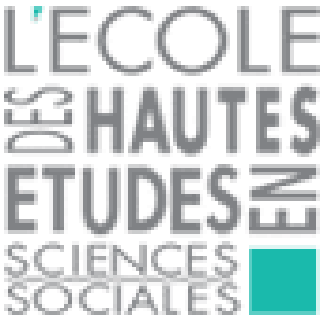
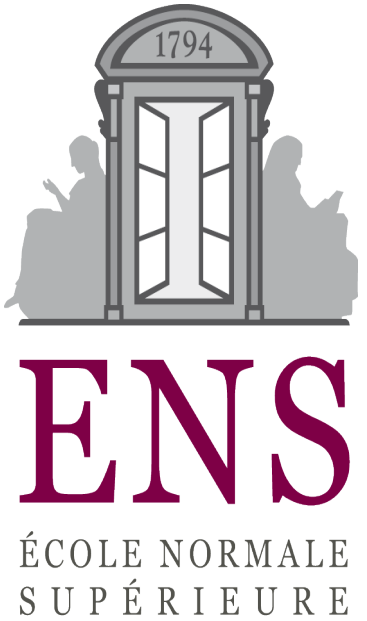
Speaker-specific behavior and acquisition

Alex Cristia
LSCP, CNRS






DANCE


THE DATING GAME

Roadmap

1. Perceiving diverse speakers

a) Linguistic form

b) Adaptation mechanisms

c) Social identity

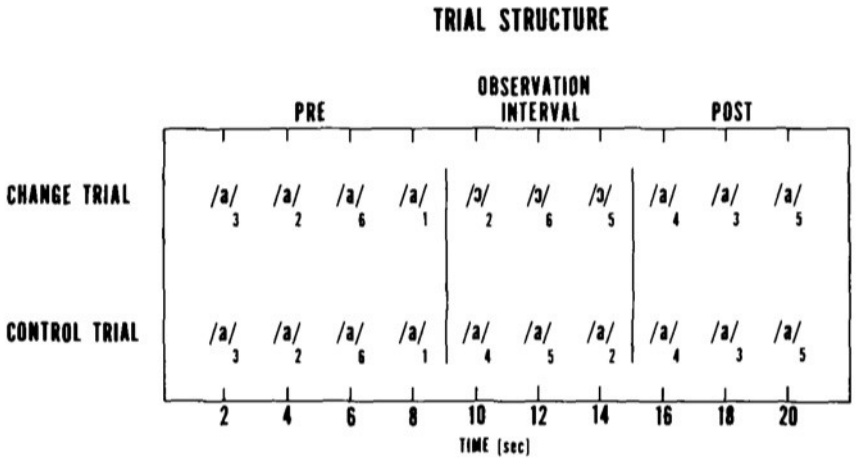
2. Producing diverse forms

- Multivariate
- Input
- Ease

3. Individual variation in perception & production

- Multiple sources (input, cognition)
- Some stability

Babies & indexical cues: The basics



The Stimulus Ensembles for the Background and Comparison Categories for All Stages of the Experiment. The Talker and Pitch-Contour Values for Each Stimulus are Given in Parentheses.

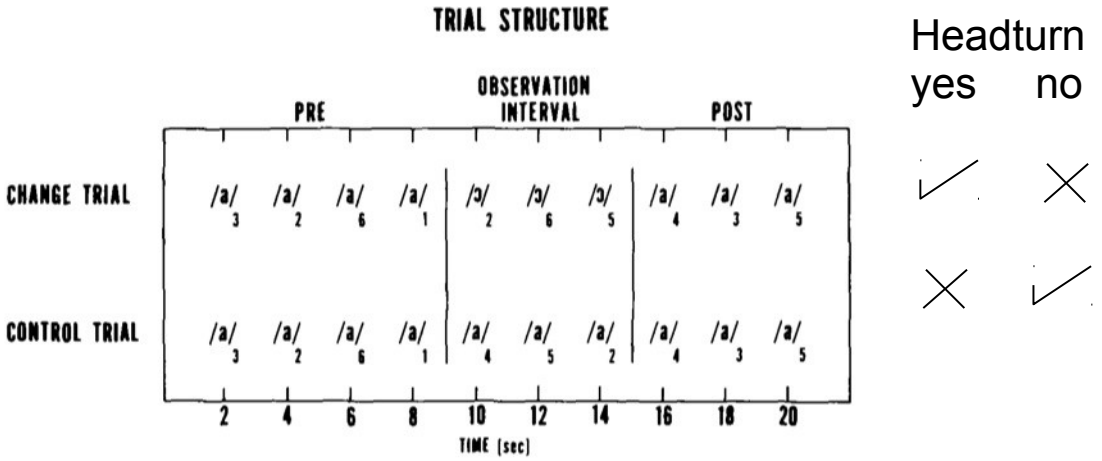
	Experimental stages	
	Background	*Comparison*
Conditioning	/a/ (Male, fall)	/ɔ/ (Male, fall)
Initial training	/a/ (Male, fall)	/ɔ/ (Male, fall)
Pitch variation	/a/ (Male, fall)	/ɔ/ (Male, fall)
	/a/ (Male, rise)	/ɔ/ (Male, rise)
Talker variation	/a/ (Male, fall)	/ɔ/ (Male, fall)
	/a/ (Female, fall)	/ɔ/ (Female, fall)
Talker × pitch variation	/a/ (Male, fall)	/ɔ/ (Male, fall)
	/a/ (Male, rise)	/ɔ/ (Male, rise)
	/a/ (Female, fall)	/ɔ/ (Female, fall)
	/a/ (Female, rise)	/ɔ/ (Female, rise)
Entire ensemble	/a/ (Male, fall)	/ɔ/ (Male, fall)
	/a/ (Male, rise)	/ɔ/ (Male, rise)
	/a/ (Female, fall)	/ɔ/ (Female, fall)
	/a/ (Female, rise)	/ɔ/ (Female, rise)
	/a/ (Child, fall)	/ɔ/ (Child, fall)
	/a/ (Child, rise)	/ɔ/ (Child, rise)

Kuhl, 1983 *IBAD*

Headturn
yes no

image from Hayes et al., 2000, *Cog Dev*

Babies & indexical cues: The basics



Headturn

yes no

The Stimulus Ensembles for the Background and Comparison Categories for All Stages of the Experiment. The Talker and Pitch-Contour Values for Each Stimulus are Given in Parentheses.

	Experimental stages	
	Background	*Comparison*
Conditioning	/a/ (Male, fall)	/ɔ/ (Male, fall)
Initial training	/a/ (Male, fall)	/ɔ/ (Male, fall)
Pitch variation	/a/ (Male, fall)	/ɔ/ (Male, fall)
	/a/ (Male, rise)	/ɔ/ (Male, rise)
Talker variation	/a/ (Male, fall)	/ɔ/ (Male, fall)
	/a/ (Female, fall)	/ɔ/ (Female, fall)
Talker × pitch variation	/a/ (Male, fall)	/ɔ/ (Male, fall)
	/a/ (Male, rise)	/ɔ/ (Male, rise)
	/a/ (Female, fall)	/ɔ/ (Female, fall)
	/a/ (Female, rise)	/ɔ/ (Female, rise)
Entire ensemble	/a/ (Male, fall)	/ɔ/ (Male, fall)
	/a/ (Male, rise)	/ɔ/ (Male, rise)
	/a/ (Female, fall)	/ɔ/ (Female, fall)
	/a/ (Female, rise)	/ɔ/ (Female, rise)
	/a/ (Child, fall)	/ɔ/ (Child, fall)
	/a/ (Child, rise)	/ɔ/ (Child, rise)

Kuhl, 1983 *IBAD*

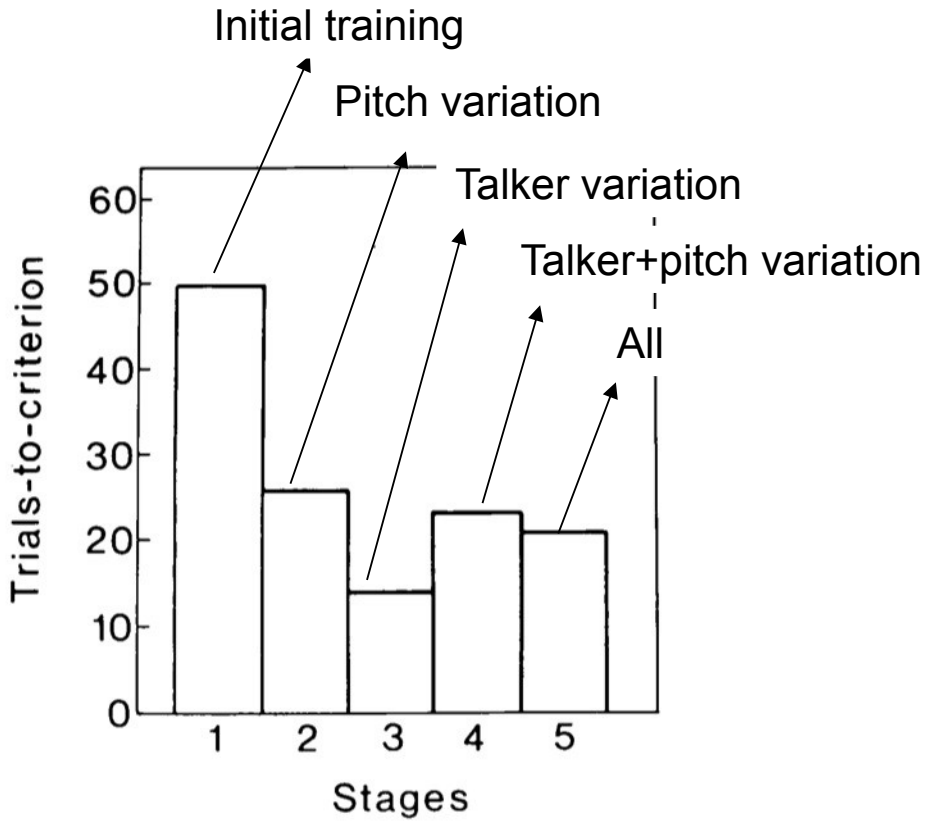
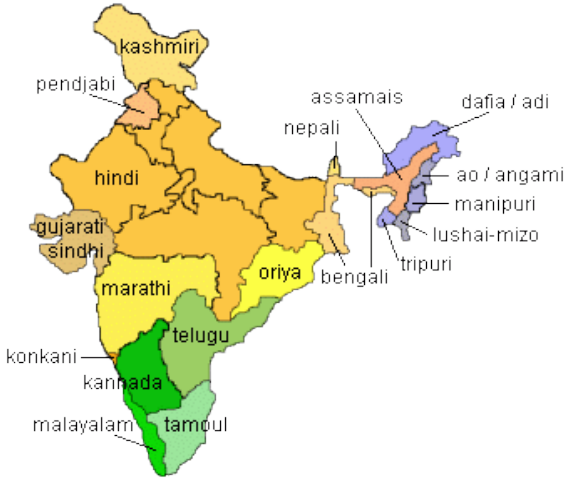
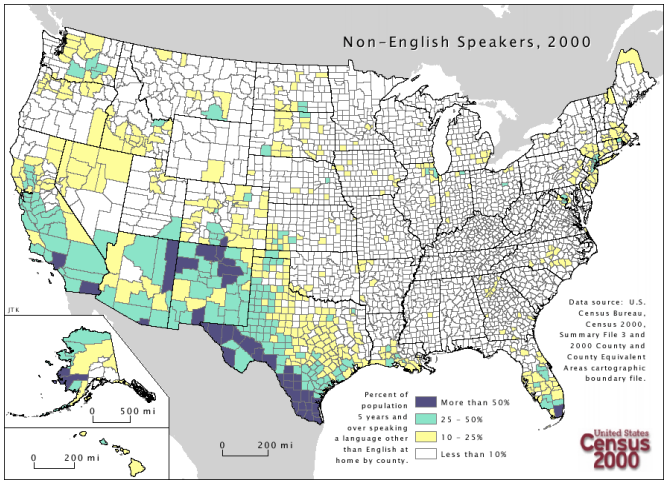


Figure 2. The mean number of trials to meet criterion (9 out of 10 consecutive trials correct) in the five stages of Experiment 1 (Initial Training, Pitch Variation, Talker Variation, Talker × Pitch Variation, and Entire Ensemble). Standard errors were 17.95, 6.16, 1.44, 6.62, and 6.58, respectively.

Criterion 9/10 correct



Signal



'Easy': Sex, age, mood...

Accent* is different

***Sociolect**:

boss≠bus

young working-class female

College students 20% correct;
63% errors due to sociolect

*accent = variation in the phonetic & phonological form driven by the talker's social identity
*Labov 2012 *Principles of language change*

End state: adults

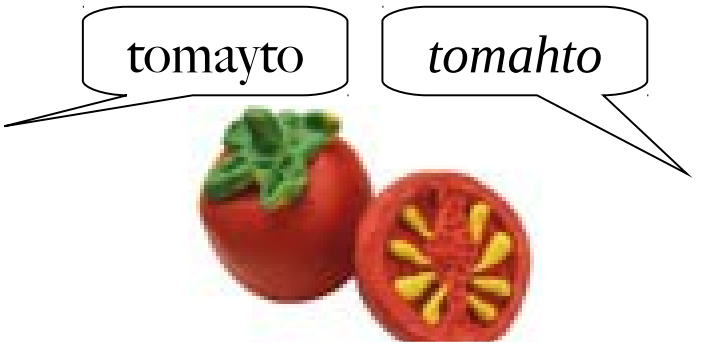
1 Retrieve linguistic form (with enough context & experience)

with a cost

2 Quickly adapt to novel accents

rule-based

3 Retrieve social identity (with enough context & experience)



1: Brunellière et al. 2011 *Brain Lang;* Floccia et al. 2006 *JEP:HPP* // **2:** Norris et al. 2003 *Cog Psy;* Maye et al. 2008 *Cog Sci* // **3:** Thomas 2007 *Lang Ling Compass*; van Berkum et al. 2008 *J Cogn Neurosci*

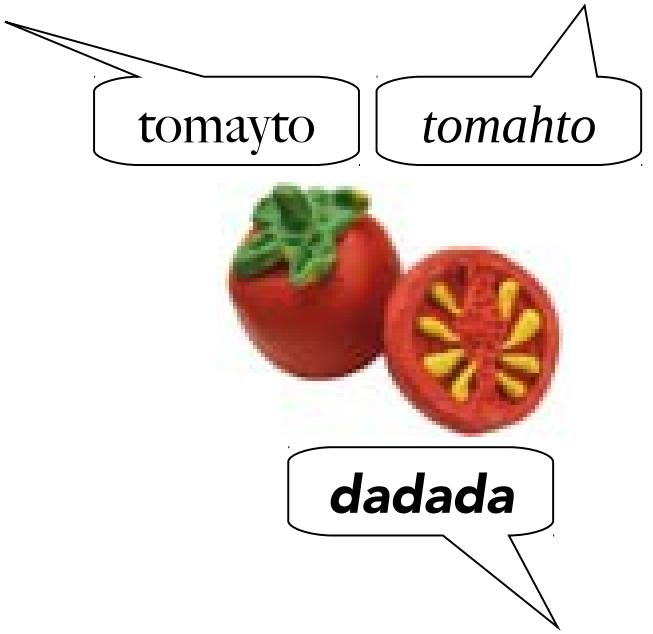
Young children

Limited lexicon & (world) knowledge

Slower speech processors *as is*

Spoken stream affected!

Q: Are 1-3 below evident in infants?



1. Retrieve linguistic form (with enough context & experience)

with a cost

2. Quickly adapt to novel accents

rule-based

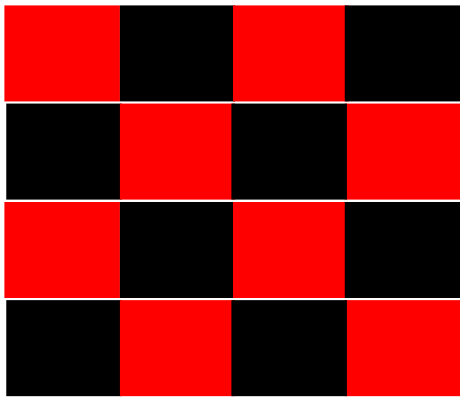
3. Retrieve social identity (with enough context & experience)

Lexicon → 'phonemic constancy'

1 Can retrieve linguistic form in unfamiliar accents

15 months: n/s

19 months: *



baby bottle nappy ...

lady totter savvy ...

Best et al. 2010 *Psyc Sci*

NB: word list hypothesized from methods

Where is the birdy?

Where is the doggy?

