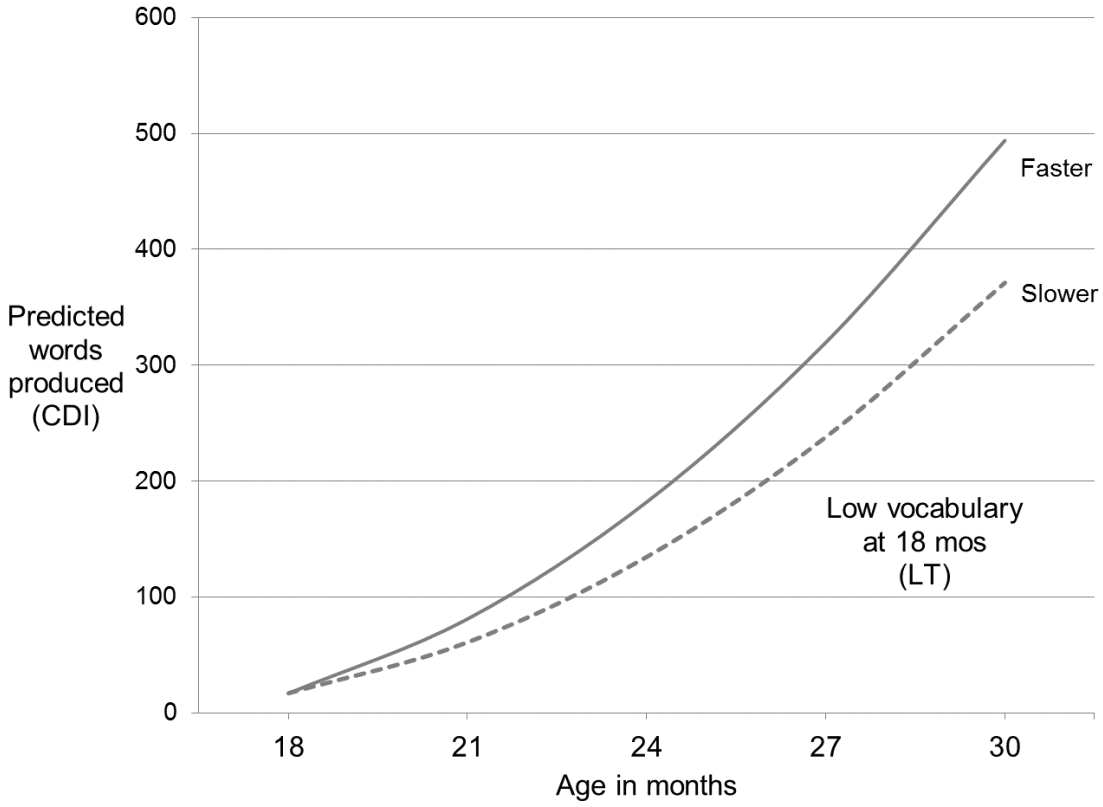
# Speech processing efficiency and vocabulary growth




Children with faster mean RT at 25 months were those who had accelerated more rapidly in vocabulary growth across the 2nd year

Fernald, Perfors & Marchman (2006)

# RT predicts which late talkers will “catch up”



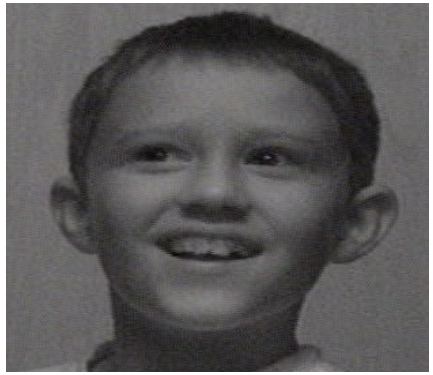

Fernald & Marchman (2012)

# Continuity into school age…

Language processing efficiency at 25 months predicts performance on standardized tests of IQ and working memory at 3rd grade

25 months

8 years



Marchman & Fernald (2008)

# How does having greater skill & speed in language processing help the child?

The child who can interpret familiar words more quickly

→ has more capacity to attend to words that follow

→ and thus to learn new words through inference

*Look at that kitty on the bench!*

*He's bigger than your teddy bear.*

# How does having greater skill & speed in language processing help the child?

The child who can interpret familiar words more quickly

- has more capacity to attend to words that follow
- and thus to learn new words through inference

*Look at that kitty on the bench!*

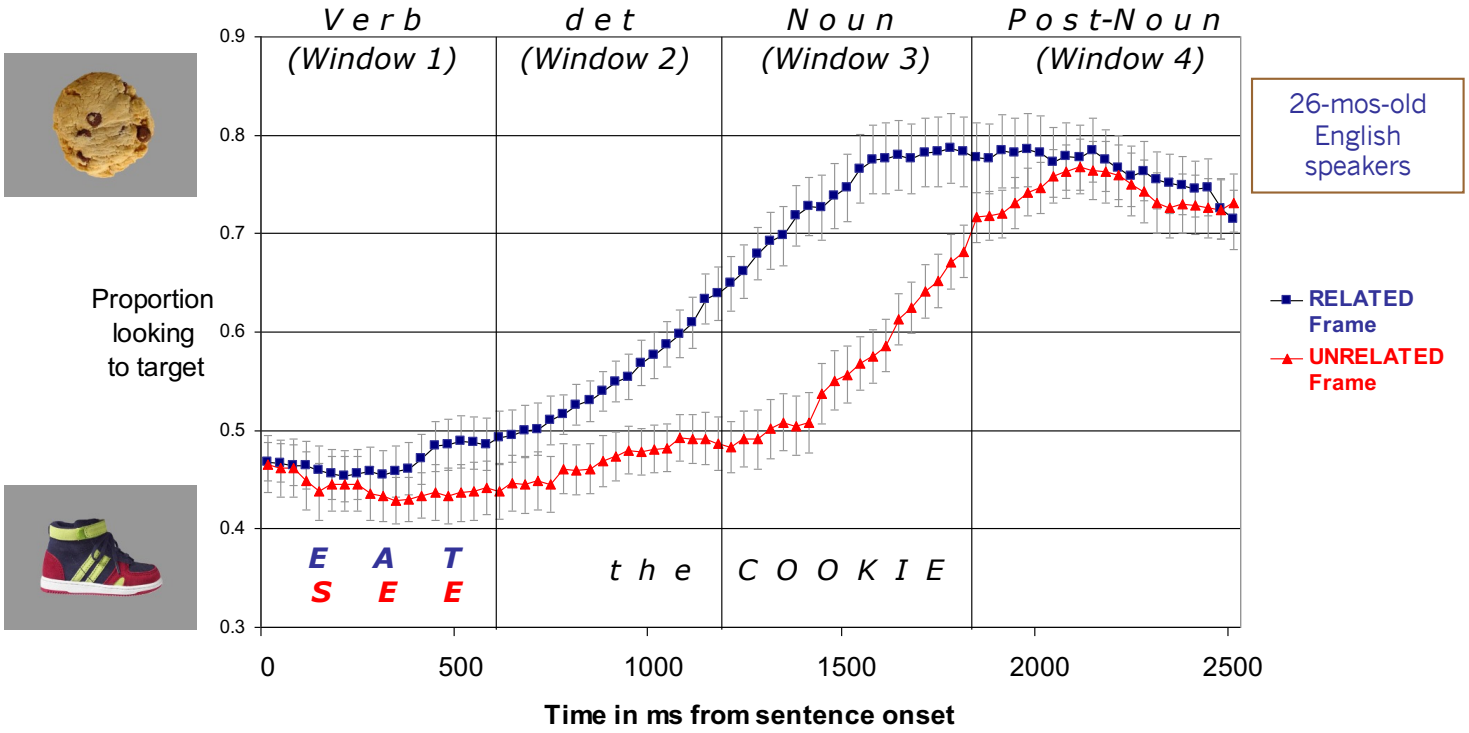
*He's bigger than your teddy bear.*

- strengthening working memory, inferential reasoning, and conceptual abilities

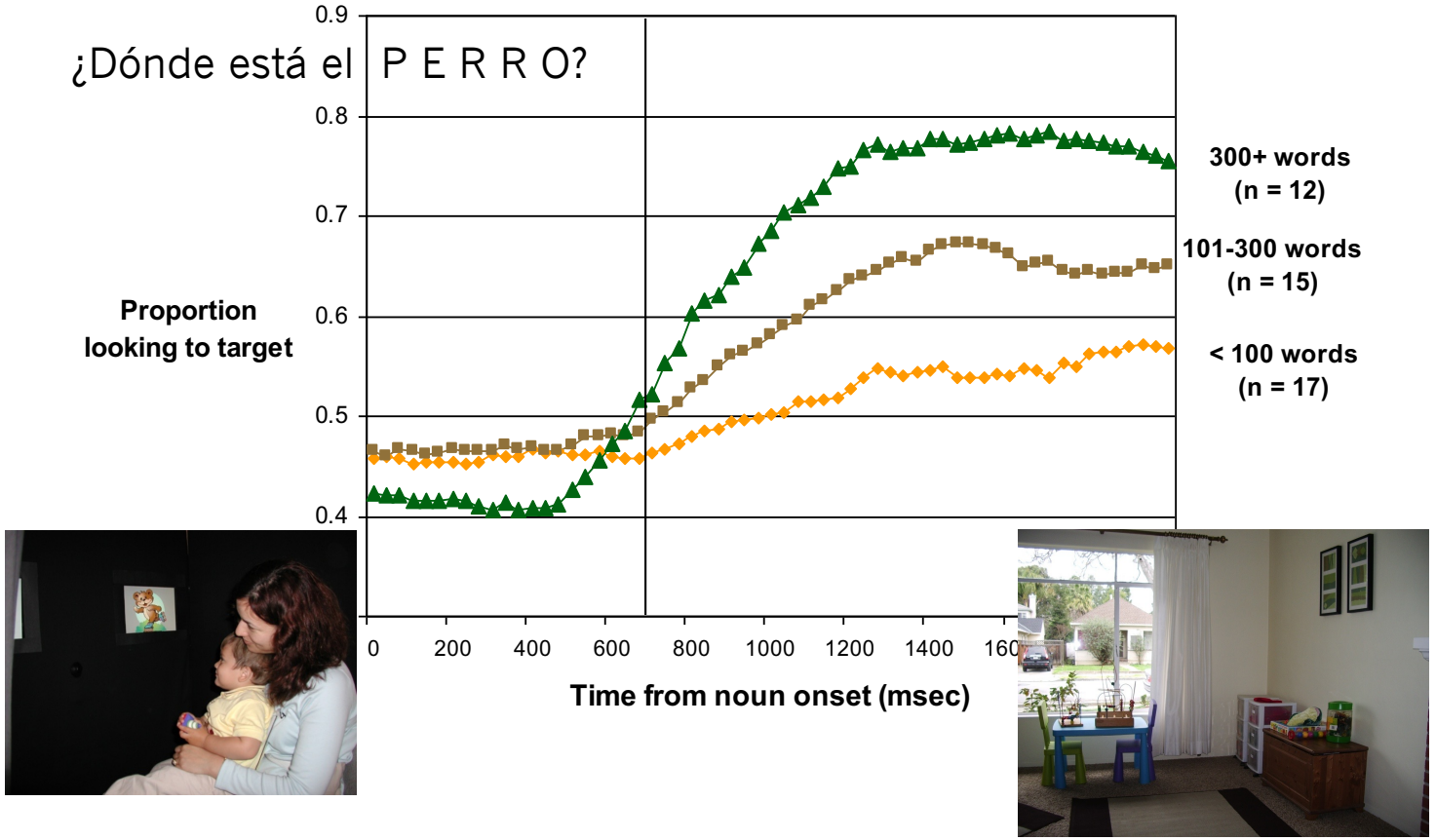
# Can children use *meaning* to "listen ahead"?

| UNRELATED frames | RELATED frames |
| --- | --- |
| Take the JUICE | Drink the JUICE |
| See the BALL | Throw the BALL |
| That's the BOOK | Read the BOOK |
| Where's the CAR | Drive the CAR |
| Find the BUNNY | Pat the BUNNY |
| It's the COOKIE | Eat the COOKIE |

# English-speakers begin to shift to the correct picture sooner when the verb is informative



# Spanish-learning children with larger vocabularies are more accurate



Spanish provides morphosyntactic information
that English does not:

Encuentra la galleta! Find the cookie!

Encuentra el zapato! Find the shoe!

Can young children learning Spanish take advantage
of such language-specific cues to reference in
processing continuous speech?

¿Dónde está el pájaro?

SAME gender

*el pájaro*

*el caballo*

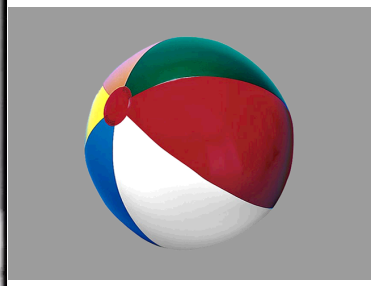
DIFFERENT gender

*el pájaro*

*la vaca*

## SAME-GENDER TRIALS

la galleta

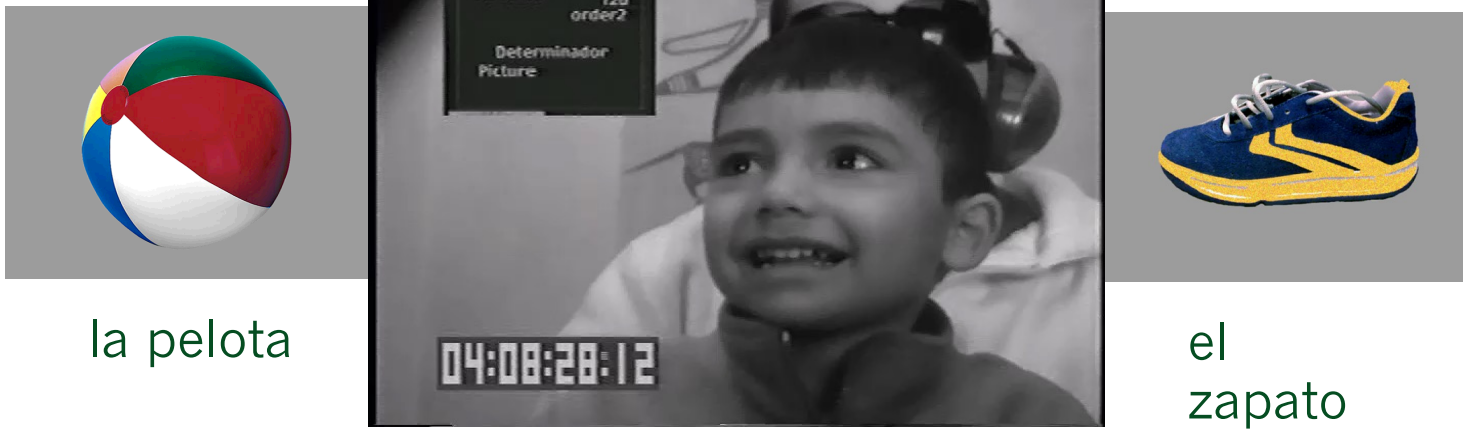
la pelota

"Encuentra la pelota"

Mean RT: 935 msec

Lew-Williams & Fernald (2007)

