Meta-analysis and mispronunciations: An introduction to meta-analyses

Katie Von Holzen
Technical University of Braunschweig, Germany

What can we gain from meta-analyses?

AMERICAN PSYCHOLOGICAL ASSOCIATION

Developmental Psychology

© 2021 American Psychological Association
ISSN: 0012-1649

2021, Vol. 57, No. 1, 1–18
https://doi.org/10.1037/dev0001141

The Development of Infants' Responses to Mispronunciations: A Meta-Analysis

Katie Von Holzen[1, 2, 3] and Christina Bergmann[4, 5]

[1] Lehrstuhl Linguistik des Deutschen, Schwerpunkt Deutsch als Fremdsprache/Deutsch als Zweitsprache, Technische Universität Dortmund
[2] Department of Hearing and Speech Sciences, University of Maryland, College Park
[3] CNRS (Integrative Neuroscience and Cognition Center, UMR 8002), Université Paris Descartes, Sorbonne Paris Cité
[4] Max Planck Institute for Psycholinguistics, Nijmegen, the Netherlands
[5] Laboratoire de Sciences Cognitives et de Psycholinguistique, Département d'études cognitives, Ecole normale supérieure, EHESS, CNRS, PSL University

UNIVERSITÉ PARIS DESCARTES

HOCHSCHULE OSNABRÜCK
UNIVERSITY OF APPLIED SCIENCES

1794
ENS
ÉCOLE NORMALE SUPÉRIEURE

What can we gain from meta-analyses?

- Experimental planning
  - Aggregated across studies, what is the true effect?
  - Statistical power and sample size
  - Publication bias
- Theoretical insights
  - Is this modulated by different variables?

What is a meta-analysis?

"… the set of statistical tools for aggregating quantitative results across studies"

Bergmann et al., 2018, *Child Development*

Phonology 32 (2015) 207–239. © Cambridge University Press 2015
doi:10.1017/S0952675715000135

*Building phonological lexical representations**

Suzanne V. H. van der Feest
University of Texas at Austin

Paula Fikkert
Radboud University Nijmegen

This paper contributes to the ongoing debate on h
dren's word representations contain. We investig
place of articulation and voicing contrasts, inspired
metrical patterns in children's early word productio
20- and 24-month-olds' perception of these fundam
mispronunciation-detection paradigm. Our results
directions of phonological changes yield differen
month-olds noticed coronal mispronunciations of
The 24-month-olds noticed voiced mispronunciat
not vice versa, while the 20-month-olds failed to noti
tions. We argue that early lexical representations ar
ways, that not all phonological contrasts are encod
the phonological system of a language determines
first in the representations of early words.

Developmental Sci
Developmental Science 17:1 (2014), pp 94–109

PAPER

**Influences of vowel and tone va
knowledge: a cross-linguistic inv**

**Leher Singh,[1] Tam Jun Hui,[1] Calista
Roberta Michnick Golinkoff[2]**

1. Department of Psychology, National University of Singapore, S
2. School of Education, University of Delaware, USA

Abstract

J. Child Lang. **38** (2011), 606–627. © Cambridge University Press 2010
doi:10.1017/S0305000910000243

Does size matter? Subsegmental cues to vowel mispronunciation detection*

NIVEDITA MANI
University of Göttingen

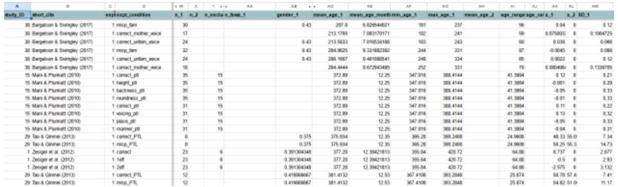
AND

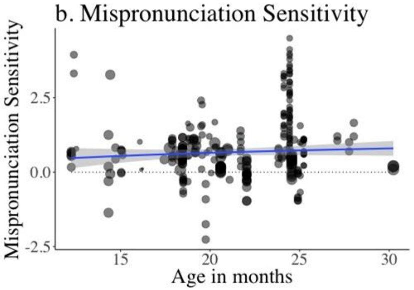
KIM PLUNKETT
University of Oxford

(*Received 17 November 2008 – Revised 28 September 2009 – Accepted 18 April 2010 – First published online 1 November 2010*)

ABSTRACT

Children look longer at a familiar object when presented with either correct pronunciations or small mispronunciations of consonants in the object's label, but not following larger mispronunciations. The current article examines whether children display a similar graded sensitivity to different degrees of mispronunciations of the vowels in familiar words, by testing children's sensitivity to 1-feature, 2-feature and 3-feature mispronunciations of the vowels of familiar labels: Children aged 1;6 did not show a graded sensitivity to vowel mispronunciations, even when the trial length was increased to allow them more time to form a response. Two-year-olds displayed a robust sensitivity to increases in vowel mispronunciation size, differentiating between small and large mispronunciations. While this suggests that early lexical representations contain information about the features contributing to vocalic identity, we present evidence that this graded sensitivity is better explained by the acoustic characteristics of the different mispronunciation types presented to children.





What can we gain from meta-analyses?

- Experimental planning
  - Aggregated across studies, what is the effect size?
  - Statistical power and sample size
    - How many participants do I need to test in order to be able to detect this effect?

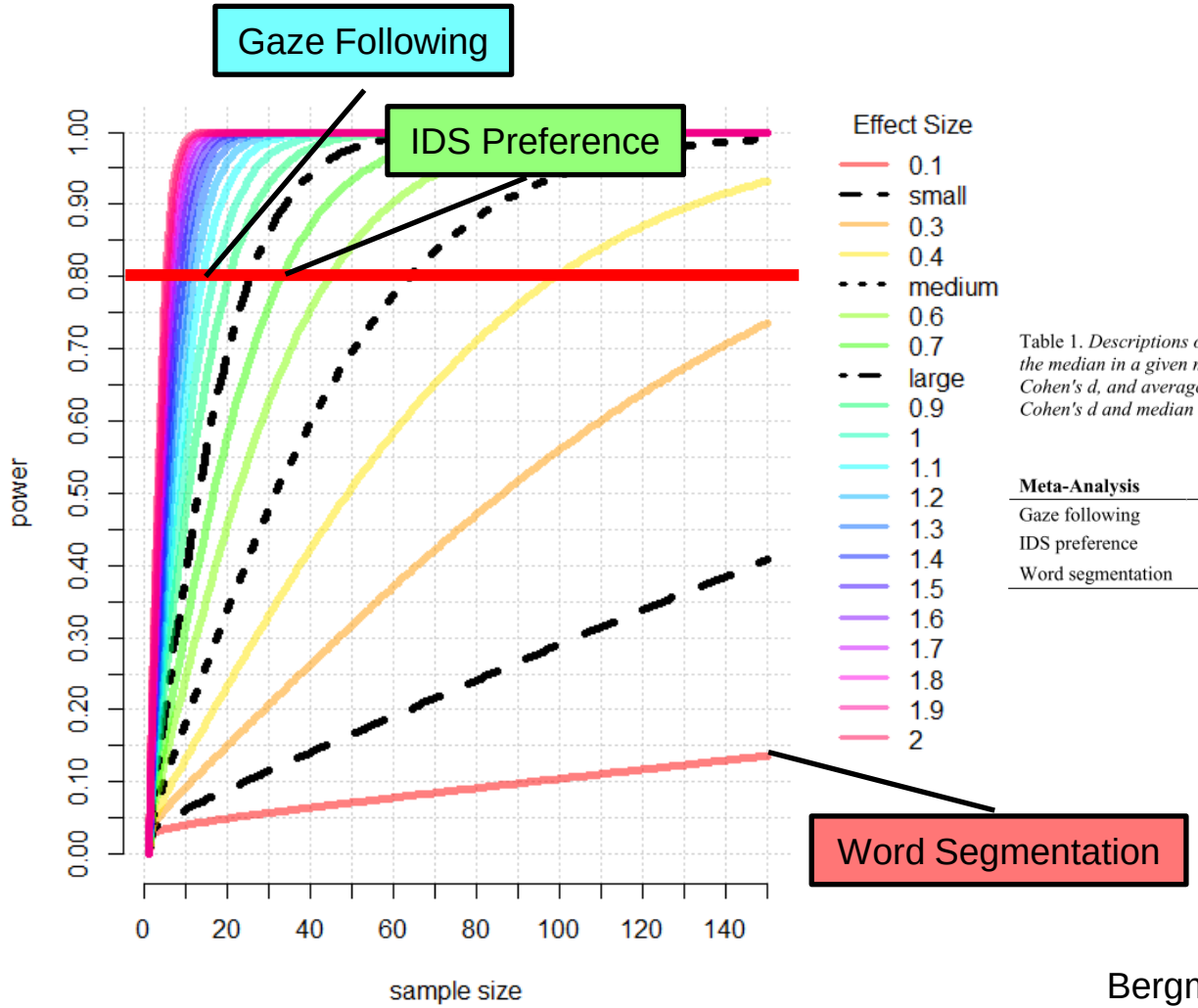


Table 1. *Descriptions of the meta-analyses. Age is reported in months, sample size is based on the median in a given meta-analysis, effect size is reported as meta-analytic weighted median Cohen's d, and average power is computed based on meta-analytic effect size estimate Cohen's d and median sample size.*