Mulak et al. in press *Child Dev*

1’ *Cost in linguistic processing*

Salience

Training

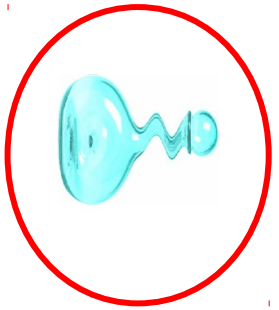
It’s a neech!

Test

Trained label

Do you see the neech?

Novel label



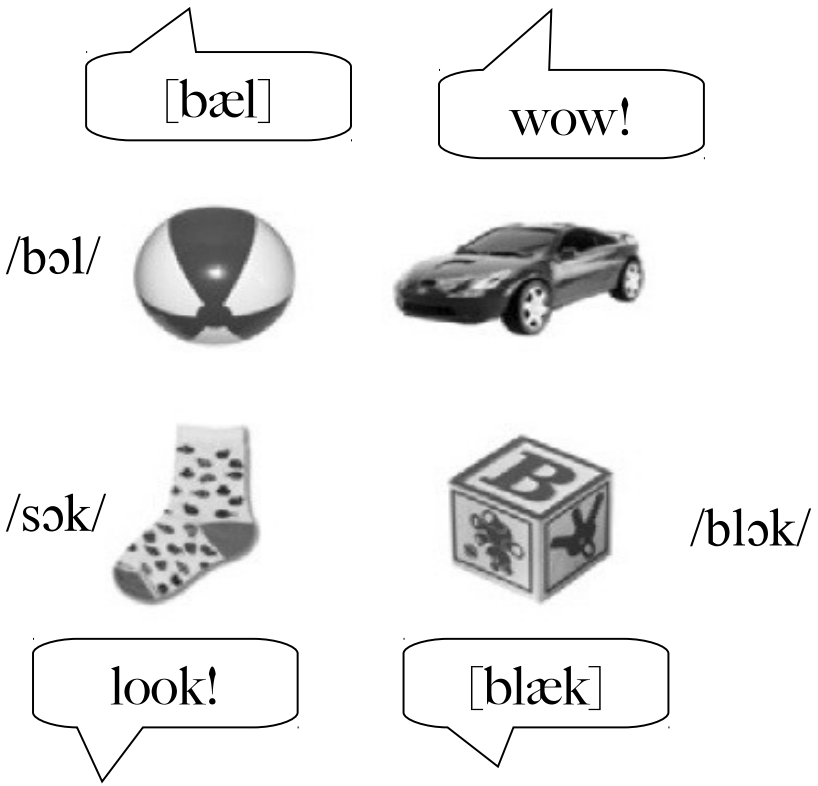
Do you see the moof?

Schmale, Hollich, & Seidl 2011 *J Child Lang*

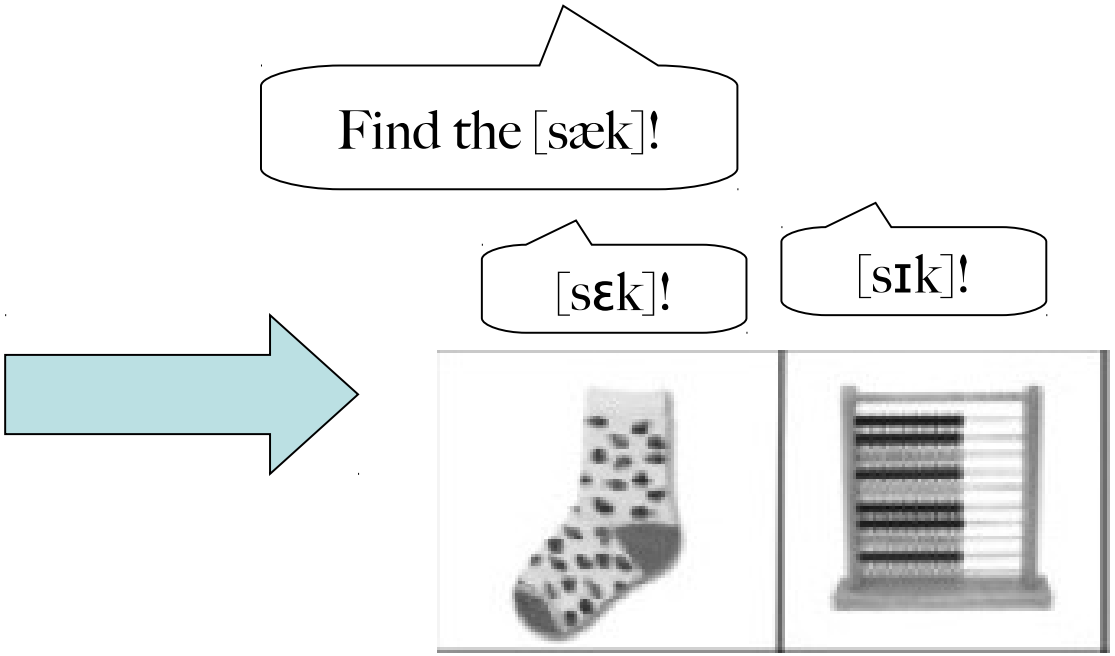
24 months *fail*

Lexicon → semantic bootstrapping

2 Can adapt to novel accents, rule-based generalization



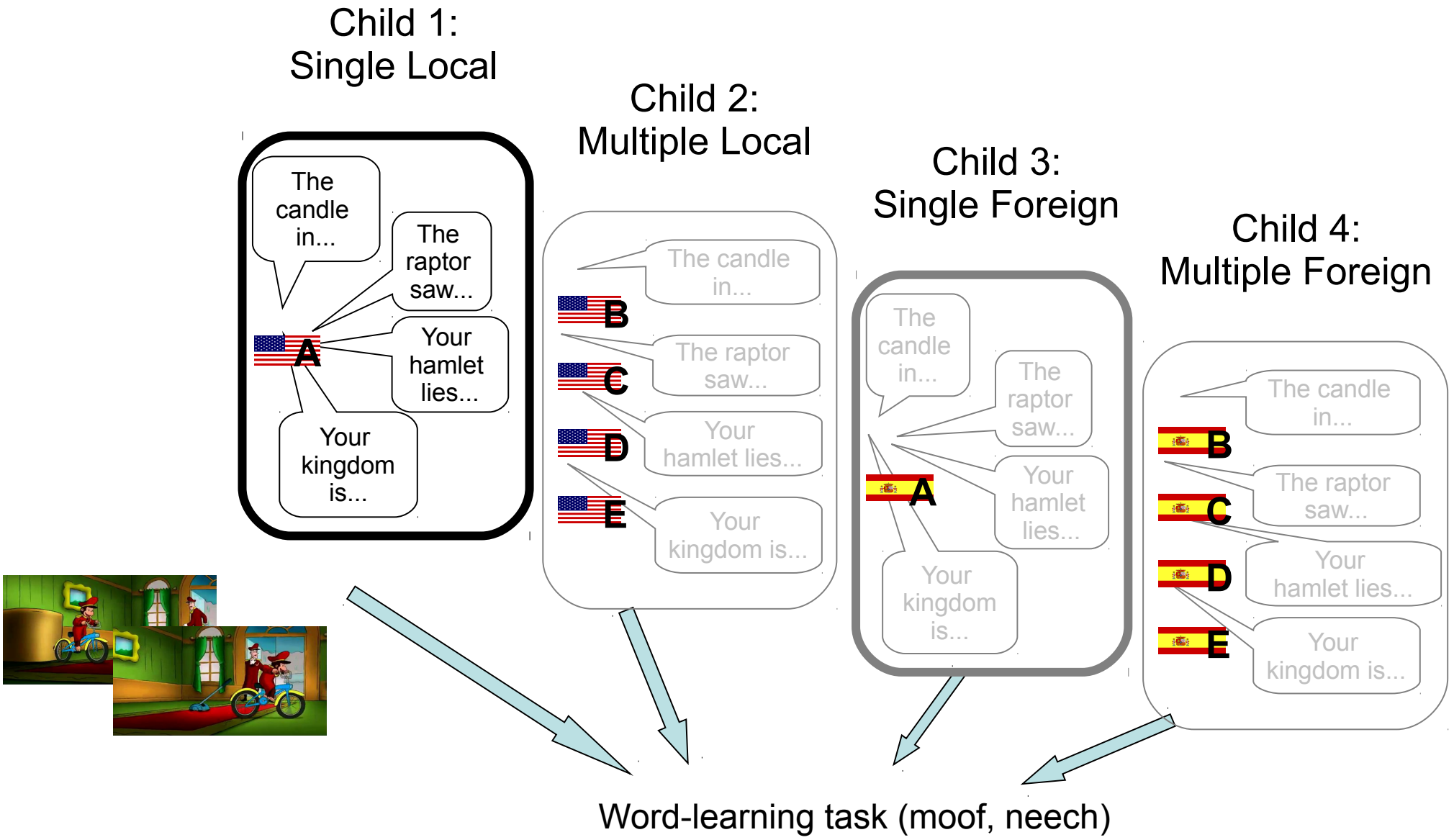
18 months



White & Aslin 2011 *Dev Sci*

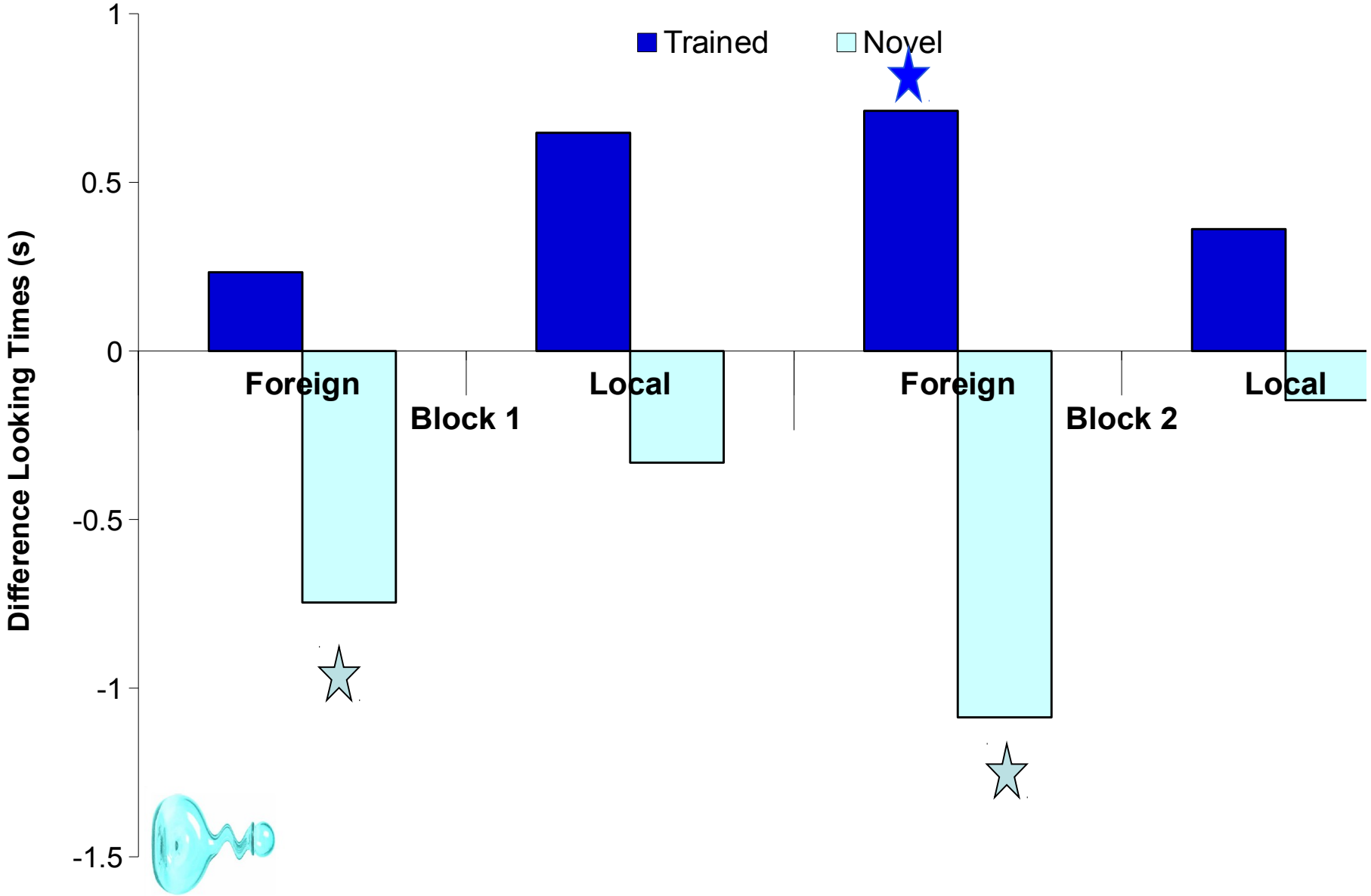
McQueen et al., 2012 *LLD* ; van Linden & Vroomen, 2008, *JCL*

Bootstrapping problem

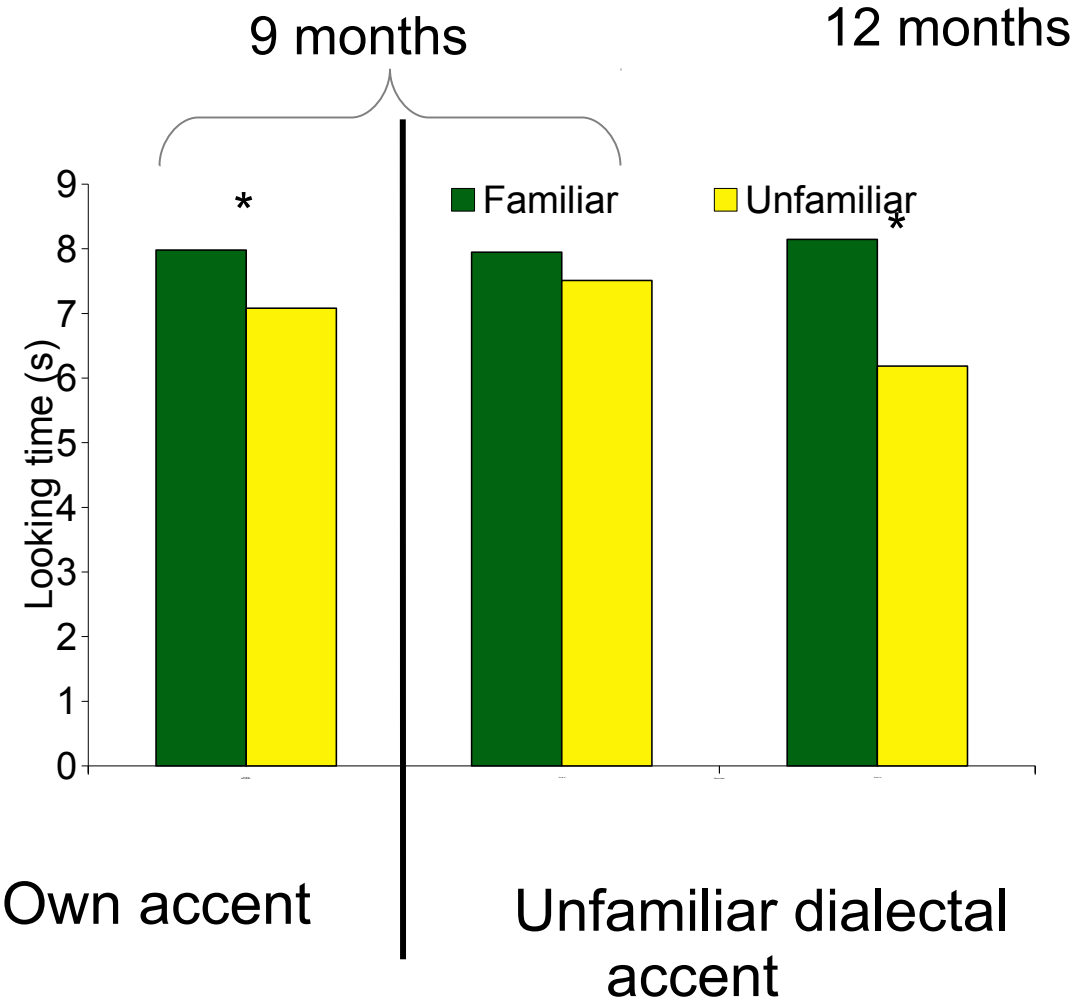


Schmale, Cristia, & Seidl, 2013, *Dev Sci*

A new mechanism: Without semantic bootstrapping



A new mechanism: Without semantic bootstrapping

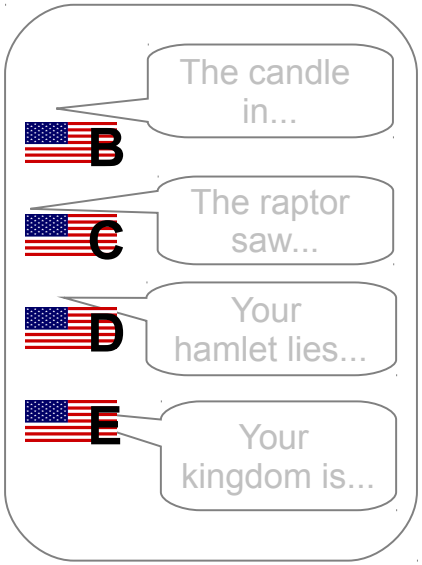


Own accent: Schmale & Seidl 2009 *Child Dev*

Unfamiliar: Schmale, Cristià, Seidl, & Johnson 2010 *Infancy*

Even without linguistic evidence!