# DIFFERENT-GENDER TRIALS



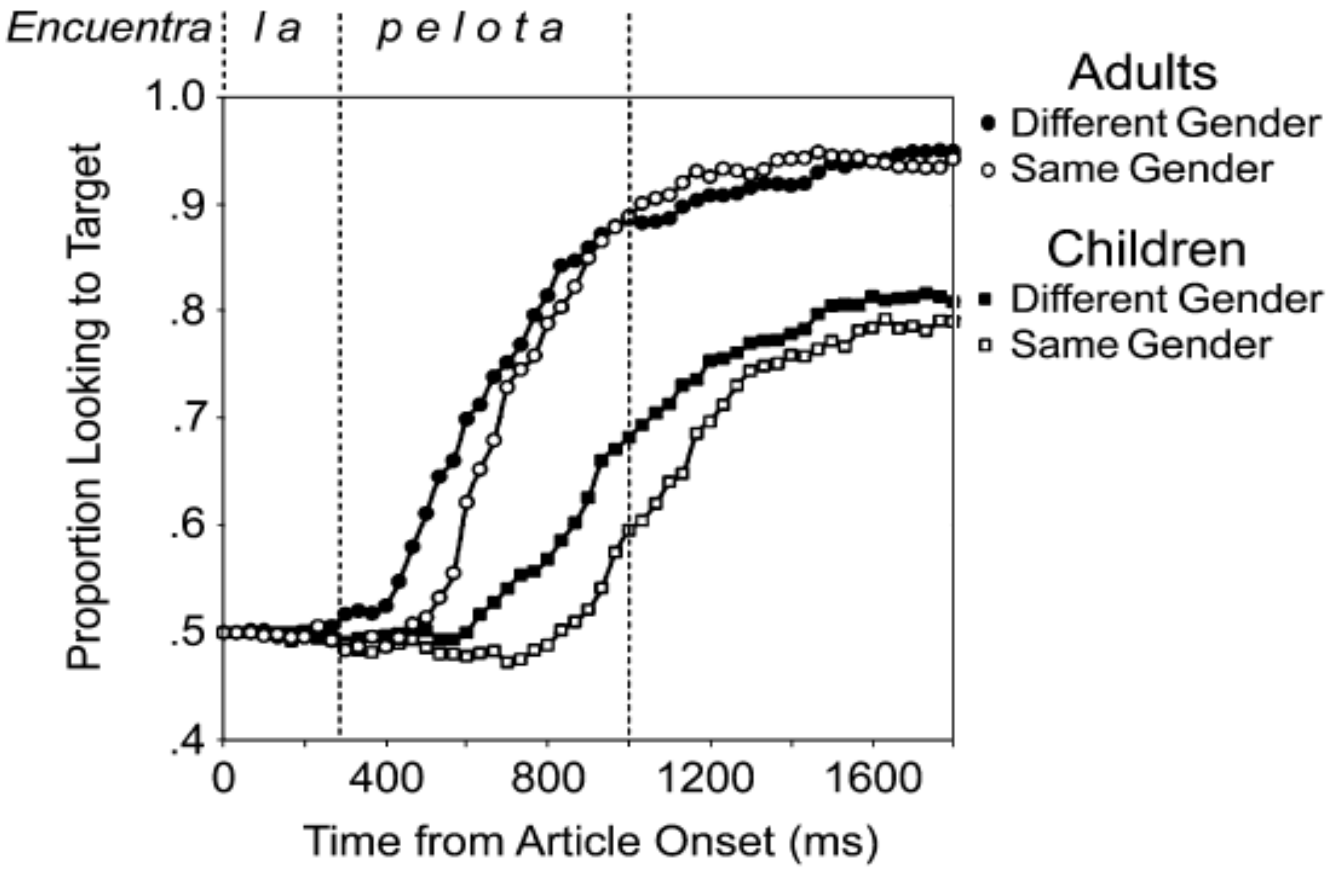

la pelota

el
zapato

“Encuentra la pelota”

Mean RT: 842 msec

Lew-Williams & Fernald (2007)




Lew-Williams & Fernald (2007)

# Outline

1. Introducing the looking-while-listening method
2. **Peekbank, a new resource**
3. Explorations of LWL methods using Peekbank
4. Insights into online word recognition using Peekbank

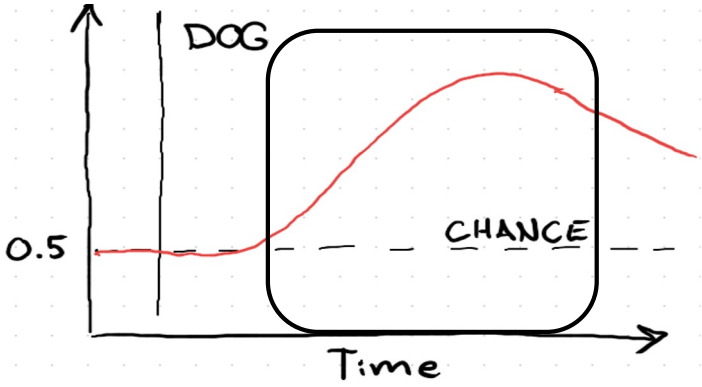
## My first word recognition study

Approximate rendering of me as a grad student

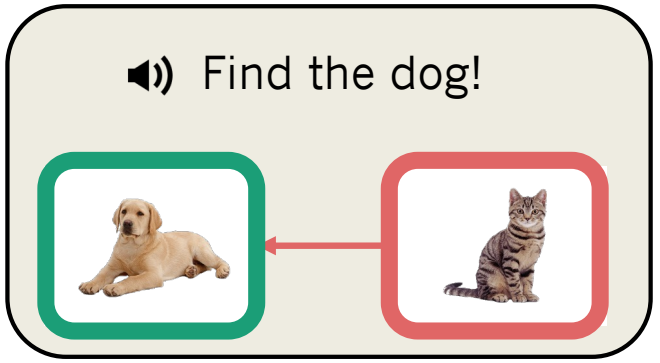

looking-while-listening paradigm

(Fernald et al., 2008; intermodal preferential looking paradigm, Hirsh-Pasek et al., 1987; visual world paradigm)

Accuracy




Reaction Time
(speed of processing)



## My first word recognition study: Many questions

What kinds of effects can I expect?

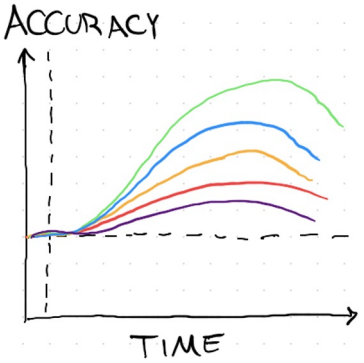
What items should I choose?

How should I analyze my data?

## My first word recognition study: The dream

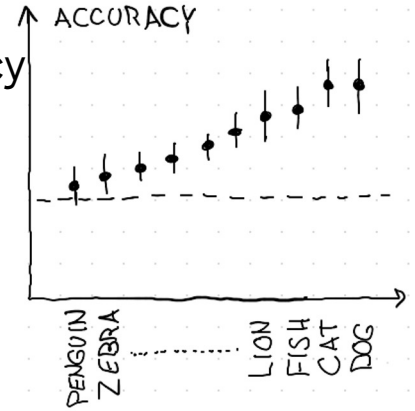
What kinds of effects can I expect?

Centralized information about effect sizes and recognition curves at different ages as well as variability



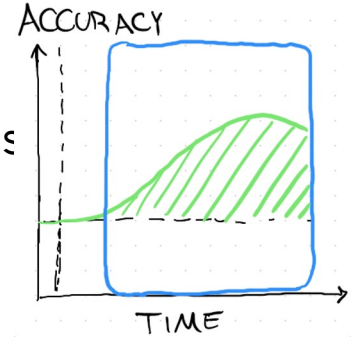

What items should I choose?

Information about item-level accuracy and variability across age




How should I analyze my data?

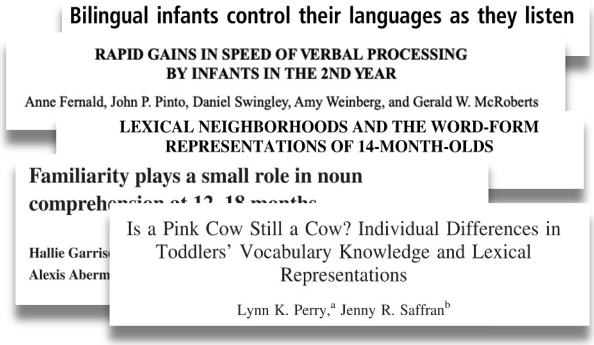
Data-based methodological consensus
Tools for understanding different analysis choices



# My first word recognition study: Reality

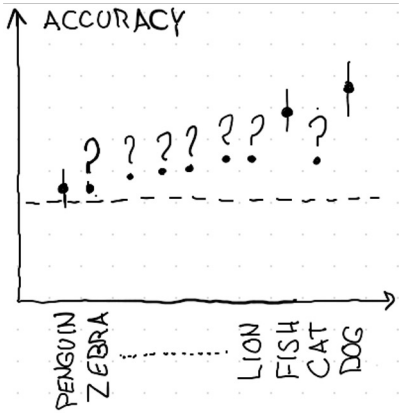
What kinds of effects can I expect?

Information about effect sizes and typical responses scattered across many (relatively small) studies



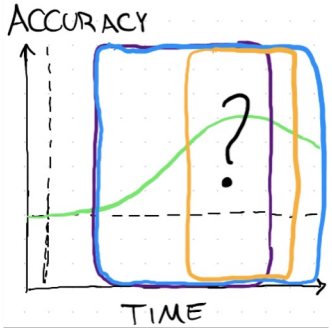

What items should I choose?

Limited information about item-level effects from individual studies




How should I analyze my data?

Widely varying recommendations and practices



peekbank Data Access Tutorial Visualizations Contributors Documentation Releases About

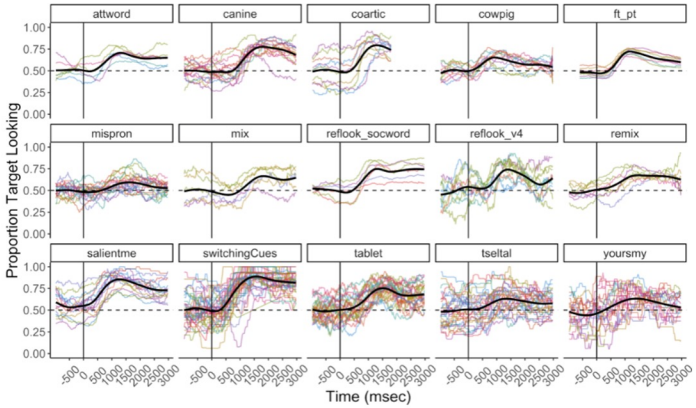
# peekbank

A flexible and reproducible interface to developmental eyetracking datasets

## What is peekbank?

peekbank is a flexible and reproducible interface to developmental eyetracking datasets.

The Peekbank project is an open database storing eye-tracking datasets on children's word recognition in a well-documented, easily accessible, tabular format. It also provides processing tools for standardizing eye-tracking data across data sources ( `peekds` R package), interfaces for accessing the database ( `peekbankr` R package), and applications for visualizing the data (Peekbank Shiny App).




### Data Access Tutorial

Get started accessing the data here

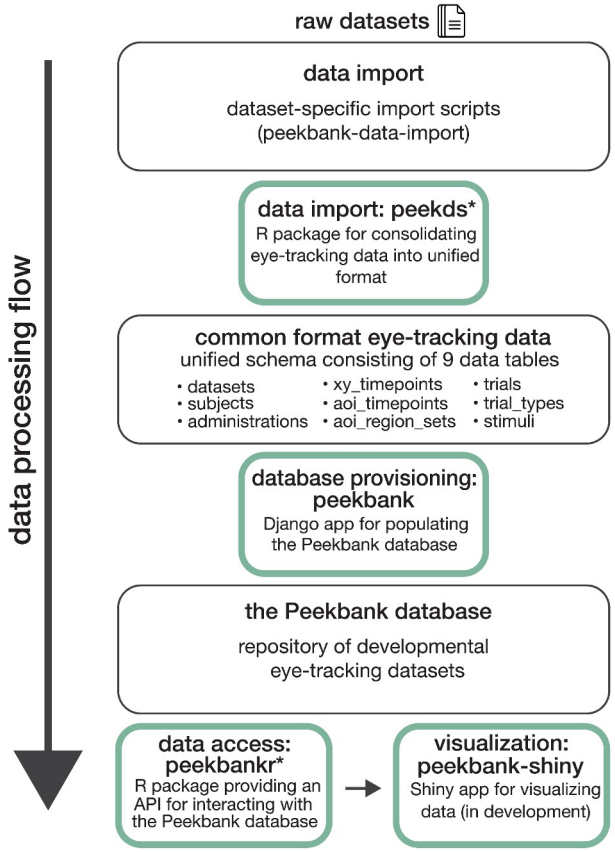
### Interactive Visualizations

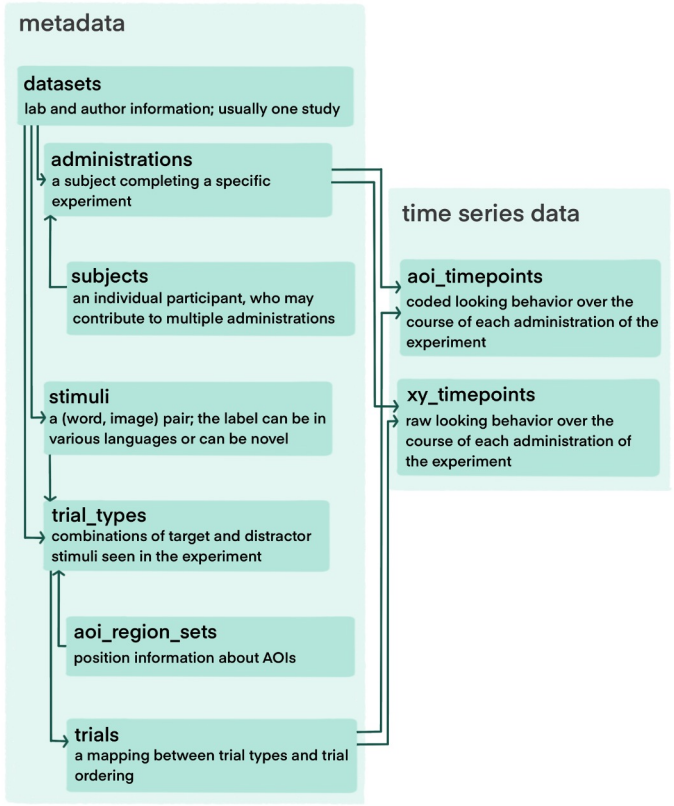
Dynamically generate visualizations of the data

### Documentation

Further information about the database

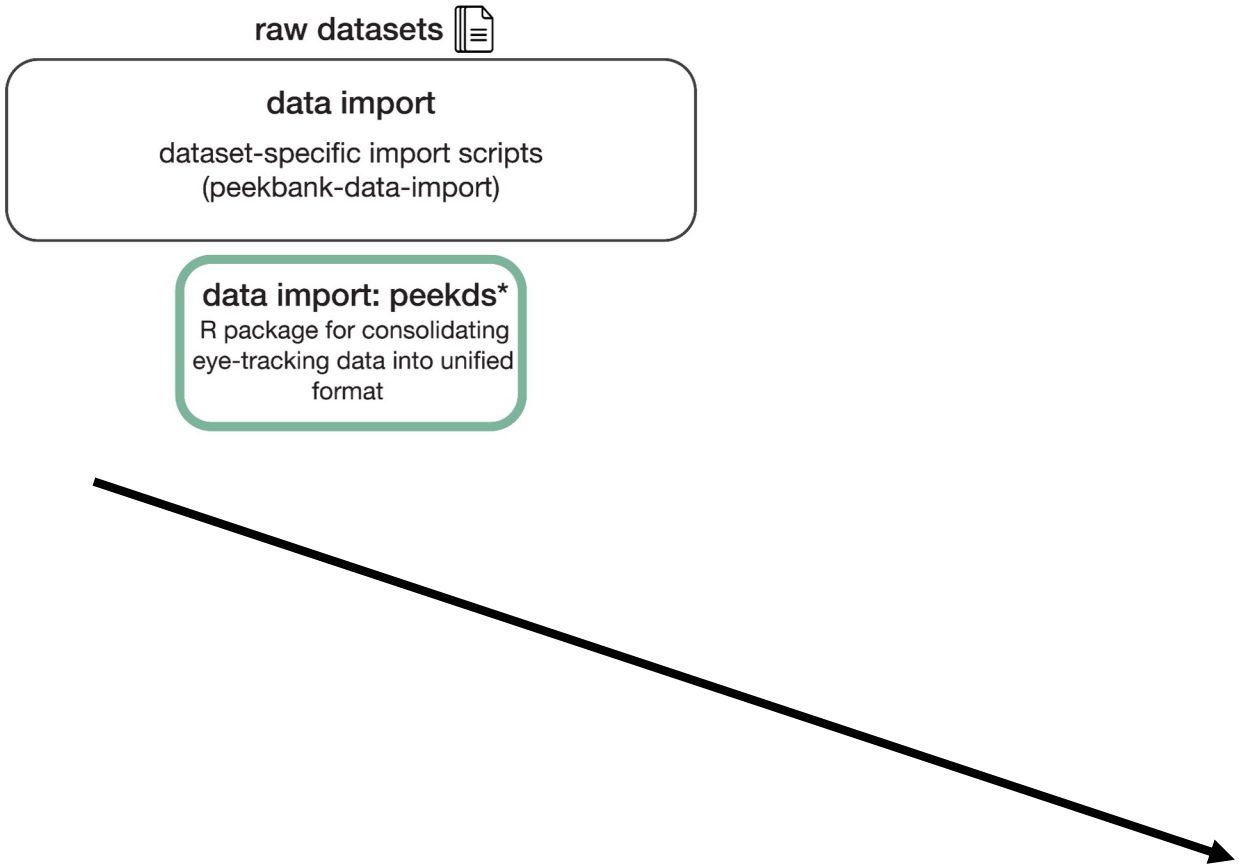
# Peekbank: An open, large-scale database of looking-while-listening studies



