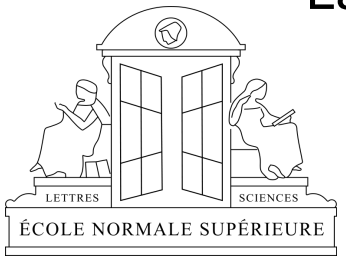
New technologies for the study of child cognition: The case of language

Alex Cristia

alecristia@gmail.com

Centre National de la Recherche Scientifique

Laboratoire de Sciences Cognitives et Psycholinguistique



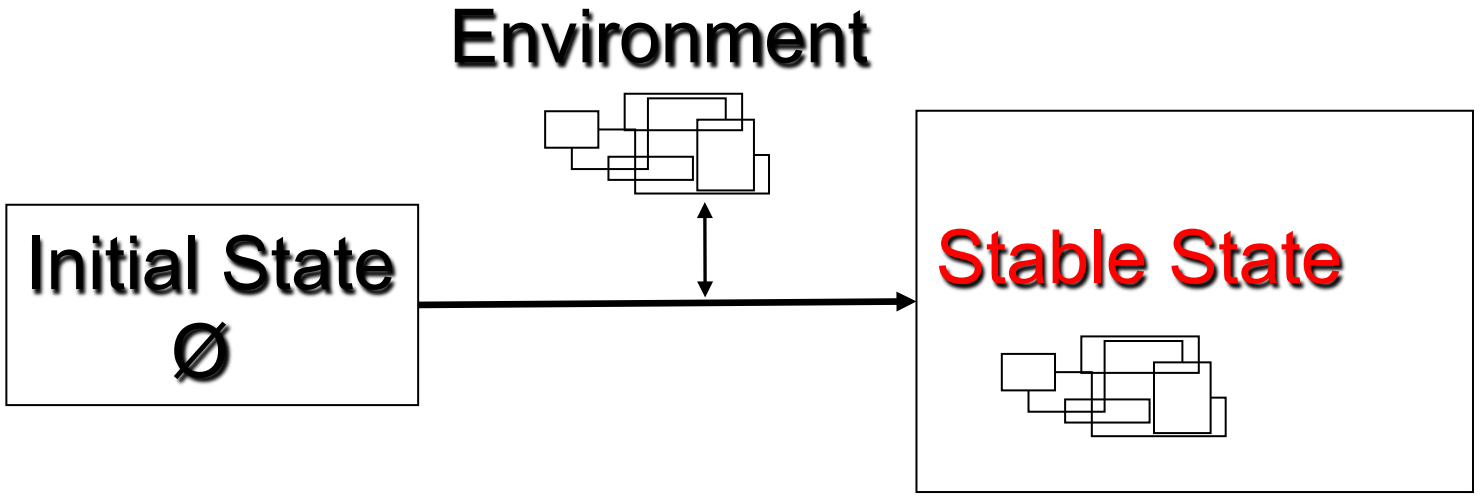
Some ground rules

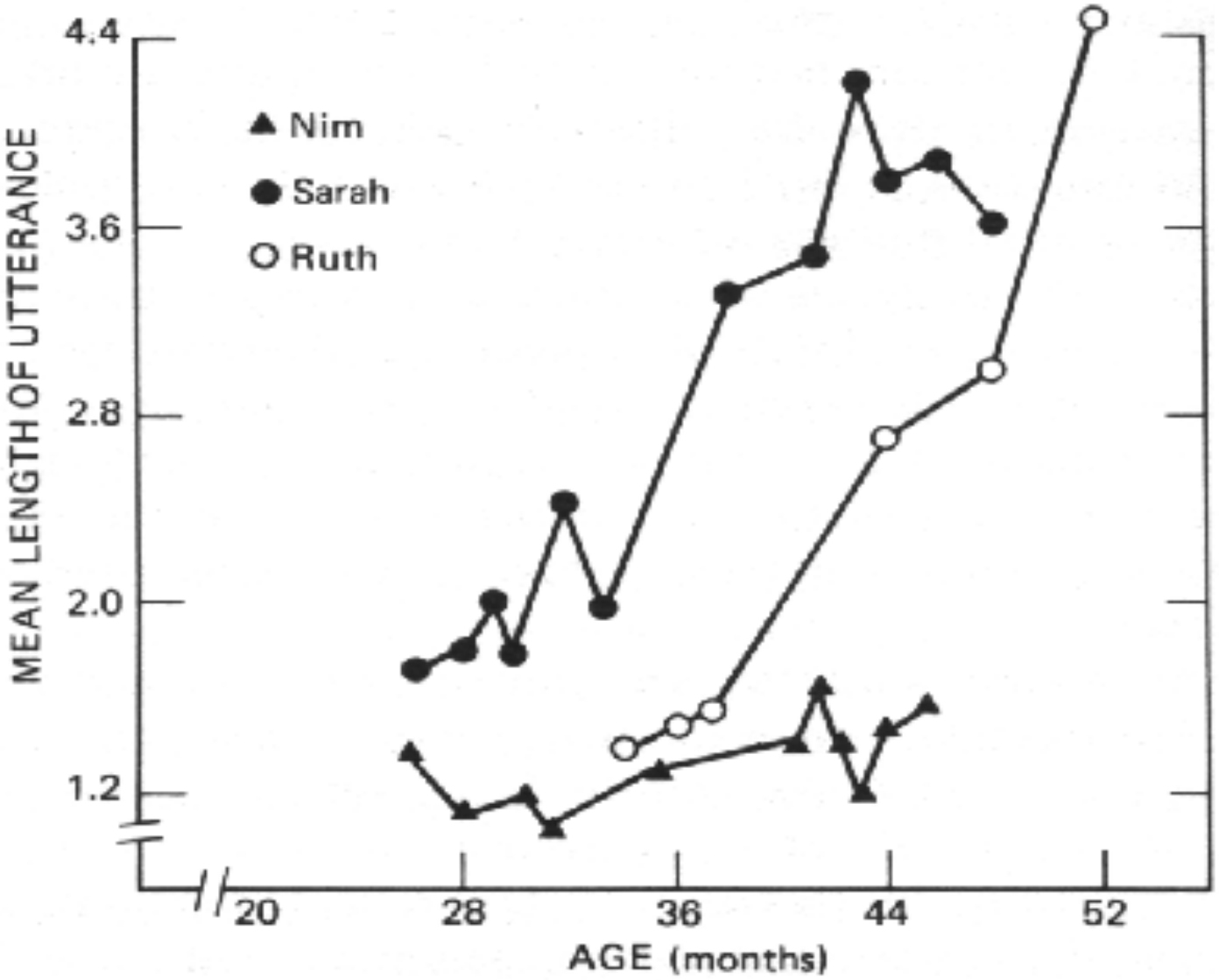
1. Interrupt me!
2. Keep a question/advice sheet

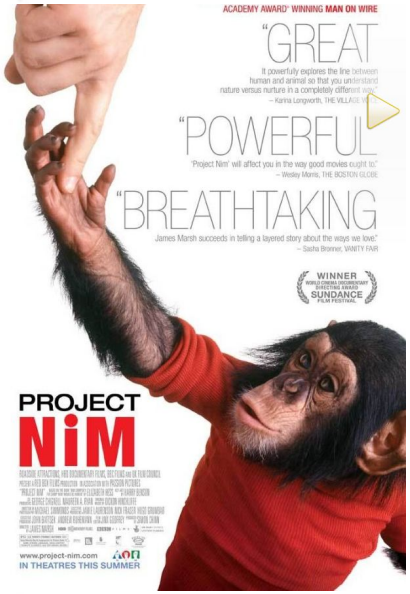
Why is early language acquisition a good case study?

-3 0 3 6 9 12

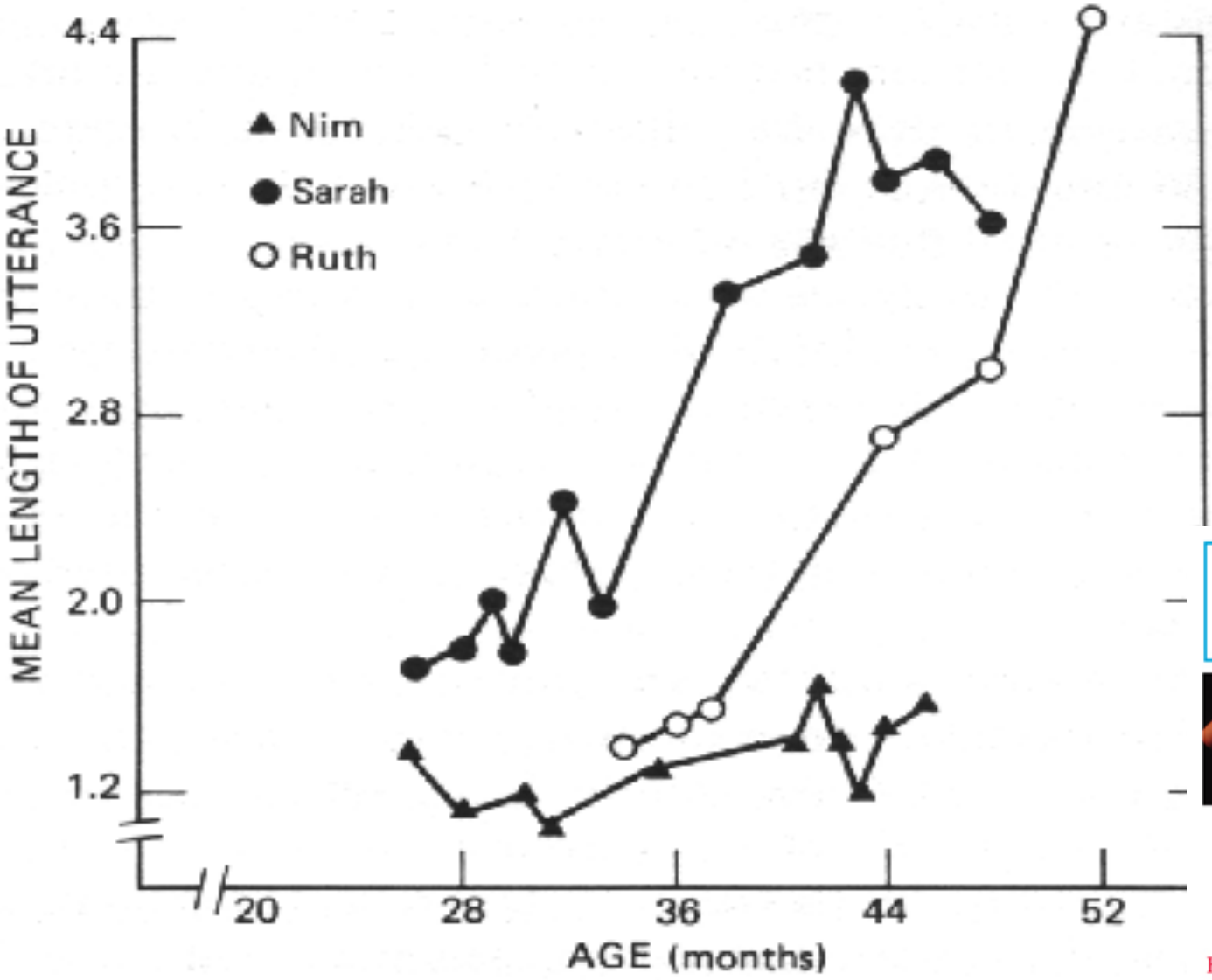
Age (months)