Diverse speakers



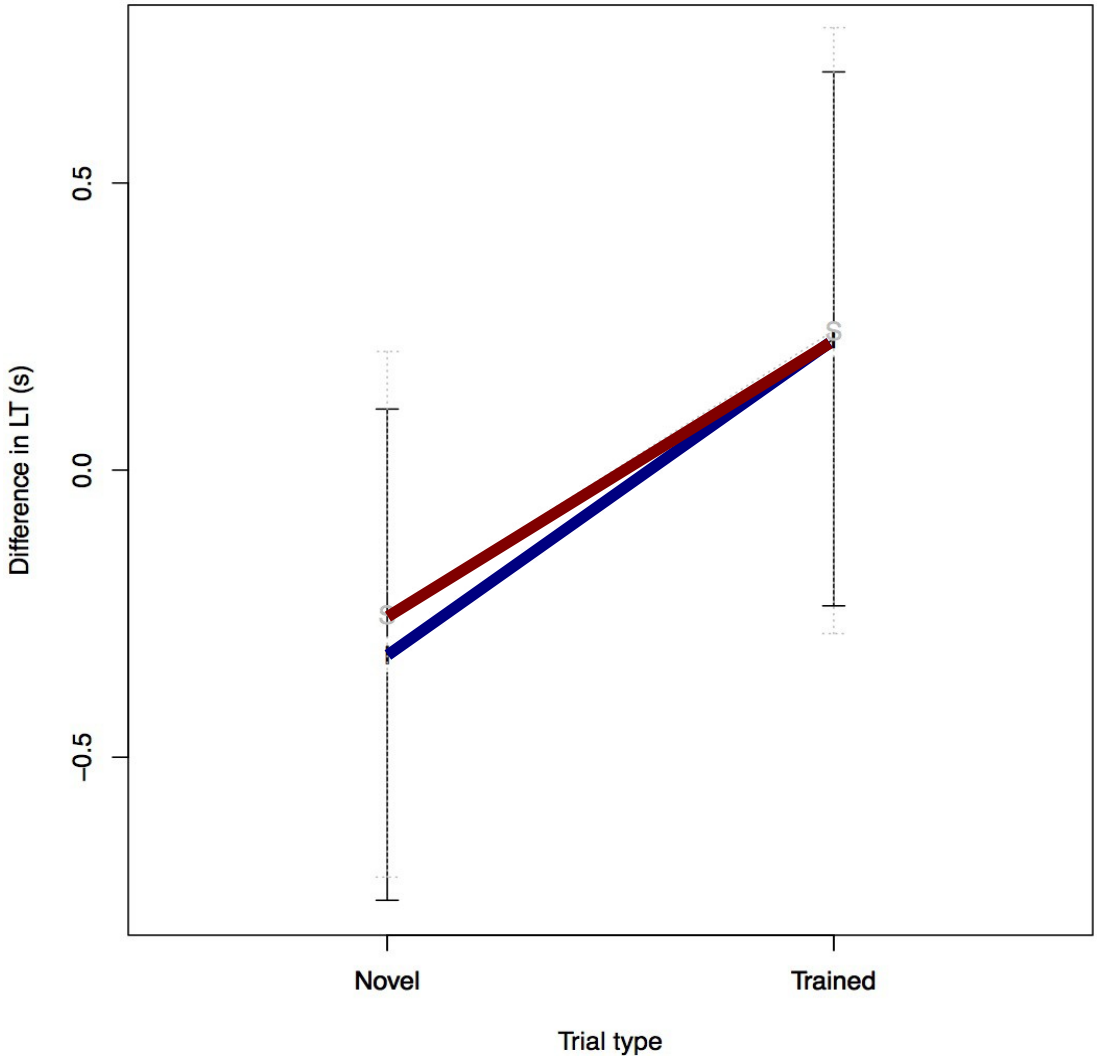
Diverse faces

B

C

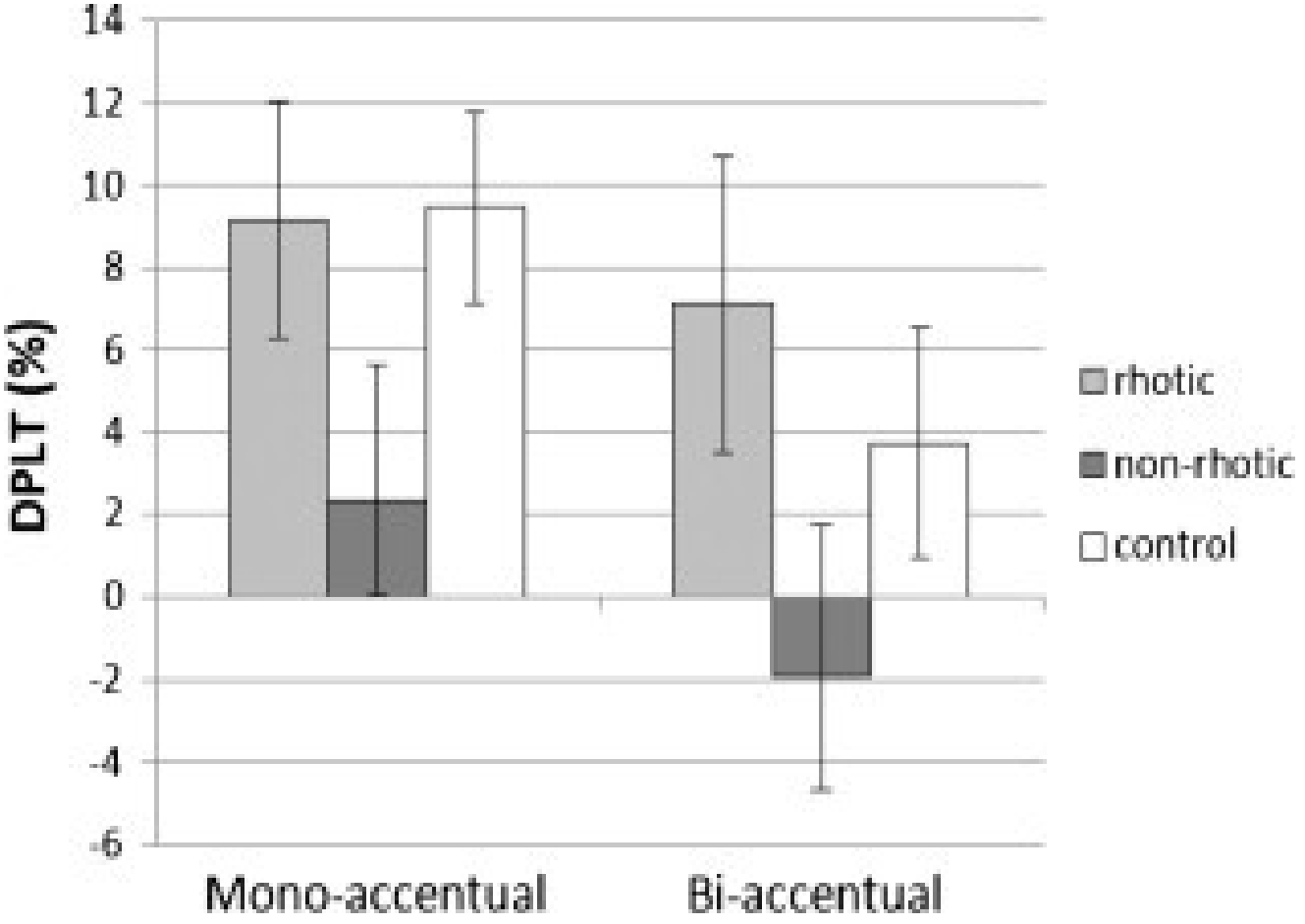
D

E



Schmale, Cristia, & Seidl, submittied

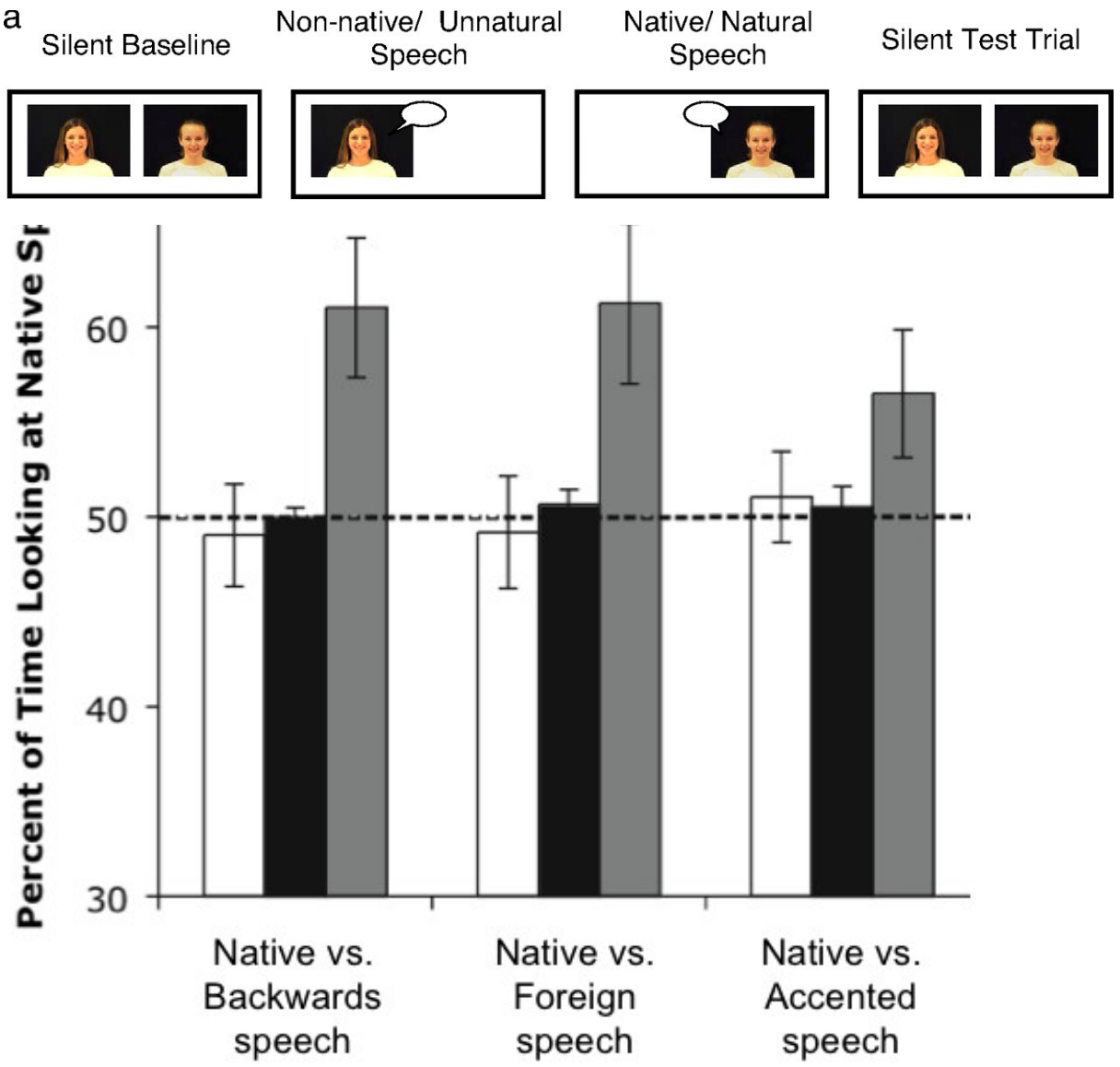
The role of exposure again



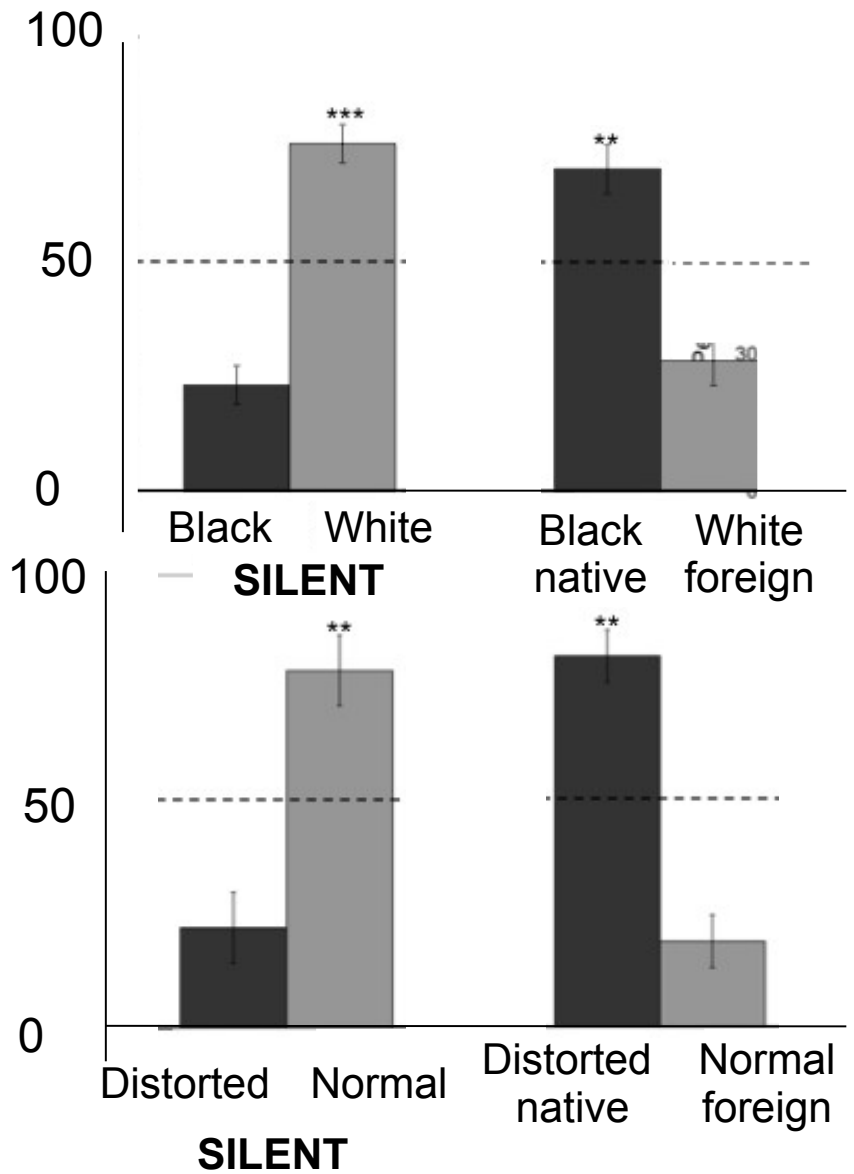
Floccia et al., 2012, *Cognition*

③ Infants' social judgments

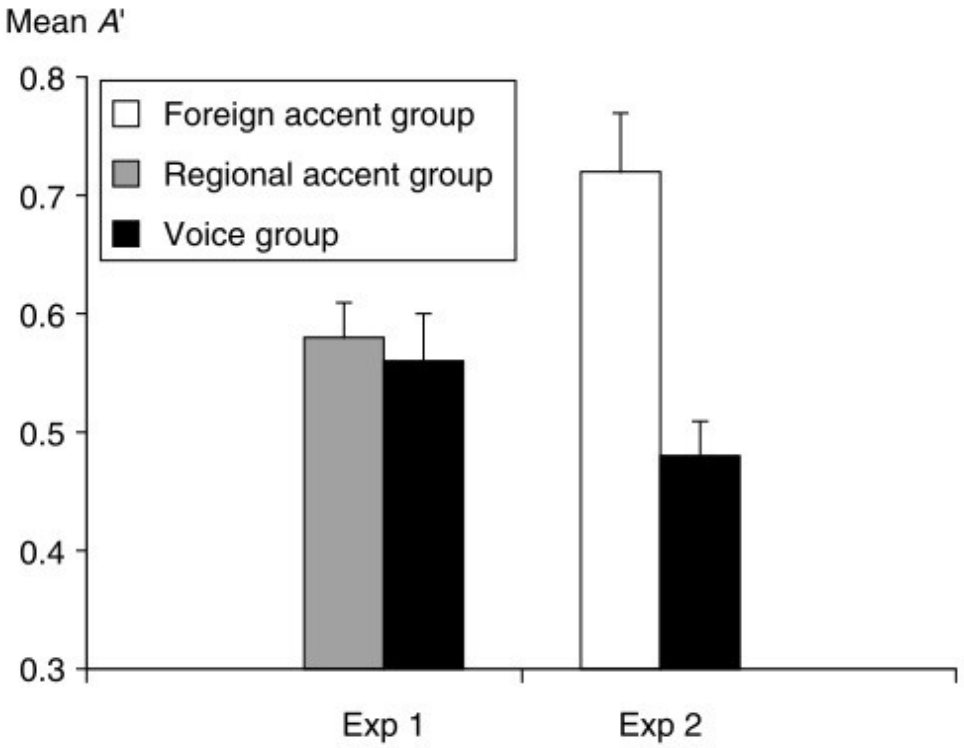
Kinzler et al., 2007, *PNAS*



Kinzler et al., 2009, *Social Cog*

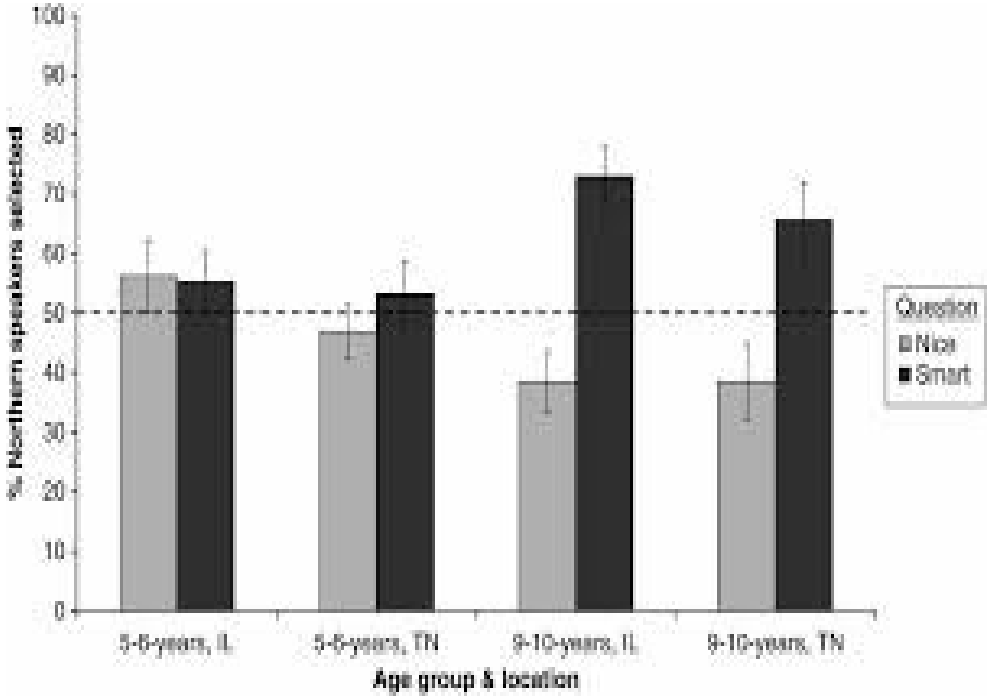


Categorization & attributes



Girard et al., 2008, *British J Dev Psy*

Floccia et al., 2009, *Intnl J Beh Dev*



Kinzler & DeJesus, 2012, *Quarterly J Exptl Psy*

Wagner et al., 2013, *JCL...*

Perceiving diverse speakers

1 Retrieve linguistic form (with enough context & experience)

with a cost

Success depends on age, precision required in the task, and task difficulty

2 Quickly adapt to novel accents

rule-based

Use of shortcuts
Variety of triggers
Expectations

3 Retrieve social identity (with enough context & experience)

Meta-linguistic knowledge & expectations evident in 'easy'/natural tasks

Roadmap

1. Perceiving diverse speakers

a) Linguistic form

b) Adaptation mechanisms

c) Social identity

2. Producing diverse forms

- Multivariate
- Input
- Ease

3. Individual variation in perception & production

- Multiple sources (input, cognition)
- Some stability

Early approaches: Labov's 4 stages

Basic grammar (<5y)