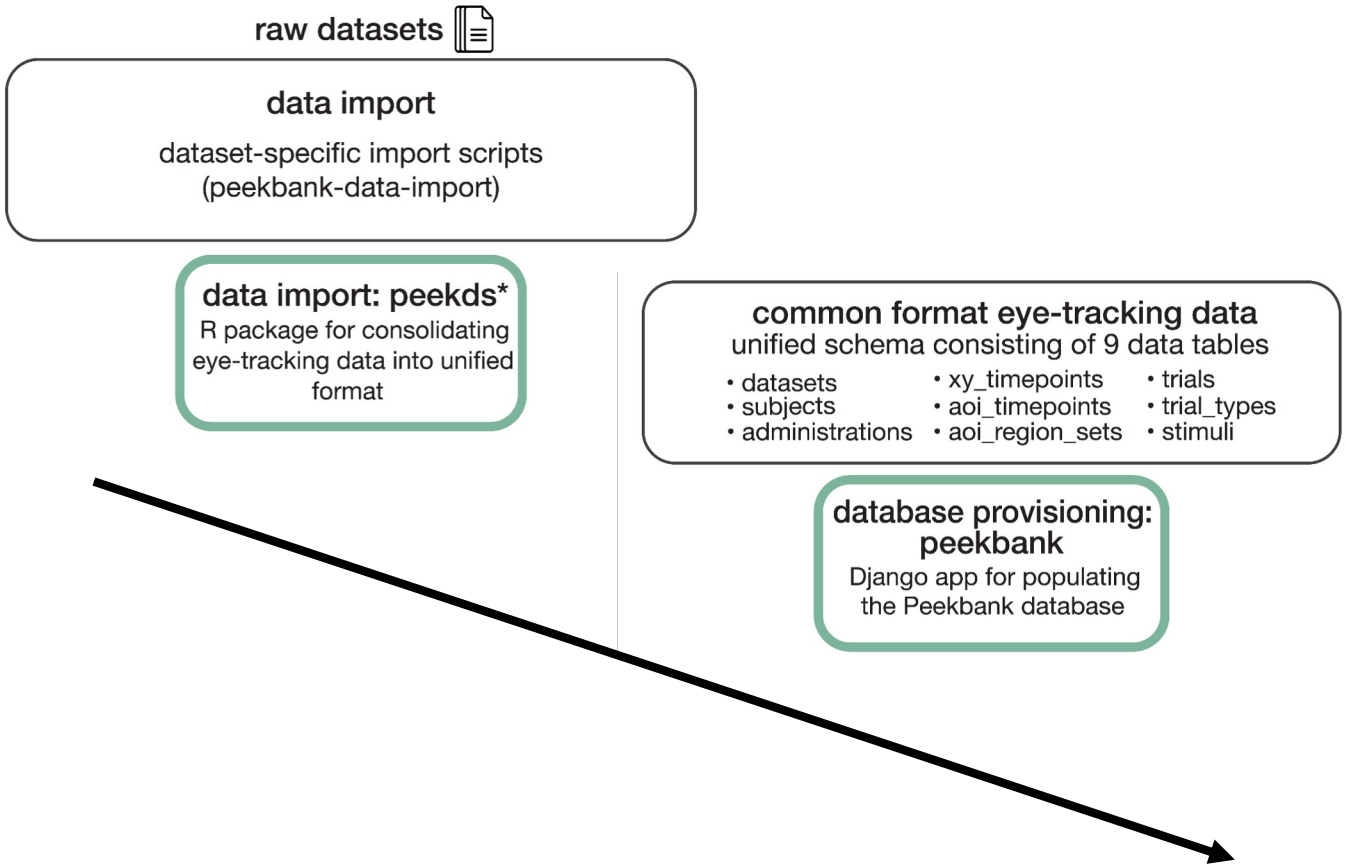




Zettersten et al. (2021), CogSci; Zettersten et al. (2022), Behavior Research Methods

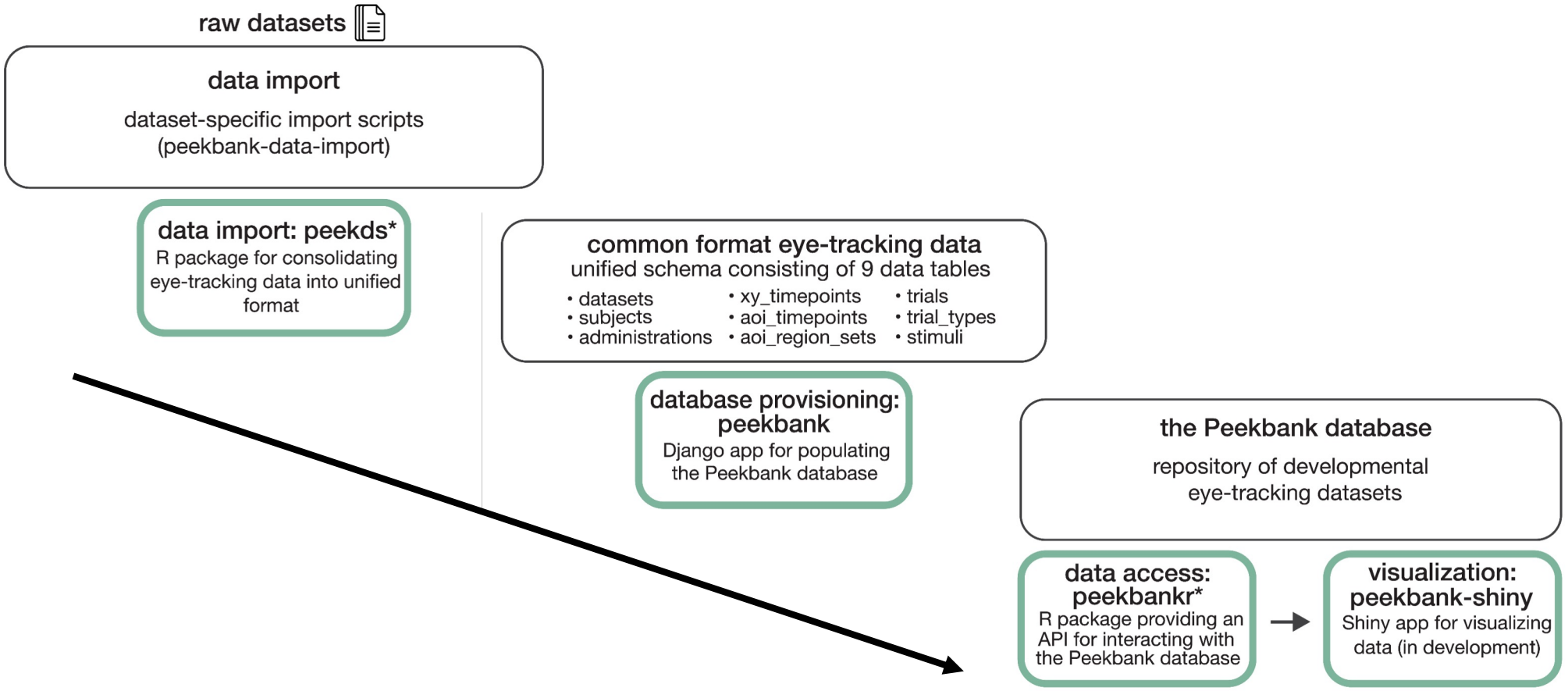
# Peekbank Data Processing Flow



# Peekbank Data Processing Flow



# Peekbank Data Processing Flow

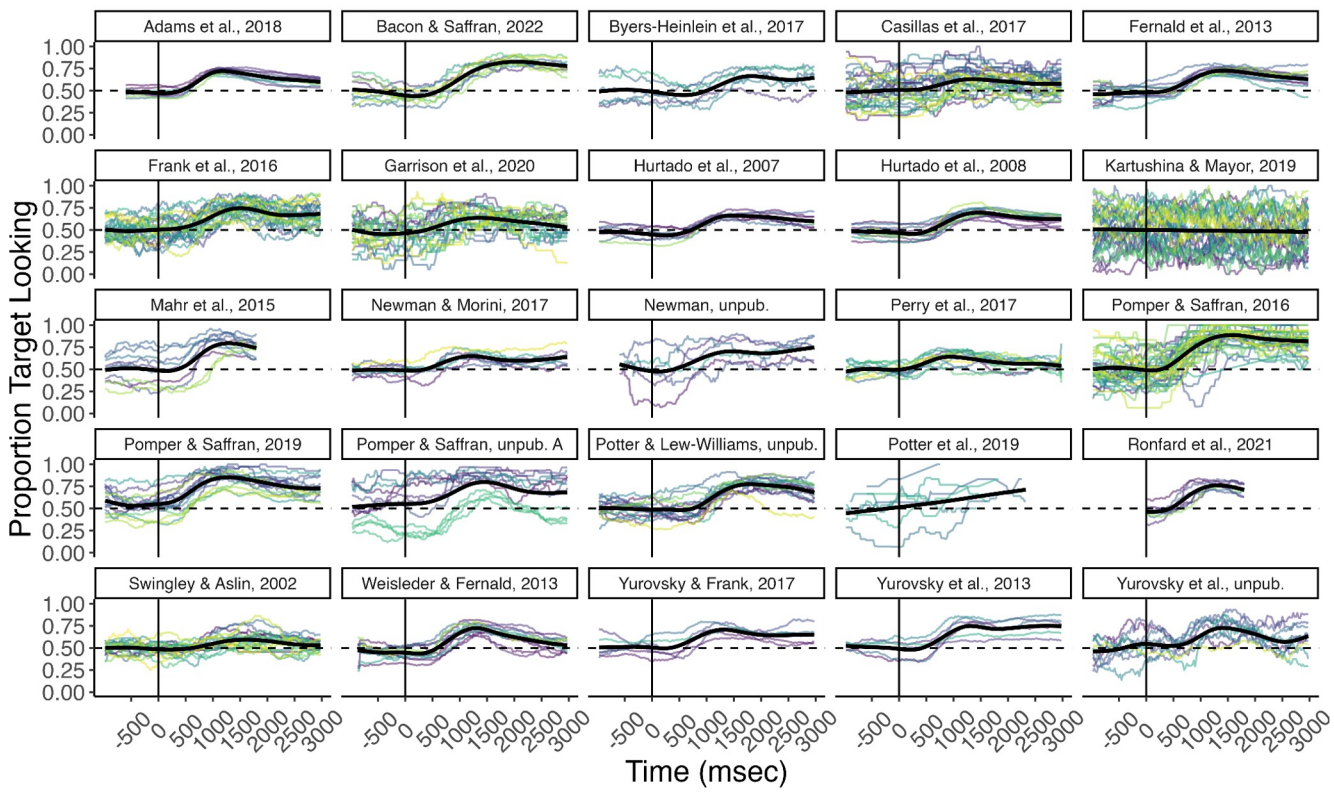


# The Peekbank database

Seed data aggregated from team-internal datasets and informal requests

Current database:

- 25 datasets
- N = 1807
- wide age range (6-70 mos.)
- Mostly English monolingual participants (19/25 datasets)



# Outline

1. Introducing the looking-while-listening method
2. Peekbank, a new resource
3. **Explorations of LWL methods using Peekbank**
4. Insights into online word recognition using Peekbank

# Swingley & Aslin (2002)

PSYCHOLOGICAL SCIENCE

## Research Report

### LEXICAL NEIGHBORHOODS AND THE WORD-FORM REPRESENTATIONS OF 14-MONTH-OLDS

Daniel Swingley[1] and Richard N. Aslin[2]

*[1]Max-Planck-Institute for Psycholinguistics, Nijmegen, the Netherlands, and [2]Department of Brain and Cognitive University of Rochester*

**Abstract**—*The degree to which infants represent phonetic detail in words has been a source of controversy in phonology and developmental psychology. One prominent hypothesis holds that infants store words in a vague or inaccurate form until the learning of similar-sounding* neighbors *forces attention to subtle phonetic distinctions. In the experiment reported here, we used a visual fixation task to assess word recognition. We present the first evidence indicating that, in fact, the lexical representations of 14- and 15-month-olds are encoded in fine detail, even when this detail is not functionally necessary for distinguishing similar words in the infant's vocabulary. Exposure to words is sufficient for well-specified lexical representations, even well before the vocabulary spurt. These results suggest developmental continuity in infants' representations of speech: As infants begin to build a vocabulary and learn word meanings, they use the perceptual abilities previously demonstrated in tasks testing the discrimination and categorization of meaningless syllables.*

**Table 1.** *Correctly pronounced (CP) target words and their mispronounced (MP) versions*

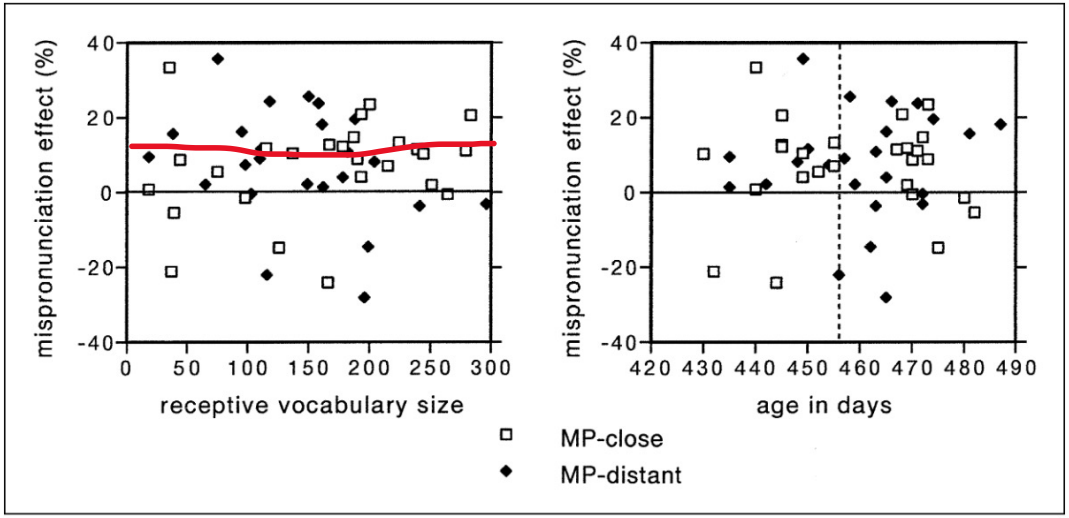
| CP | MP-close | MP-distant |
|---|---|---|
| apple (/æpl̩/) | opple (/apl̩/) | opal (/opl̩/) |
| baby (/beᴵbi/) | vaby (/veᴵbi/) | raby (/ɹeᴵbi/) |
| ball (/bɔl/) | gall (/gɔl/) | shawl (/ʃɔl/) |
| car (/kaɹ/) | cur (/kɜ˞/) | kier (/kiɹ/) |
| dog (/dɔg/) | tog (/tɔg/) | mog (/mɔg/) |
| kitty (/kɪti/) | pity (/pɪti/) | yitty (/jɪti/) |

*Note*. International Phonetic Alphabet transcriptions are provided in parentheses. Bisyllabic words were stressed on the first syllable.

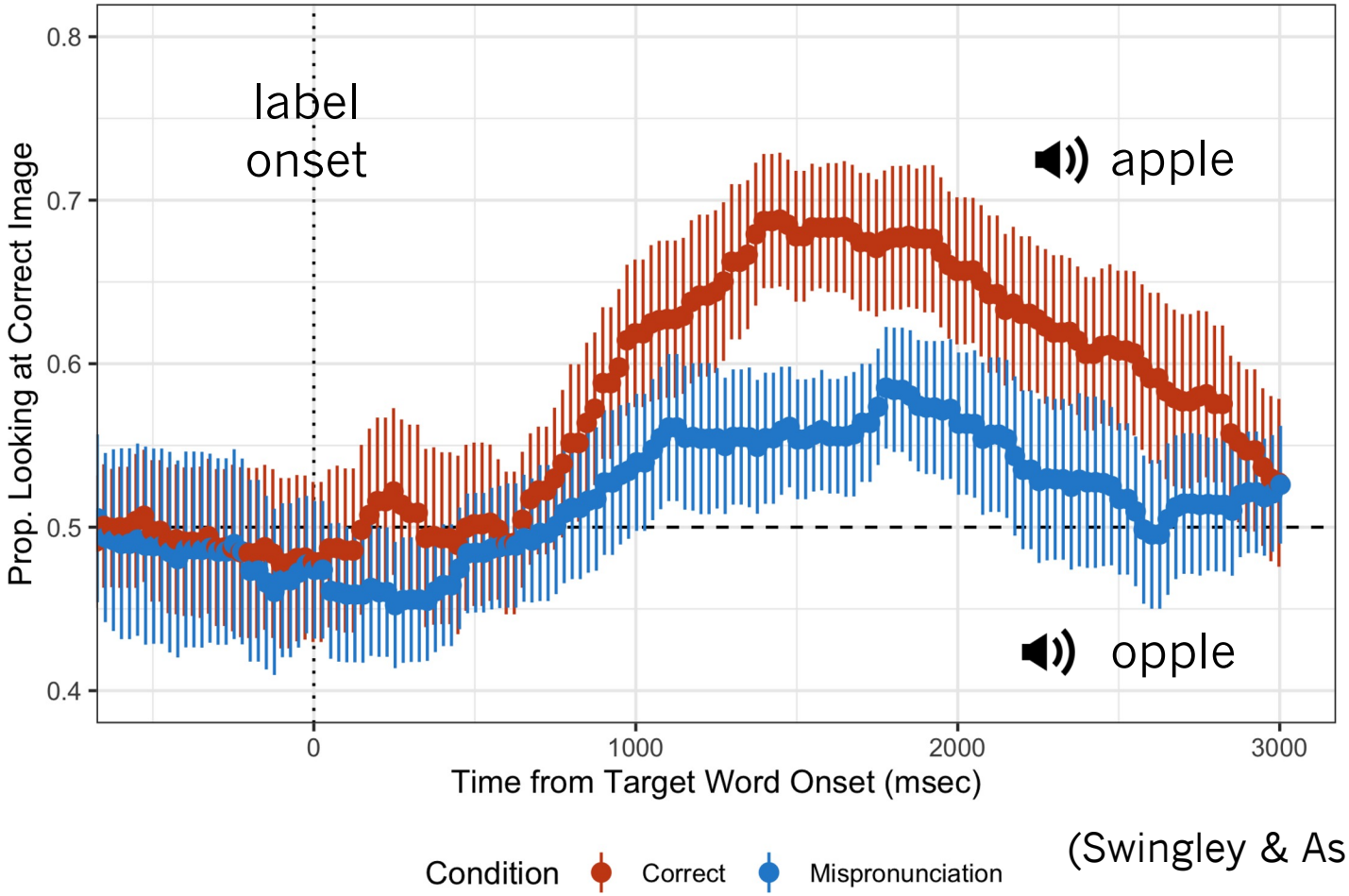
months of age frequently failed to choose the corr performance was good when the words were not similar. This pattern of results suggests that words are not encoded with full phonetic de tail.