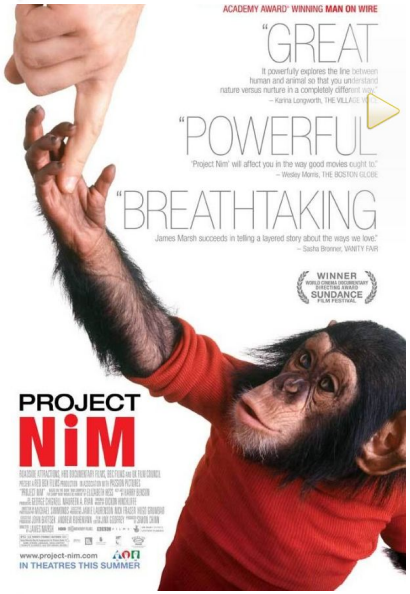

Environment
Initial State
Ø
Stable State



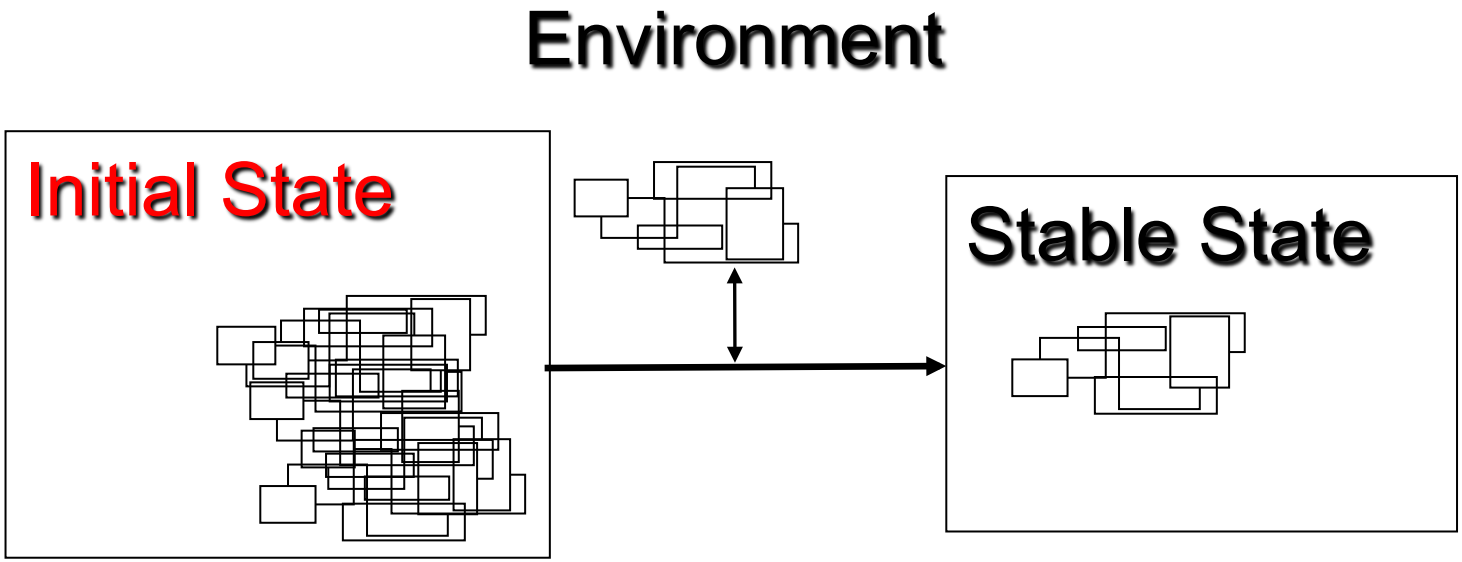


Terrace 1979
Science

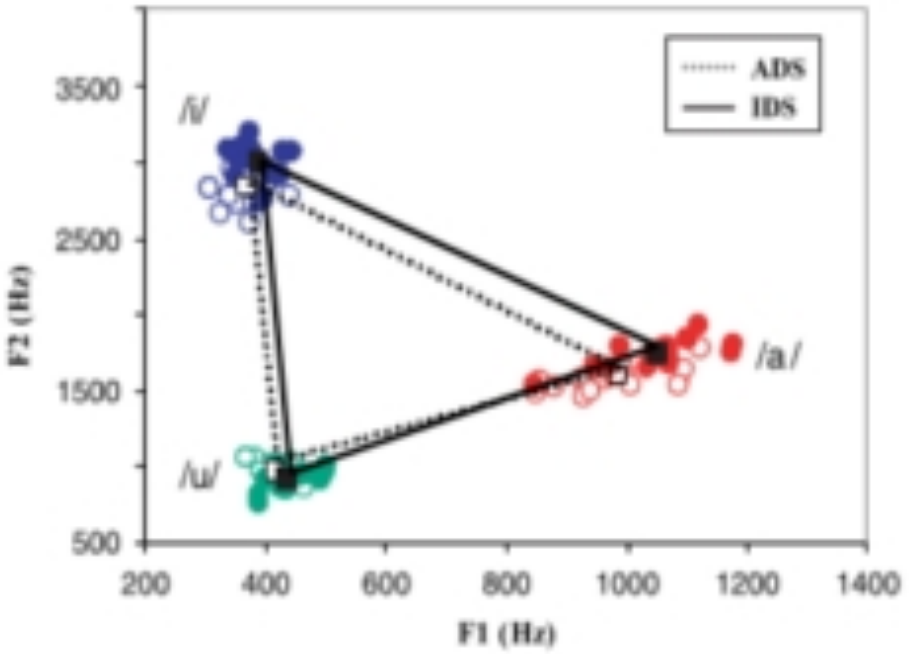


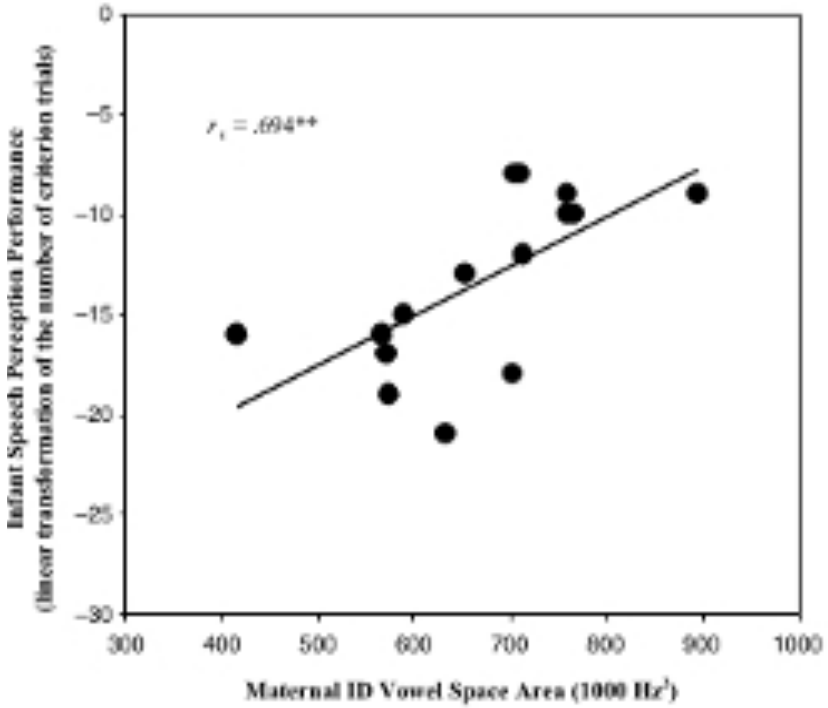


Terrace 1979
Science


Environment
Initial State
Stable State

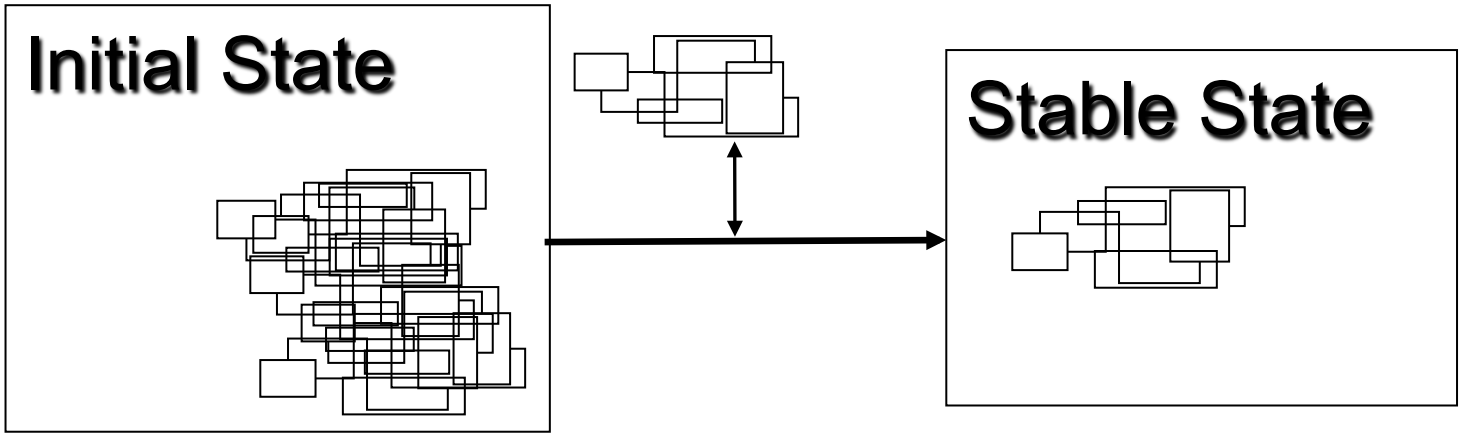
Effects of caregiver speech on infant language





Liu, Tsao & Kuhl 2003 *Dev Sci*

Environment

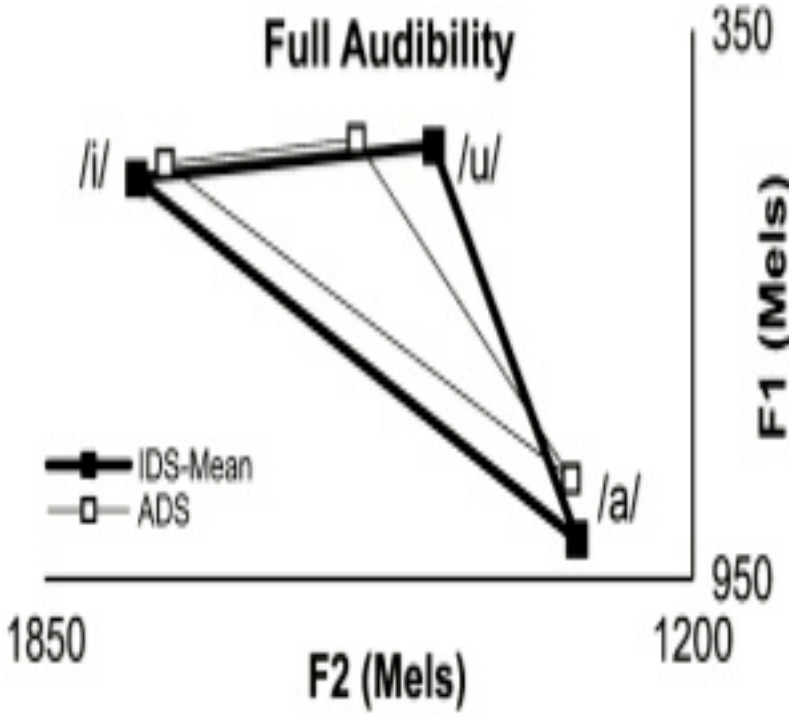


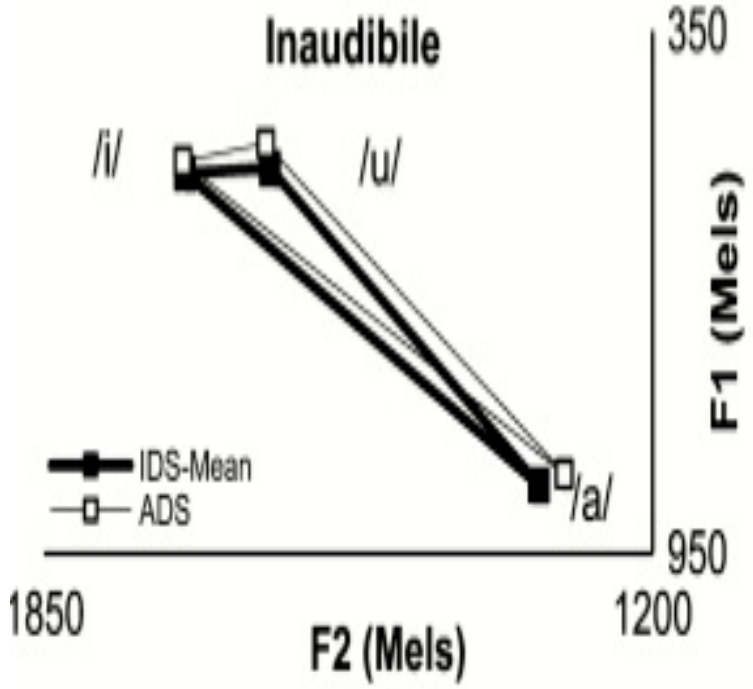
… or …?

	Believed kid could hear	Believed kid could not hear
Kid could in fact hear		
Kid could not in fact hear		

Lam & Kitamura 2012 *Dev Sci*

… or effects of the child on the caregiver?





Lam & Kitamura 2012 *Dev Sci*

How can we study the development of a complex system, which is **robust to variation** but **not fully independent** of the environment?

Goals of this class

- To showcase one “recent & potentially ecologically valid” method for each below:

1. Describe language input and output
2. Measure child behavior
3. Assess neural bases of language processing

- **Contribute to your informed methodological choices**

Students, to your sheets

Slide 16:

A) How many years of research experience do you have?

B) Which of these sections are relevant to you? (Yes/no + further info/comments/questions welcome!)

1. Describe language input and output
2. Measure child behavior
3. Assess neural bases of language processing

Describing verbal input and output

Desiderata

In small groups (1-3 people) come up with two criteria that you believe a method for describing children's input & output should meet

A brief history of input-output

Diary studies

- XVIII-XXth: Stern; Leopold …
- ❑ Focused on child output
- ❑ Not free from bias, can miss key events

CHILDES & other transcribed corpora

- Includes some experimentally-elicited items
- Home recordings capture slice of development, both input & output
- ❑ 'Observer' in the room → ecological validity?
- ❑ Hand-transcription is costly
- ❑ Additional work needed for levels below human consciousness


What we say to babies...
Well, Johnny, you've spilled your milk for the last time! I'll not tolerate that behavior any longer!...
What they hear
dawn