The vernacular (5-12y) → Standard variants

Replacement? Addition?

Social perception (early teens)

Stylistic variation (late teens)

Labov 1964

Use of standard variants

Main effect: social background

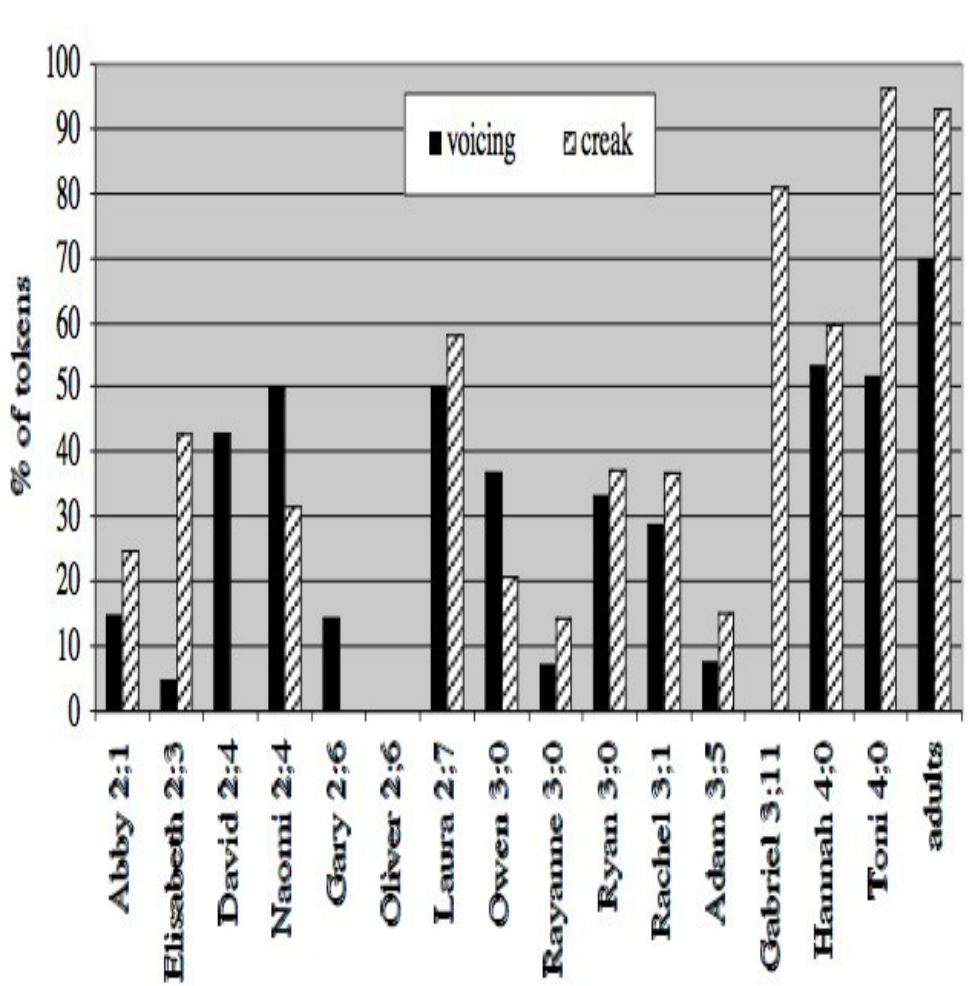
- Age interaction: differences between low and high SES increase with age (Chevrot et al. 2011 *Lang Sci*)

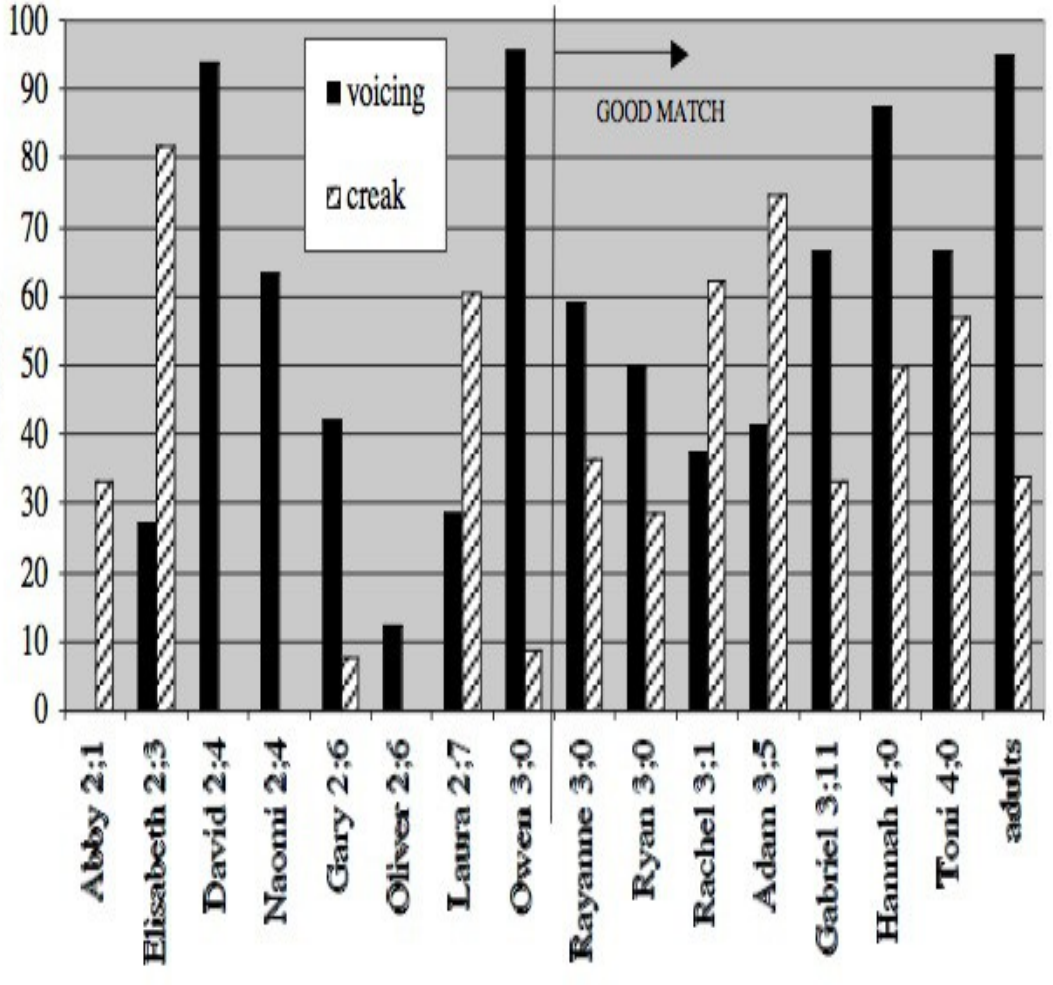
Context/register effects even at ~3y (Smith et al. 2011)

Age effects are variable (Smith et al. 2007 *Lang Var Change*; vs. Romaine, 1984)

Acquisition of variable allophones

Foulkes et al. 2001 *Working Paper*





Producing diverse forms

Multivariate approach inescapable

Role of the input

- Itself variable

Difficulty of controlling gestures

- Some features apparent earlier than others

Mechanisms in place?

Roadmap

1. Perceiving diverse speakers

a) Linguistic form

b) Adaptation mechanisms

c) Social identity

2. Producing diverse forms

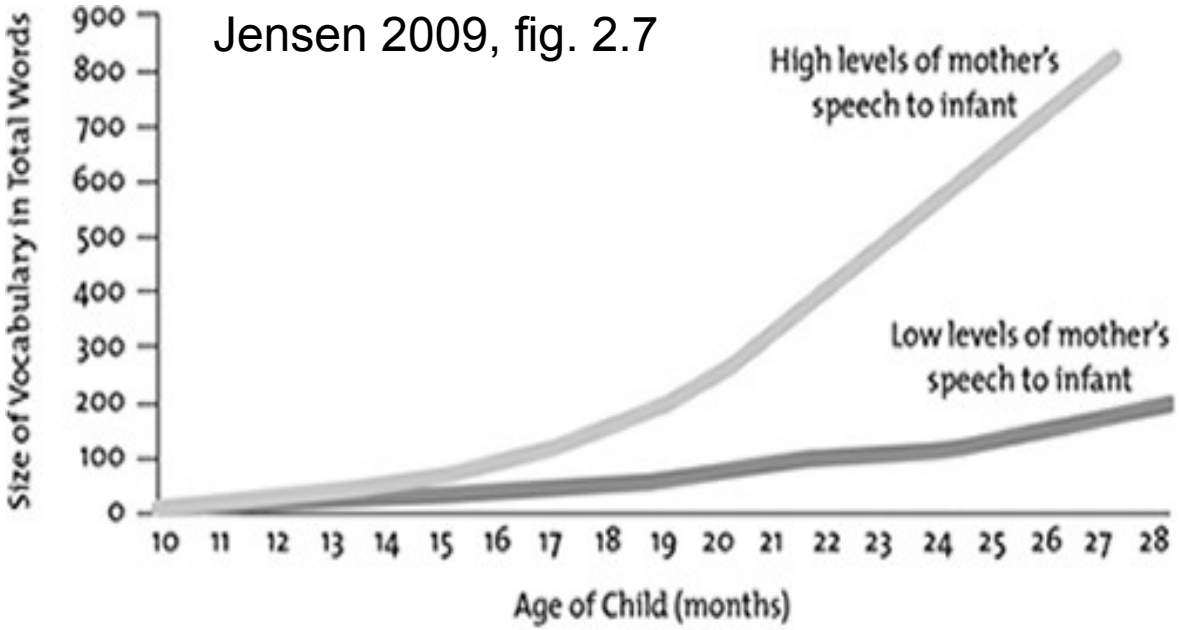
- Multivariate
- Input
- Ease

3. Individual variation in perception & production

- Multiple sources (input, cognition)
- Some stability

Quantity & quality

Jensen 2009, fig. 2.7

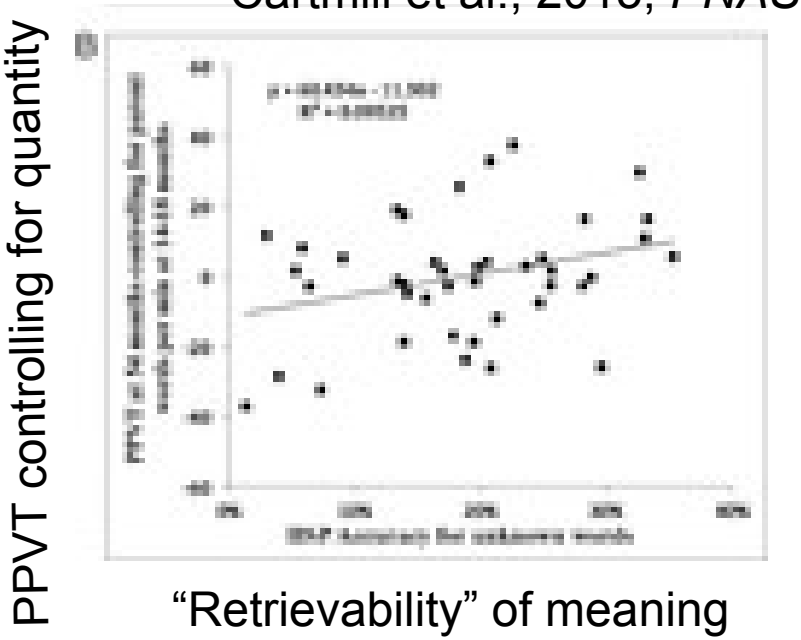


See also:

http://www.danielwillingham.com/1/post/2013/09/more-on-the-vocabulary-development-of-toddlers.html

http://www.cognitionandculture.net/home/blog/75-alexs-blog/2363-whats-the-point-of-talking-to-your-child