The discrepancy betv mance and young childre has been attributed to tas discrimination procedure referential tasks requirin (e.g., Locke, 1988). Of c for the discrepancy in fir nature of children's lexi assumed (partly on the dren's perceptual repres in press), while others ha well, 1975; Menyuk, Me

Several researchers ha

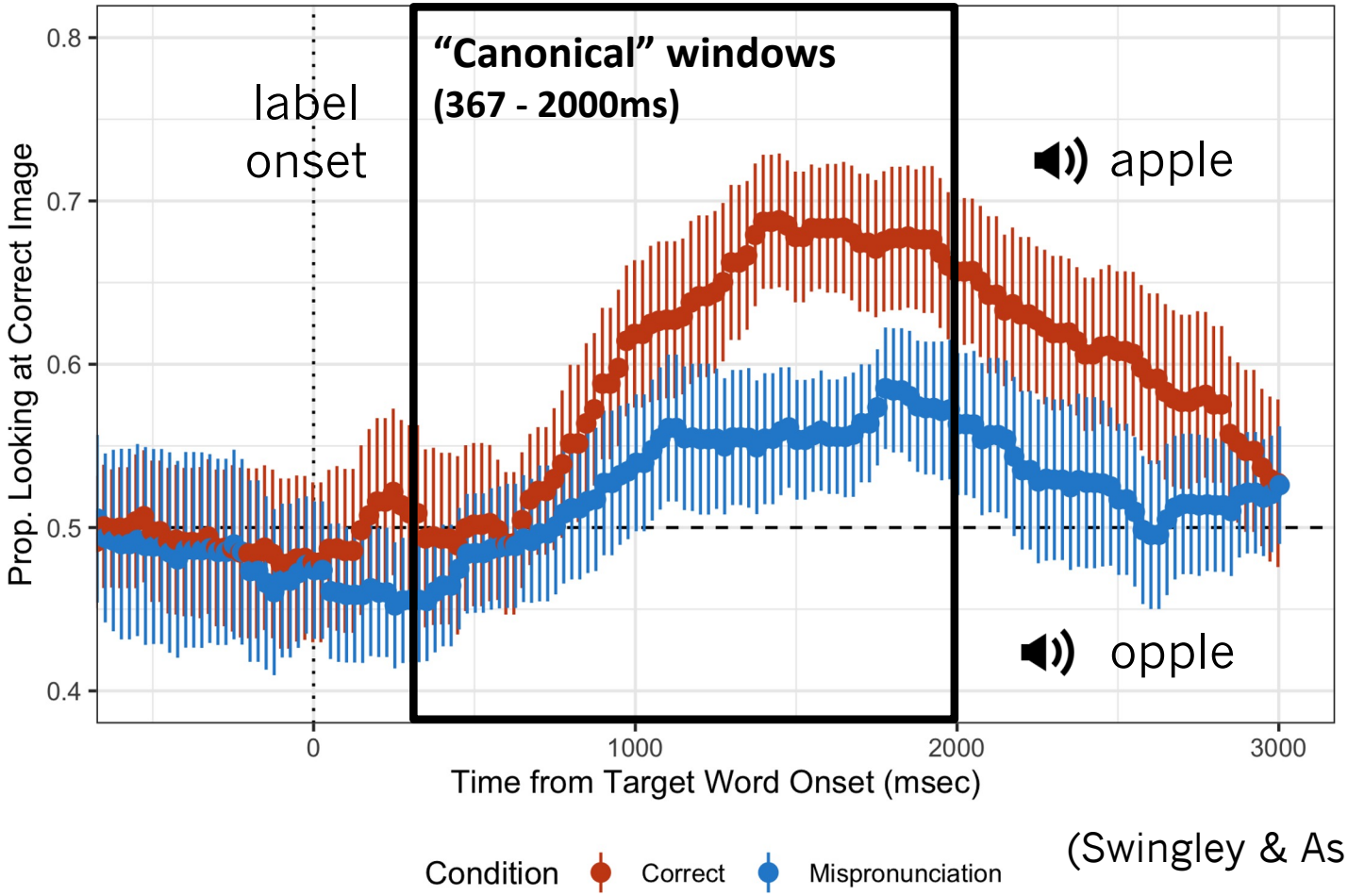


# Measures: The problem of selecting time windows of analysis



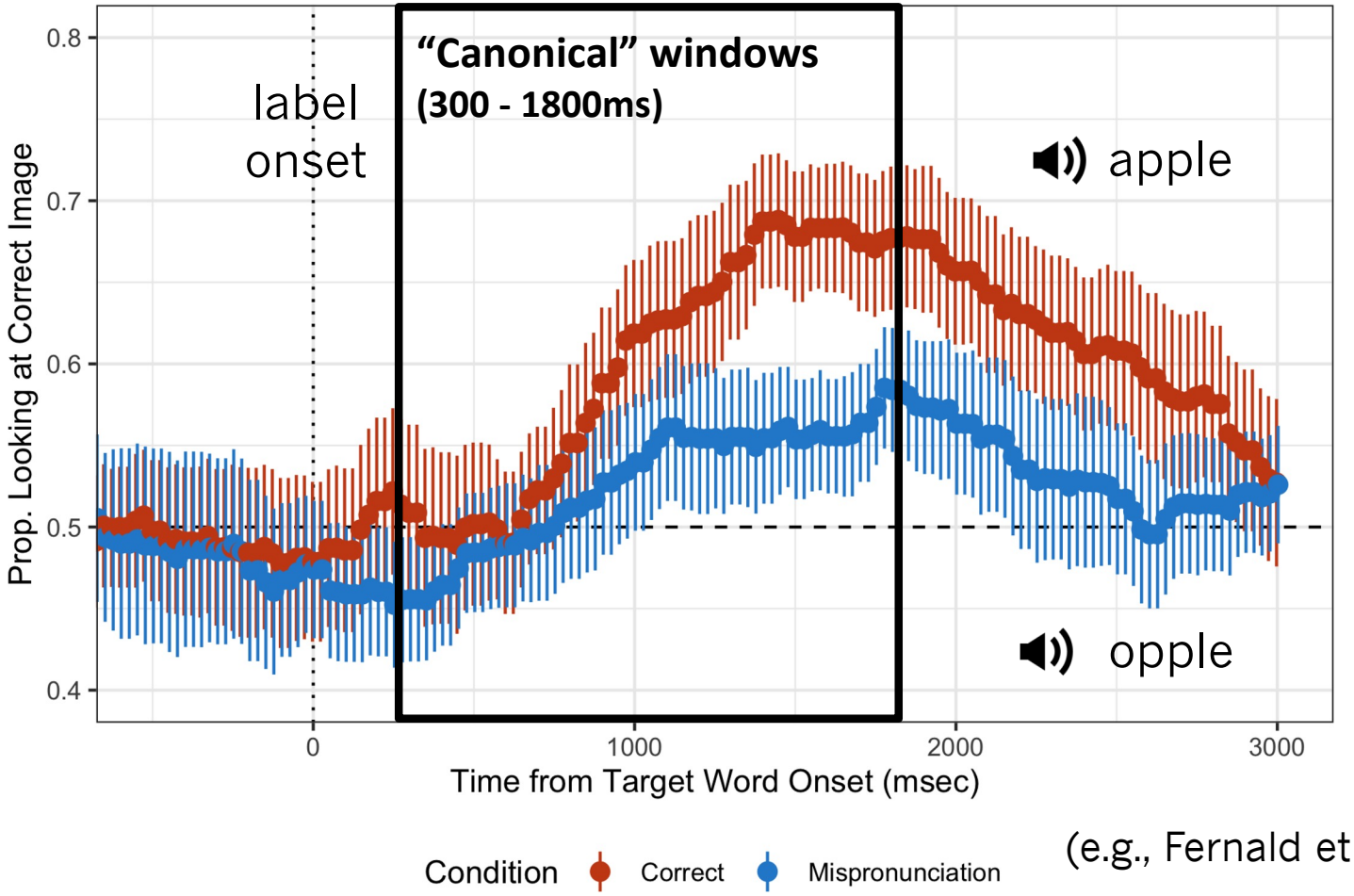

(Swingley & Aslin, 2002)

# Measures: The problem of selecting time windows of analysis



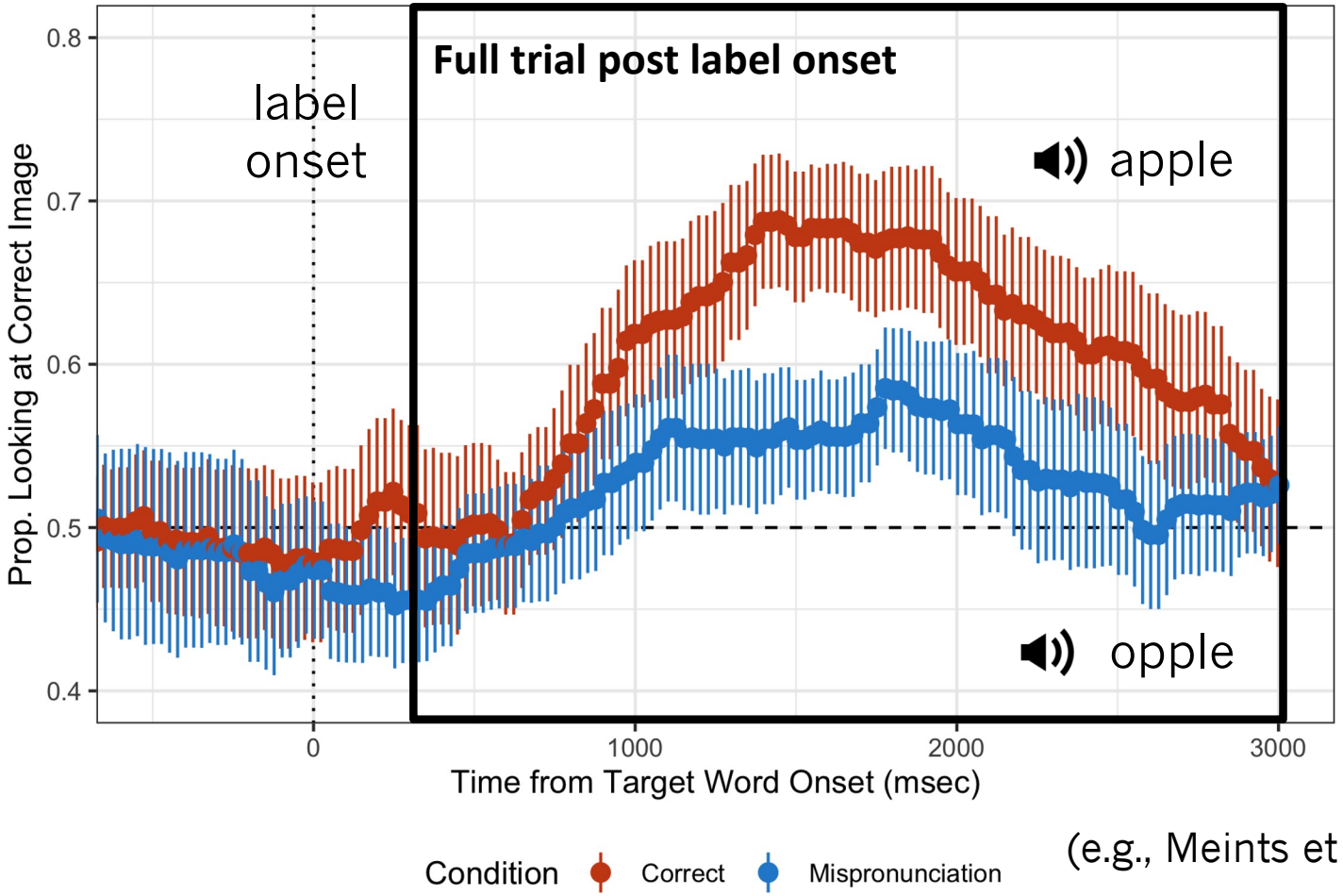

(Swingley & Aslin, 2002)

# Measures: The problem of selecting time windows of analysis



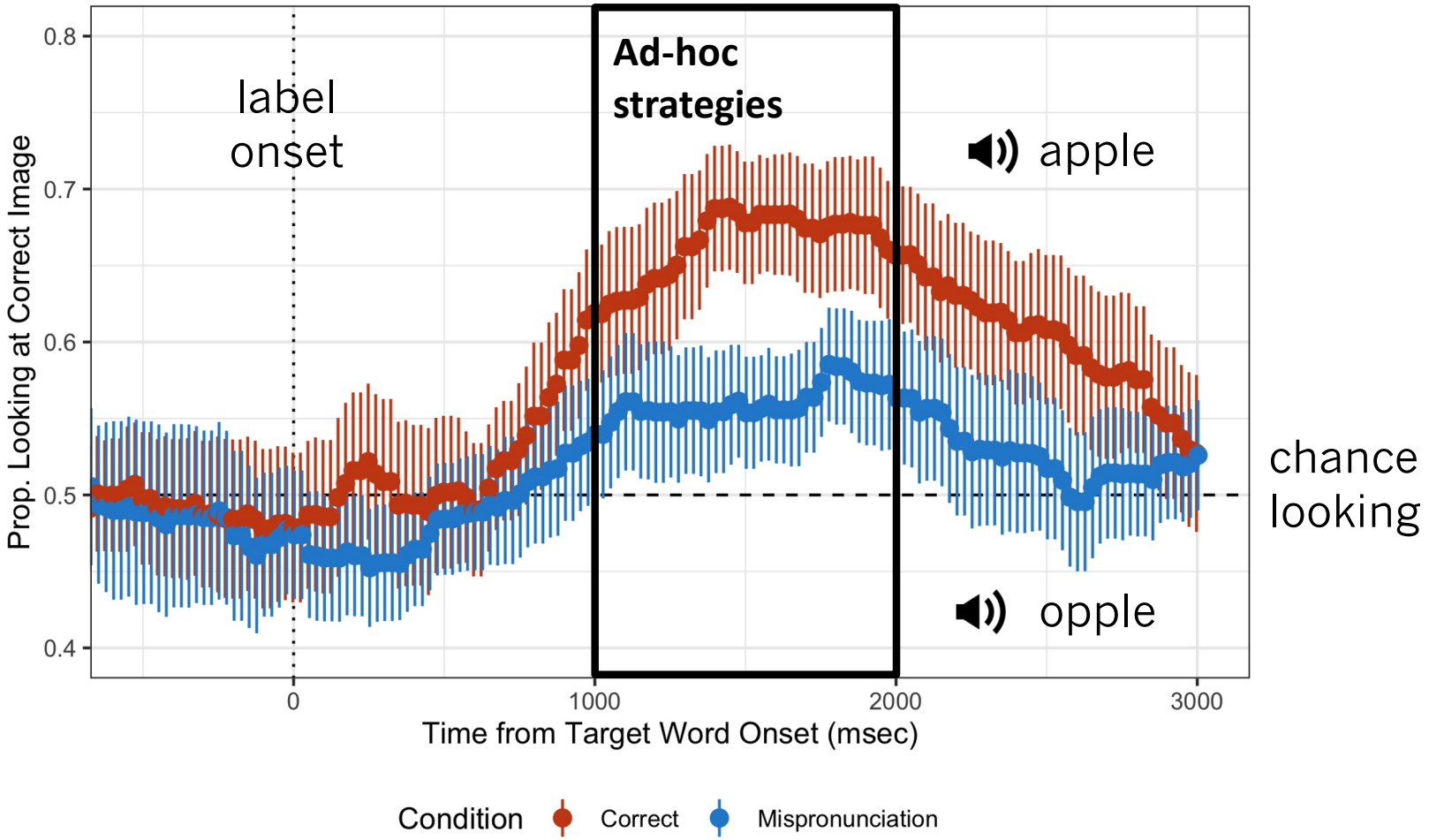

(e.g., Fernald et al., 2008)