Meta-Analysis	Age	Sample Size	N Effect Sizes	N Papers	Effect Size (SE)	Power
Gaze following	14 (3-24)	23 (12-63)	32	11	1.08 (0.16)	0.95
IDS preference	4 (0-9)	20 (10-60)	48	16	0.73 (0.13)	0.61
Word segmentation	8 (6-25)	20 (4-64)	284	68	0.16 (0.03)	0.08

Bergmann et al., 2018, *Child Development*

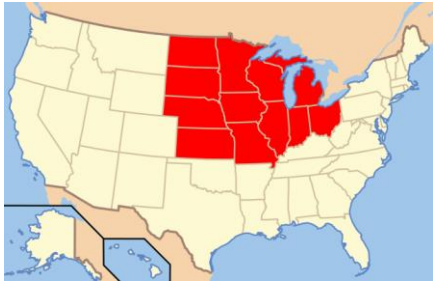
What can we gain from meta-analyses?

- Experimental planning
  - Aggregated across studies, what is the effect size?
  - Statistical power and sample size

Milk
/mɪlk/

Mulk
/mʌlk/

Bilk
/bɪlk/

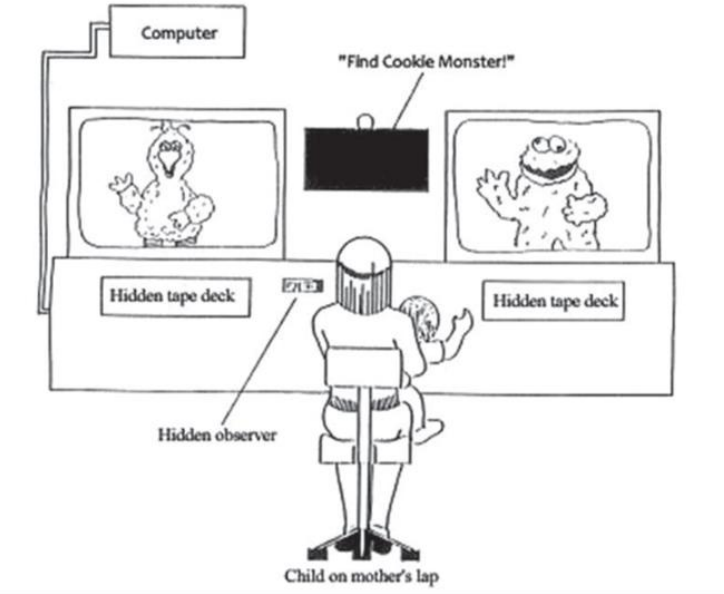


Milk
/mɛlk/

Mispronunciation sensitivity: Sensitivity to sound changes in familiar words during word recognition