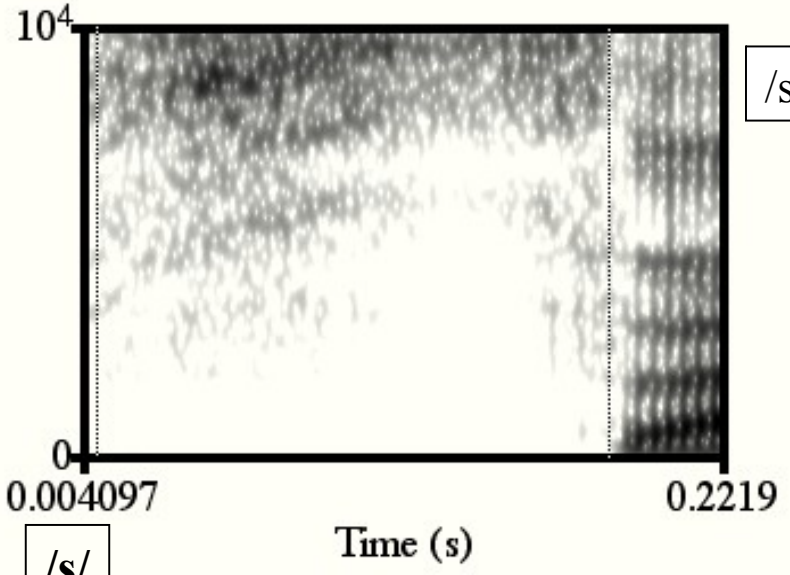
Cartmill et al., 2013, *PNAS*



"Retrievability" of meaning

Cristià 2011 *JASA*

Phonetic detail



/s/ pole ~8kHz

/s/

Mom speech

Baby /s-ʃ/ discrimination

42 moms+ babies

- 4-6m: 18
- 12-14m: 24

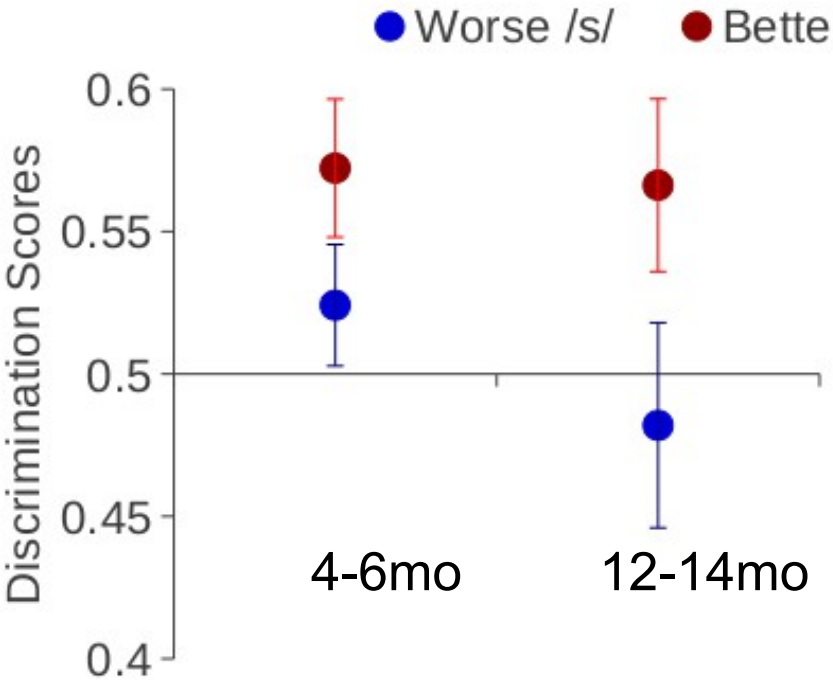
/ʃɔ .. ʃa/

…

/ʃɔ .. ʃa/

/sɔ .. sa/

/ʃɔ .. ʃa/



Infant speech & toddler language

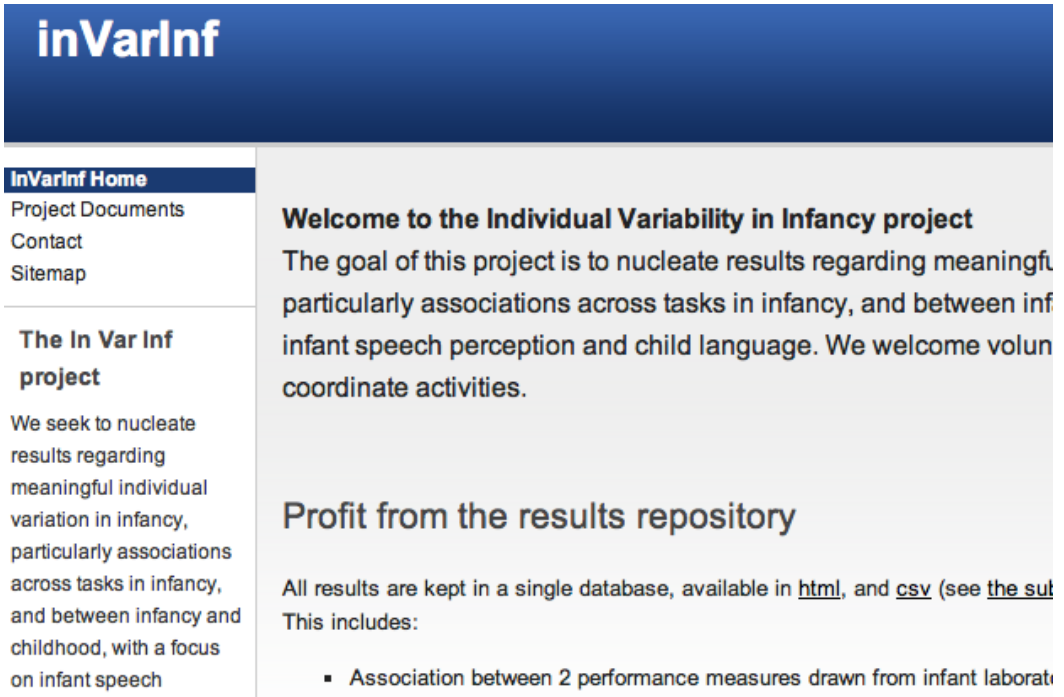
Searches in scholar.google.com, Pubmed, Science Direct, and Proquest

Identified 20 articles and theses

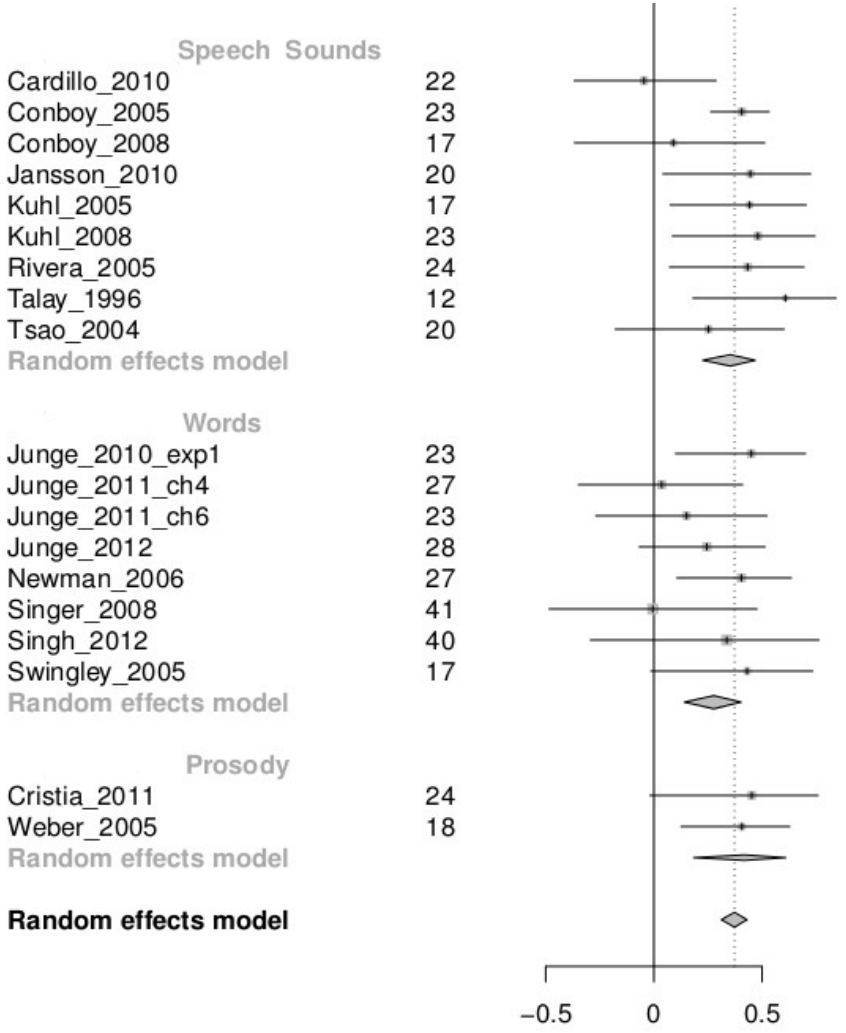
- 43 effect sizes [r] (+ 8 without comparable ES)
- Median r within infant group

Grouped into:

- phones,
- words,
- prosody



Speech perception measures predict language



r=.31 [.22, .4]

Speech perception at 4-12 months
explains 5-15% variance
in vocabulary size at 11-48 months

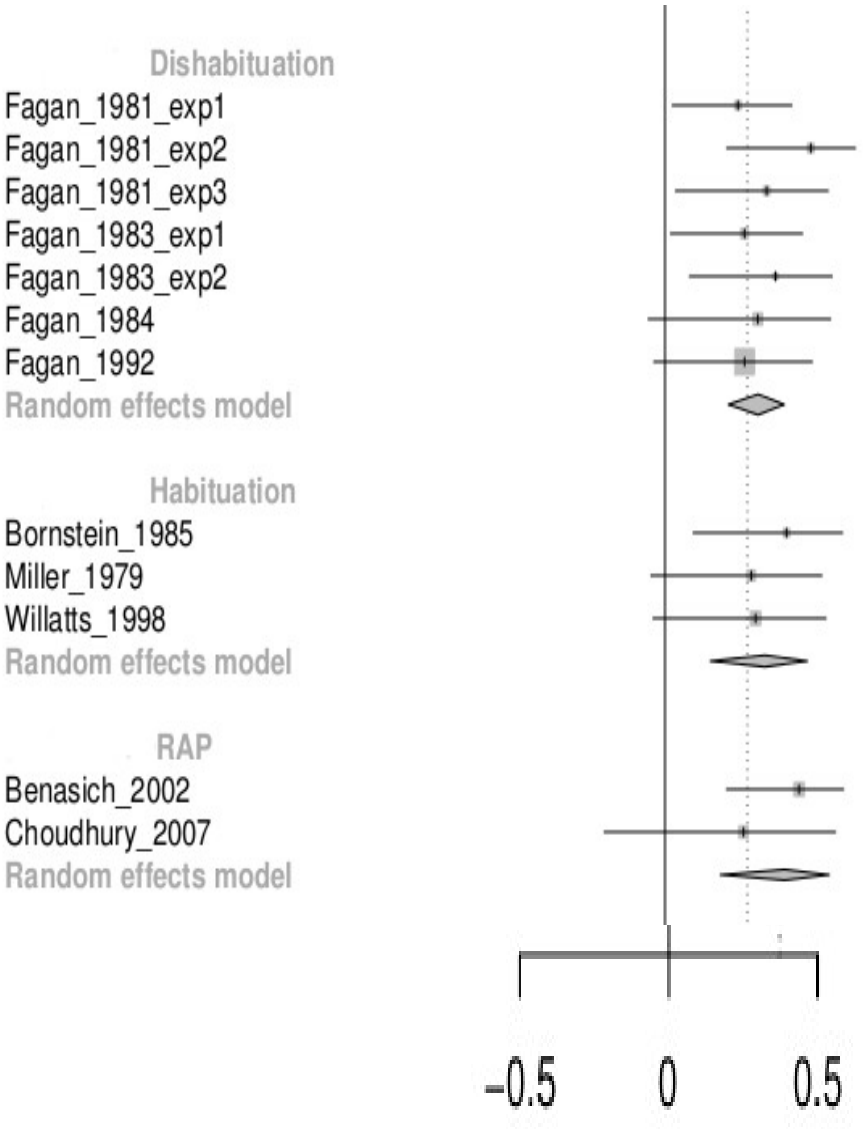
Speech sounds r=.35 [.22, .47]
Words r=.28 [.14,.4]
Prosody r=.42 [.18, .61]

Cristia, Seidl, Junge, Soderstrom, & Hagoort, in press, *Dev Sci*

Database available at sites.google.com/site/invarinf (Individual Variation in Infancy)

Non-linguistic measures do too

Information/auditory processing measures gathered at 4-12 months explain >15% variance in vocabulary size at 11-48 months!!



Two speech sound tasks r=.43 (.22--.6)

Two word processing tasks r=.29 (.02--.52)

Speech-non-linguistic task r=.25 (.08--.4)

Multivariate studies

N=45	5-6.5 months	6.5-8.4 months
Speech	Prosody (Preference trochees>iambs– cf. Herold et al. 2008)	Sounds (Discrimination 'ship'-'sheep' – cf Kuhl et al., 2008)
Non-linguistic	Dishabituation (Visual Recognition Memory – cf. Rose, Feldman, & Jankowski, 2009)	Cognitive control (A-not-B – cf. Lalonde & Werker, 1995; Conboy et al., 2008)

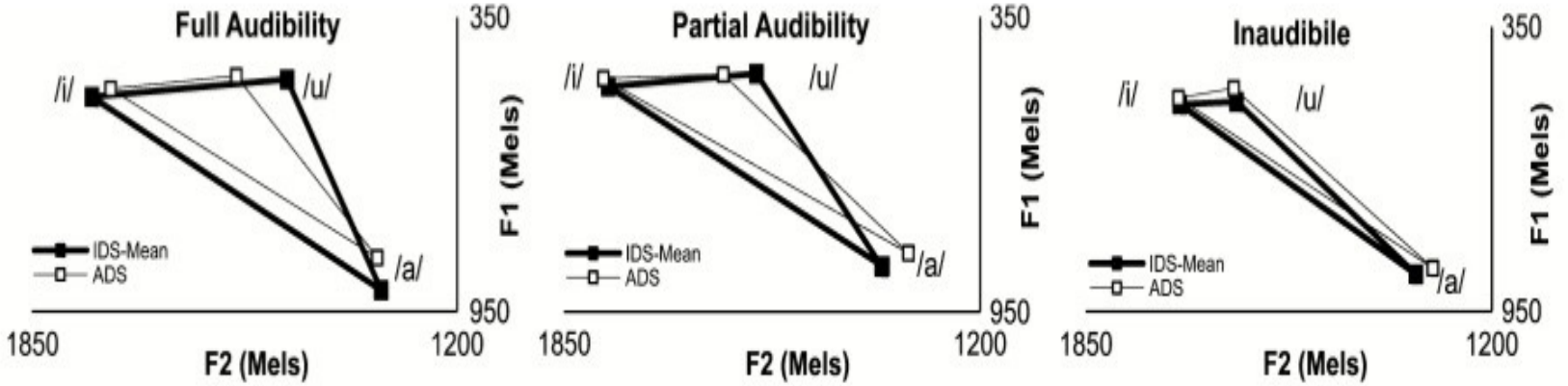
	Trochees	Vowels	VRM
Vowels	0.30*		
VRM	0.05	-0.07	
A/B	0.1	-0.2	0.09

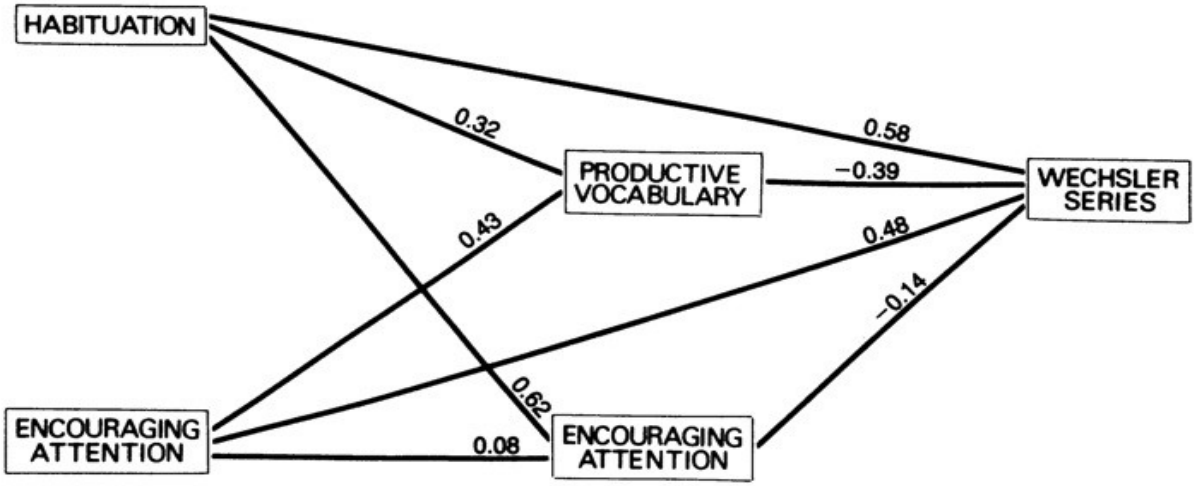
N=26 CDI at 24 months
Trochees-Vocab size r=.56
All other n/s

Seidl, Cristia, Wang, & French (in progress)

Multivariate studies (cont.)