(Semi-)unsupervised analyses

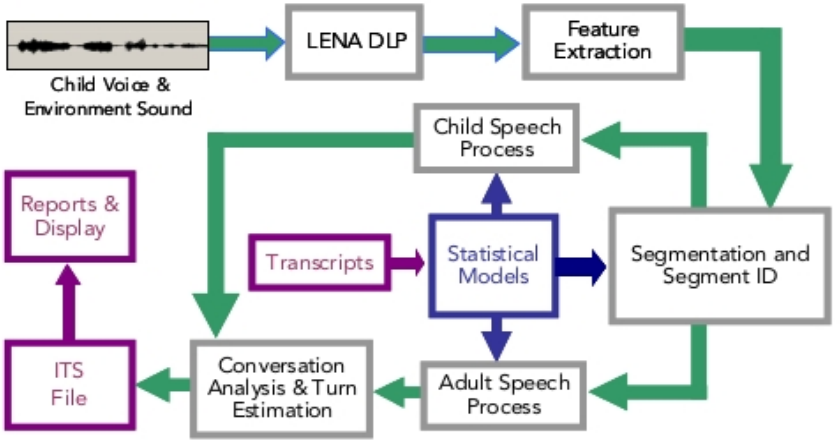


Figure 1. LENA Language Environmental Analysis Audio Processing System

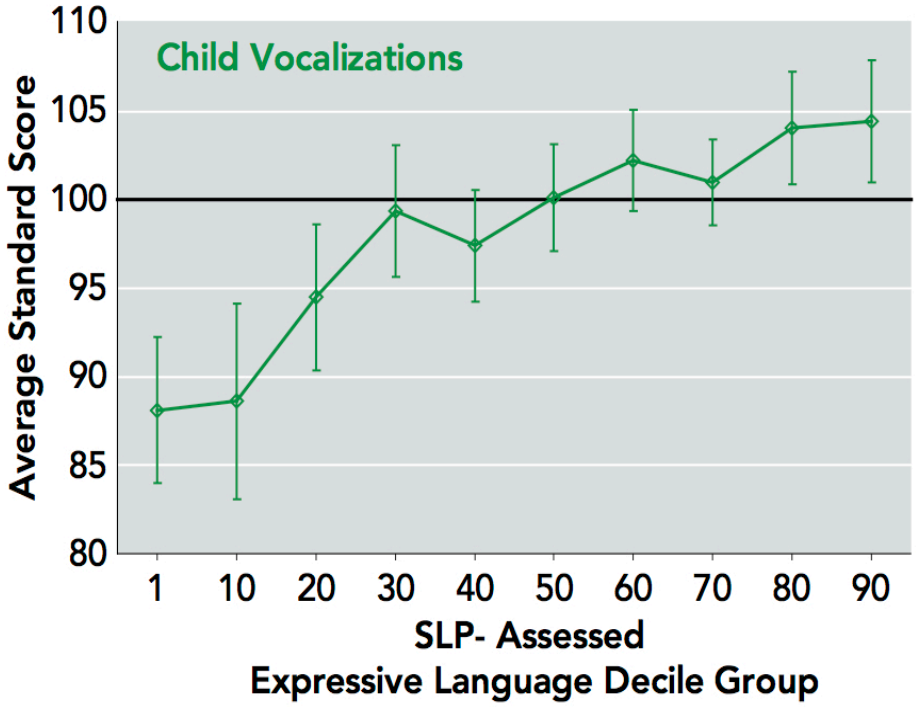
LENA™ Every Word Counts

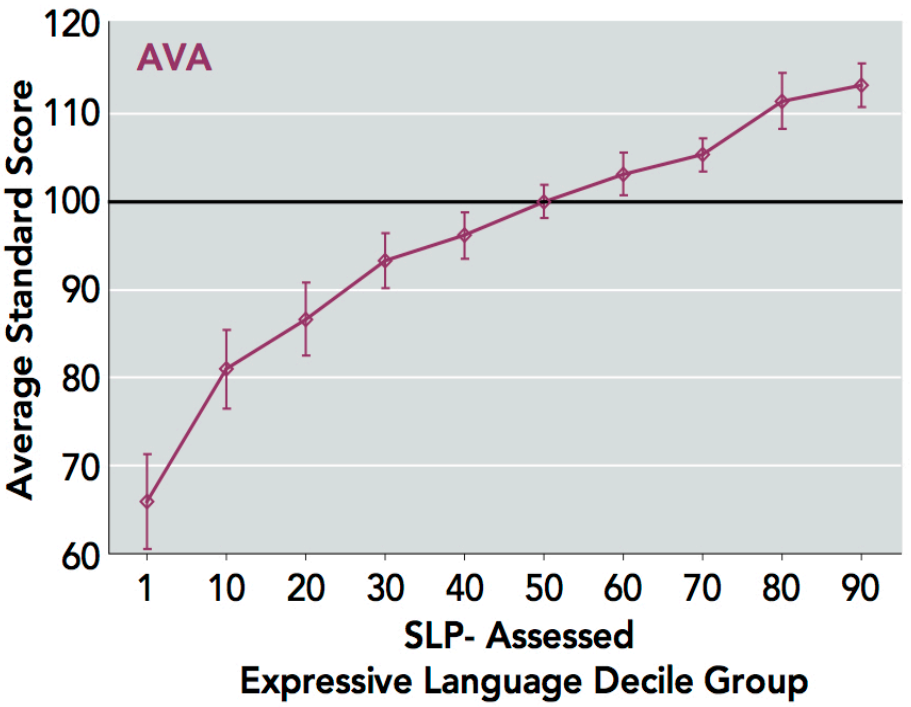
Fundamental reads:
Oller et al. 2010 PNAS
LENA technical reports http://lenafoundation.org/Research/TechnicalReports.aspx

Segment ID Code	Segment Description
MAN / MAF	Male Adult / Male Adult - Faint
FAN / FAF	Female Adult / Female Adult - Faint
CHN / CHF	Key Child / Key Child - Faint
CXN / CXF	Other Child / Other Child - Faint
NON / NOF	Noise / Noise - Faint
OLN / OLF	Overlap / Overlap - Faint
TVN / TVF	Electronic / Electronic - Faint
SIL	Silence

Example output

Example potential descriptive validity





LENA tech report 01-2 “Power of talk”

Further info on LENA

- Cost
  - Software ~12k US$
    - Student version ~2k US$
    - Many people use it as a recording device – not a great idea
  - Recorders .3k US$
  - Clothes...
- Sometimes hard to be the first in the country to buy it
- Foundation, but works like a company
  - Software is a closed box
- Emergent group (now beyond it): DARCLE.org

DARCLE in progress

- Unbiased estimates of reliability
- Open-source daylong recordings analyses
  - Compatible with alternative recording devices
- {Weakly | un}supervised analyses on the pipeline
  - Classifying adult speech into:
    - key-child-directed
    - other-child-directed
    - adult-directed
  - Classifying languages in multilingual recordings, bottom-up detection of situations...
  - But not fully-automatic transcriptions ☹

Ask me about this summer's data collection on traditional societies!

Students, to your sheets

Comments/questions on this section?

Measuring child behavioral

Traditional methods

Pros

- Quite widely applied
- Acceptance of their use
- Can benefit from others' expertise or experiences directly or indirectly

Cons

- Field does not always utilize previous knowledge
- Often design taps language skills indirectly
- Most often, not well-understood
- Imprecise measures, particularly of individuals

Traditional methods

Pros

- Quite widely applied
- Acceptance of their use
- Can benefit from others' expertise or experiences directly or indirectly

Cons

- **Field does not always utilize previous knowledge**
- Often design taps language skills indirectly
- Most often, not well-understood
- Imprecise measures, particularly of individuals

Goals when reading previous work

Peruse merrily

Have a general idea on some topic

Find a worthwhile research area

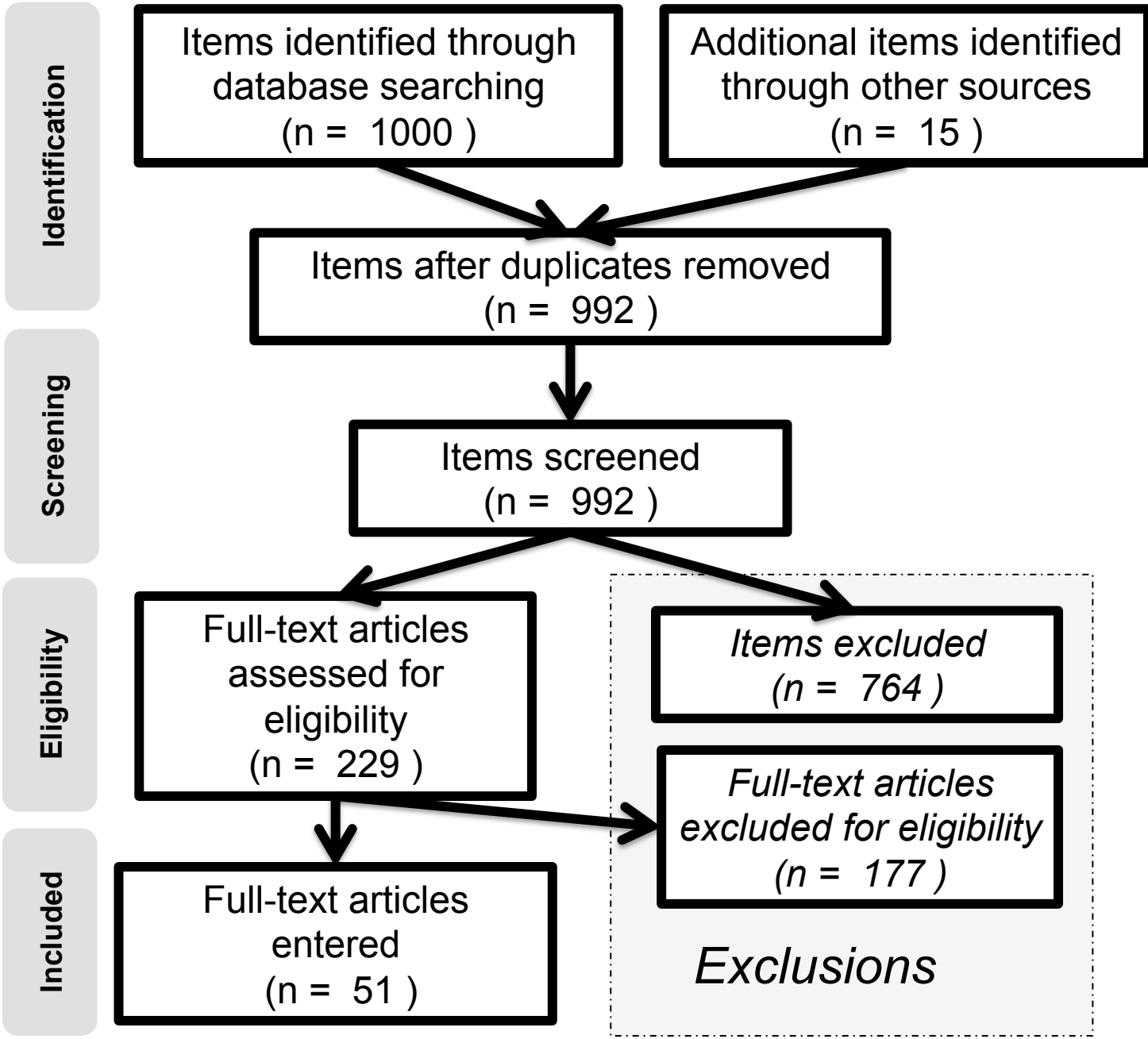
Extend current knowledge

Personal cost of being wrong =

Effort you should be ready to spend on preparation

Follow PRISMA

Sample workflow in a systematic review



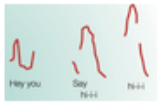
Home
Visualizations
Power Analysis
Reports
Browse Raw Data
Documentation
Contribute
Team
Source code

MetaLab

Interactive tools for community-augmented meta-analysis,
power analysis, and experimental planning in language acquisition research

12	269	981	12,029
Meta-analyses	Papers	Effect sizes	Participants

Meta-analyses currently in MetaLab



Infant directed speech prefere

Looking times as a function of whethe
directed vs. adult-directed speech is pr
stimulation.

We welcome researchers interested in contributing to Metalab. Pl

Contributions can take various forms:

1. Adding unpublished data. Ideally we can link to a report on necessary data.
2. Suggesting and/or adding published papers not yet include necessary data is either included or provided in your email.
3. Creating new meta-analyses. You can either collaborate on see the section listing resources further down.

New meta-analyses looking for contributors

Currently, the following topics are in the process of becoming a m

- Segmenting words from an artificial stream of speech, an e
- Detecting mispronunciations of known words, without a lea

Resources

Here are a few resources on creating meta-analyses compatible w

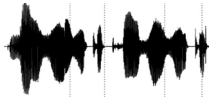
- *Metalab MA template*
- *Code book*
- *Interrogating PubMed via a script*

With
M. **Braginsky**, M. **Lewis**, M. **Frank**,
P. **Piccinini**, C. **Bergmann** & S. **Tsuji**

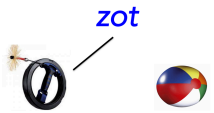
Phenomena in MetaLab

Prosody

Sounds



Words

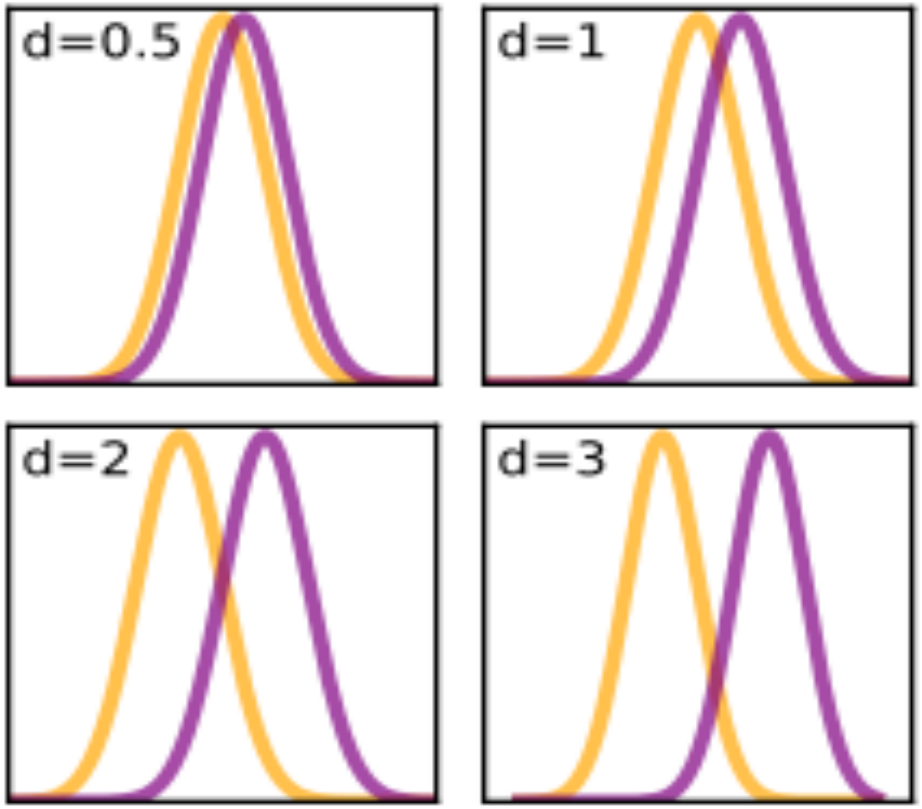


Communication

Phenomenon	N papers (conditions)
IDS preference (Dunst, Gorman, & Hamby, 2012)	16 (50)
Phonotactic learning (Cristia, in prep.)	15 (47)
Vowel discrimination (native) (Tsuji & Cristia, 2014)	40 (167)
Vowel discrimination (non-native) (Tsuji & Cristia, 2014)	21 (72)
Statistical sound learning (Cristia, in prep)	11 (40)
Word segmentation (Bergmann & Cristia, 2015)	67 (295)
Mutual exclusivity (Lewis & Frank, in prep.)	20 (60)
Concept-label advantage (Lewis & Long, unpublished)	16 (100)
Online word recognition (Frank, Lewis, & MacDonald, 2016)	12 (32)
Gaze following (Frank, Lewis, & MacDonald, 2016)	15 (45)
Pointing and vocabulary (Colonnesi et al., 2010)	25 (30)