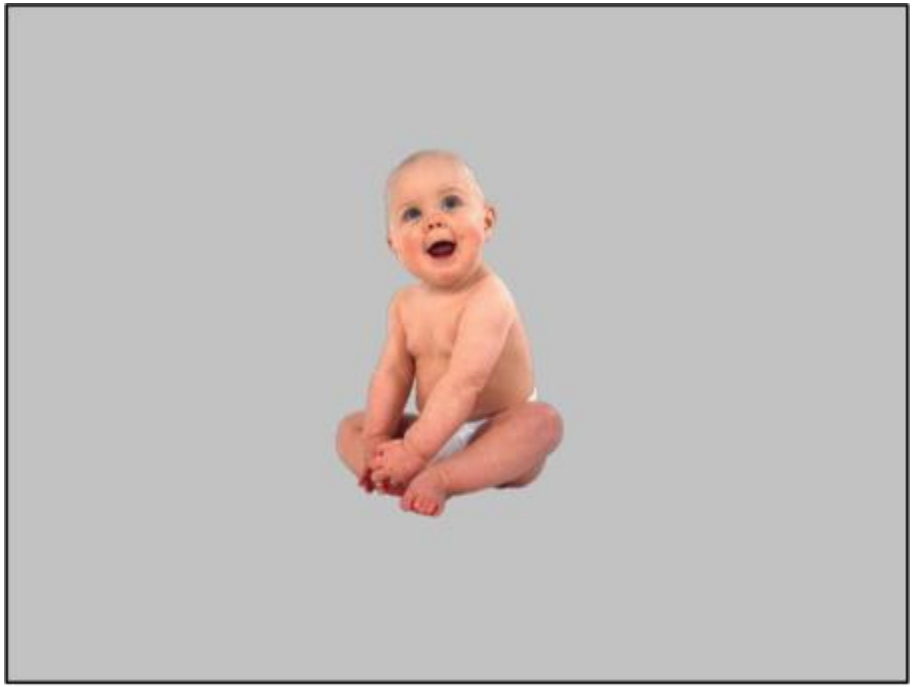
Mulk
/mʌlk/

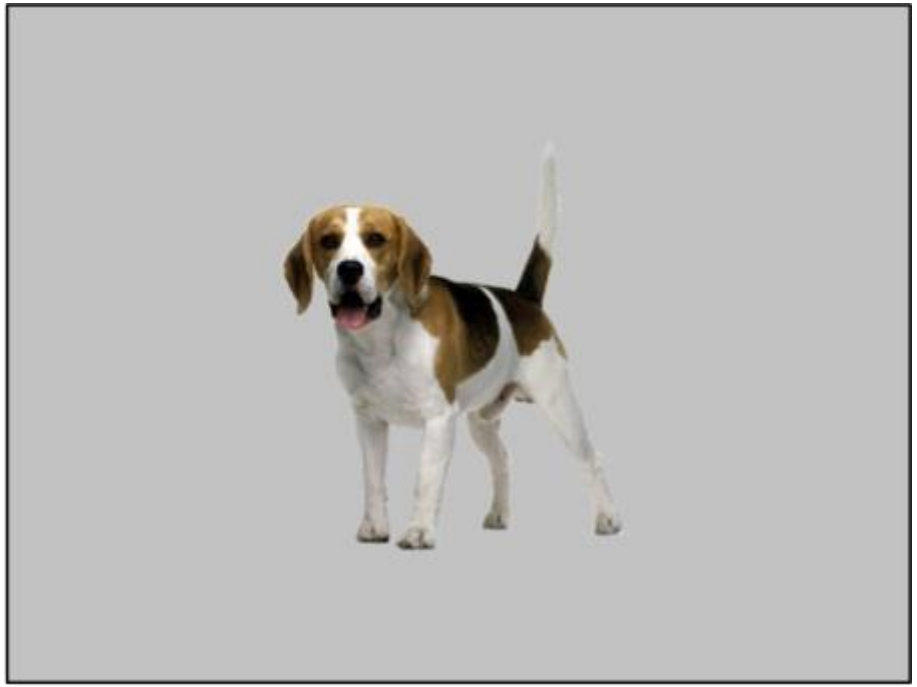
Bilk
/bɪlk/



Golinkoff, Hirsh-Pasek,
Cauley, & Gordon, 1987

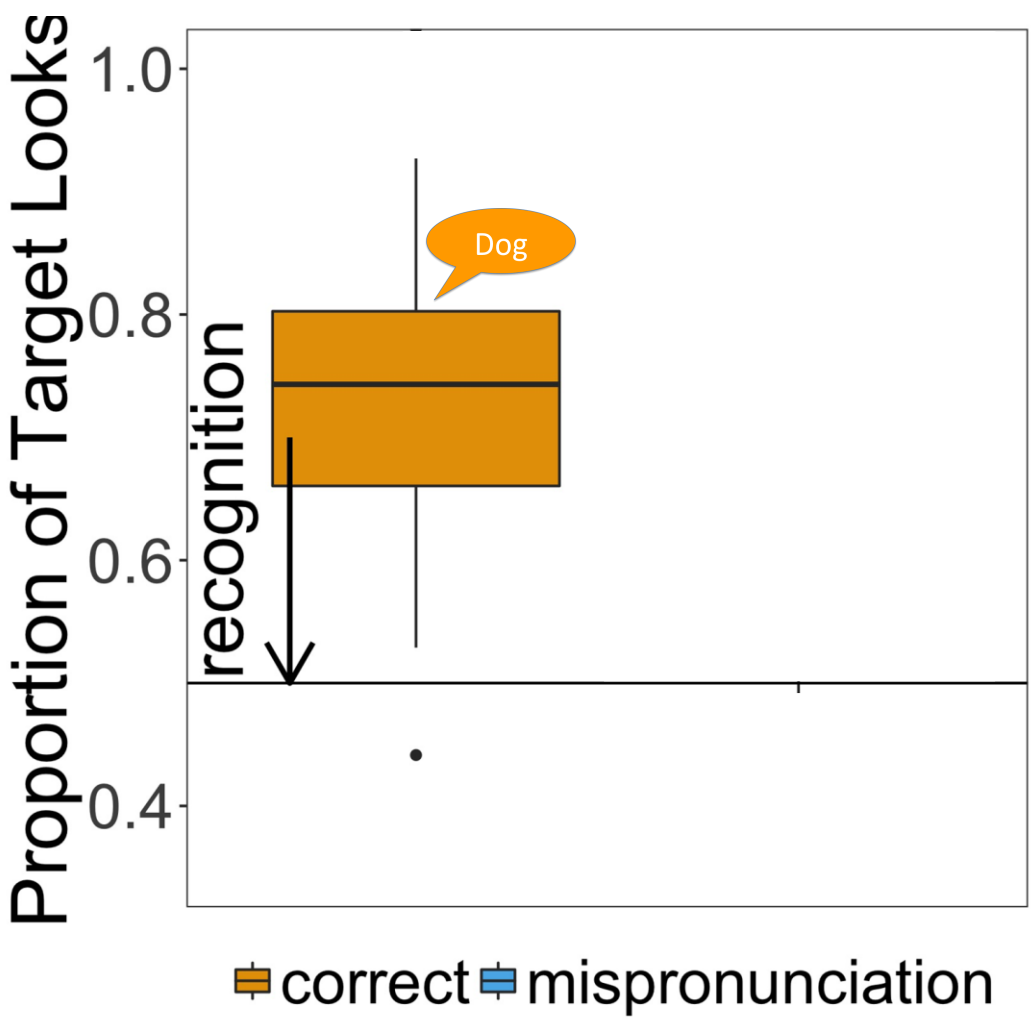
Bergelson Lab, Duke University



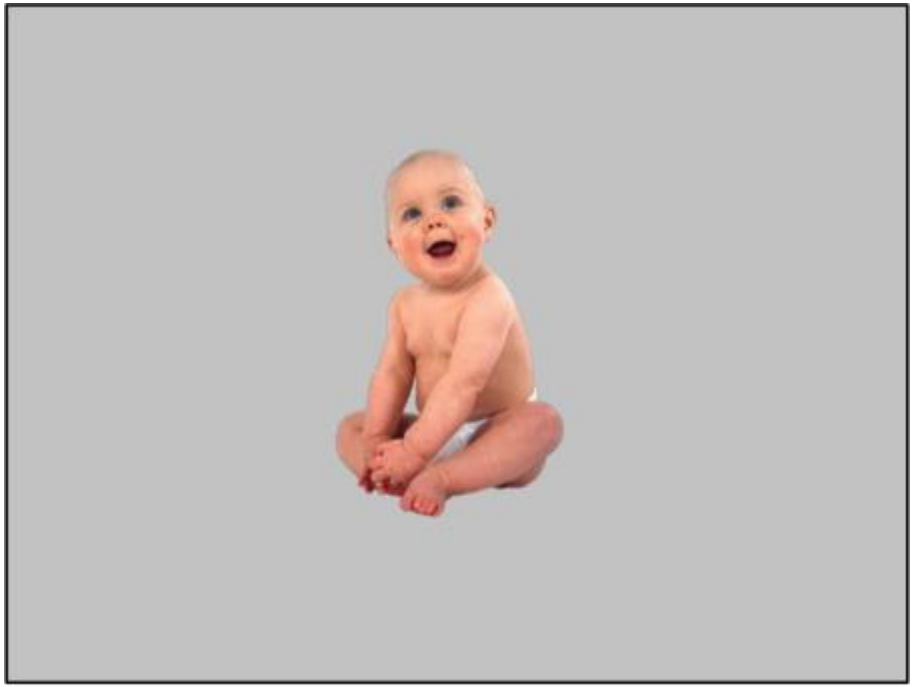


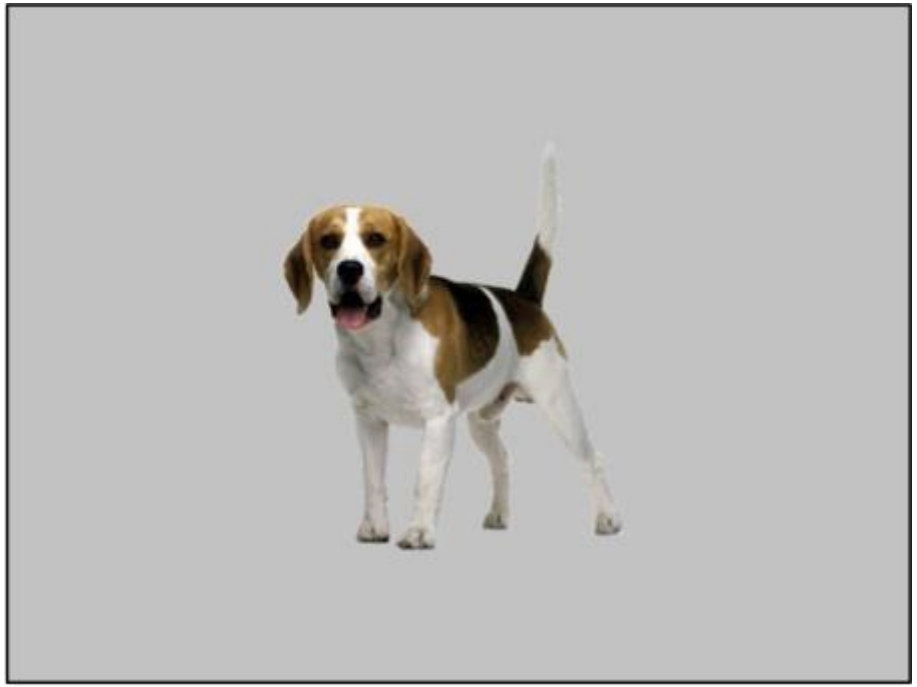
Dog

Swingley & Aslin, 2000



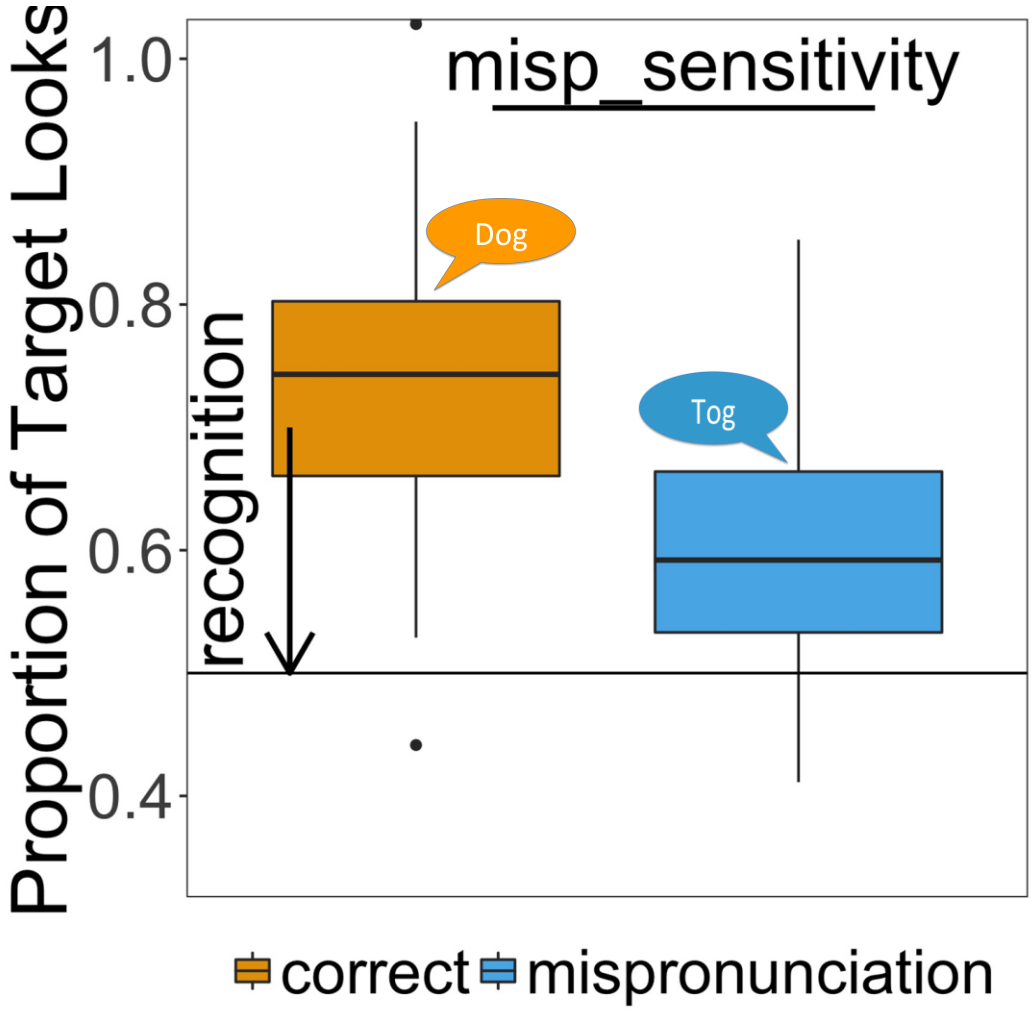
Swingley & Aslin, 2000





Tog

Swingley & Aslin, 2000



- *"sensitivity to a small, but potentially meaning-altering change in the acoustic word form"*
- 18-23-month-olds have phonologically well specified representations for familiar words

Swingley & Aslin, 2000

What can we gain from meta-analyses?

- Experimental planning
  - Aggregated across studies, what is the true effect?
  - Statistical power and sample size

Data Sample

- Screened over 400 papers
  - Original data
  - Infants younger than 31 months
  - Familiar word recognition
  - Measured eye movements
- Final sample: 32 items
  - 249 experimental conditions
  - 2252 infants