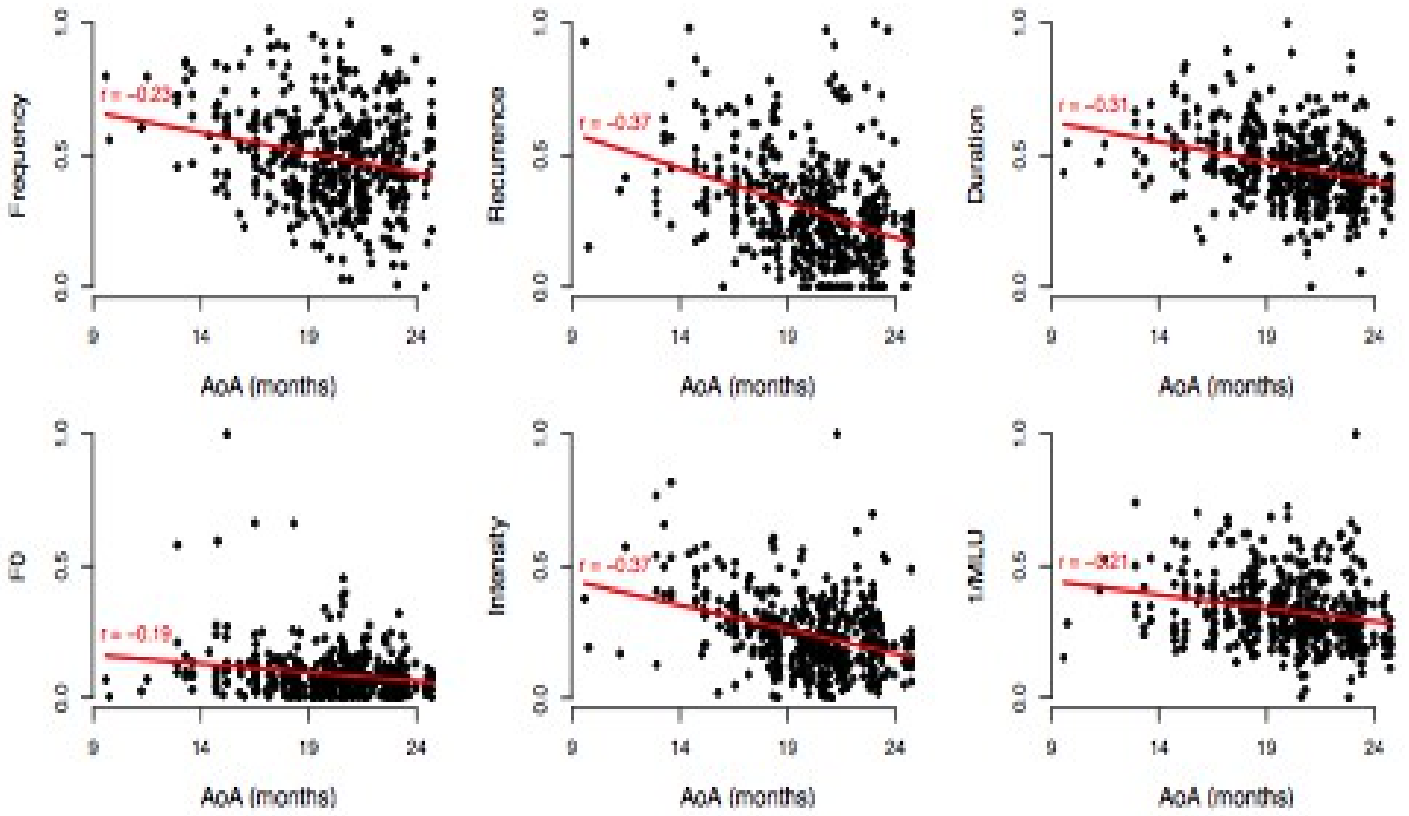
Lam & Kitamura 2012 *Dev Sci*





Bornstein 1985 *PNAS*

New methods: Speechome project



Vosoughi et al. 2010 *Proc Cog Sci*

http://www.youtube.com/watch?v=x6hxpzGHObc
http://www.media.mit.edu/cogmac/videos/feb16_20s_medium.mov

New methods: LENA

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

http://www.lenafoundation.org/Research/TechnicalReports.aspx

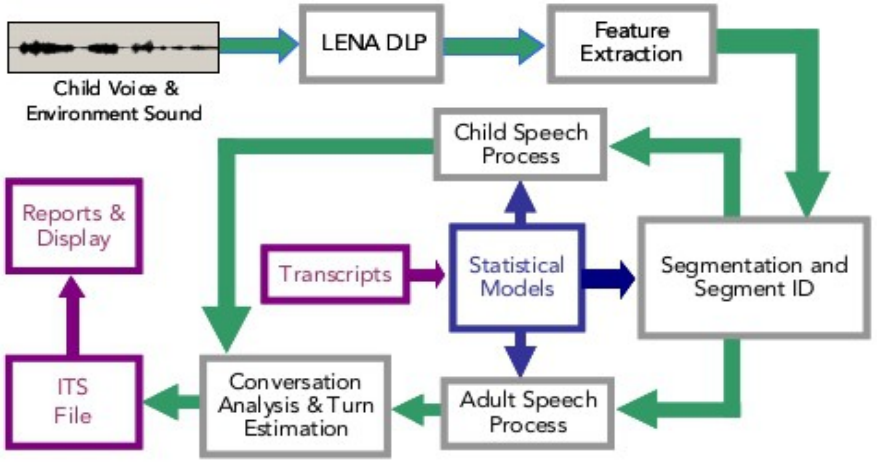
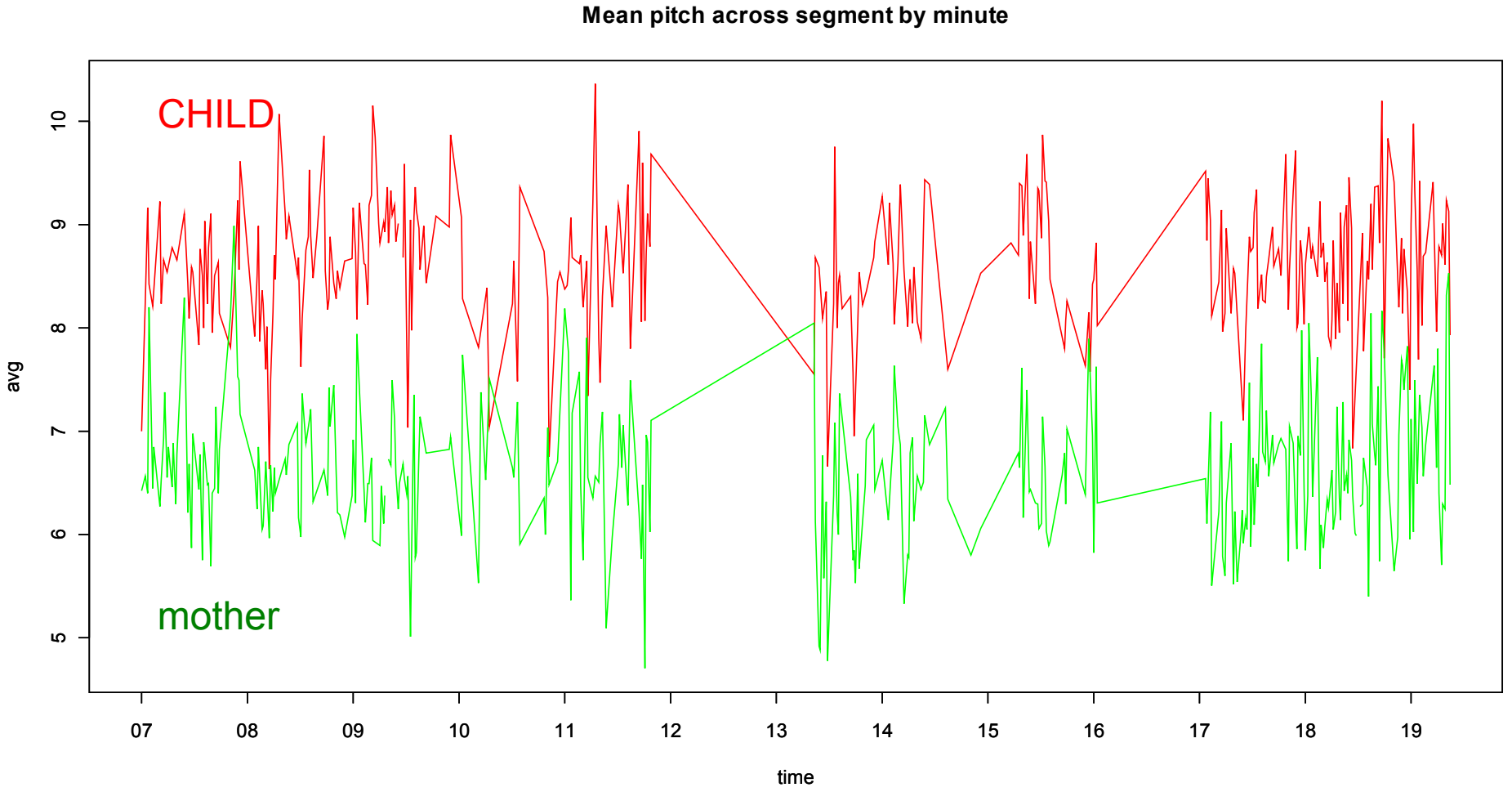


Figure 1. LENA Language Environmental Analysis Audio Processing System

Quantity
Rough quality (acoustic diversity)

Interaction parameters

Can be coupled with e.g. Praat



x axis: time of day; y axis average pitch of each vocalization

Ko, Reimchen, Cristia, Seidl, & Soderstrom, in preparation

Many thanks to Eon-Suk Ko for sharing this slide!

Individual variation in perception & production

- Stability
- Multivariate approaches necessary
- Complex causality

Open question:

Measures' sensitivity

- Cumulative dev sci
- Novel methods
- Quantity to increase precision


"Thank You"

The following slides were not discussed, but remain here for curious souls

on the meta-analysis

each study reports multiple correlations

		Infant speech perception measure			Vocabulary measure				
Author	**Year**	**Design**	**Age**		**N**	**Age**		**E**	**r**
				Speech sounds					
Talay	1996	Contrast	8 to 18	CHT: several native consonants	6 & 6	55 to 62	PPVT	+	0.66
Tsao	2004	Correlation	6	CHT: Trials to criterion /u-y/	20	13	U	-	-0.7
					16	16	U	-	-0.47
					13	24	P	-	-0.48
				CHT: Percent correct /u-y/	20	13	U	+	-0.05
					16	16	U	+	-0.17
					13	24	P	+	0.05
Conboy	2005	Correlation	11	CHT: d' non-native /t-d/	23	11	U	+	-0.37
				CHT: d' native /t-th/ minus d' non-native /d-t/	10	11	U	+	-0.37
Kuhl	2005	Correlation	7	CHT: d' native /ta-pa/	17	18	P	+	0.49
					16	24	P	+	0.49
				CHT: d' non-native /?i-t?i/	17	18	P	-	0.5
					16	24	P	-	0.22
Rivera	2005	Contrast	11	ERP: non-native /t-d/	13 & 11	18 to 30	P	+	-0.53
Kuhl	2008	Correlation	7.5	ERP: MMN native /ta-pa/	21	18	P	-	0.43
					23	24	P		-0.43
				ERP: MMN non-native /?i-t?i/ or /ta-da/	21	24	P	+	-0.61
Conboy	2008	Correlation	11	CHT: Number of conditioning trials	17	11	U	-	0.39
				CHT: d' native /ta-tha/	17	11	U	+	0.05
				CHT: d' non-native /ta-da/	17	11	U	-	0.43
Cardillo	2010	Contrast	7 to 11	CHT: d' /u-y/	9 & 8	60	PPVT	+	0.23
		Correlation	7	CHT: Trials to criterion /u-y/	22	18	P	-	0.26
					20	24	P		0.37
				CHT: Percent correct /u-y/	22	18	P	+	0.05
					20	24	P		0.18
Jansson	2010	Correlation	12	ERP: MMN non-native vowels	20	24	P	+	0.45

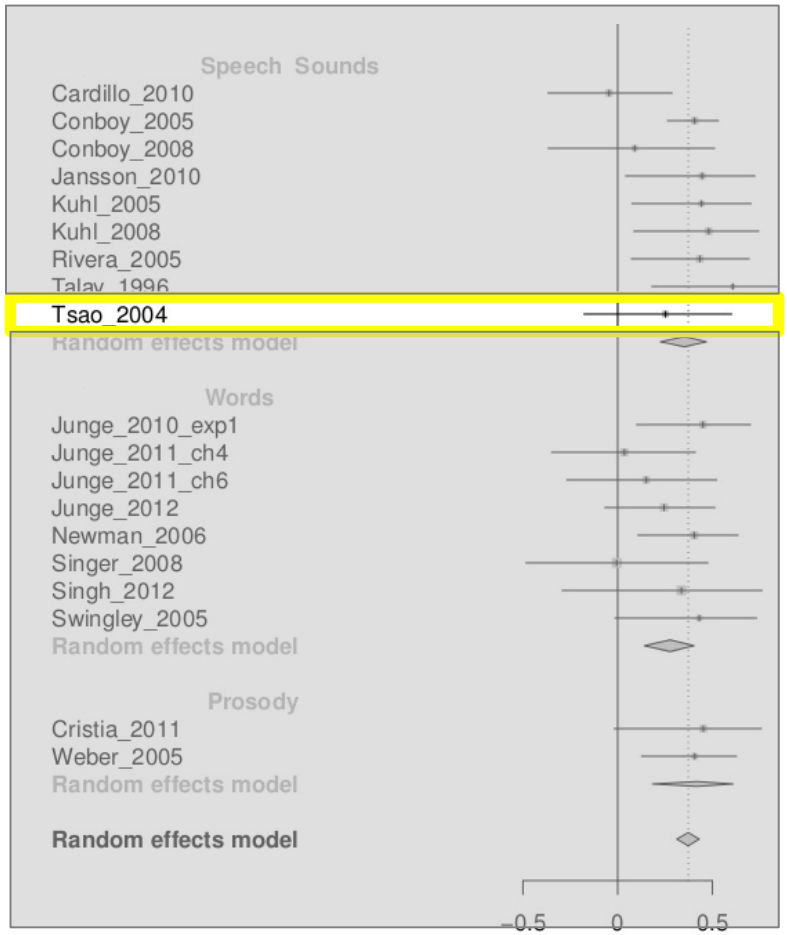
so I coded expected direction of relationship for each