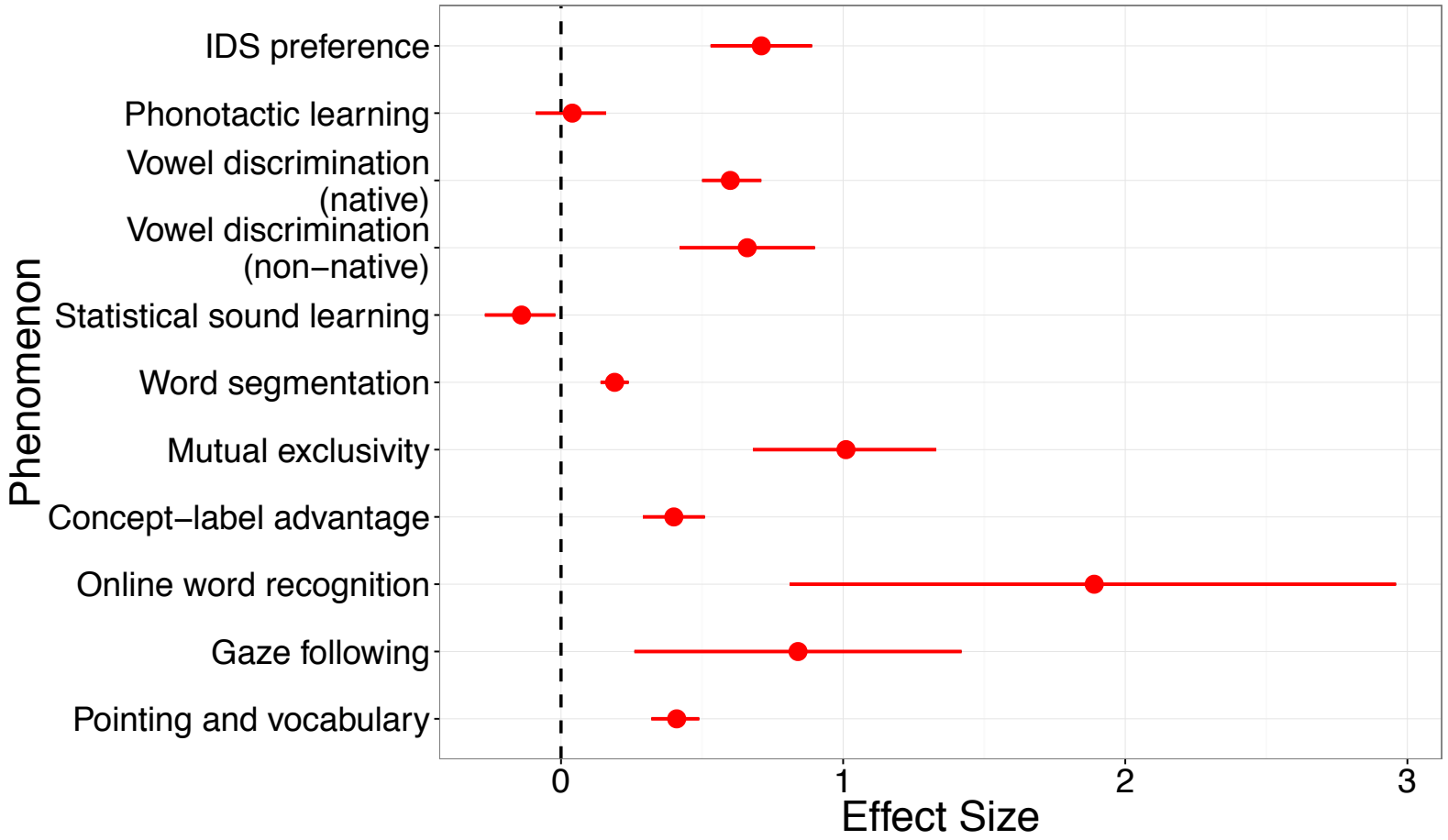
Thanks to Molly Lewis!

Why effect sizes are useful for study planners



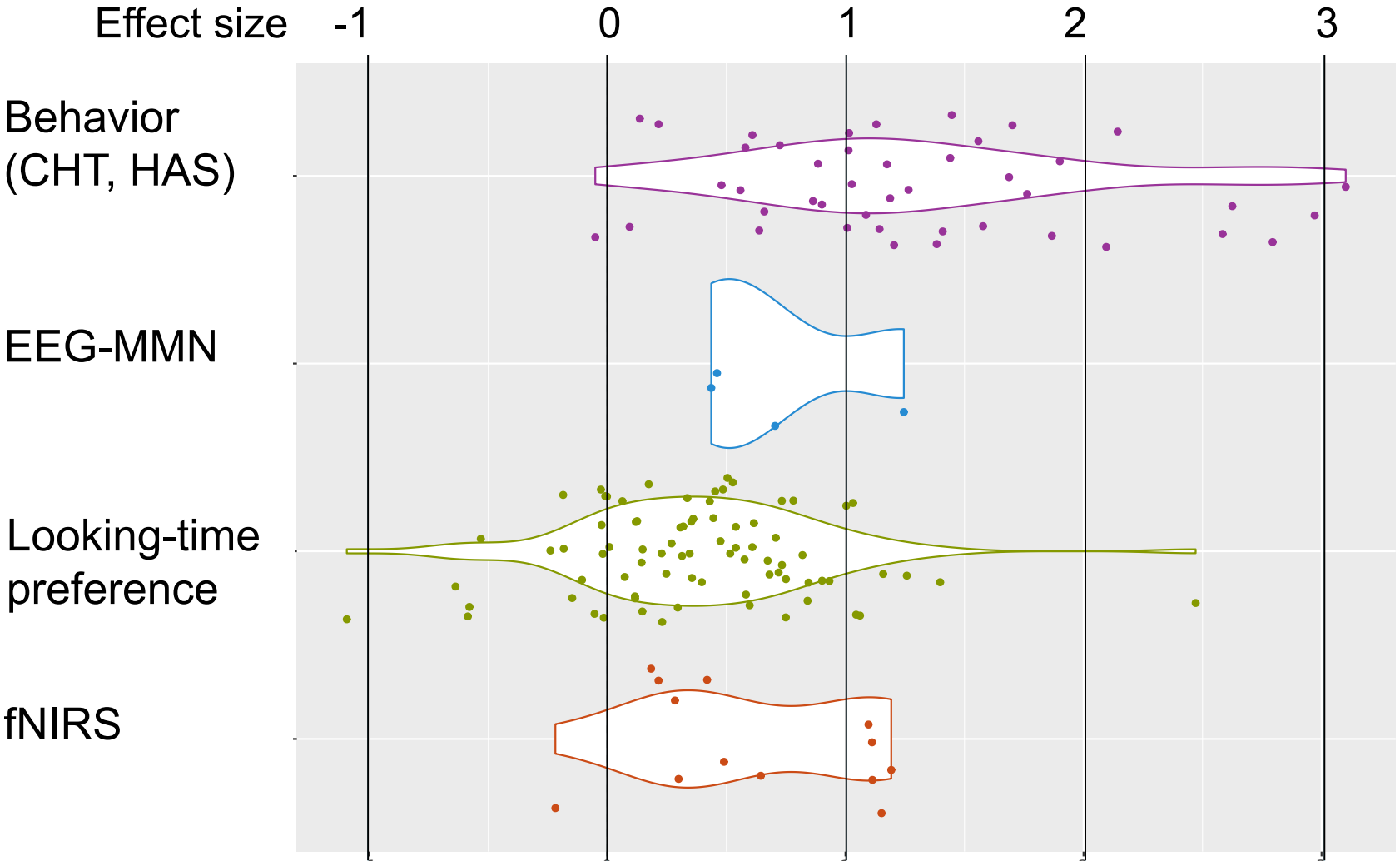
Wikipedia on Effect size

Effect sizes by phenomenon



Thanks to
Molly Lewis!

Effect sizes by method



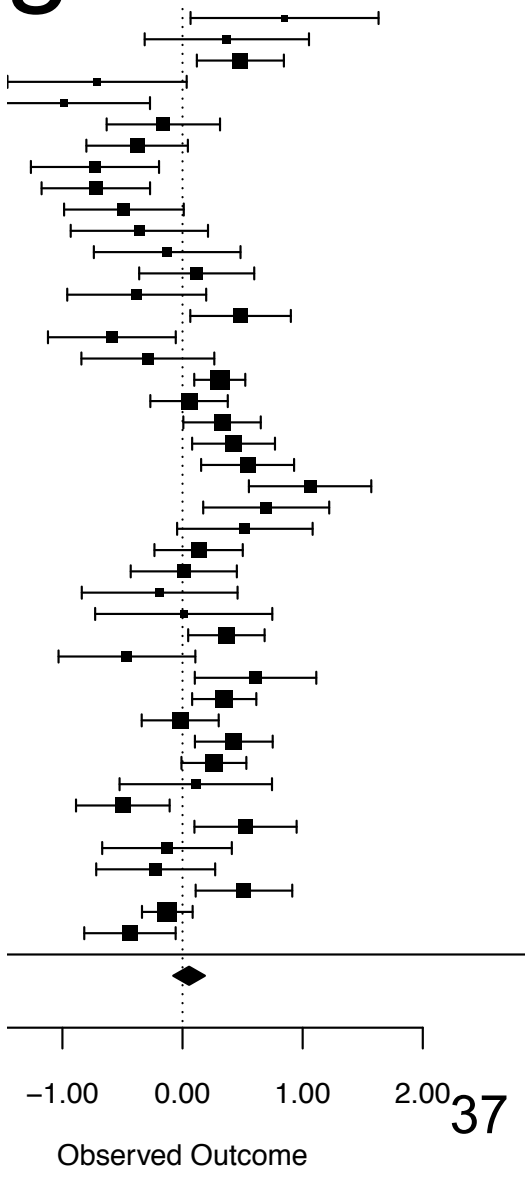
InPhonDB – native vowels (Tsuji & Cristia, 2013)

visualization from metalab.stanford.edu

Not well understood: E.g. sound pattern learning

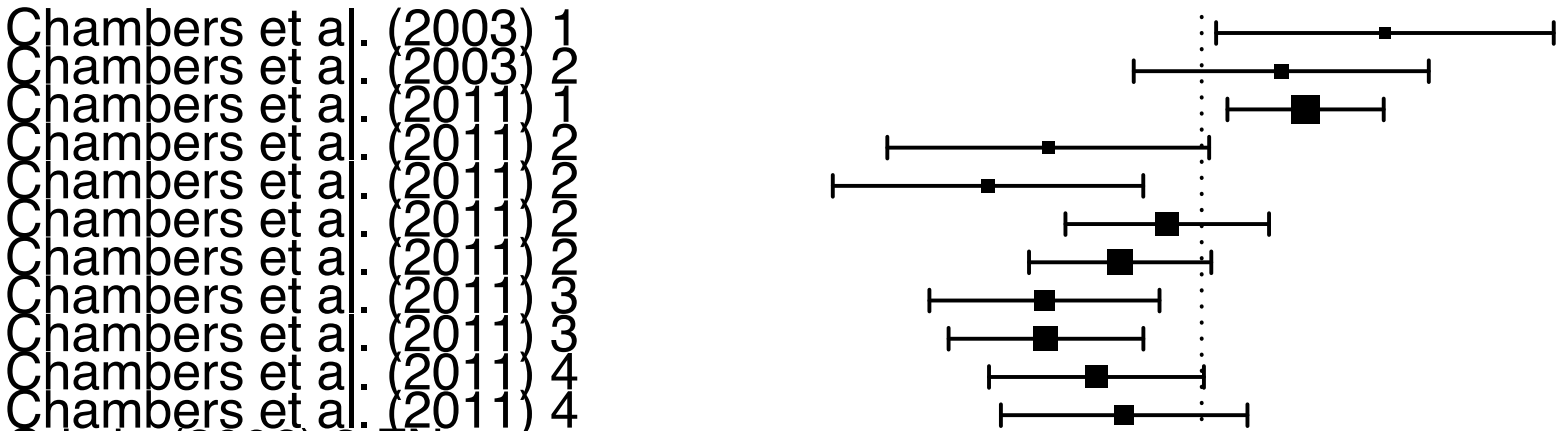
Fixed-length familiarization to a sound pattern (e.g., all the words begin with “p”)

+ preferential looking test



Not well understood: E.g. sound pattern learning

Variation even within a single lab!!