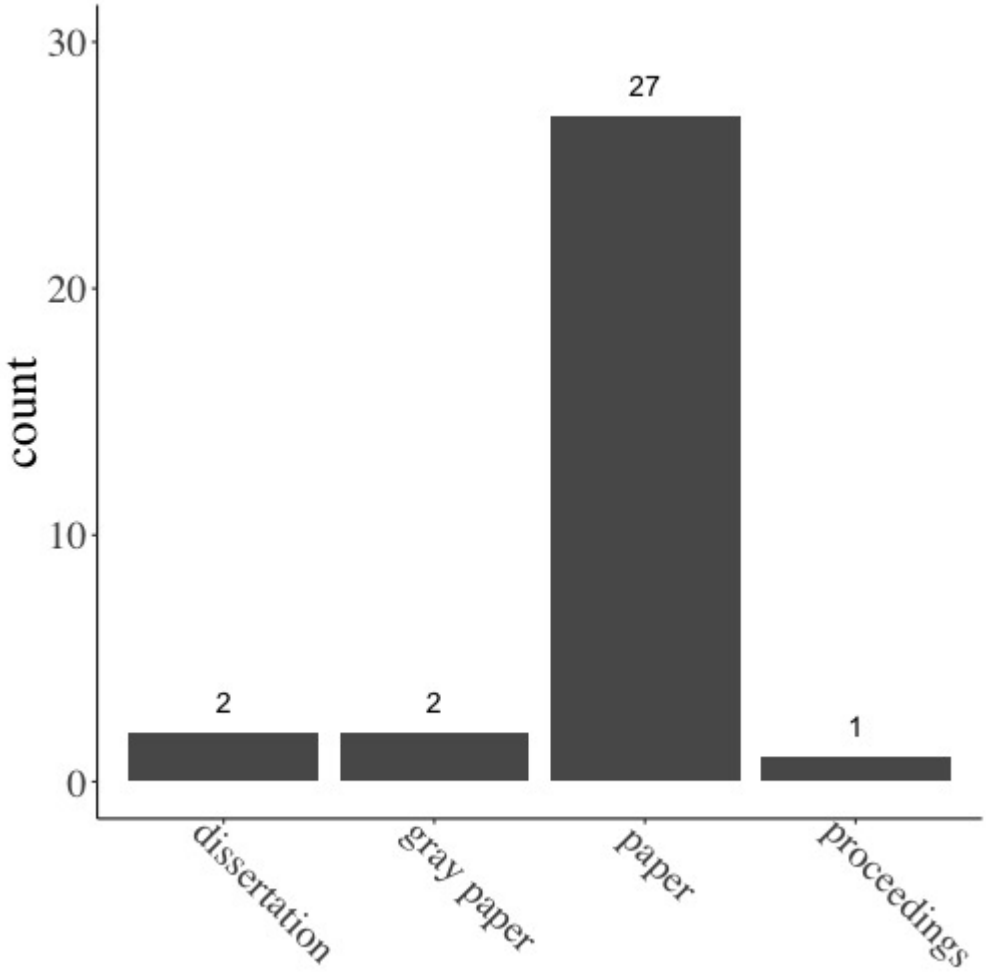


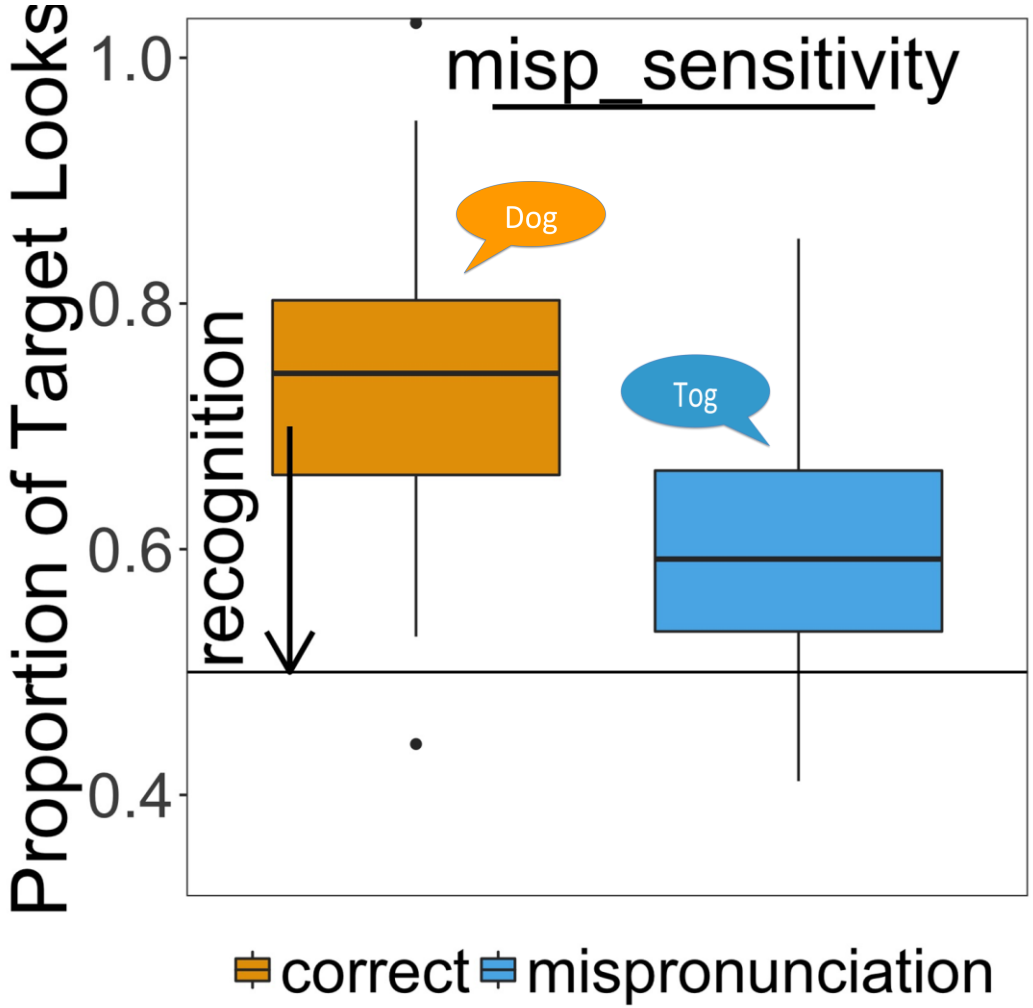
Table 1
Summary of All Papers

Paper	Format	Age	Vocabulary	Familiarity	Target overlap	Size	Position	Type	*N* effect sizes
Altvater-Mackensen (2010)	Dissertation	22, 25	None	Fam, Unfam	O, Unfam	1	O, O/M	C	13
Altvater-Mackensen et al. (2014)	Paper	18, 25	None	Fam	O	1	O	C	16
Bailey & Plunkett (2002)	Paper	18, 24	Comp	Fam	None	1, 2	O	C	12
Bergelson & Swingley (2018)	Paper	7, 9, 12, 6	None	Fam	None	Unspec	O/M	V	9
Bernier & White (2017)	Proceedings	21	None	Unfam	Unfam	1, 2, 3	O	C	4
Delle Luche et al. (2015)	Paper	20, 19	None	Fam	O	1	O	C/V	4
Durrant et al. (2015)	Paper	19, 20	None	Fam	O	1	O	C/V	4
Højen et al. (2016)	Gray paper	19, 20	Comp/Prod	Fam	C, O	2–3	O/M, C/M	C/V, V, C	6
Höhle et al. (2006)	Paper	18	None	Fam	None	1	O	C	4
Mani & Plunkett (2007)	Paper	15, 18, 24, 14, 20	Comp/Prod	Fam	O	1–2, 1	O	V, C/V, C	14
Mani & Plunkett (2010)	Paper	12	Comp	Fam	O	1	M, O	V, C	8
Mani & Plunkett (2011)	Paper	23, 17	None	Unfam	Unfam	1–3, 1, 2, 3	M	V	15
Mani et al. (2008)	Paper	18	Comp/Prod	Fam	O	1	M	V	4
Ramon-Casas & Bosch (2010)	Paper	24, 25	None	Fam	None	Unspec	M	V	4
Ramon-Casas et al. (2009)	Paper	21, 20	Prod	Fam	None	Unspec	M	V	10
Ren et al. (2019)	Paper	19	None	Unfam	None	1	O, C	C	8
Skoruppa et al. (2013)	Paper	23	None	Unfam	O/M	1	C	C	4
Swingley & Aslin (2000)	Paper	20	Comp	Fam	None	1	O	C/V	2
Swingley & Aslin (2002)	Paper	15	Comp/Prod	Fam	None	1, 2	O/M	C/V	4
Swingley (2003)	Paper	19	Comp/Prod	Fam	O	1	O, M	C	6
Swingley (2009)	Paper	17	Comp/Prod	Fam	None	1	O, C	C	4
Swingley (2016)	Paper	27, 28	Prod	Unfam	Unfam	1	O/M	C/V, C, V	9
Tamasi (2016)	Dissertation	30	None	Unfam	Unfam	1, 2, 3	O	C	4
Tao & Qinmei (2013)	Paper	12	None	Fam	None	Unspec	Unspec	T	4
Tao et al. (2012)	Paper	16	Comp	Fam	None	Unspec	Unspec	T	6
van der Feest & Fikkert, (2015)	Paper	24, 20	None	Fam	O	1	O	C	16
van der Feest & Johnson (2016)	Paper	24	None	Fam	O	1	O	C	20
Wewalaarachchi et al. (2017)	Paper	24	None	Unfam	Unfam	1	O/M/C	C/V/T, V, C, T	8
White & Aslin (2011)	Paper	18	None	Unfam	Unfam	1	M	V	4
White & Morgan (2008)	Paper	18, 19	None	Unfam	Unfam	1, 2, 3	O	C	12
Zesiger & Jöhr (2011)	Paper	14	None	Fam	None	1	O, M	C, V	7
Zesiger et al. (2012)	Paper	12, 19	Comp/Prod	Fam	None	1, 2	O	C	6

Note. Age = mean age (in months). Vocabulary: Comp = comprehension; Prod = production. Distractor familiarity: Fam = familiar; Unfam = unfamiliar. Target overlap: O = onset; M = medial; C = coda. Mispronunciation size: Number of features changed; commas indicate separate comparison, dashes indicate an aggregated range. Mispronunciation Position: O = onset; M = medial; C = coda. Mispronunciation type: C = consonant; V = vowel; T = tone. A slash separator indicates no distinction was made in the stimuli, and unspec. indicates that the value was unspecified in the paper.

Effect Size calculation

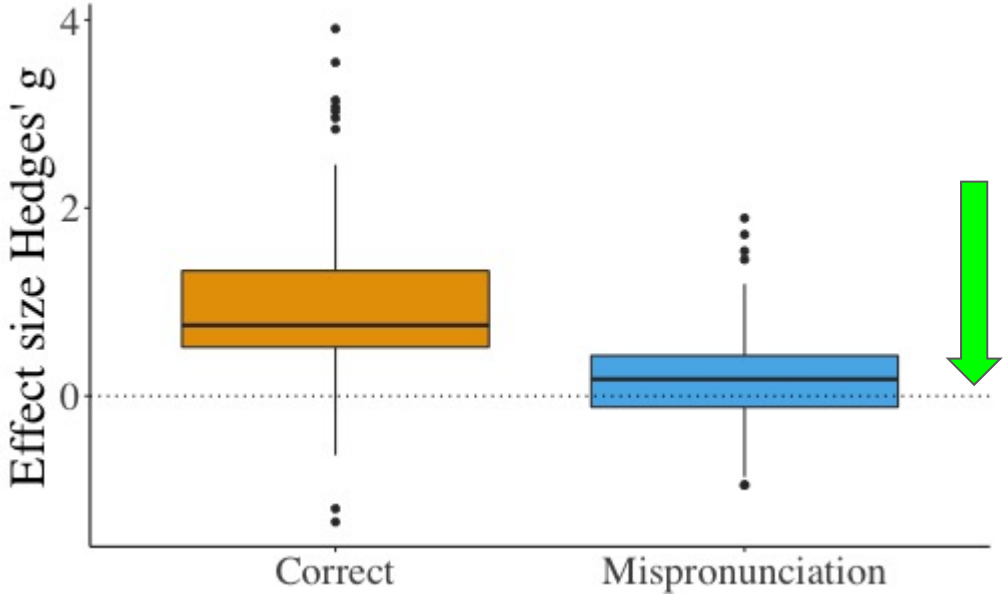
- Hedges' *g*
  - Hedges, 1981; Morris, 2000
  - Effect size that corrects for small sample sizes
  - Based on raw data (mean & sd) or test statistics reported in paper