		Infant speech perception measure			Vocabulary measure				
Author	**Year**	**Design**	**Age**		**N**	**Age**		**E**	**r**
				Speech sounds					
Talay	1996	Contrast	8 to 18	CHT: several native consonants	6 & 6	55 to 62	PPVT	+	0.66
Tsao	2004	Correlation	6	CHT: Trials to criterion /u-y/	20	13	U	-	-0.7
					16	16	U	-	-0.47
					13	24	P	-	-0.48
				CHT: Percent correct /u-y/	20	13	U	+	-0.05
					16	16	U	+	-0.17
					13	24	P	+	0.05
Conboy	2005	Correlation	11	CHT: d' non-native /t-d/	23	11	U	+	-0.37
				CHT: d' native /t-th/ minus d' non-native /d-t/	10	11	U	+	-0.37
Kuhl	2005	Correlation	7	CHT: d' native /ta-pa/	17	18	P	+	0.49
					16	24	P	+	0.49
				CHT: d' non-native /?i-t?i/	17	18	P	-	0.5
					16	24	P	-	0.22
Rivera	2005	Contrast	11	ERP: non-native /t-d/	13 & 11	18 to 30	P	+	-0.53
Kuhl	2008	Correlation	7.5	ERP: MMN native /ta-pa/	21	18	P	-	0.43
					23	24	P		-0.43
				ERP: MMN non-native /?i-t?i/ or /ta-da/	21	24	P	+	-0.61
Conboy	2008	Correlation	11	CHT: Number of conditioning trials	17	11	U	-	0.39
				CHT: d' native /ta-tha/	17	11	U	+	0.05
				CHT: d' non-native /ta-da/	17	11	U	-	0.43
Cardillo	2010	Contrast	7 to 11	CHT: d' /u-y/	9 & 8	60	PPVT	+	0.23
		Correlation	7	CHT: Trials to criterion /u-y/	22	18	P	-	0.26
					20	24	P		0.37
				CHT: Percent correct /u-y/	22	18	P	+	0.05
					20	24	P		0.18
Jansson	2010	Correlation	12	ERP: MMN non-native vowels	20	24	P	+	0.45

and calculated a weighted median within that study

		Infant speech perception measure			Vocabulary measure				
Author	**Year**	**Design**	**Age**		**N**	**Age**		**E**	**r**
				Speech sounds					
Talay	1996	Contrast	8 to 18	CHT: several native consonants	6 & 6	55 to 62	PPVT	+	0.66
Tsao	2004	Correlation	6	CHT: Trials to criterion /u-y/	20	13	U	-	-0.7
					16	16	U	-	-0.47
					13	24	P	-	-0.48
				CHT: Percent correct /u-y/	20	13	U	+	-0.05
					16	16	U	+	-0.17
					13	24	P	+	0.05
Conboy	2005	Correlation	11	CHT: d' non-native /t-d/	23	11	U	+	-0.37
				CHT: d' native /t-th/ minus d' non-native /d-t/	10	11	U	+	-0.37
Kuhl	2005	Correlation	7	CHT: d' native /ta-pa/	17	18	P	+	0.49
					16	24	P	+	0.49
				CHT: d' non-native /?i-t?i/	17	18	P	-	0.5
					16	24	P	-	0.22
Rivera	2005	Contrast	11	ERP: non-native /t-d/	13 & 11	18 to 30	P	+	-0.53
Kuhl	2008	Correlation	7.5	ERP: MMN native /ta-pa/	21	18	P	-	0.43
					23	24	P		-0.43
				ERP: MMN non-native /?i-t?i/ or /ta-da/	21	24	P	+	-0.61
Conboy	2008	Correlation	11	CHT: Number of conditioning trials	17	11	U	-	0.39
				CHT: d' native /ta-tha/	17	11	U	+	0.05
				CHT: d' non-native /ta-da/	17	11	U	-	0.43
Cardillo	2010	Contrast	7 to 11	CHT: d' /u-y/	9 & 8	60	PPVT	+	0.23
		Correlation	7	CHT: Trials to criterion /u-y/	22	18	P	-	0.26
					20	24	P		0.37
				CHT: Percent correct /u-y/	22	18	P	+	0.05
					20	24	P		0.18
Jansson	2010	Correlation	12	ERP: MMN non-native vowels	20	24	P	+	0.45

then one can calculate median effect size and weight per infant group, using those independent points

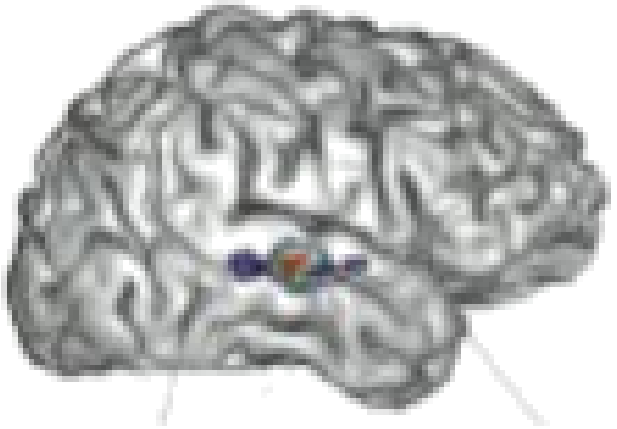


The neural bases of accent perception

The “interpretation” of accent

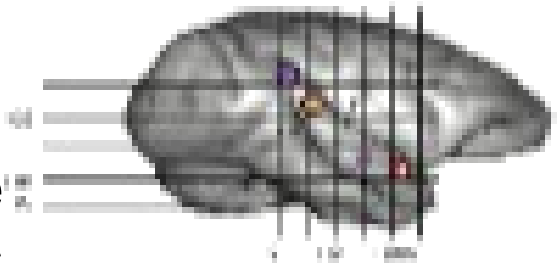
Voice perception

Right Hemisphere Humans



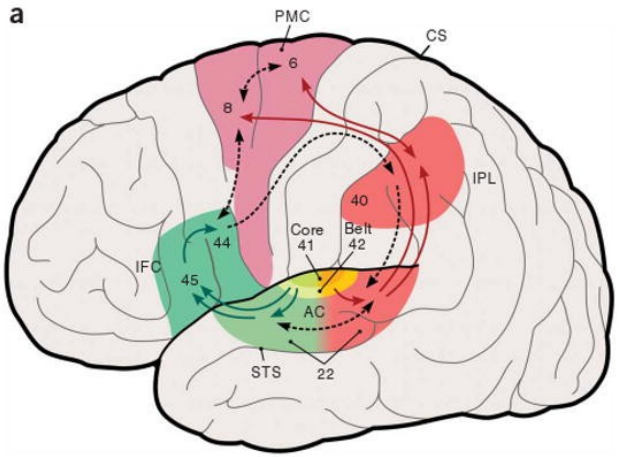
Belin & Grosbras 2010 Neuron

Right Hemisphere Macaques



Adult brain networks involved in **speech perception**

Left Hemisphere Humans



Rauschecker & Scott 2009
Nature Neuroscience

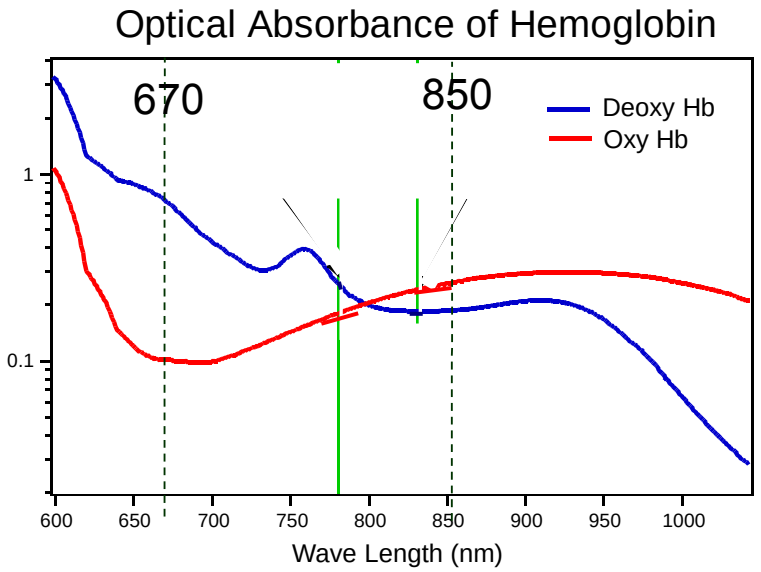
NIRS (Near Infra-red Spectroscopy)

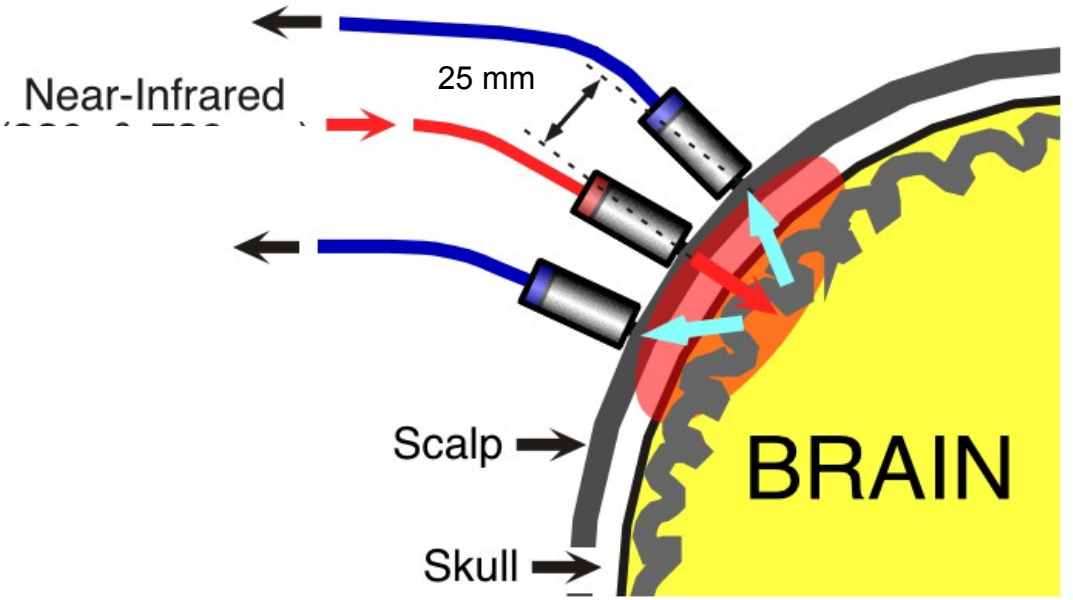
+ Non-invasive, silent, portable: perfect for babies & speech!

Local changes of hemoglobin concentration

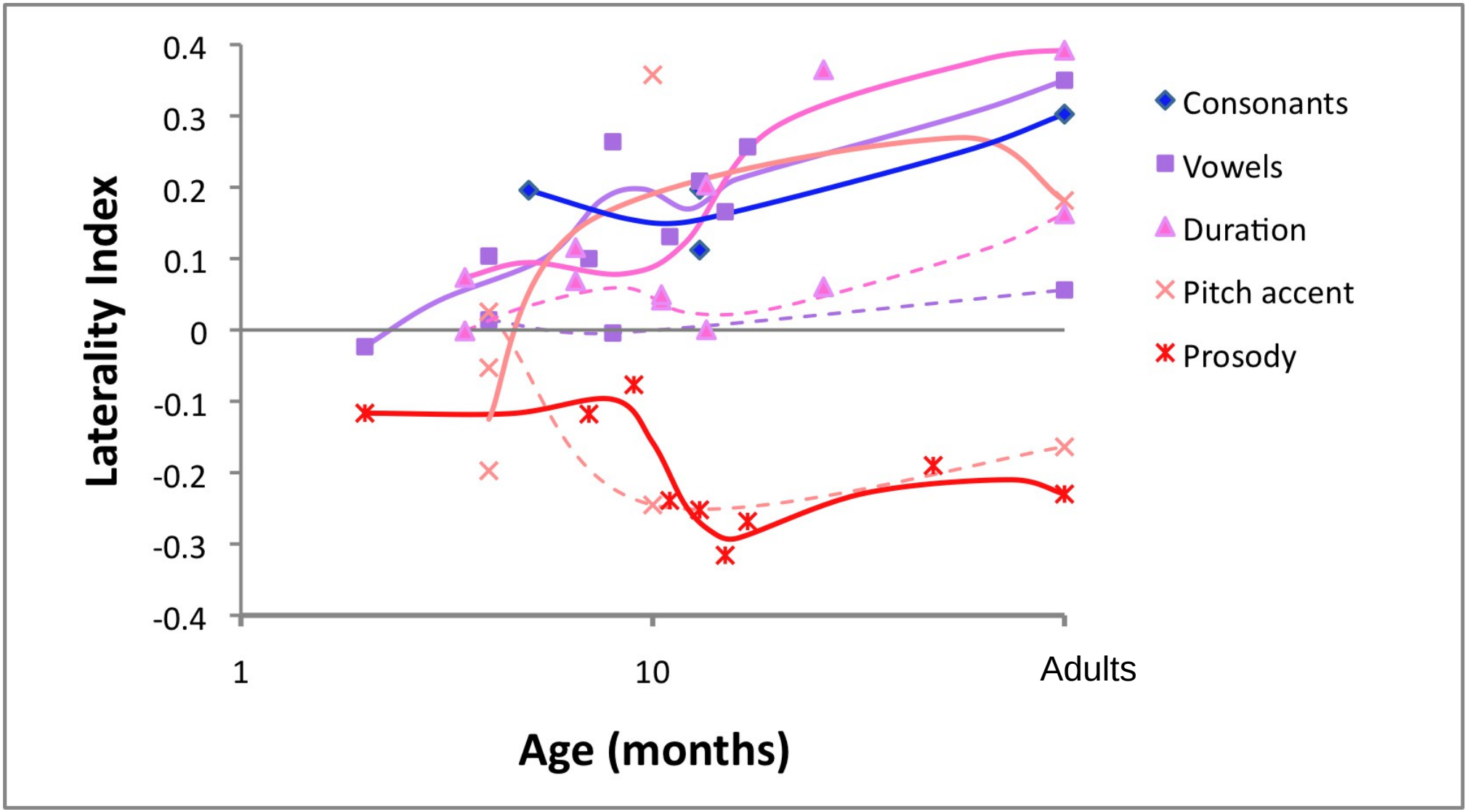
+/- temporal resolution 10 Hz

- spatial resolution ~25 mm



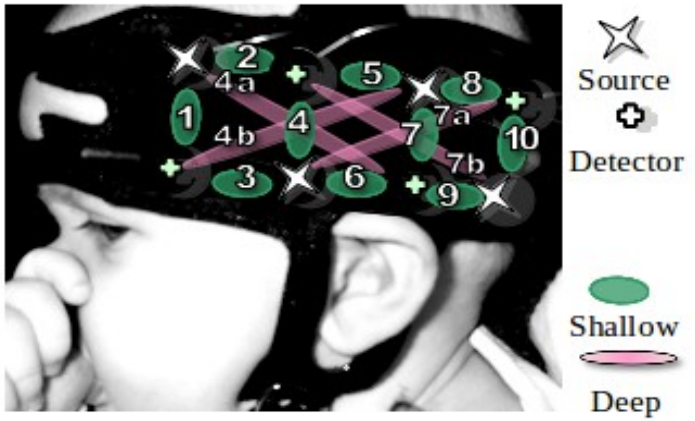


Lateralization phonetic contrasts



Minagawa, Cristià & Dupoux 2011, *J Dev Cogn Neurosci*

Predictions & instruments

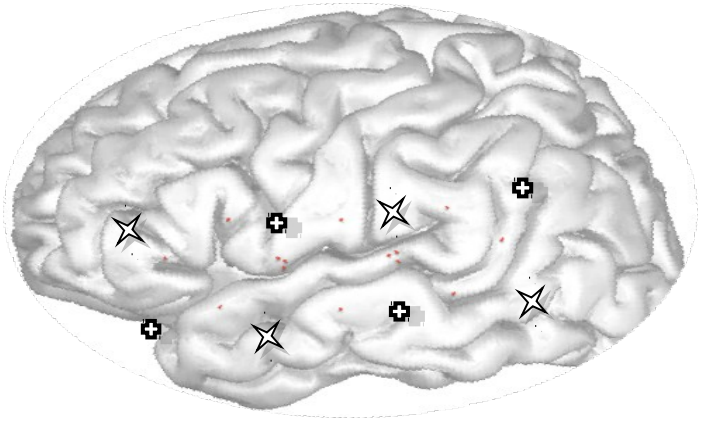


Since accent discrimination relies on *language-specific* knowledge...