If interested in individual differences, ask me about that!


limb movement
Anticipatory eye-movements
gaze-contingent eye-tracking
Oh so THAT's why baby researchers spend so long tinkering with methods!
skin conductance
heart rate
pupillometry

Desiderata

In small groups (1-3 people) come up with two criteria that you believe a method for measuring children's behavior should meet

Tests on touch screens

Some previous work

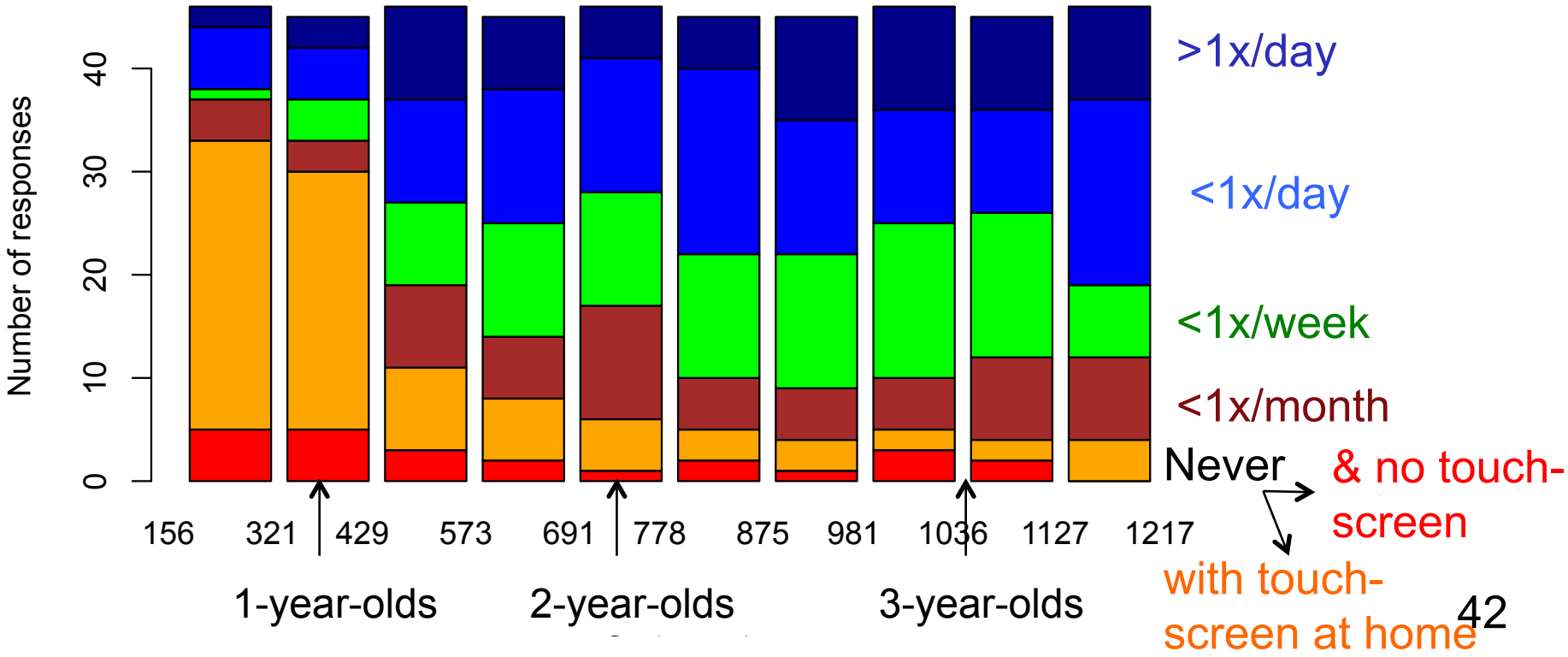
- On learning from (interactive) screens versus humans (see e.g. Heather Kirkorian's work)
- Computerized Comprehension Test (see Margaret Friend's work)

→ both in the lab, but can we take it out to the wild?

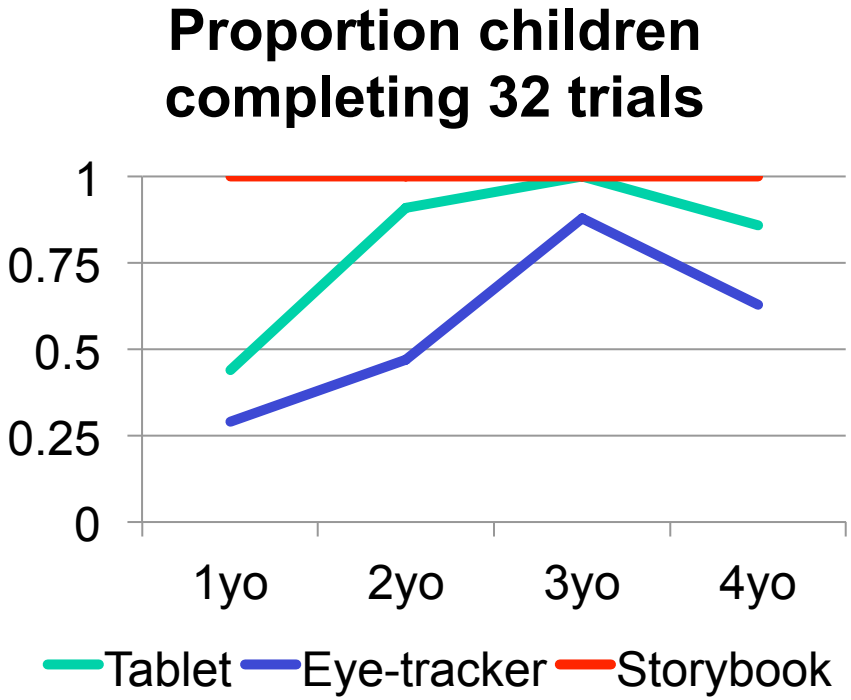
Ecological validity

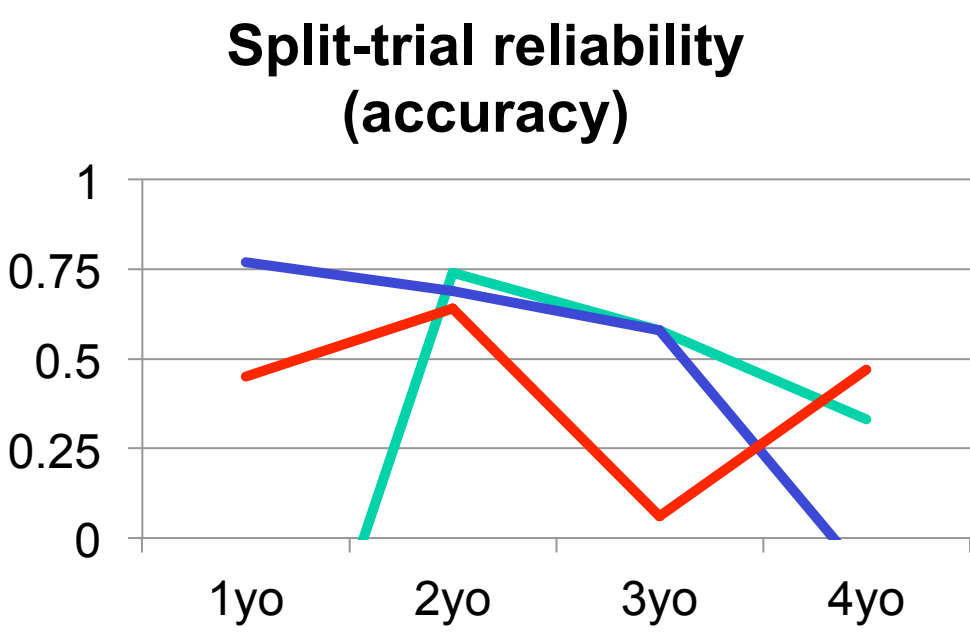
Cristia & Seidl (2014) Plos

- Online questionnaire sent to families who had previously participated in a French babylab
- 455 respondents with infants 5 to 40 mo



Direct comparison: 2AFC word comprehension in a museum





Frank, Sugarman, Horowitz, Lewis, & Yurovsky, in press
total N kids >200

A French attempt

51 2yo tested in 3 Parisian daycares

Phase I: Warm-up

Play “pop the bubble” game

Phase II: Training

3 trials with known words → repeat until correct answer

1 novel minimal pair (s/sh) → do not repeat

Phase II: Test

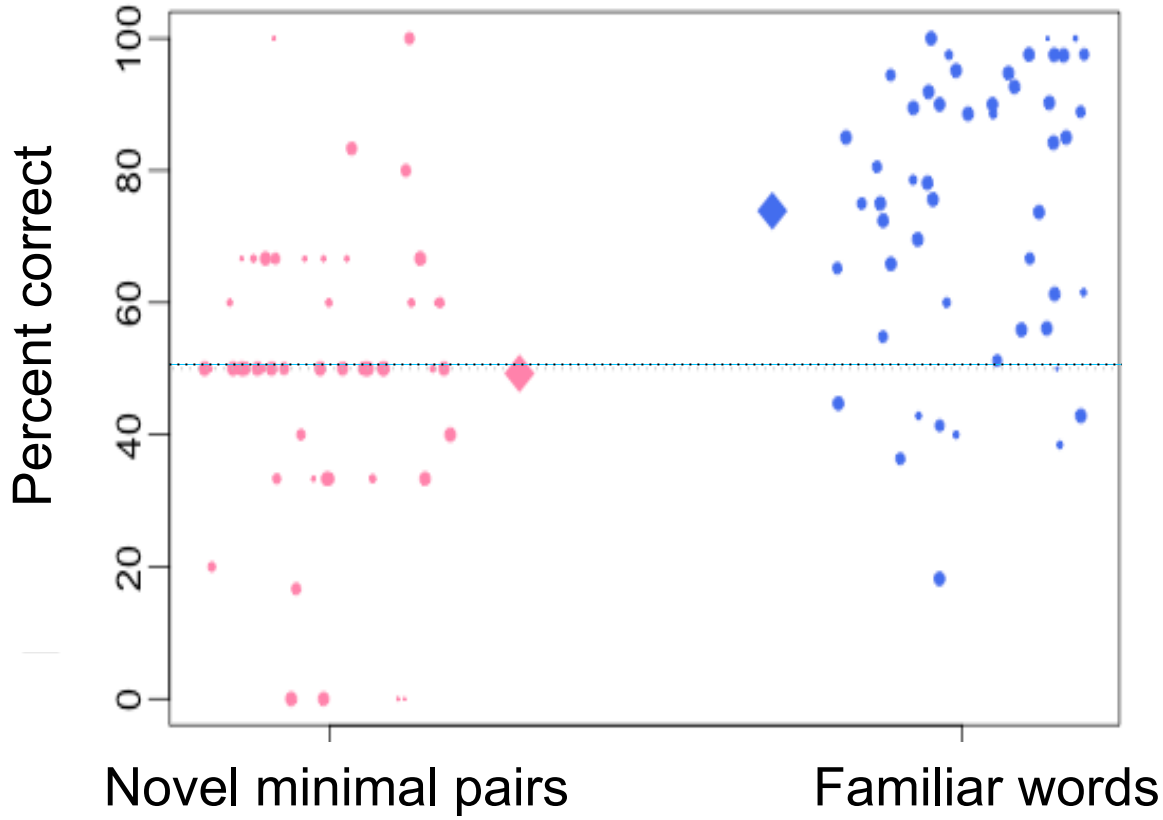
41 French pairs in pseudo-randomized order (with constraints on side of target & difficulty level repetition)

+ 2 novel minimal pairs (s/sh)

Novel minimal pairs are introduced upon appearance of the image, familiar ones are not

