

ARE BABY-TALK WORDS SHAPED BY BIOMECHANICAL CONSTRAINTS ON ARTICULATION?

Mits Ota, Barbora Skarabela, Judit Fazekas, & Lovisa Wihlborg
University of Edinburgh

ICPC 2015



Baby-talk words

Register-specific conventionalized lexical items in infant/child-directed speech



Lexical redundancy



rabbit



bunny

train



choo-choo

stomach



tummy



Common formal characteristics

- Reduplication – full or partial
boo-boo, night-night; daddy, poopy
- Paucity of onset/coda clusters
Cf. stomach vs. tummy
- Recurrent edge patterns (i.e., ‘diminutives’)
horsie, piggy; teddy, bunny
- Convergence on dominant prosodic pattern
English: ‘σσ
- Reduction in segmental inventory?
Stops/nasals > Liquids (Ferguson, 1977).

(Ferguson, 1964, 1977, 1978; Soderstrom, 2007; Turpin et al., 2013)

Common formal characteristics

Reduplication in BTWs across languages

Arabic: *bubbu* 'baby', *zēze* 'breast/milk'

Czech: *nini* 'sleep', *lulu* 'wee', *papu* 'food'

Gilyak: *ypyp* 'bath', *amqamq* 'walk', *nena* 'doll'

Japanese: *pompon* 'stomach', *kukku* 'shoe'

Marathi: *toto* 'bath', *nuni* 'penis', *gargar* 'hot'

Spanish: *nene* 'baby', *tuto* 'sleep', *nino* 'nice'