Swingley & Aslin, 2000

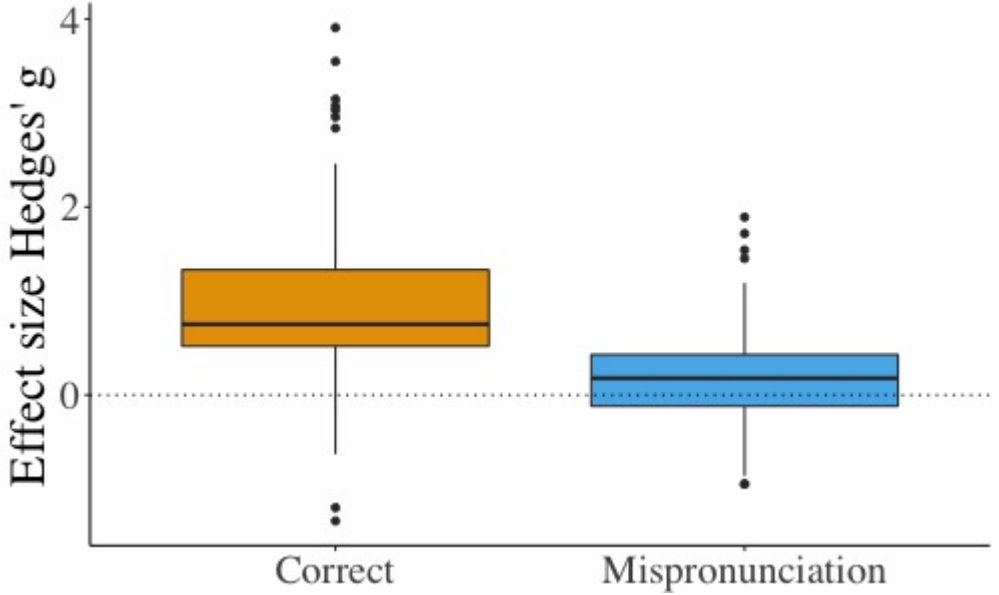
Object Identification / Recognition

- Effect sizes above chance
- Correct Pronunciation
  - $g = 0.91$
  - 95% CI [0.63, 1.14]
  - (SE = 0.12, $p < .0001$)
- Mispronunciation
  - $g = 0.25$
  - 95% CI [0.13, 0.37]
  - SE = 0.06, $p < .0001$



Mispronunciation Sensitivity

- Effect sizes different between correct and mispronunciation
  - β = 0.61
  - 95% CI [0.48, 0.72]
  - SE = 0.06, $p < .0001$

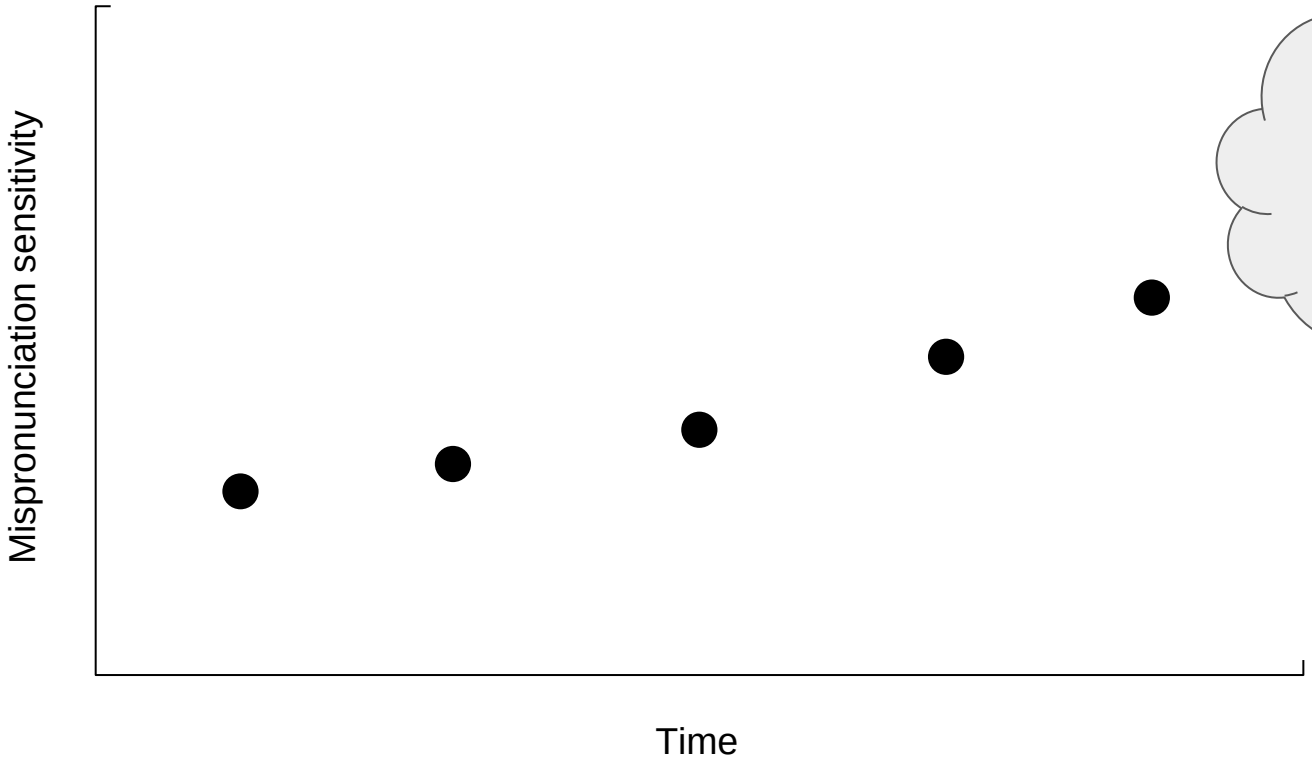


What can we gain from meta-analyses?

- Experimental planning
  - Aggregated across studies, what is the true effect?
    - Mispronunciation effect: β = 0.61
  - Statistical power and sample size
    - Power: 54%
    - Median sample size: 24 infants
    - To achieve 80% power (recommended), need to test 44 infants

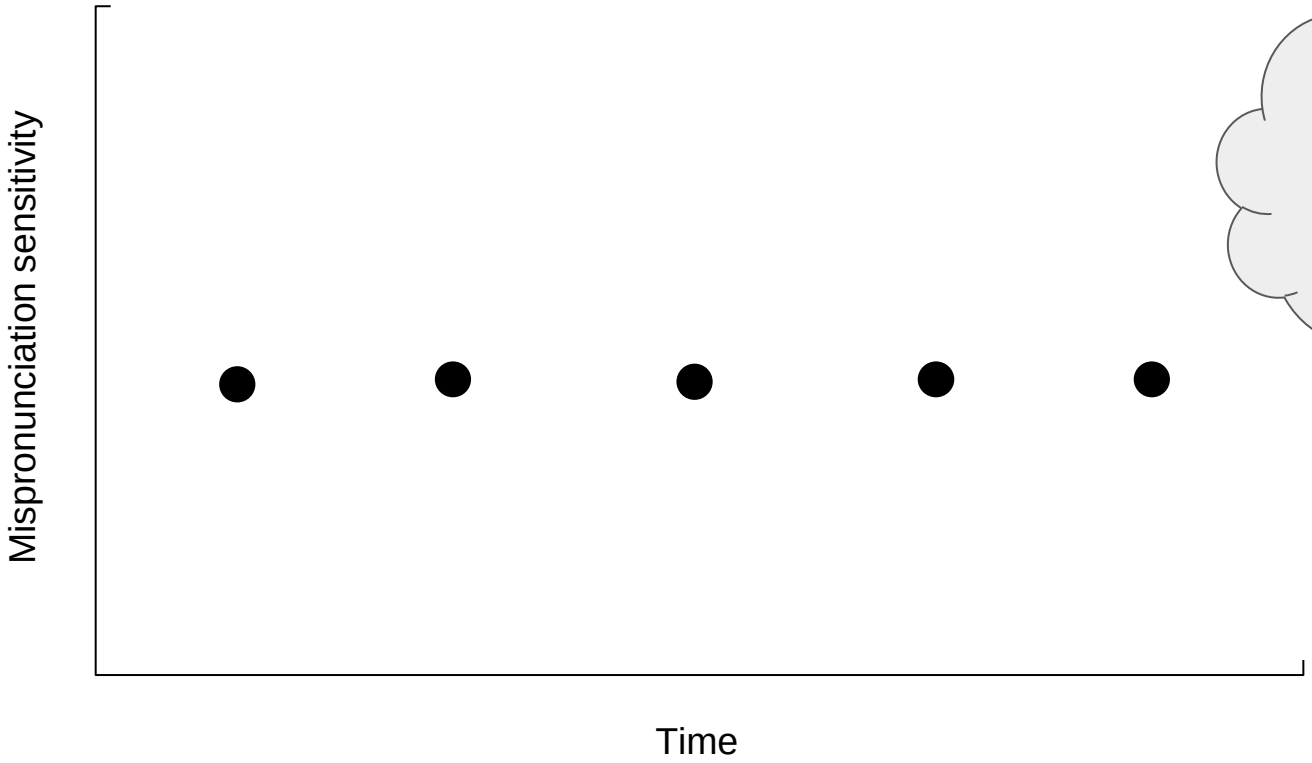
What can we gain from meta-analyses?