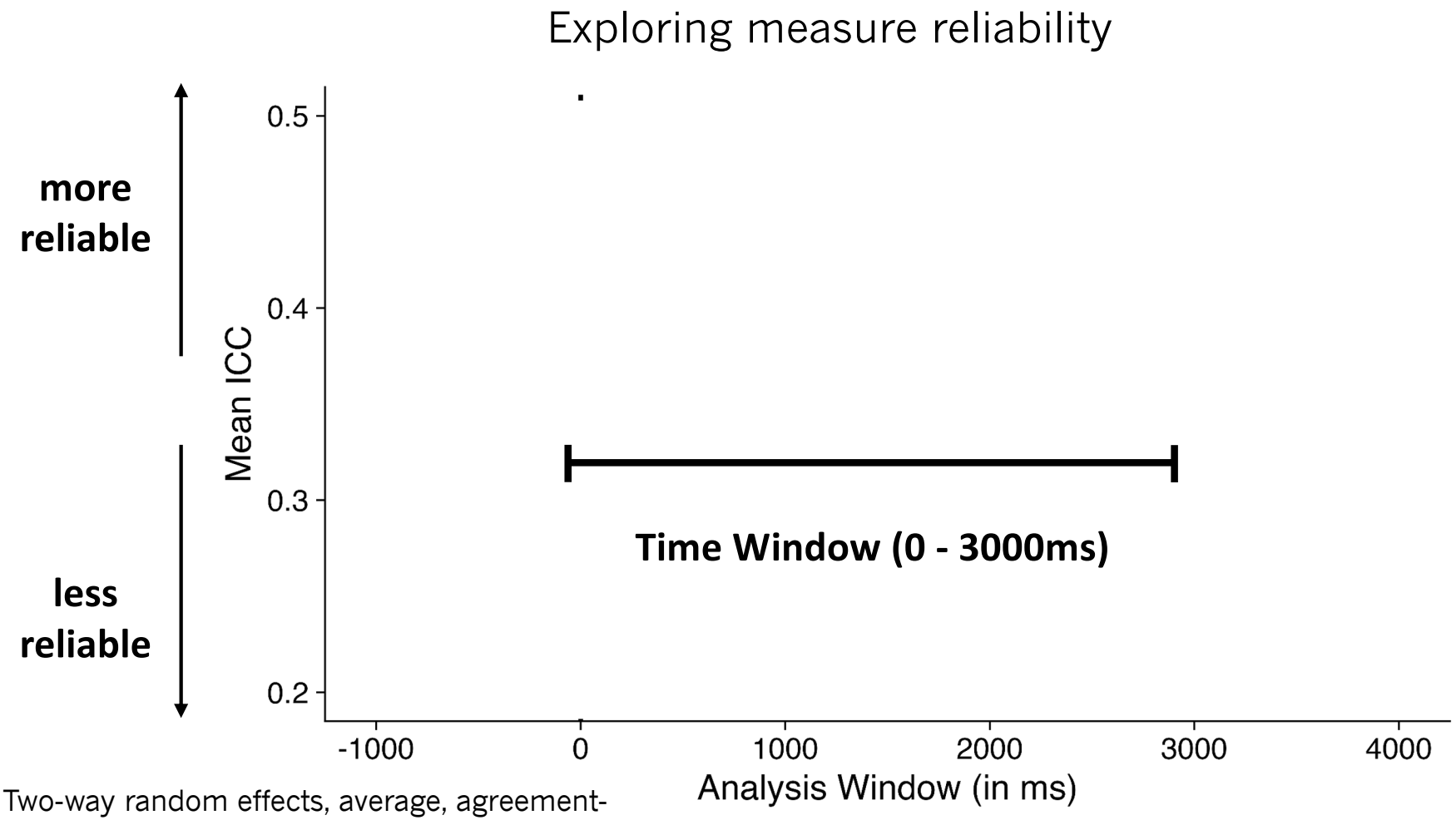
# Measures: The problem of selecting time windows of analysis



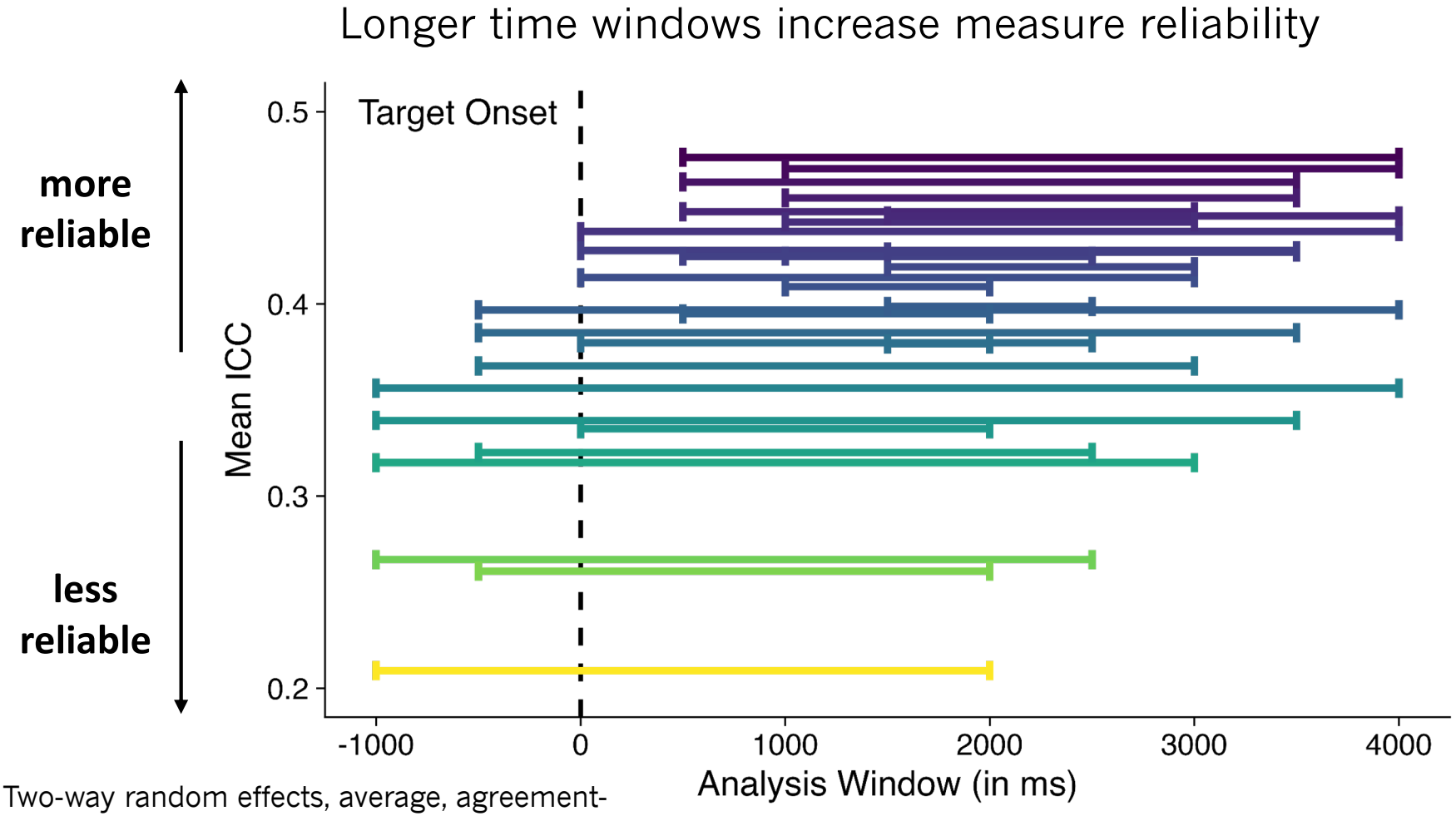

(e.g., Meints et al., 1999)

# Measures: The problem of selecting time windows of analysis

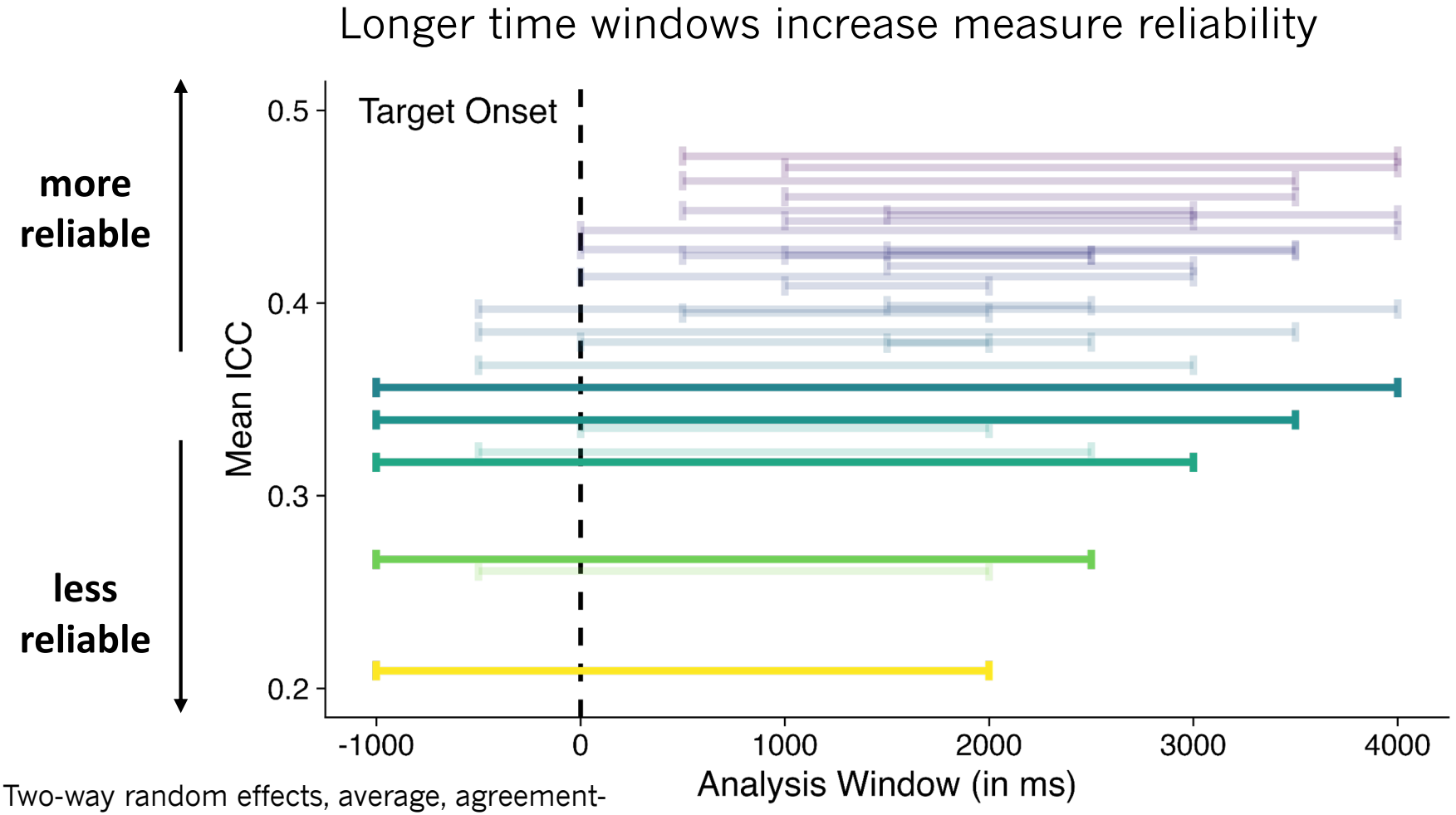




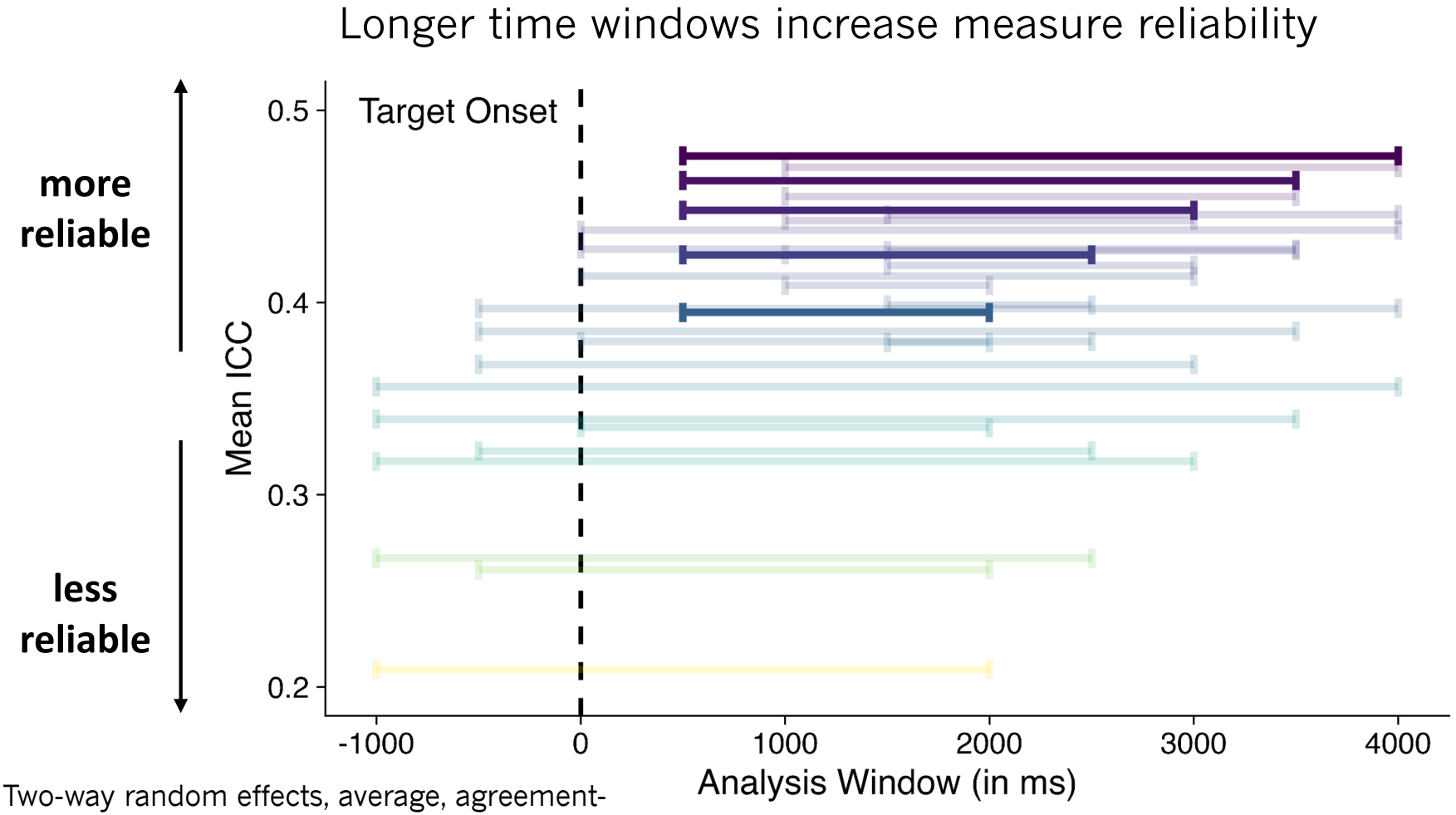

Two-way random effects, average, agreement-based ICC, by dataset (Case 2A; McGraw & Wong, 1996)