With Laia Fibla & Charlotte Maniel

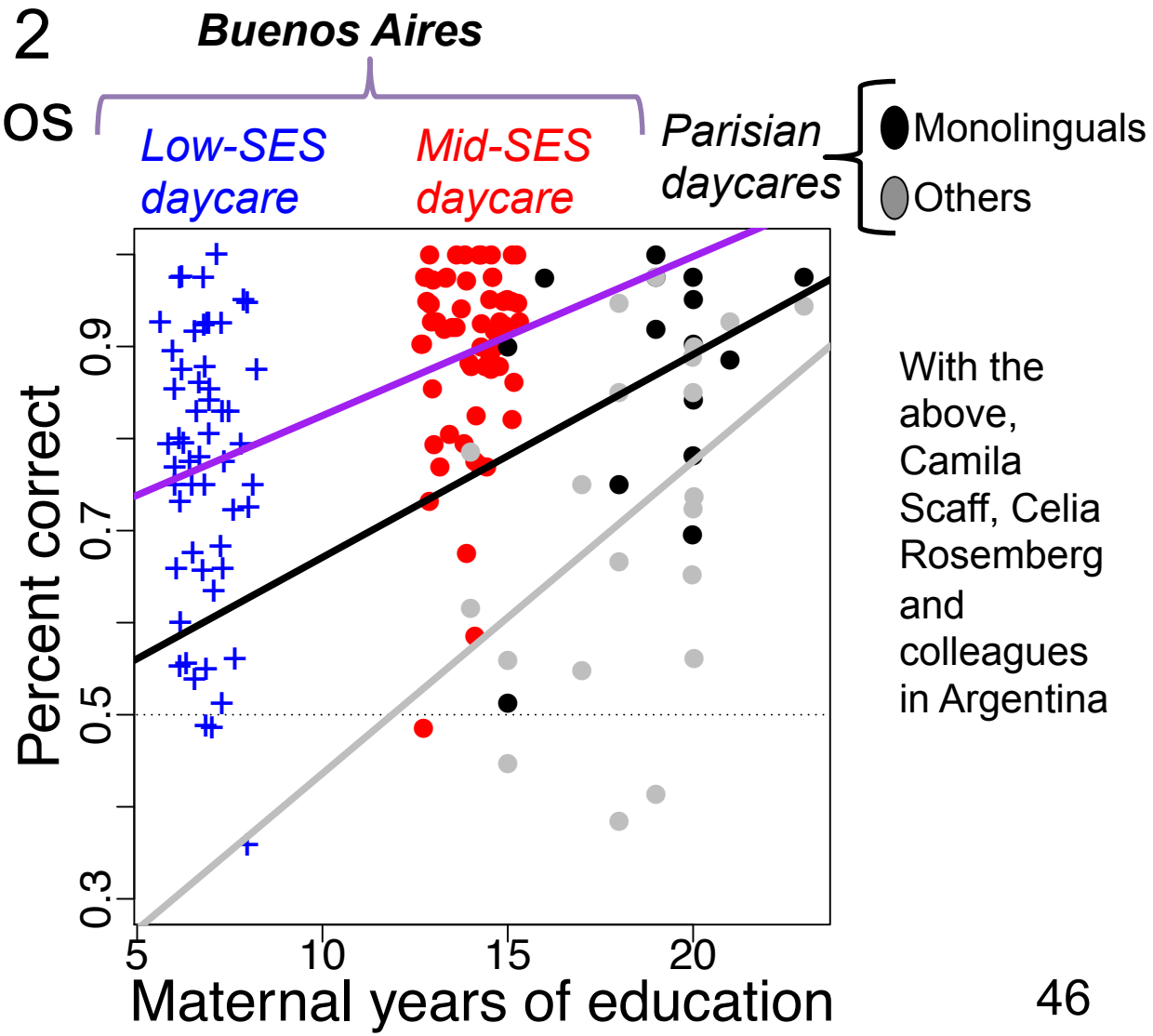
Overall accuracy



With Laia Fibla & Charlotte Maniel

Cross-cultural extension

120 3yo tested in 2 daycares in Buenos Aires

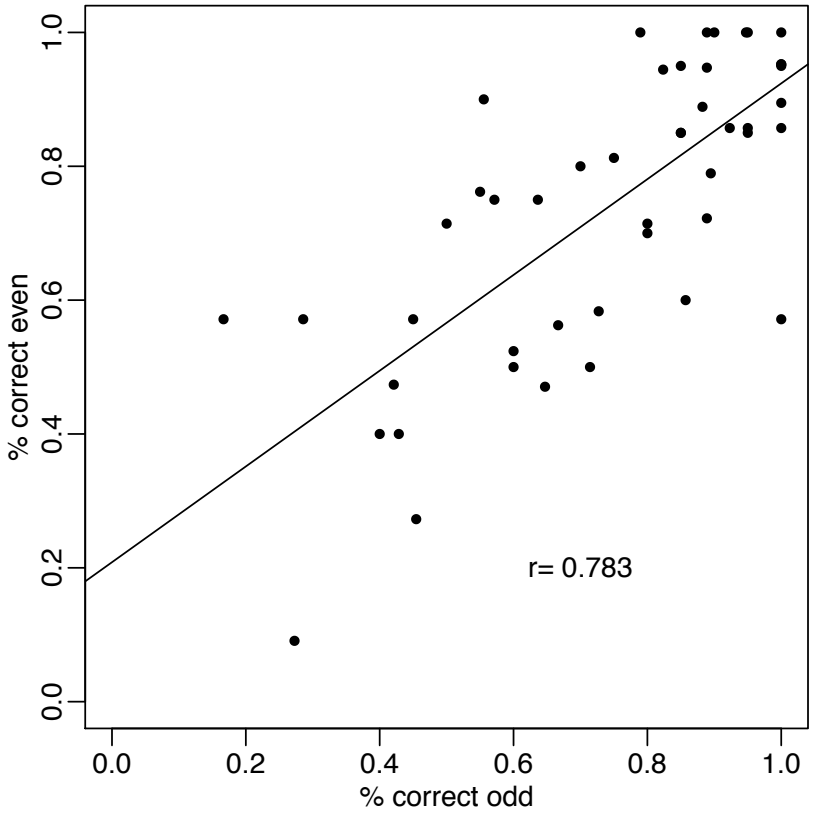


With the above, Camila Scaff, Celia Rosemberg and colleagues in Argentina

Ask me about current extension testing at home!