Common formal characteristics

Diminutives in BTWs across languages

Arabic: *-o*

Comanche: *-ci*

Czech: *-ček, -ka*

Dutch: *-je*

German: *-chen*

Gilyak: *-k/q*

Japanese: *o-*

Marathi: *-[k]ula*

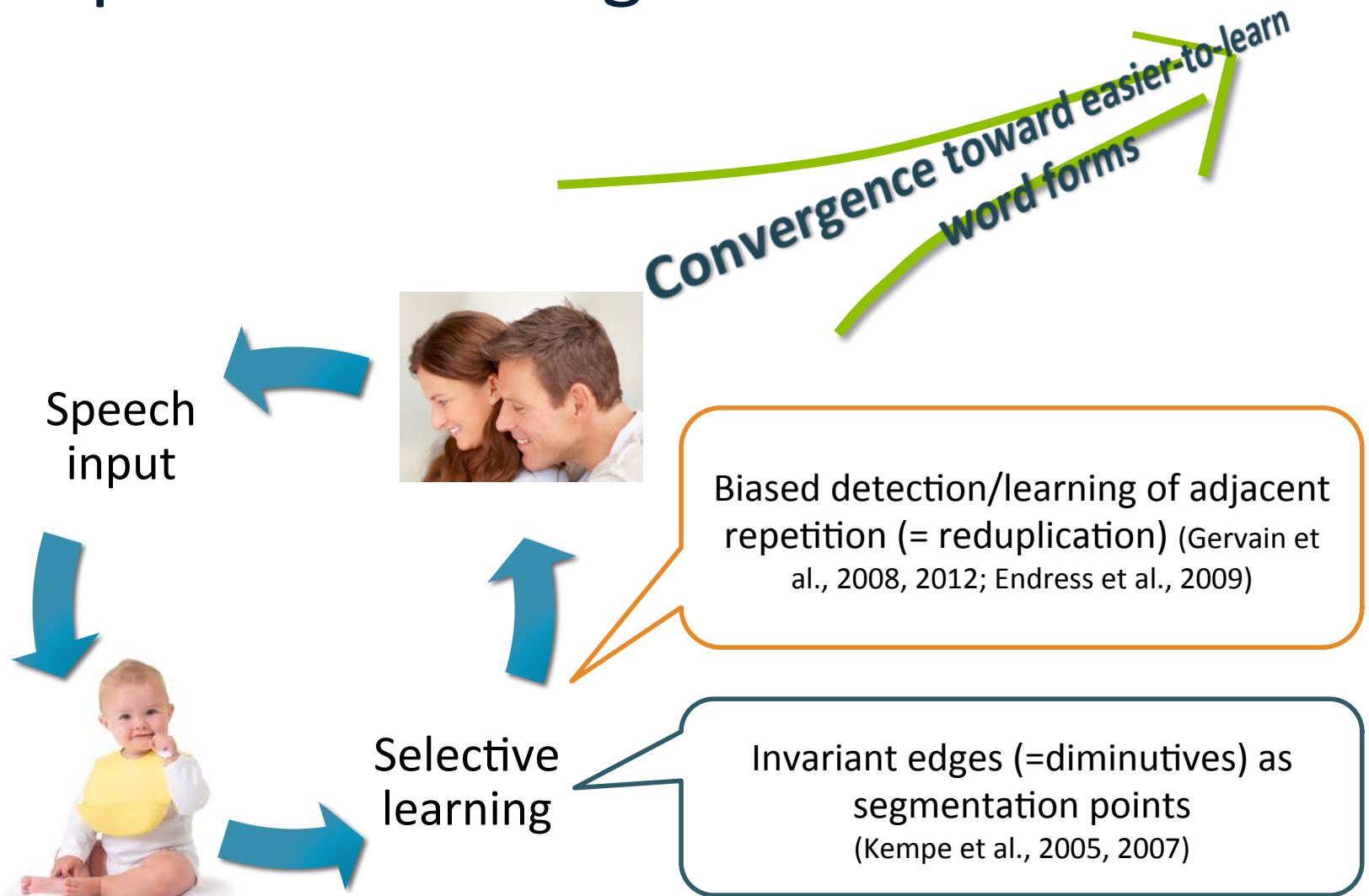
Spanish: *-ito/ita*

Turkish: *-cik*

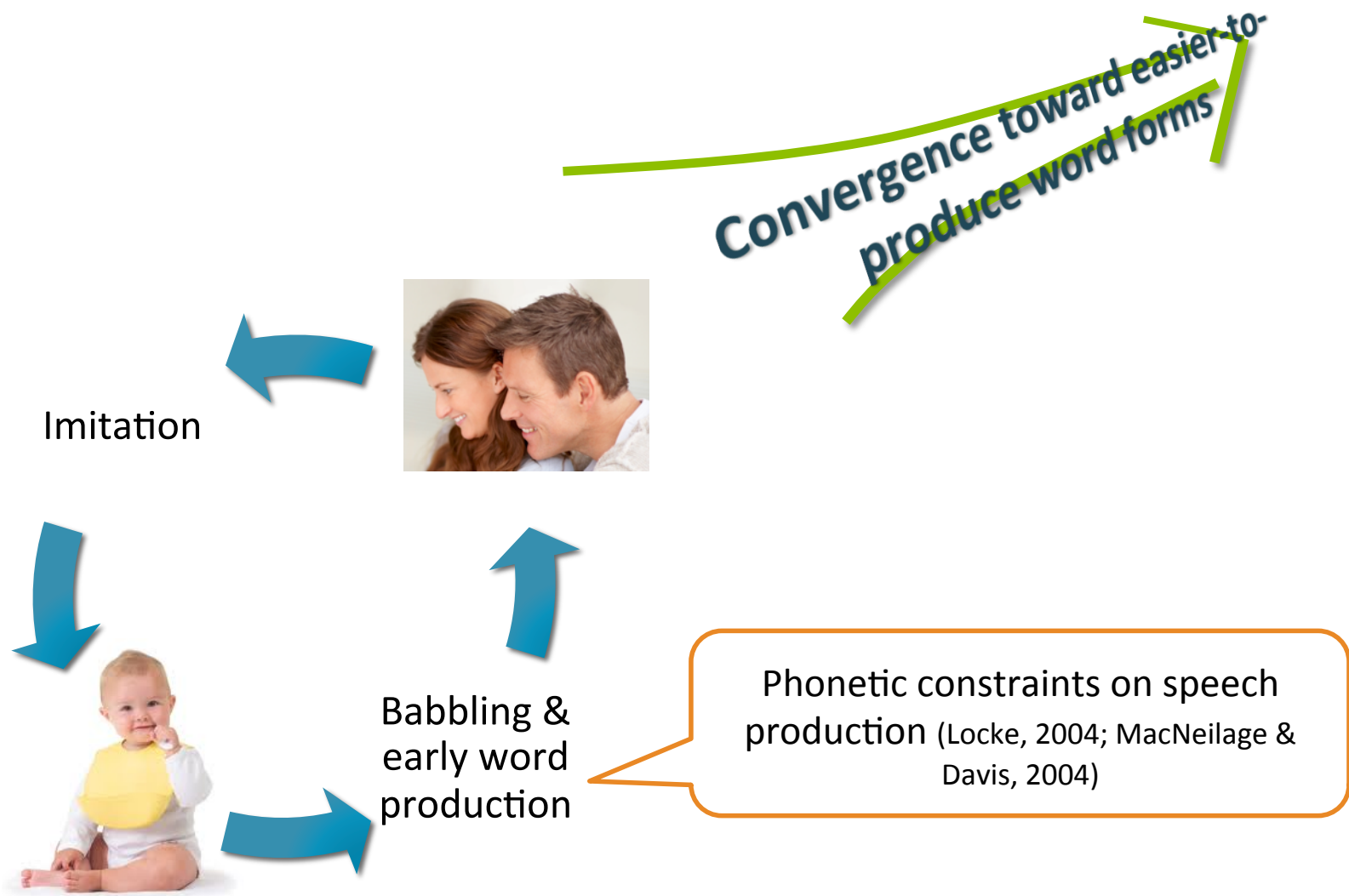
Background questions

1. Why have baby-talk words at all?
2. Why do they have similar phonological shapes?

Perceptual & learning biases?