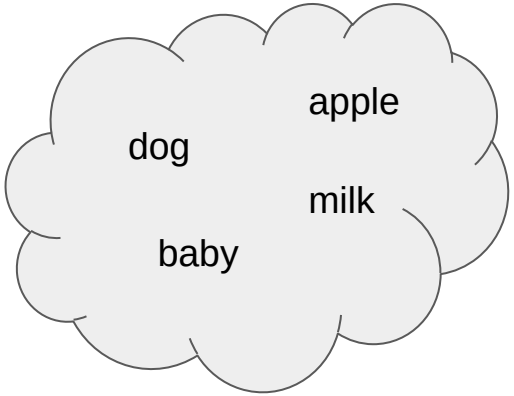
- Experimental planning
  - Aggregated across studies, what is the true effect?
  - Statistical power and sample size
- Theoretical insights
  - Is this modulated by different variables?



Theory 1: Mispronunciation sensitivity increases over time

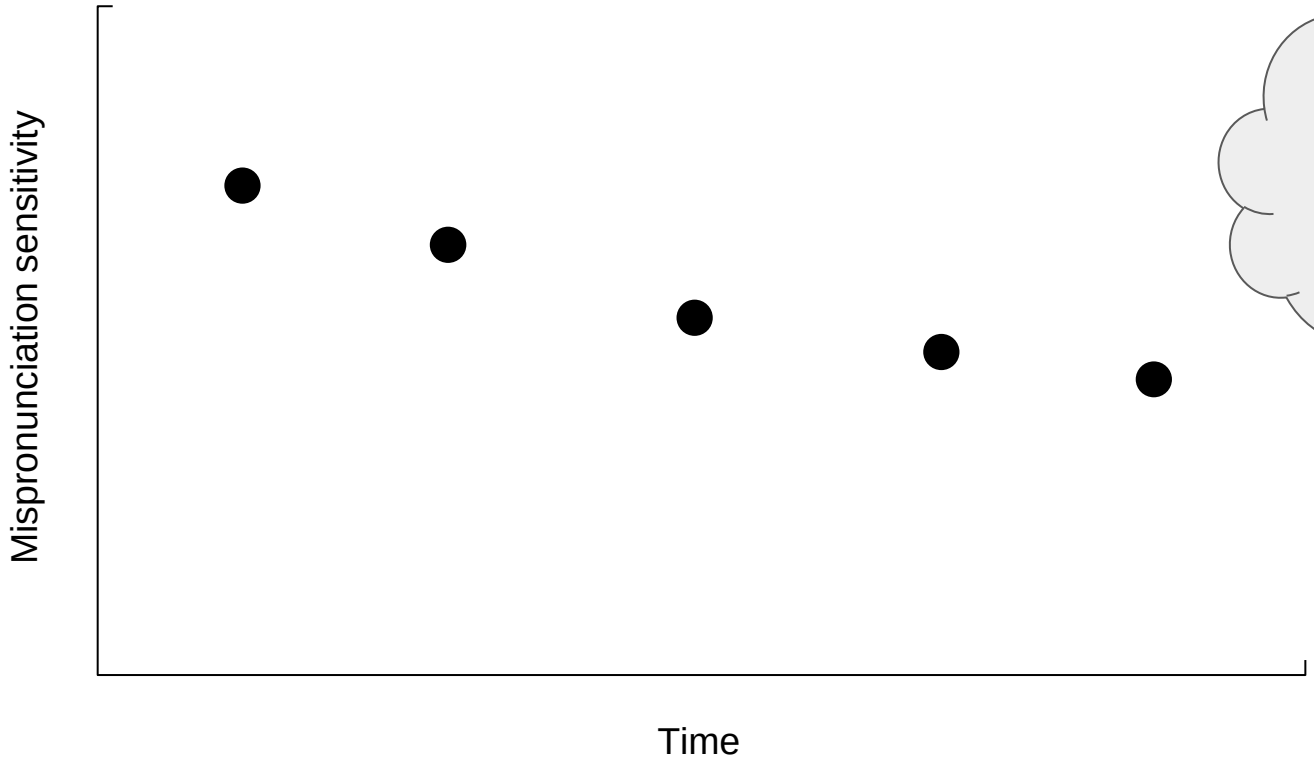
- From holistic to more detailed representations
- Altvater-Mackensen & Mani, 2013; Mani & Plunkett, 2007; van der Feest & Fikkert, 2015

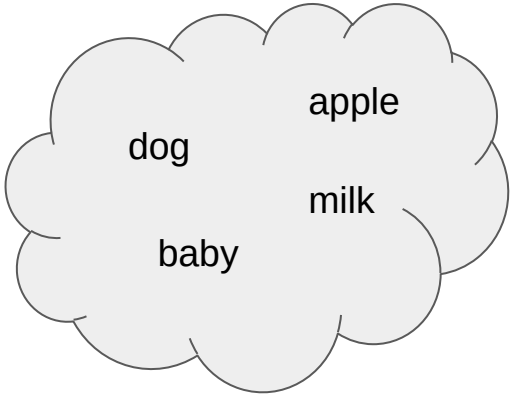




Theory 2: Mispronunciation sensitivity stays the same over time

- Early specificity
- Swingley & Aslin, 2000; Bailey & Plunkett, 2002; Zesiger et al., 2012