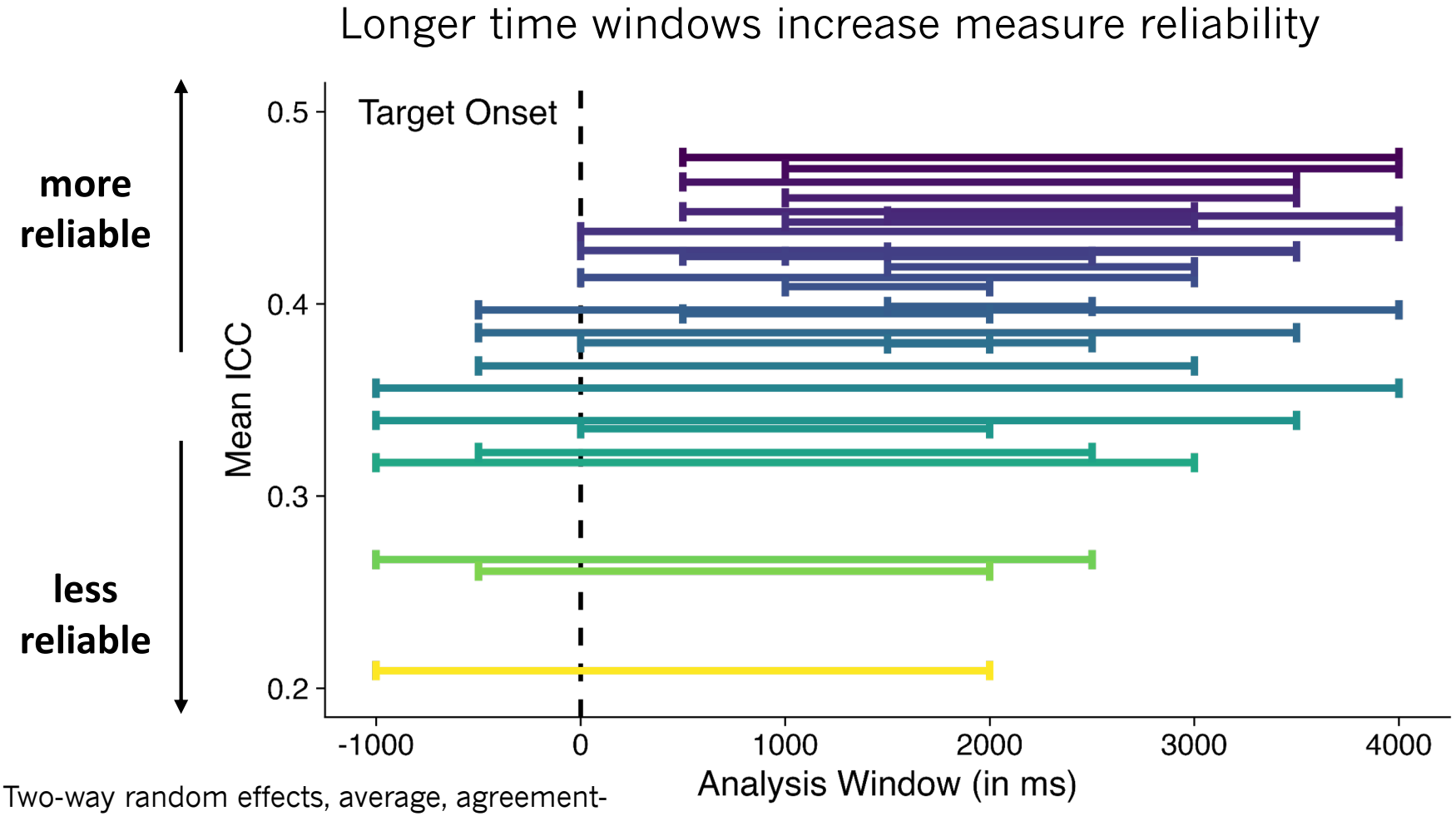

Two-way random effects, average, agreement-based ICC, by dataset (Case 2A; McGraw & Wong, 1996)




Two-way random effects, average, agreement-based ICC, by dataset (Case 2A; McGraw & Wong, 1996)




Two-way random effects, average, agreement-based ICC, by dataset (Case 2A; McGraw & Wong, 1996)




Two-way random effects, average, agreement-based ICC, by dataset (Case 2A; McGraw & Wong, 1996)

# Outline

1. Introducing the looking-while-listening method
2. Peekbank, a new resource
3. Explorations of LWL methods using Peekbank
4. **Insights into online word recognition using Peekbank**

Estimating developmental change in online word recognition accuracy & speed

Analytic Challenges:

- Individual studies are often constrained to small item sets and age ranges
- Items are typically selected based on child age
    - dog may be tricky for 12-mo-olds, but “too easy” for 24-mo-olds

→ Peekbank allows us to estimate change across items, ages, & participants

Questions:

- Are there overarching patterns (“laws”) in developmental change in word recognition?

→ One candidate: the “power law of practice”

Estimating developmental change in online word recognition accuracy & speed

The log-linear/power law of practice

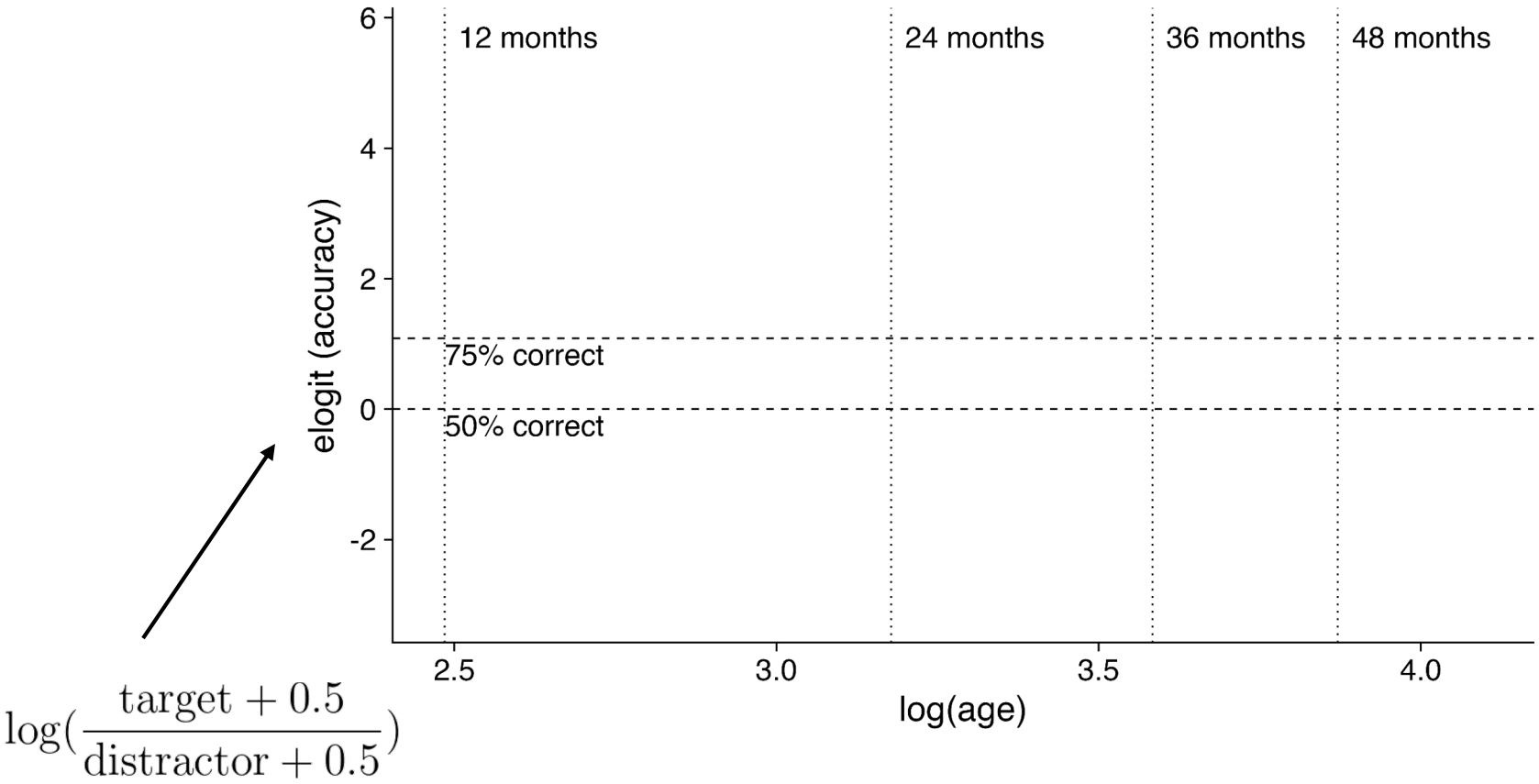
log(task performance) ~ log(amount of practice)

"A plot of the logarithm of time to perform a task against the logarithm of amount of practice approximates a straight line."
(Anderson, 1982; also in Fitts & Posner, 1967; Newell & Rosenbloom, 1981)
→ classic law of skill acquisition

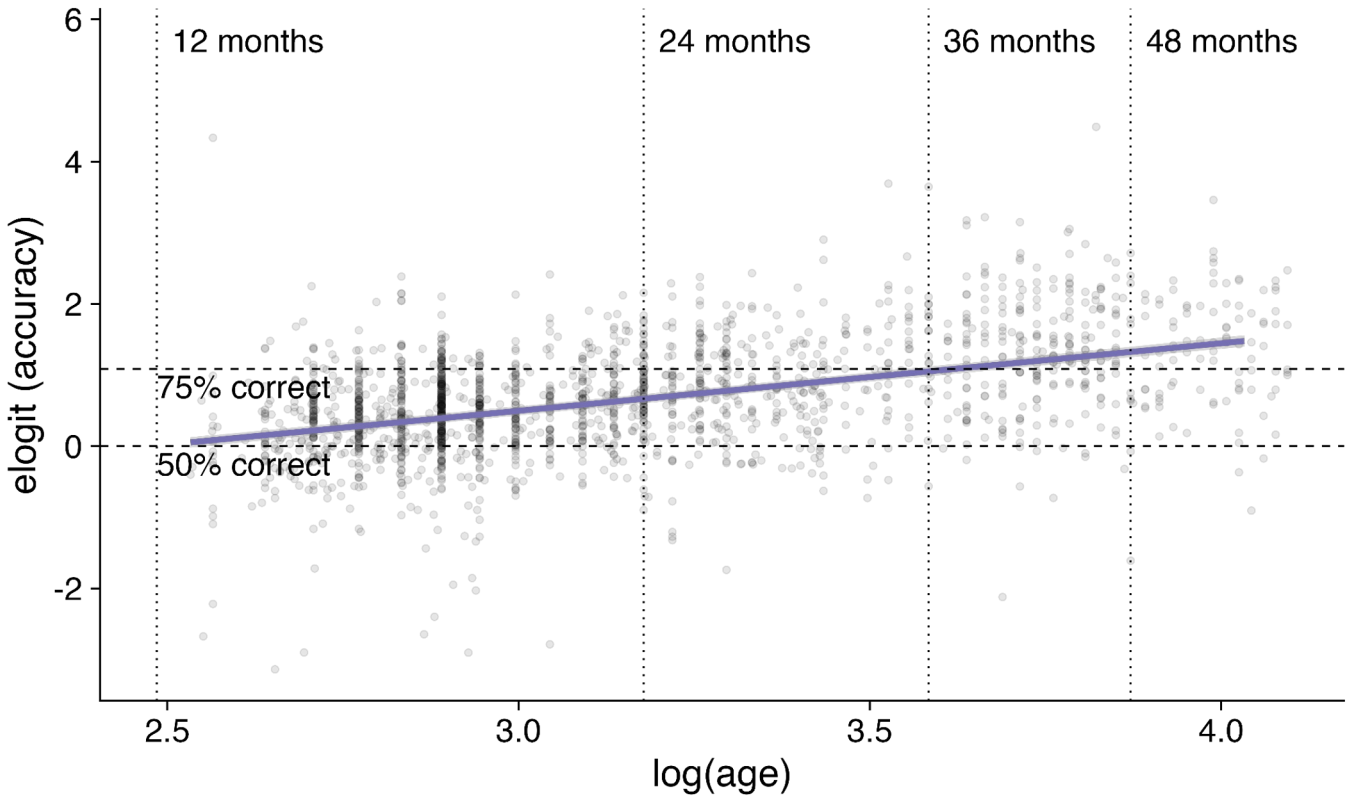
Do we see a similar log-linear relationship in word recognition?

log(task performance) ~ log(age)

# Estimating developmental change in online word recognition accuracy & speed

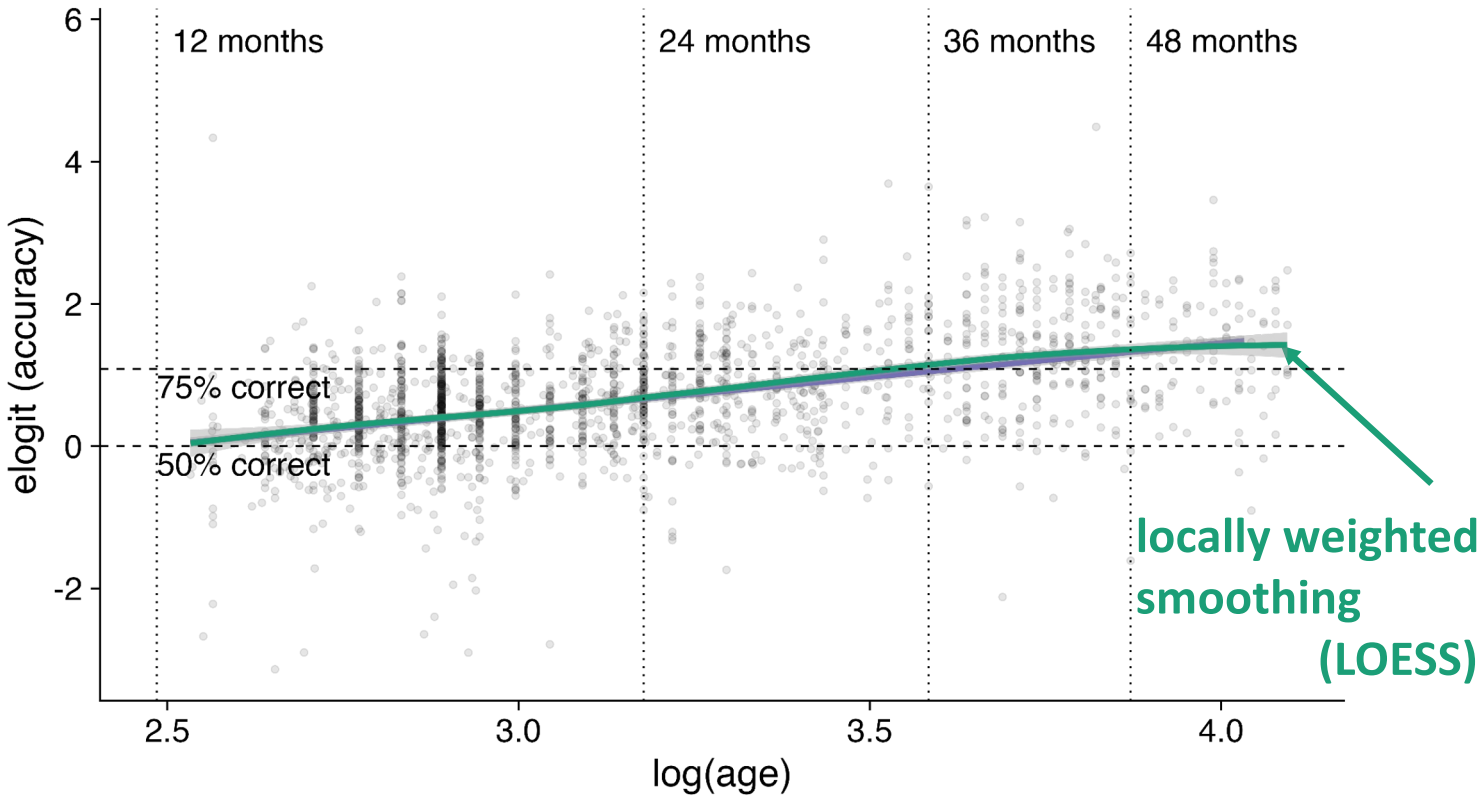


Estimating developmental change in online word recognition accuracy & speed



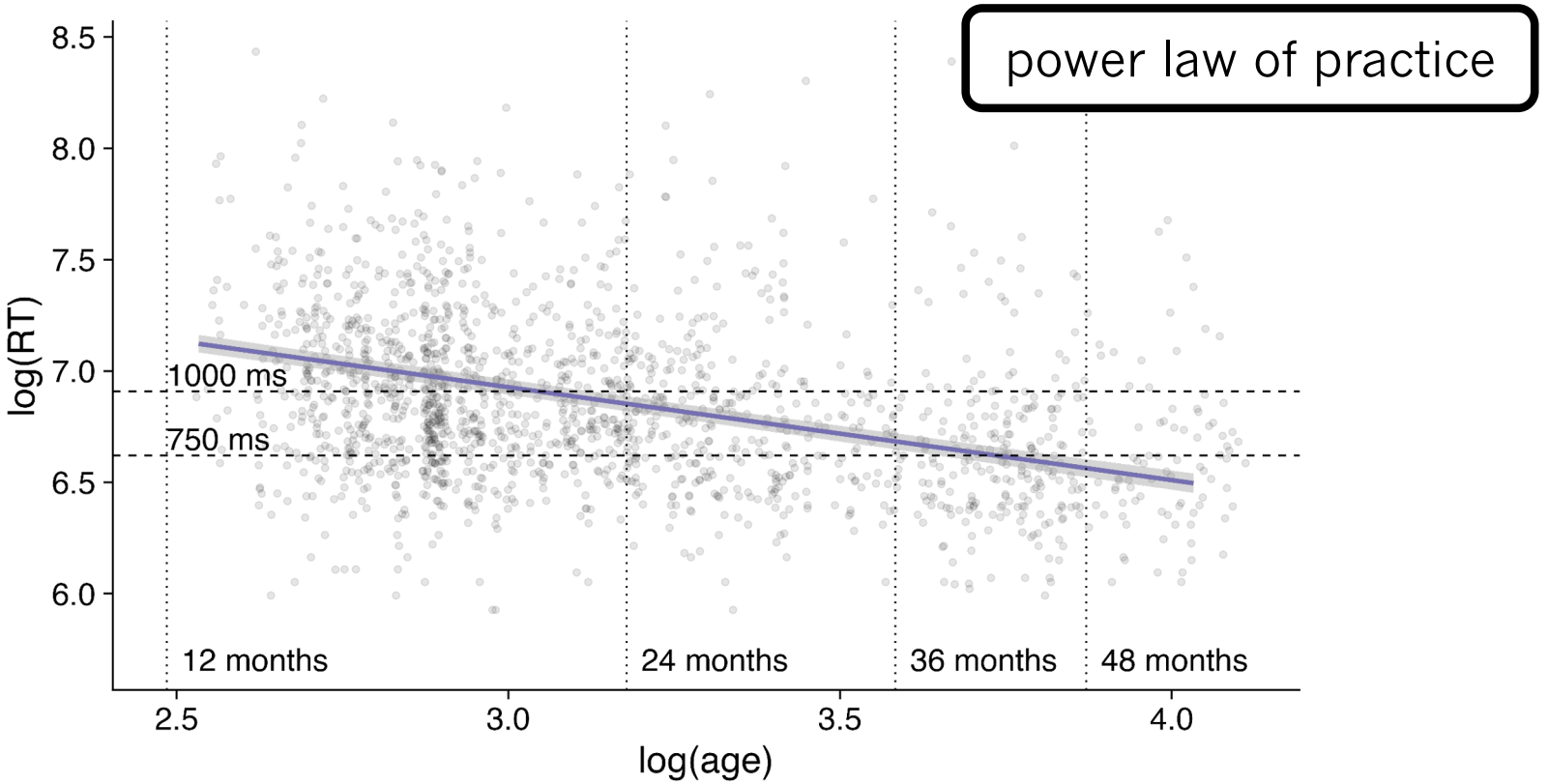

Linear mixed-effects model accounting for non-independence due to participants, items, and datasets