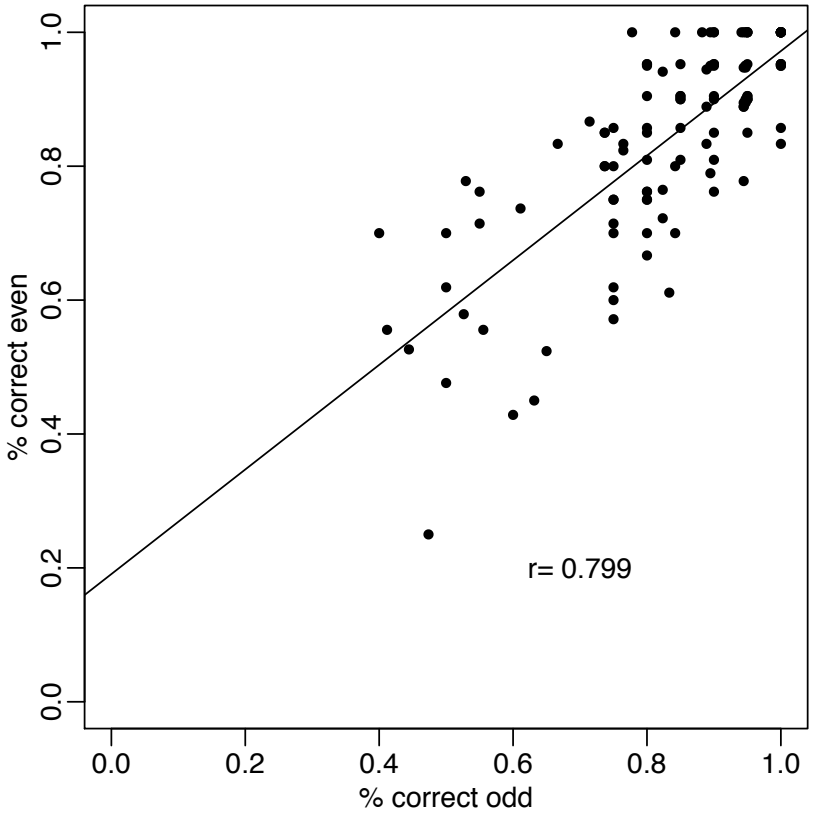
Split-trial reliability: Accuracy

Paris 1.9- to 3.3-year-olds

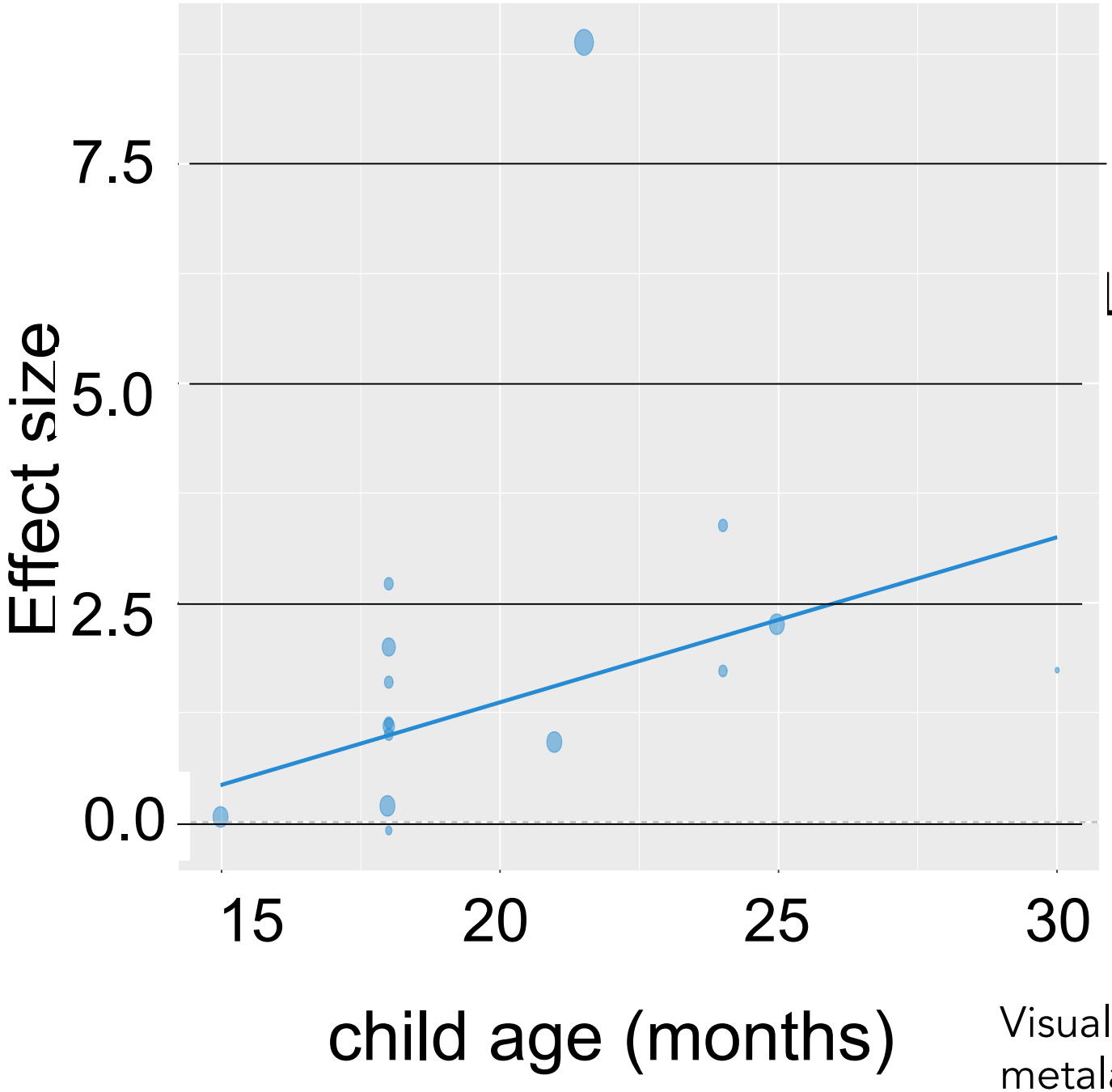


raw Cronbach's alpha = .85

Bs. As. 2.9- to 4.2-yo

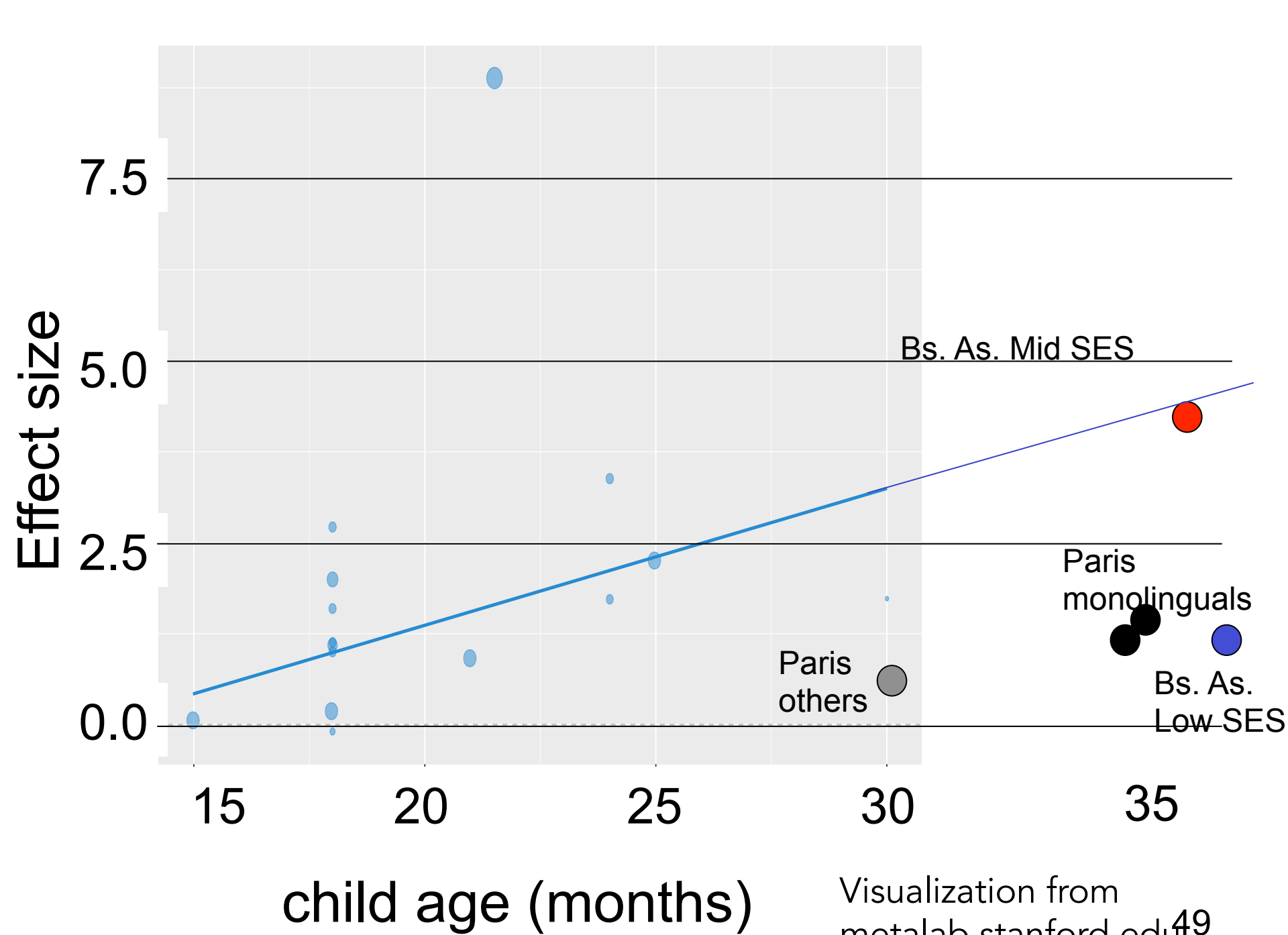


raw Cronbach's alpha = .87



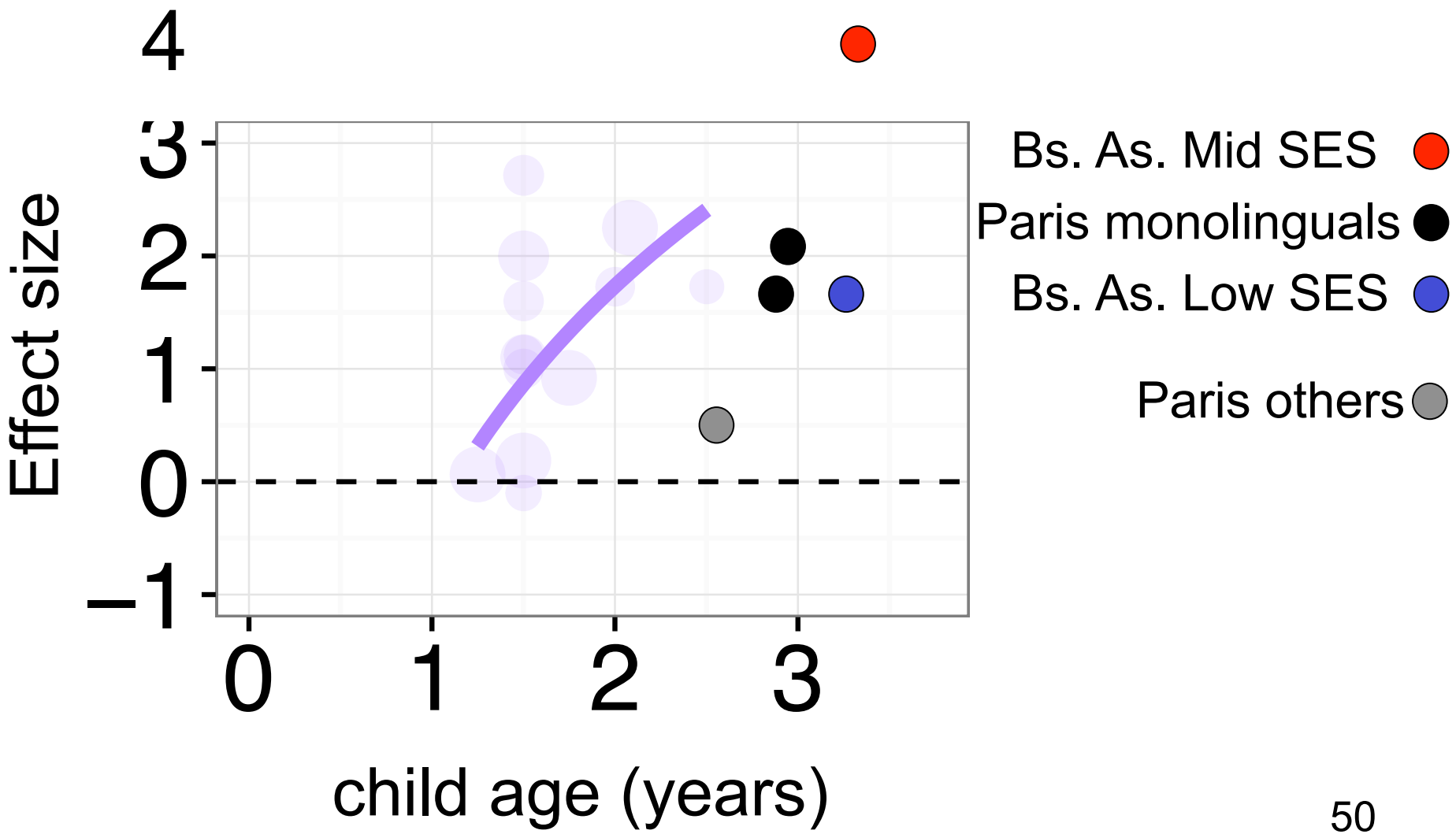
MetaLab
Looking-while-listening lab studies

Visualization from metalab.stanford.edu



Visualization from
metalab.stanford.edu

Effect sizes by phenomenon



Lots behind the scenes!

Resources now starting to appear

- Our app http://github.com/alecristia/mandy_ipadVocab/
- Mike Frank http://github.com/langcog/tablet/
- Python tb https://github.com/piantado/kelpy

Ask me about:

- *Interactive gestures you could build upon (drag & drop?)*
- *Hardware (apple vs. android, phone vs. tablet)*
- *Working with in-lab engineers versus programmers (free-lance & companies) versus self*

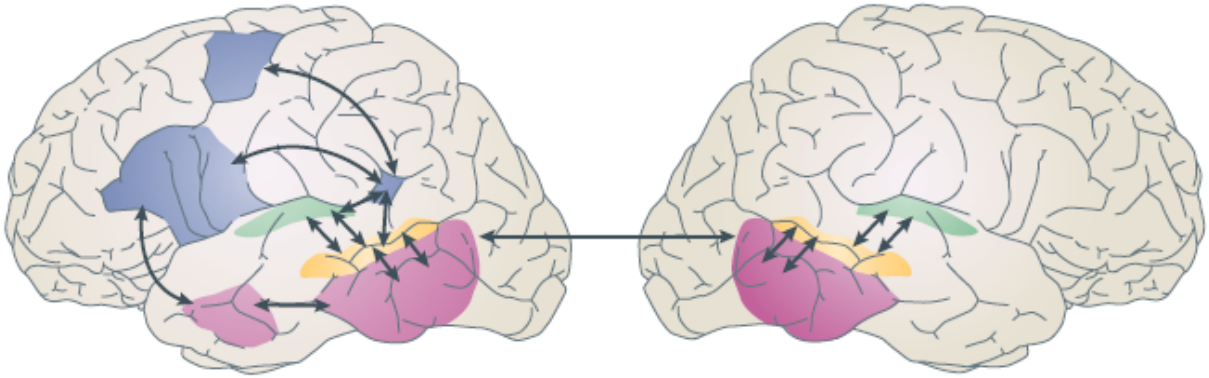
Students, to your sheets

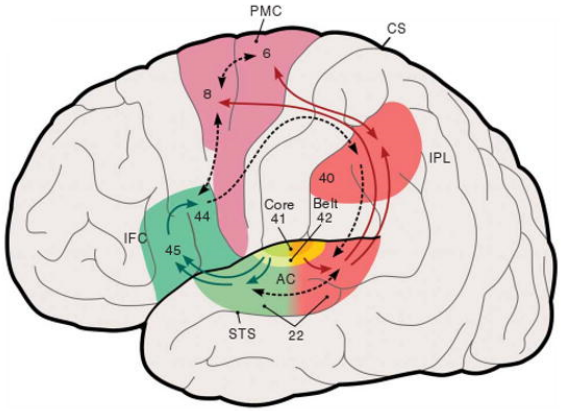
Comments/questions on this section?

Neural bases of language

How does the human brain sustain language acquisition and processing?

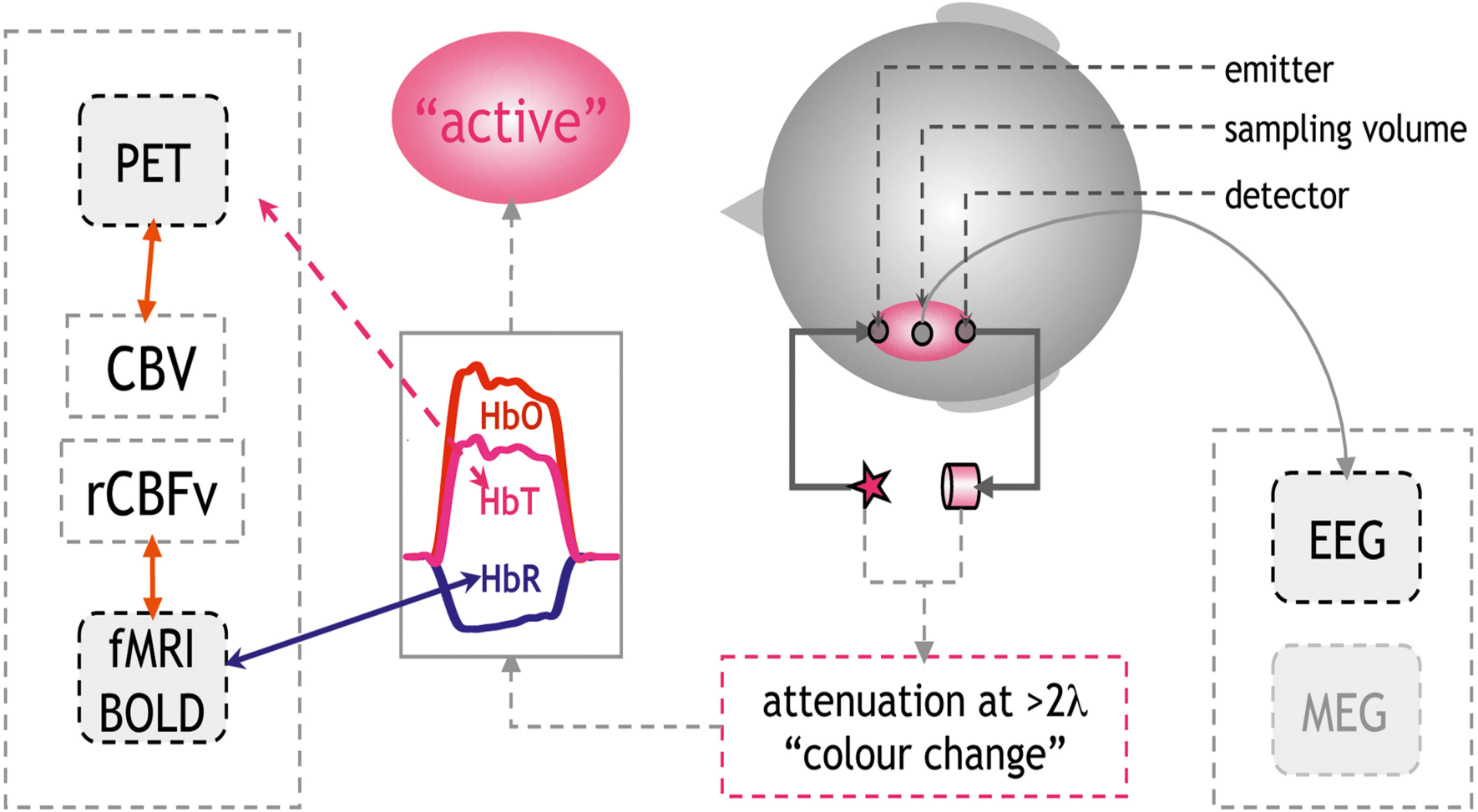
Hickok and Poeppel 2007





Rauschecker and Scott 2009

Local activation, neurons firing, and neurovascular consequences



Obrig et al. 2010 *Frontiers*

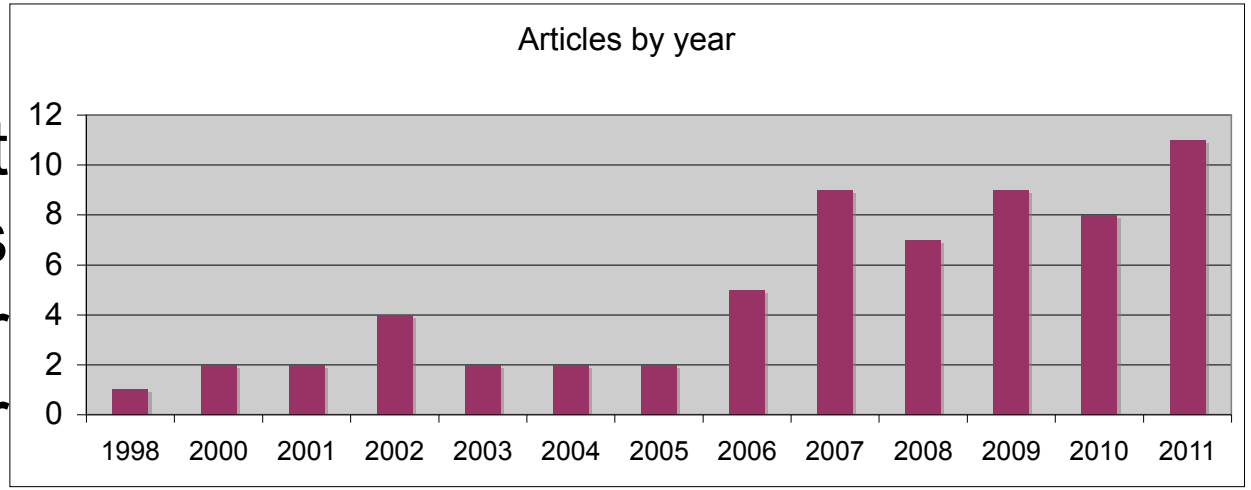
Desiderata

In small groups (1-3 people) come up with two criteria that you believe a method for measuring children's neural responses should meet

Enter Near InfraRed Spectroscopy

	fMRI	(f)NIRS	EEG
awake	(no)	yes	yes
visual stims	(no)	yes	yes
movement	highly sensitive	(relatively insensitive)	very sensitive
spatial	mm	cm+uncertainty?	(reconstructed)
temporal	seconds	seconds	milliseconds