Theory 3: Mispronunciation sensitivity decreases over time

- Mani & Plunkett, 2011

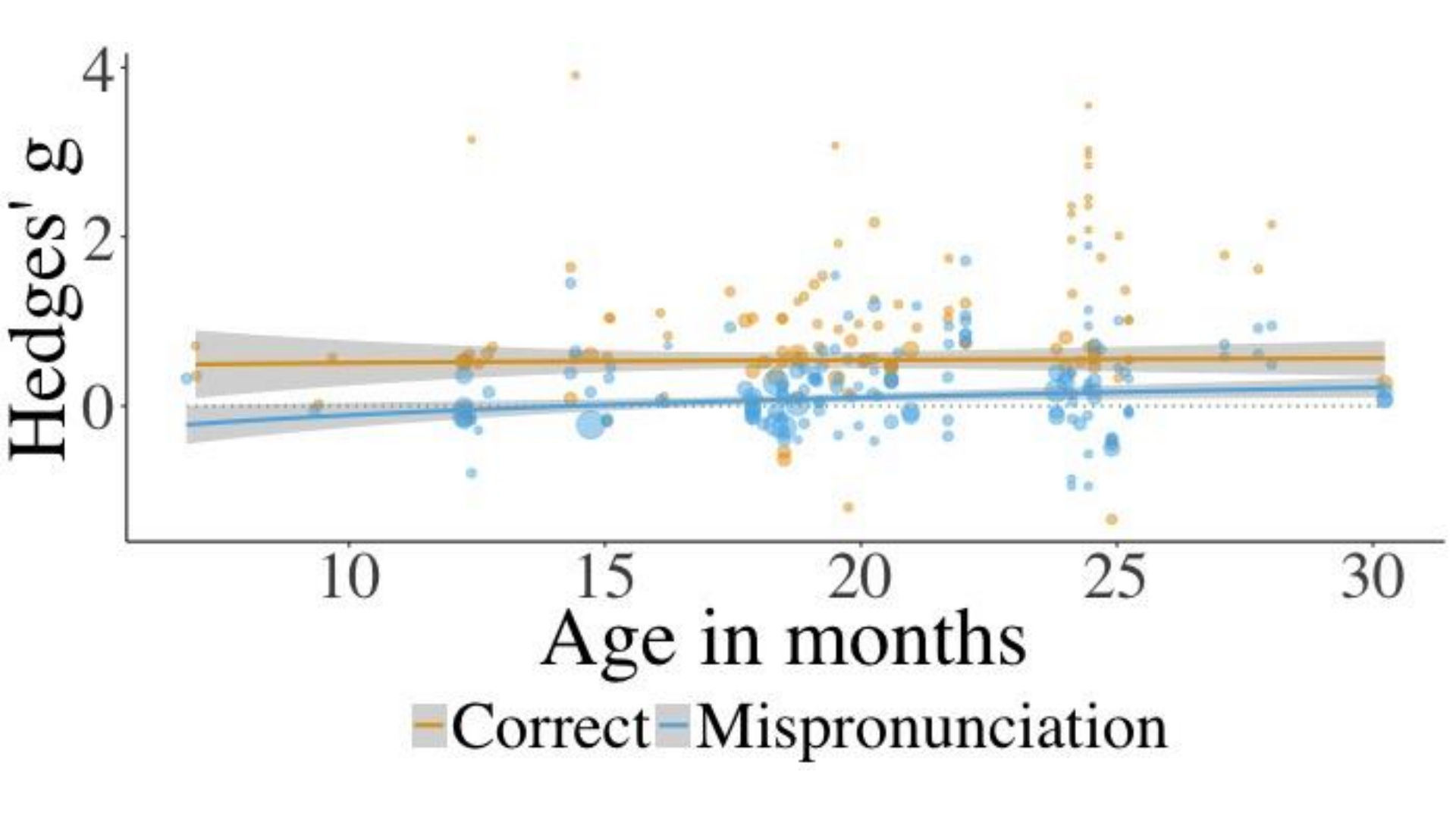
Altvater-Mackensen & Mani, 2013; Mani & Plunkett, 2007; van der Feest & Fikkert, 2015

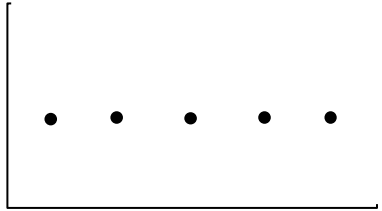
Mani & Plunkett, 2011

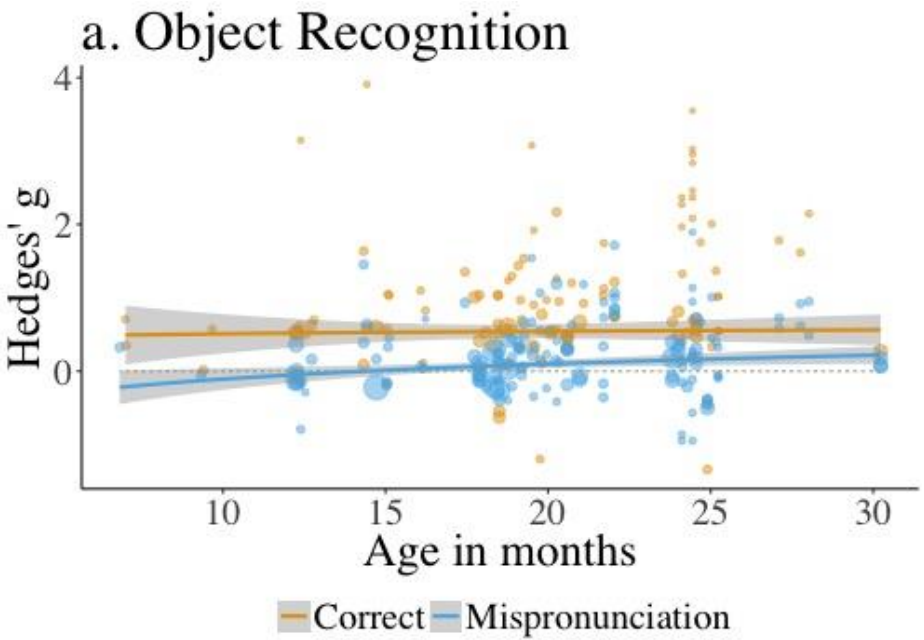
Swingley & Aslin, 2000; Bailey & Plunkett, 2002; Zesiger et al., 2012

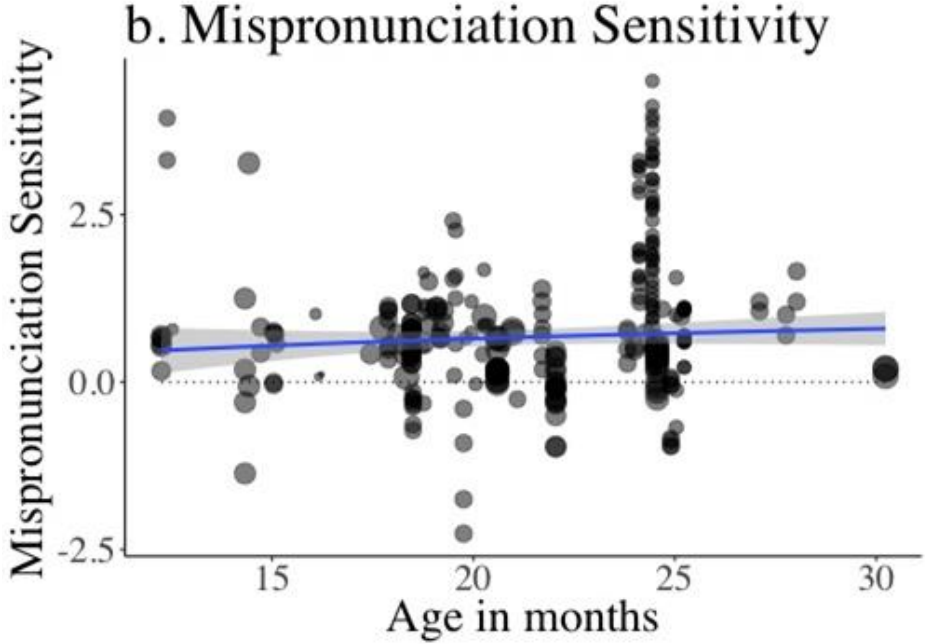
What can we gain from meta-analyses?

- Experimental planning
  - Aggregated across studies, what is the true effect?
  - Statistical power and sample size
- Theoretical insights
  - Is this modulated by different variables?
    - Age/vocabulary


Hedges' g
4
2
0
10
15
20
25
30
Age in months
Correct
Mispronunciation







Early specificity: Mispronunciation sensitivity stays the same with age

=

=

Early specificity: Mispronunciation sensitivity stays the same with age

What can we gain from meta-analyses?

- Experimental planning
  - Aggregated across studies, what is the effect size?
    - Medium sized effect
  - Statistical power and sample size
    - Most studies are probably underpowered
    - To achieve 80% power (recommended), need to test 44 infants
- Theoretical insights
  - Is this modulated by different variables?
    - Age/vocabulary
      - No: support for the early specificity hypothesis

What can we gain from meta-analyses?

- Theoretical insights
  - Is this modulated by different variables?
    - Vocabulary size
    - Mispronunciation size, distractor characteristics, language background, etc.
- Publication bias
  - Are significant results overrepresented in the literature?
- Unforeseen insights
  - Exploratory analyses that come to light during data entry
  - The choices we make for data analysis may influence the conclusions we draw in the field

Thank you!

Dr. Christina Bergmann

Student Assistant: Emelyne

Available on PsyArXiv!
psyarxiv.com/dze29

OSF
osf.io/nvc8m

MetaLab
metalab.stanford.edu/