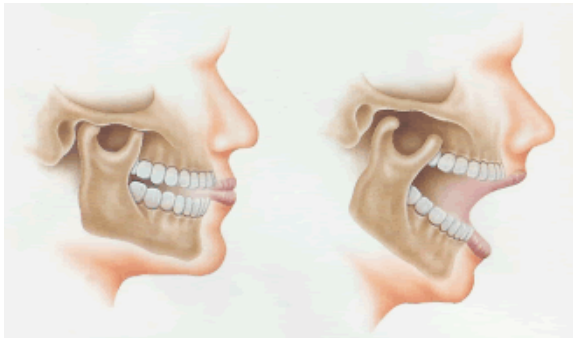
Production constraints?



Frame/Content (F/C) theory

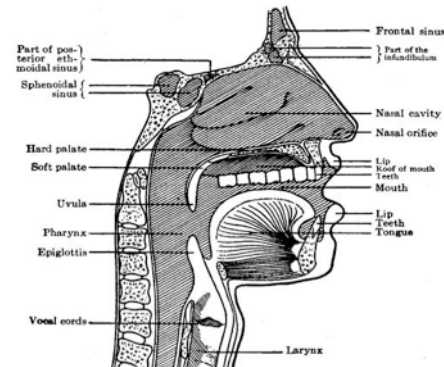
(MacNeilage, 1998; MacNeilage & Davis, 1990, 2000)

'Frames'



- Mandibular oscillation
- Pre-motor
- Syllable-related structures

'Content'



- Lips, tongue & jaw
- Motor planning
- Segment-related structures

Evolution

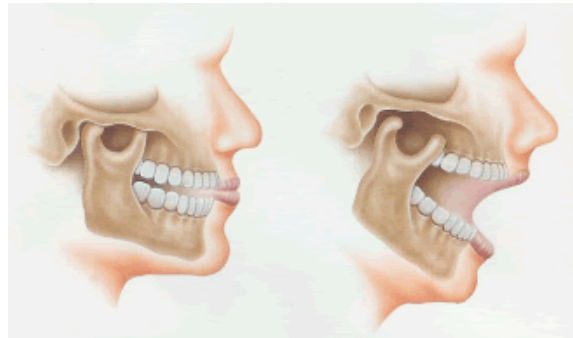


Development



Biomechanically favored patterns (1)

The canonical CV syllable = Frames + basic content



Complete oral
closure

Open jaw +
phonation

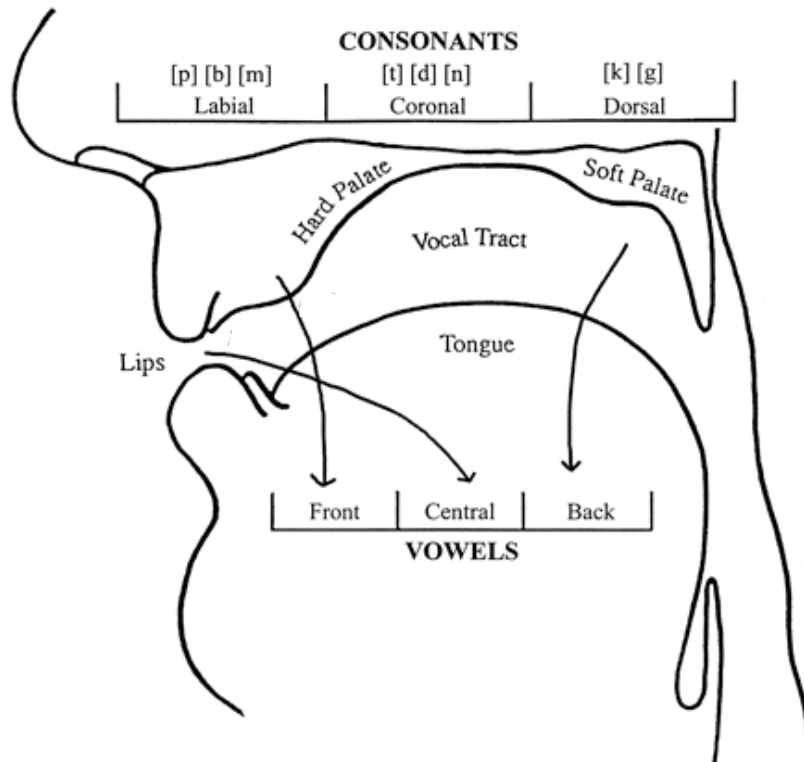
Oral/nasal
stop

Vowel

e.g., [dæ], [ma]