Number of infant fNIRS articles published per year



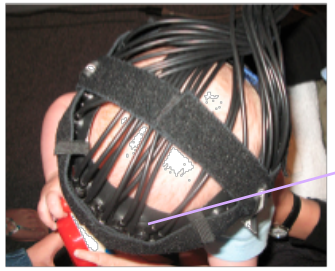
A systematic review in Cristia et al. 2012 *Plos One*

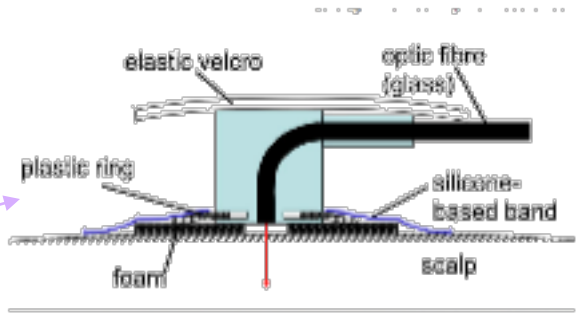
Basics of fNIRS

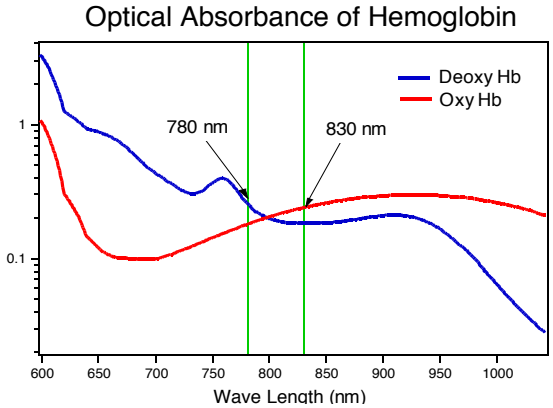
Fundamental reads:
Lloyd-Fox et al. 2010
Neurosci Biobeh Rev

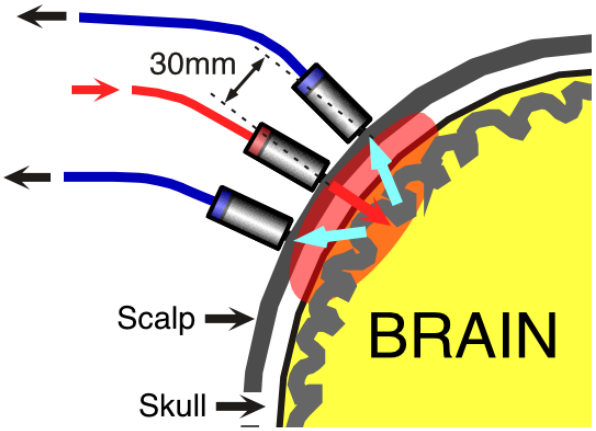
Obrig et al. 2010 Frontiers

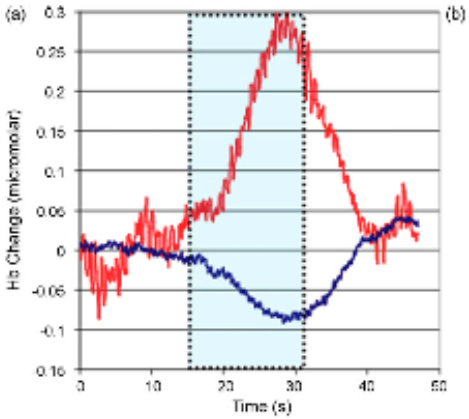
Quaresima, Bisconti, & Ferrari 2012 Brain & Lang

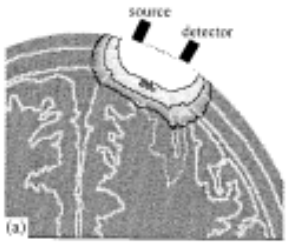


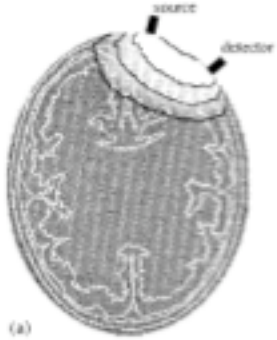


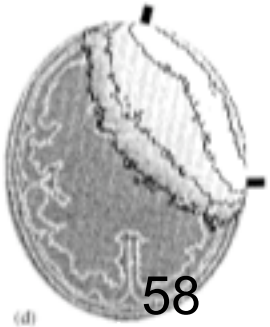


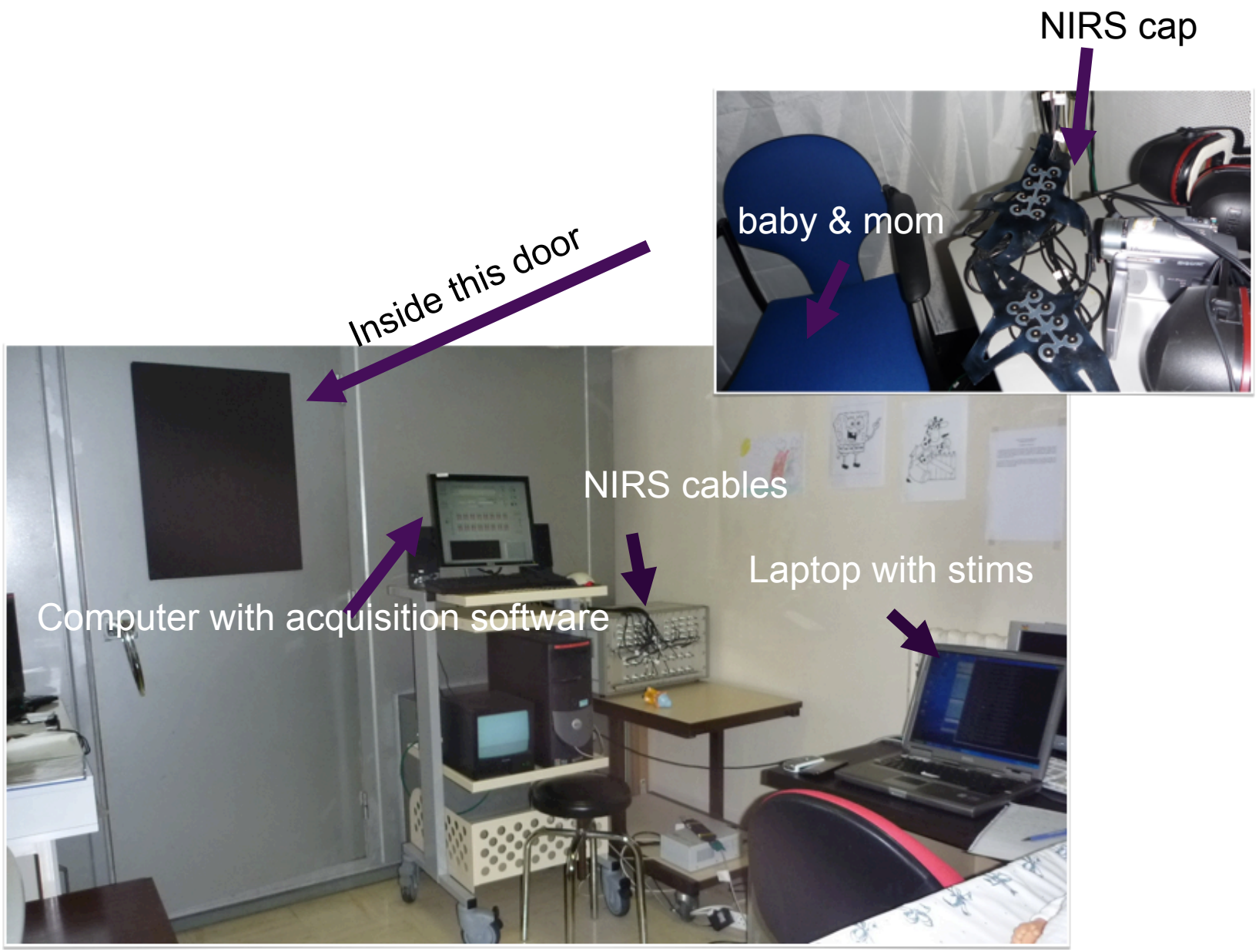






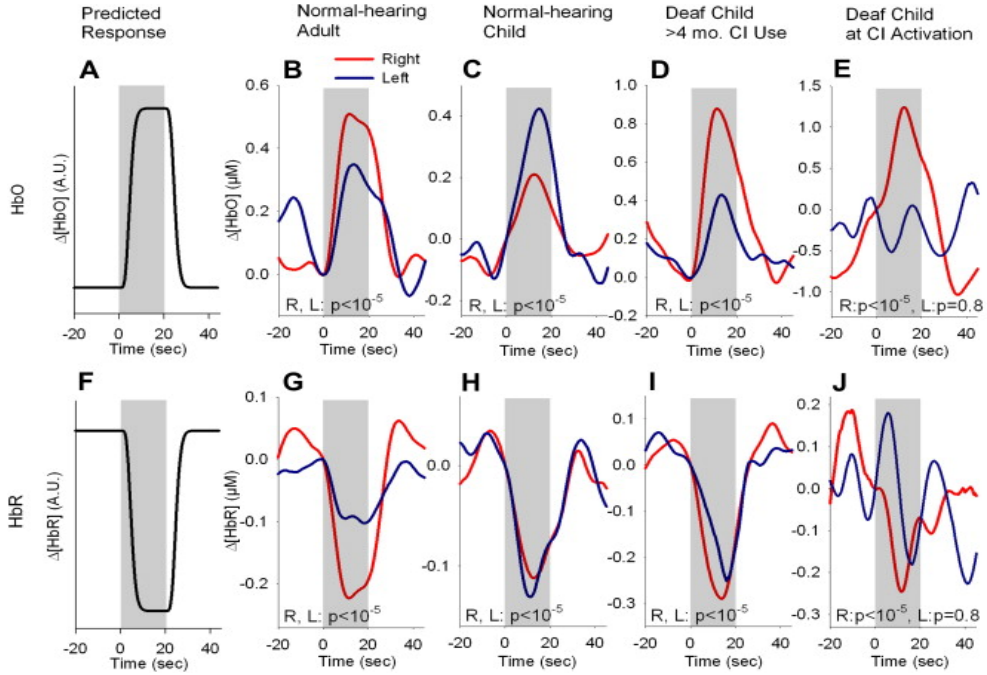


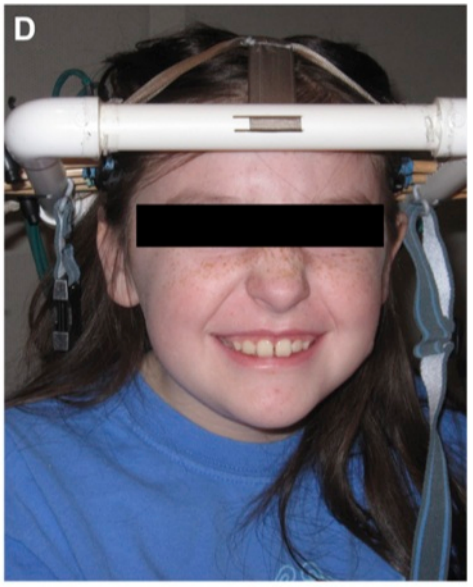



NIRS cap
baby & mom
Inside this door
NIRS cables
Laptop with stims
Computer with acquisition software

Children with a cochlear implant

	Normal-hearing adults (n=11/11)	Normal-hearing children (n=11/12)	Deaf children with > 4 mo. CI use (n=37/40)	Deaf children at CI activation (n=9/13)
Significant HbO Cortical Response	9 (82%)	8 (73%)	19 (51%)	7 (78%)
Significant HbR Cortical Response	11 (100%)	8 (73%)	26 (70%)	5 (56%)
Significant HbO or HbR Cortical Response	11 (100%)	9 (82%)	28 (76%)	7 (78%)
Age of Responders[1] **[Range]**	30.4 ± 8.3 [24 - 48]	9.4 ± 3.4 [4 - 15]	7.9 ± 3.8 [2 - 19]	4.7 ± 1.6 [2 - 8]





Sevy et al. 2010 Hear Res

Effect sizes, reliability?

- Effect sizes small
- P-hacking likely
  - Little use of correction for multiple comparisons
- Fortunately, ongoing efforts (Blase, Lloyd-Fox & colleagues) to study...
  - ... test-retest reliability of localization & strength of responses
  - ... convergence with fMRI in localization

Salient weaknesses & strategies

Imprecise localization

- fNIRS+fMRI: mapping surface to structure in infancy (Lloyd-Fox, in progress)

Analyses not standardized + lax significance criteria = increased risk of false positives

- Databases to integrate results (Cristia et al. 2012 *Plos One*), possibly repositories & common data format
- Community forming around an analysis program (HomER, Tuppert et al., 2009 *Applied Optics*)

Instability of results

- Study of experimental features leading to low signal-to-noise ratios (Minagawa et al. 2014; Cristia, Minagawa, & Dupoux, 2015)
- Methodological developments: better caps, subtraction of 'skin' noise (Takahashi, et al., 2011 *Neuroimage*)

Conclusions

- fNIRS is a useful technique
  - Respectable spatial resolution (not deep structures, requires careful planning)
  - Infants & children: portable & easier to fit than EEG cap
  - Adults: cheaper than fMRI/MEG
  - May allow testing in more ecological conditions

- It's also a relatively *young* technique
  - Instrumentation under development
  - Post-processing & analyses not yet crystalized

Students, to your sheets

Comments/questions on this section?

Melanie Soderstrom

DARCLE

Amanda Seidl

Laia Fibla & Charlotte Maniel

Emmanuel Dupoux

Yasuyo Minagawa

Thank you.

alecristia@gmail.com

- PROJET FINANCÉ PAR L'ANR - PROJECT FUNDED BY THE ANR