Estimating developmental change in online word recognition accuracy & speed



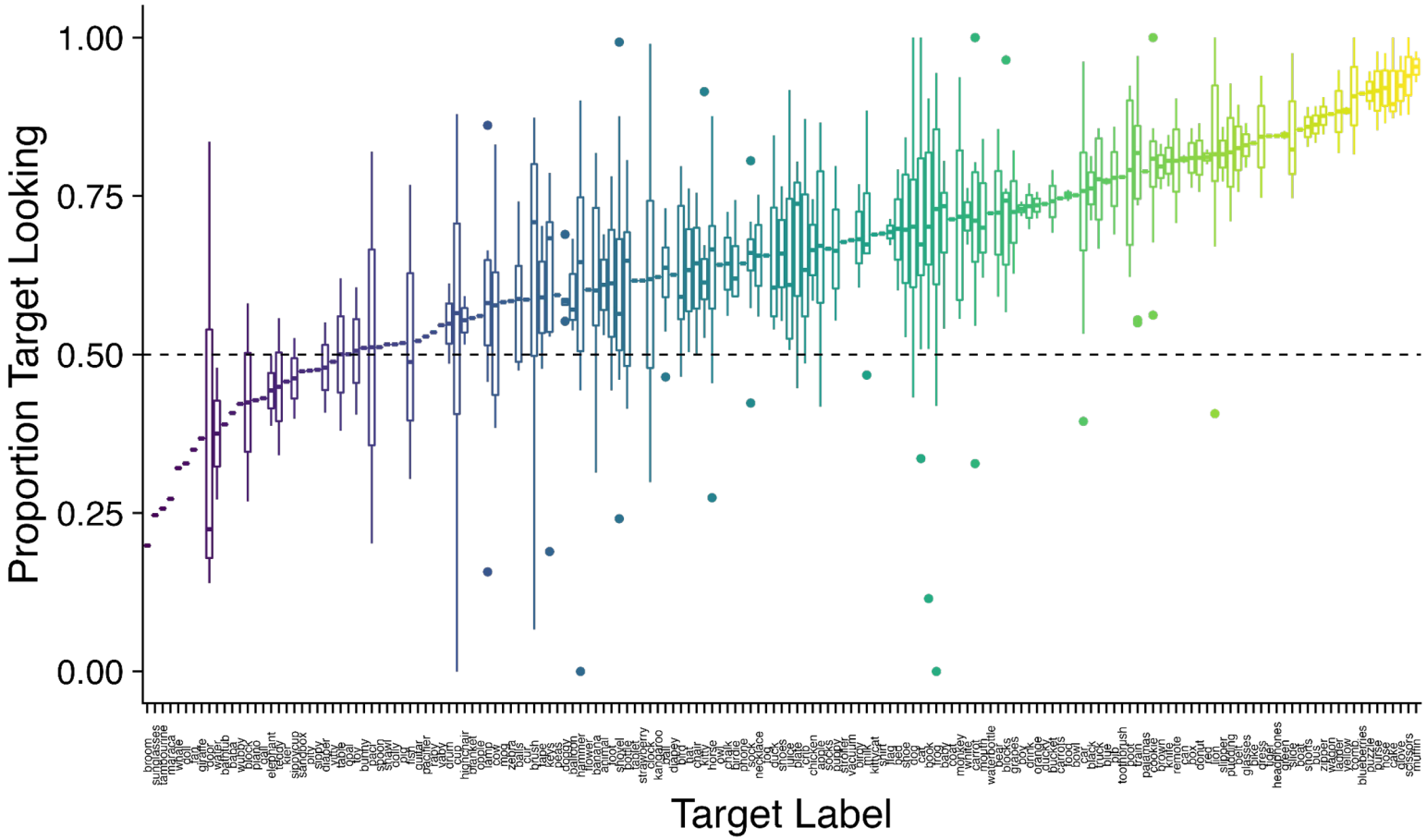

Linear mixed-effects model accounting for non-independence due to participants, items, and datasets

Estimating developmental change in online word recognition accuracy & speed

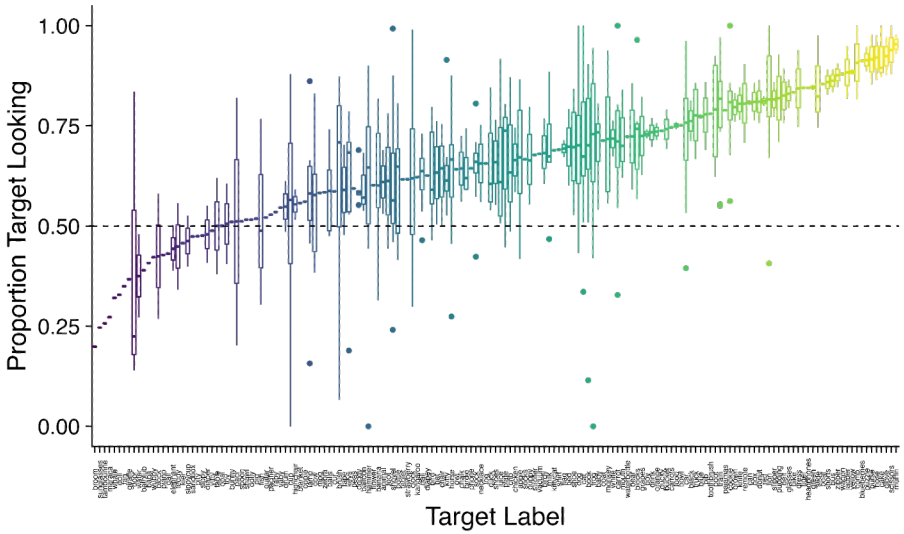



Linear mixed-effects model accounting for non-independence due to participants, items, and datasets

Large amount of variability across items in Peekbank



# Explaining item-level variability in Peekbank

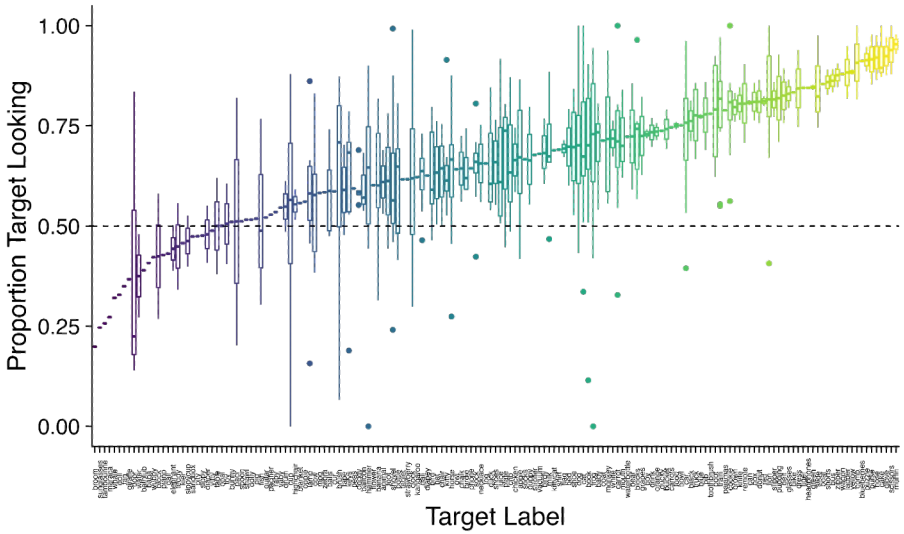



**Age of Acquisition**
(e.g., Wordbank)

**Word Frequency**
(e.g., CHILDES)

**Individual-level Vocabulary Measures**
(e.g., child-specific MCDI)

# Explaining item-level variability in Peekbank

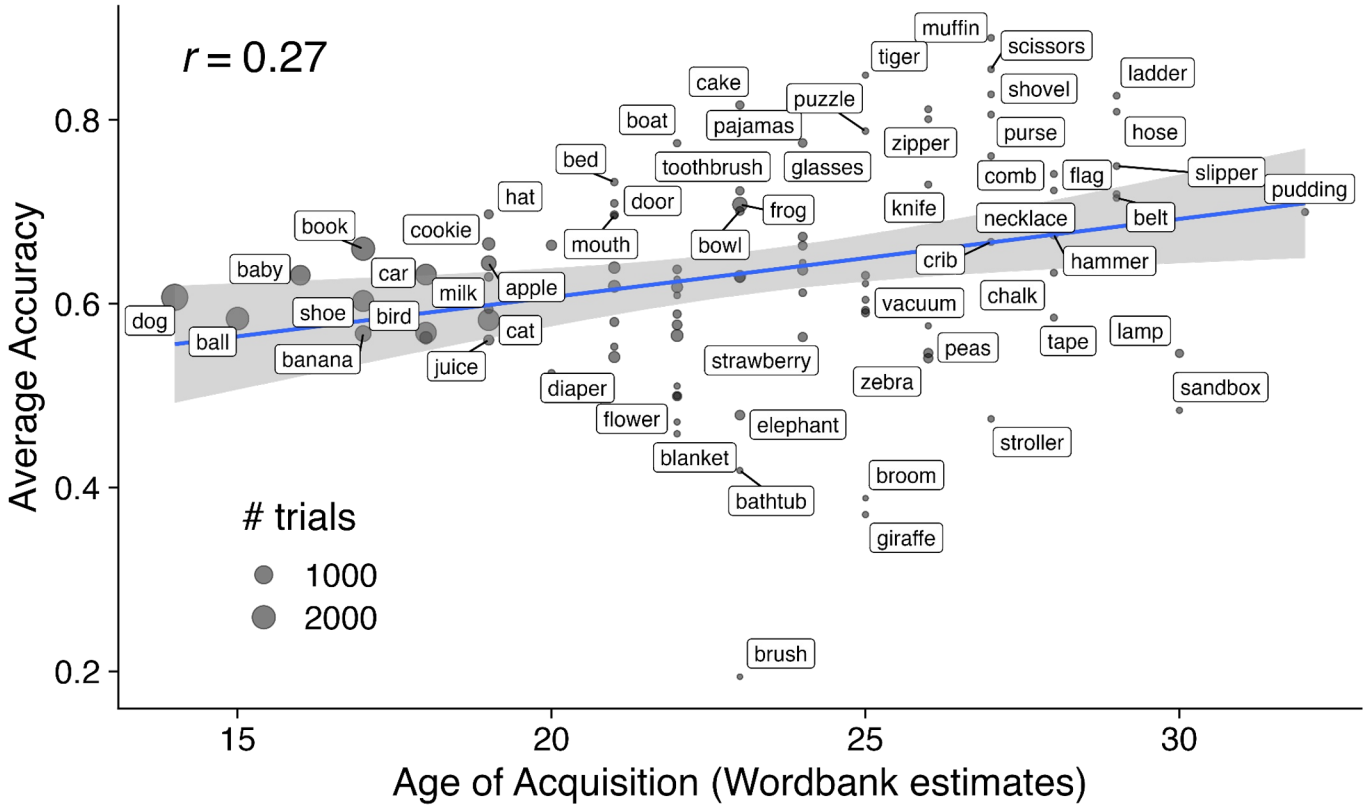



**Age of Acquisition**
(e.g., Wordbank)

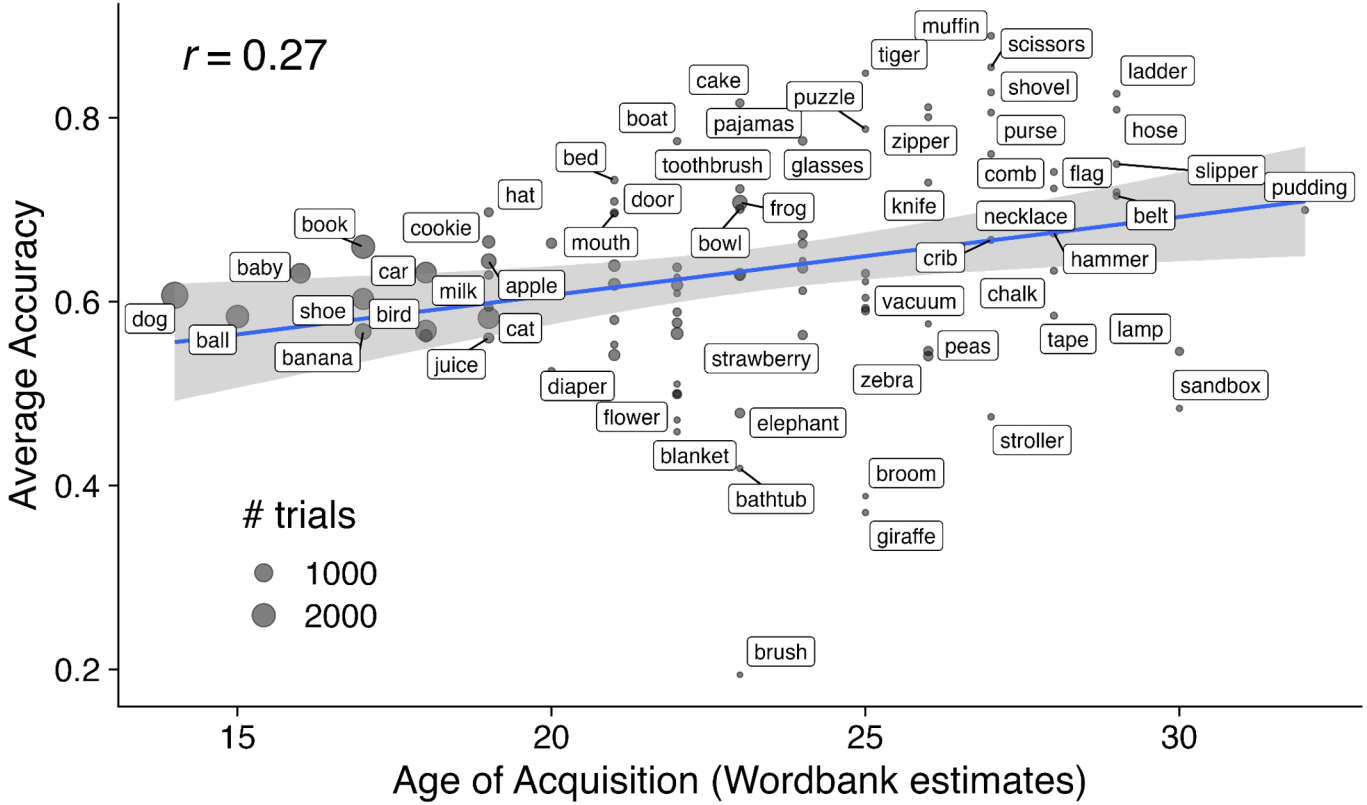
**Word Frequency**
(e.g., CHILDES)

**Vocabulary Measures**
(e.g., MCDI)

# Age of Acquisition predicts item-level accuracy from Peekbank



# Age of Acquisition predicts item-level accuracy from Peekbank ...but in the wrong direction!



# Predicting AoA from item-level accuracy in Peekbank

What might be obscuring relationships between AoA and accuracy?