Accent alternation will elicit higher activation in left perisylvian cortices

Difference greater for older infants*

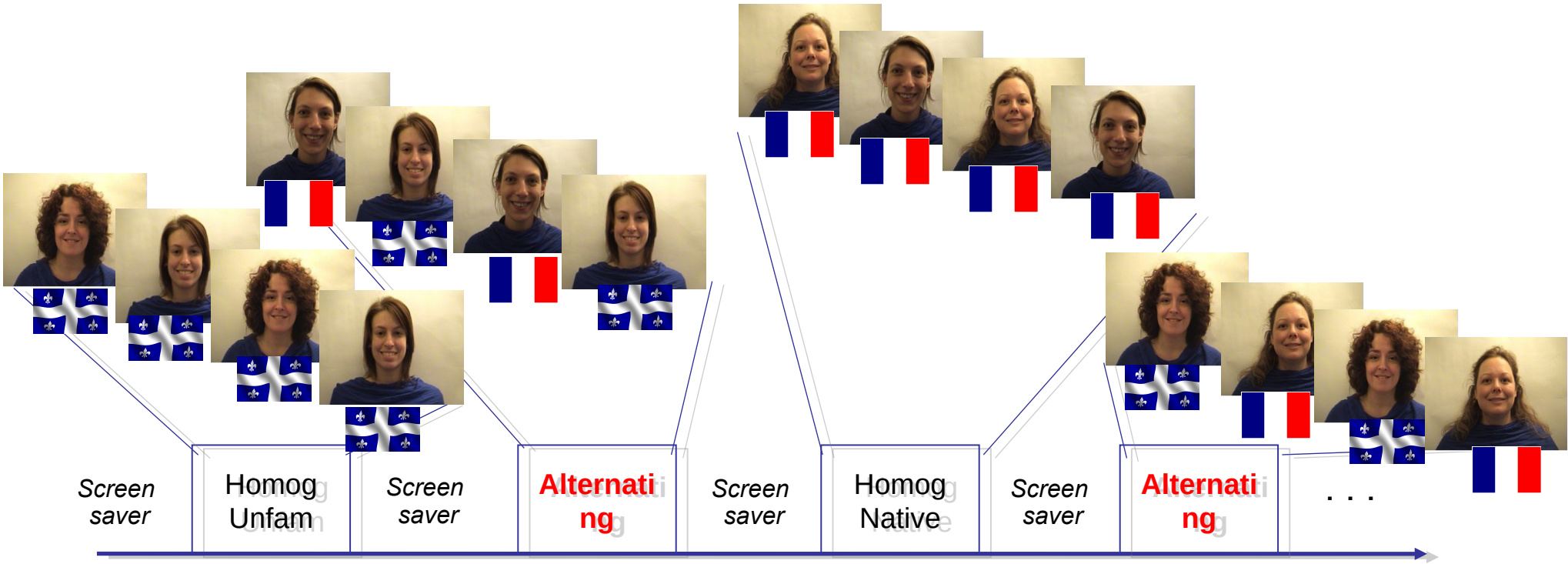
3.5-4mo versus 5mo



*cf. Butler et al. 2011 & Kitamura et al. in press

Accent discrimination: paradigm

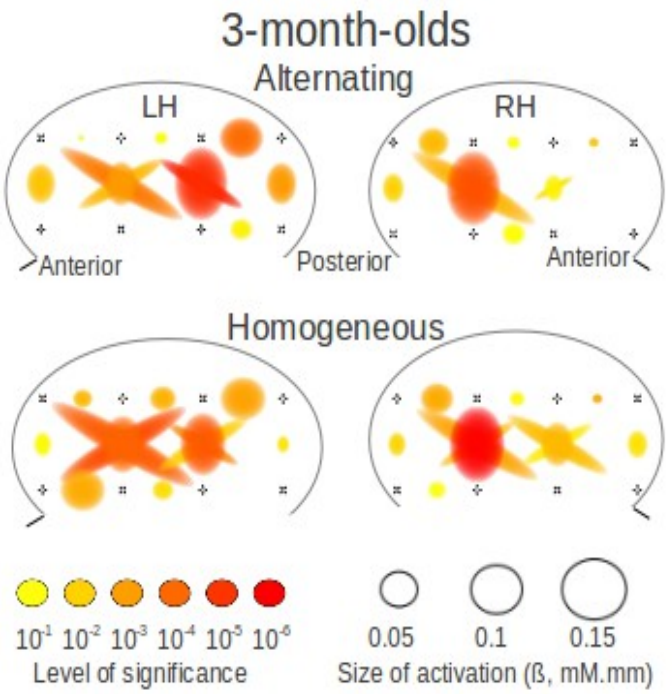
Accent discrimination: paradigm

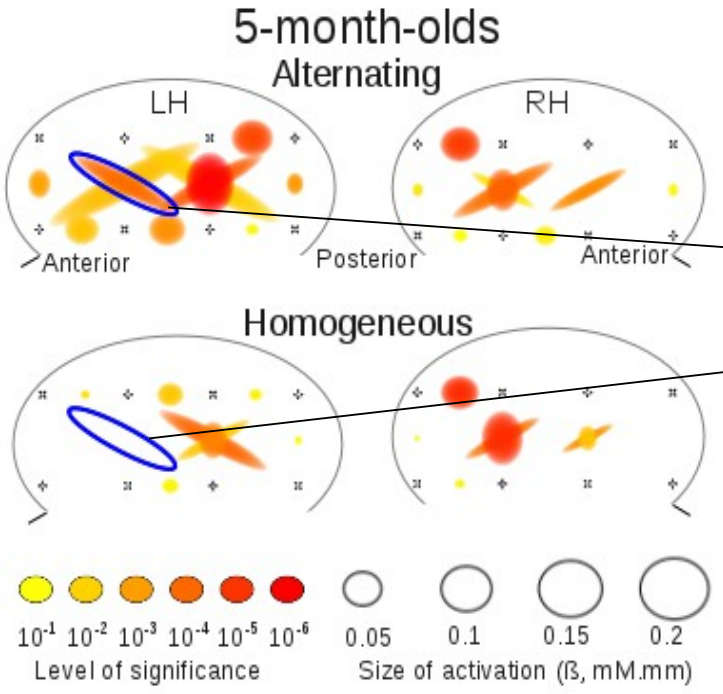


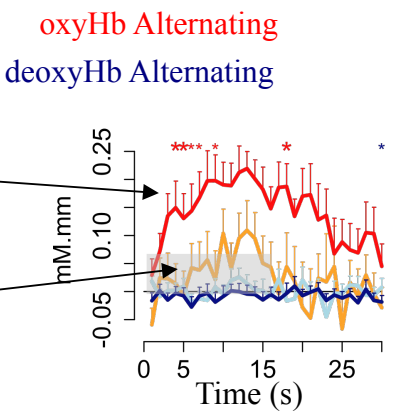
Change in:

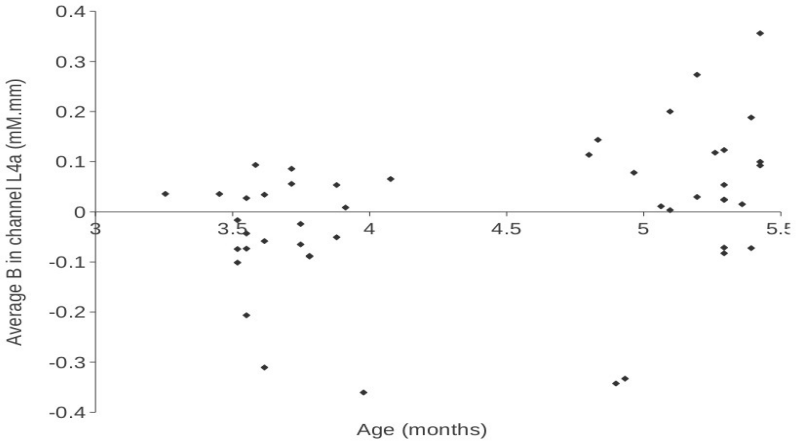
Talker	+	+	+	+
Language	-	+	-	+

Results









Cristia et al., under review

Open questions

Variation with task?

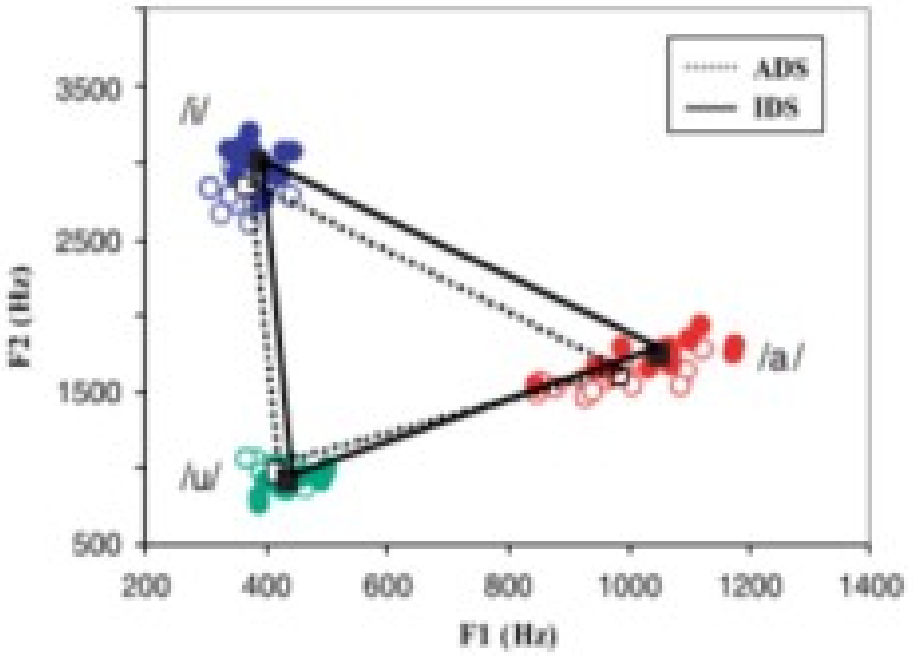
- e.g. Social “judgments”?

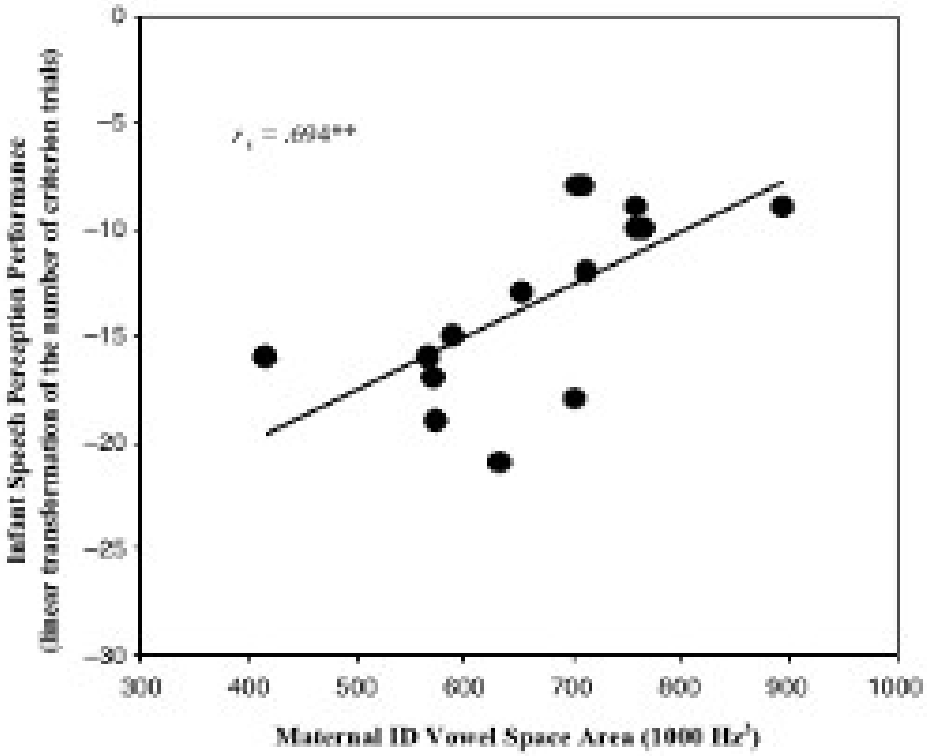
Changes with age?

- i.e. accent becomes a “talker” feature?

extra slides on multivariate studies

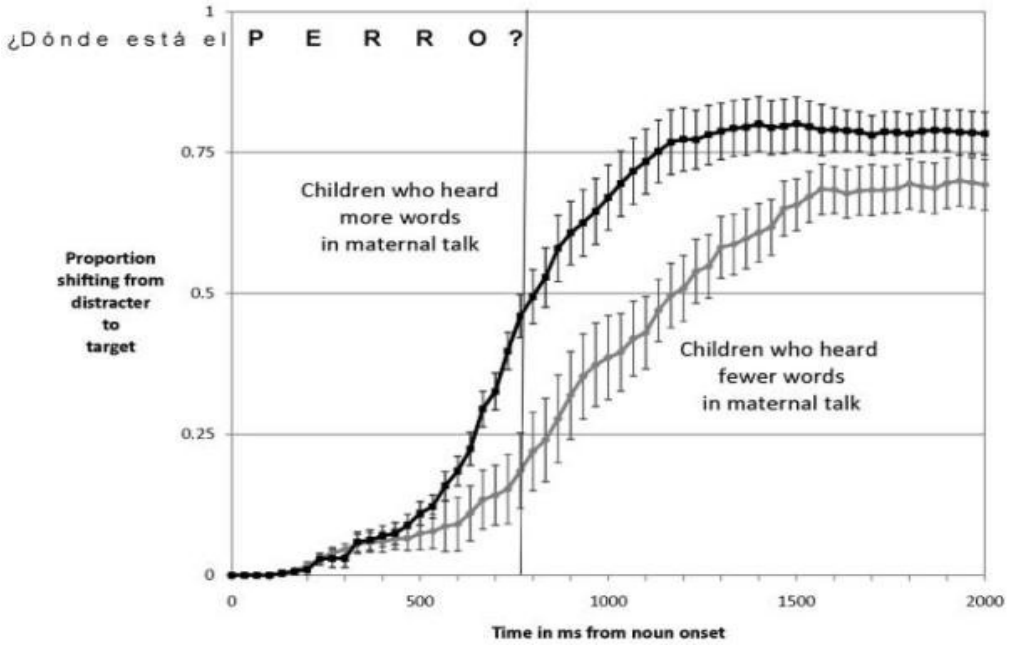
Phonetic detail*

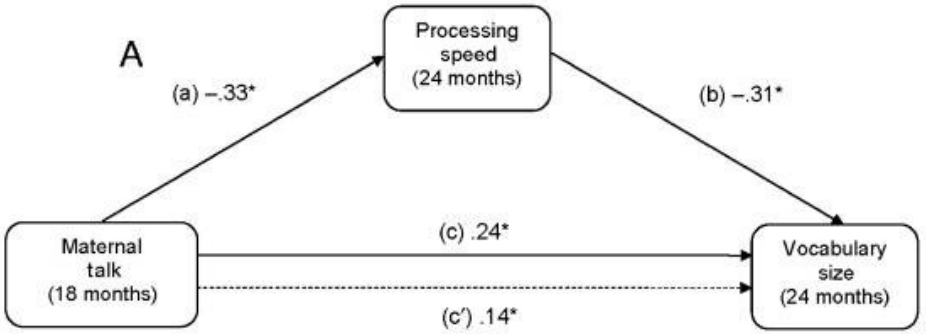




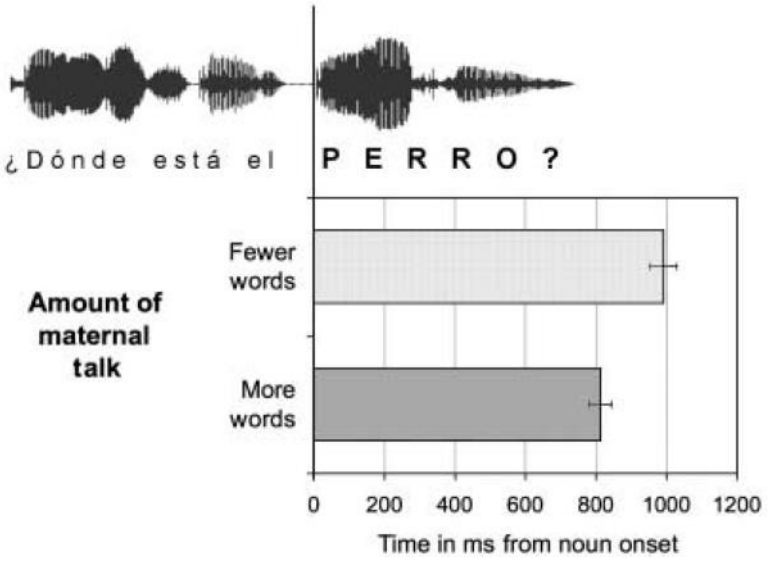
Liu, Tsao & Kuhl 2003 *Dev Sci*

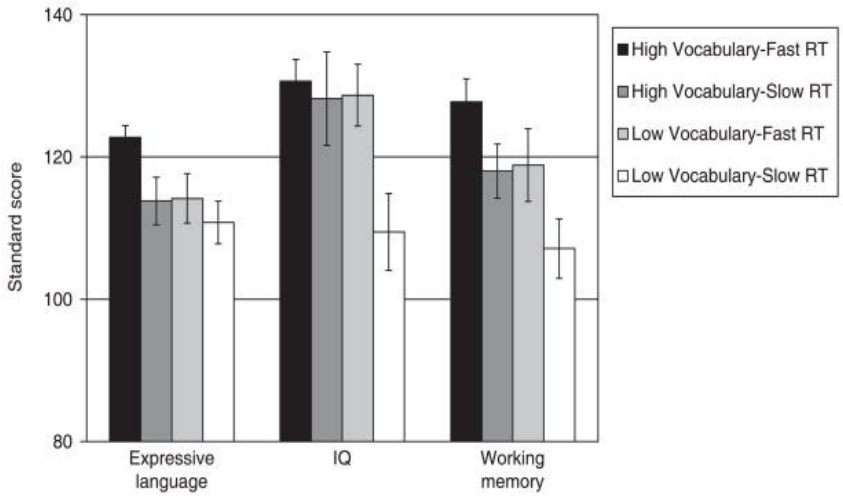
Multivariate studies (cont.)





Hurtado et al. 2008, *Dev Sci*





Marchman & Fernald 2008, *Dev Sci*