- Age-dependent item selection (i.e., harder words be seen by older kids)
- Salience of the visual items matters

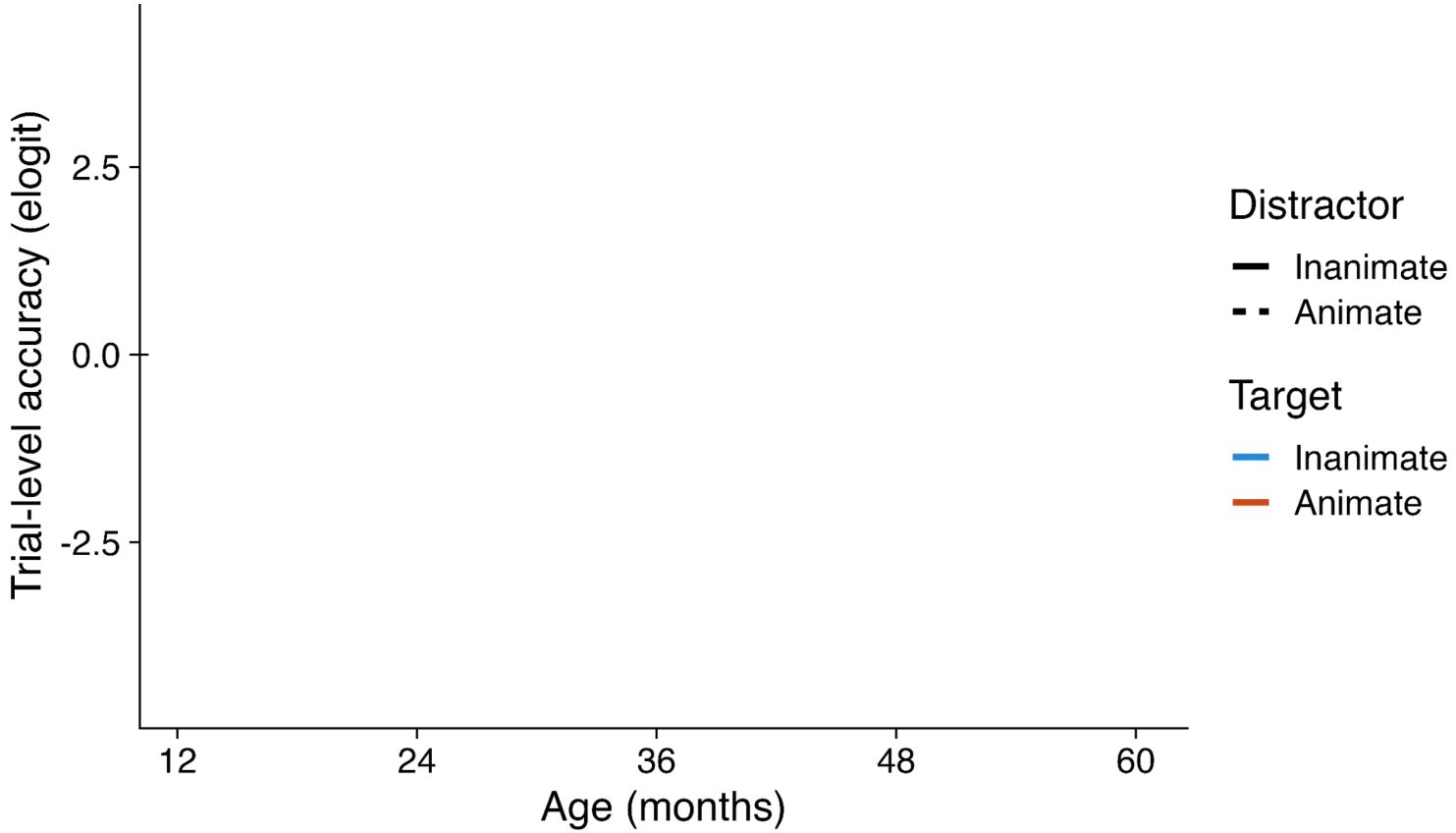

target
animate

distractor
animate

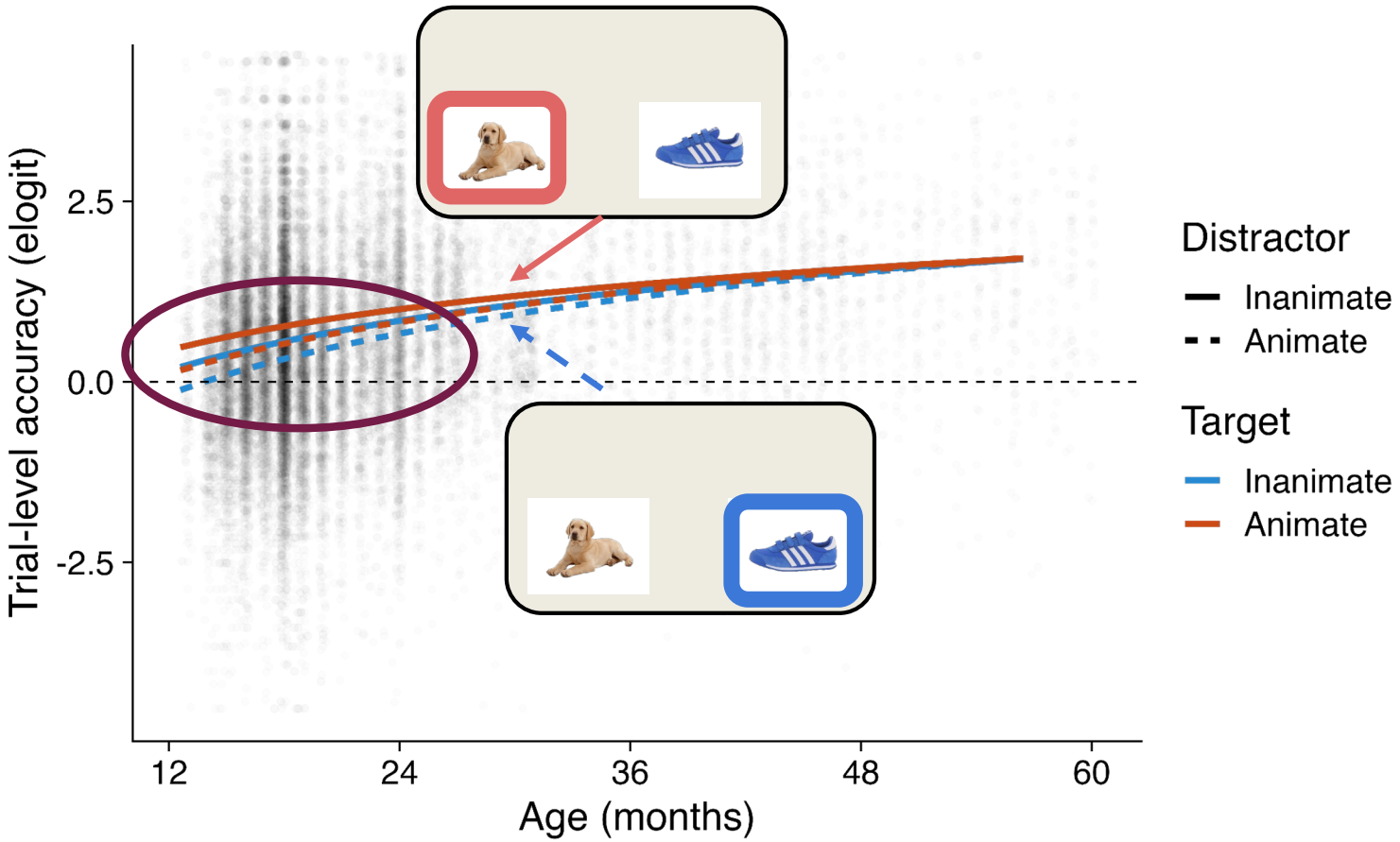

target
animate

distractor
inanimate

Animacy (salience) affects accuracy - especially among younger kids



Animacy (salience) affects accuracy - especially among younger kids

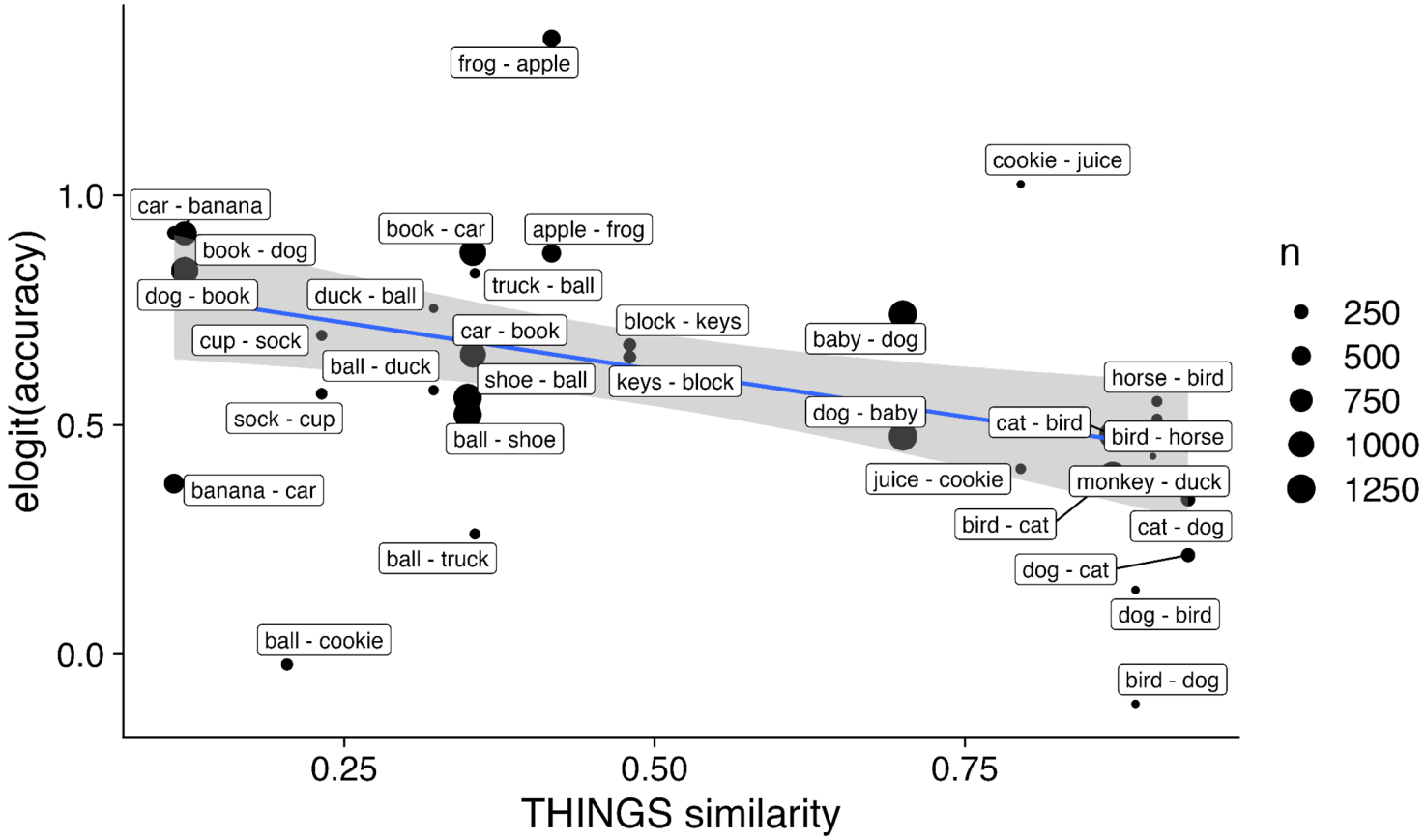


Predicting AoA from item-level accuracy in Peekbank

What might be obscuring relationships between AoA and accuracy?

- Age-dependent item selection (i.e., harder words tend to be seen by older kids)
- Salience of the visual items matters
- The target-distractor relationship matters too

Accuracy is higher for more dissimilar target-distractor pairs




THINGS data: Hebart et al. (2020)

Predicting AoA from item-level accuracy in Peekbank

What might be obscuring relationships between AoA and accuracy?

- Age-dependent item selection (i.e., harder words tend to be seen by older kids)
- Salience of the visual items matters
- The target-distractor relationship matters too

Predicting AoA from item-level accuracy in Peekbank

What might be obscuring relationships between AoA and accuracy?

- Age-dependent item selection (i.e., harder words tend to be seen by older kids)
- Salience of the visual items matters
- The target-distractor relationship matters too

→ revisiting measures and models

Predicting AoA from item-level accuracy in Peekbank

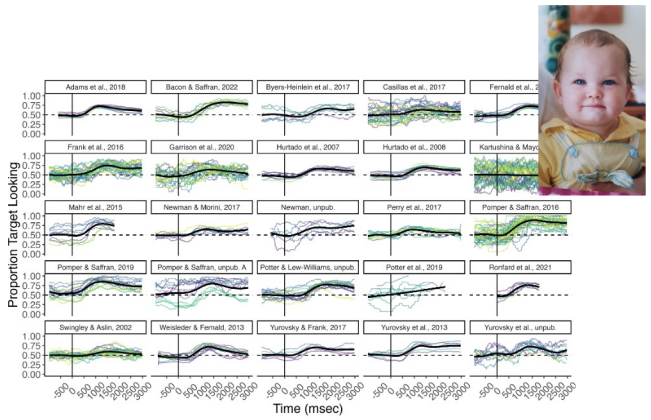
What might be obscuring relationships between AoA and accuracy?

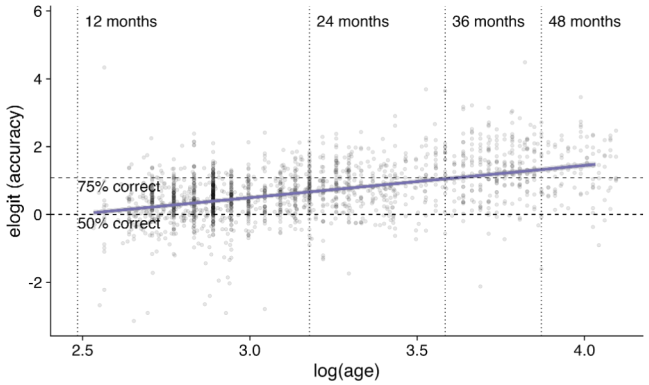
- Age-dependent item selection (i.e., harder words tend to be seen by older kids)
- Salience of the visual items matters
- The target-distractor relationship matters too

→ revisiting measures and models
→ important to focus on the looking process

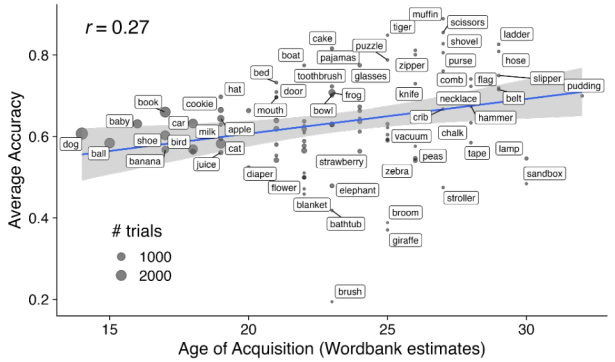
# Summing up

Aggregating data through Peekbank helps us improve methods and ask theoretical questions about lexical development “at scale”



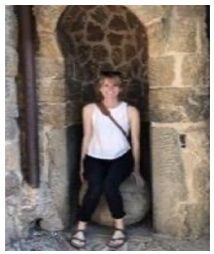

Peekbank allows us to model developmental change in word recognition consistent with skill acquisition




Predicting item-level patterns is still hard + messy, even almost 2000 kids later!

# The Peekbank Team


Sarp Uner

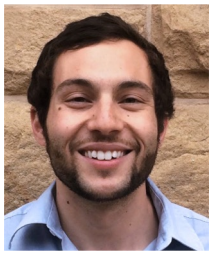

Rose M. Schneider

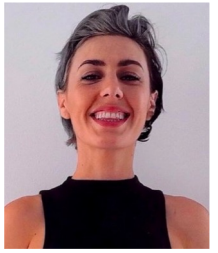

Stephan Meylan


Kunal Handa


Naiti S. Bhatt


Benny deMayo


Daniel Yurovsky


Ruthe Foushee


Molly Lewis


Michael C. Frank


Mika Braginsky

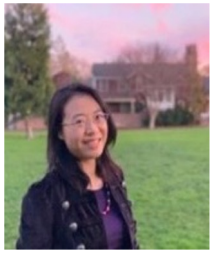

Kyle MacDonald


Tian Linger Xu


Bria Long


Virginia Marchman


Esat Boucaud


Angeline S. M. Tsui


Jessica Mankewitz


George Kachergis


Annissa N. Saleh


Alexandra Carstensen


Veronica Boyce


Daniel Swingley


Heidi Baumgartner

# Practicum: digging into Peekbank