Biomechanically favored patterns (2)

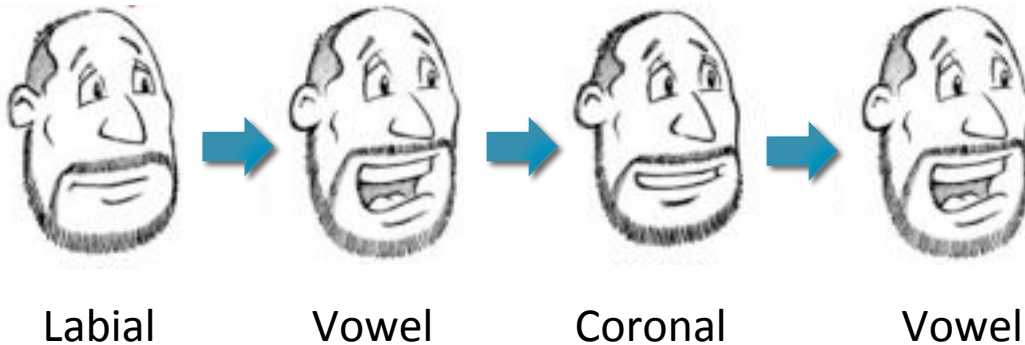
3 CV co-occurrence obeying lingual inertia



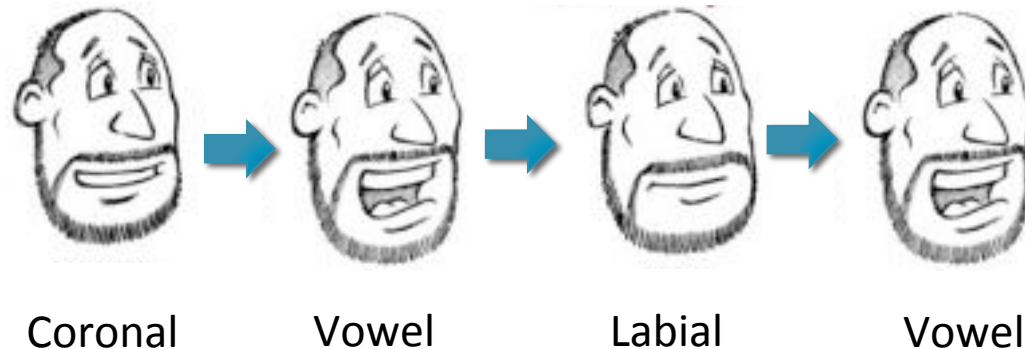
1. Coronal-Front: e.g., [di]
2. Labial-Central: e.g., [ba]
3. Dorsal-Back: e.g., [gu]

Biomechanically favored patterns (3)

Labial-(V)-coronal ('LC') intersyllabic sequence



Lab V Cor V
Easier (Lab) first



Cor V Lab V
Difficult (Cor) first

Evidence for F/C predictions

	Babbling	First words	Adult lexicon	Baby-talk words
Predominance of canonical CVs	D&M90	Many	Maddieson 1984	
Preference for Cor-Front, Lab-Central & Dor-back CV combinations	D&M95	DMM02, DM02	MDKM00	M&D04
Preference for LC (vs. CL) intersyllabic sequences		M&D00, M&00	M&00	MD04

D&M95 = Davis & MacNeilage (1995). DMM02 = Davis, MacNeilage, & Matyear (2002). M&D90 = MacNeilage & Davis (1990). M&D00 = MacNeilage & Davis (2000). M&D04 = MacNeilage & Davis (2004). M&00 = MacNeilage, Davis, Kinney, & Matyear (2000)

Are BTWs biomechanically friendlier than adult words?



[ræbit]

[bʌni]

Canonical CVs?	0 out of 2	2 out of 2
CV combinations?	Cor-Front	Lab-Central and Cor-Front
LC vs. CL?	CL	LC

Predictions

Compared to their adult word equivalents, baby-talk words should have higher proportions of:

1. canonical CVs
2. Cor-front, Lab-central, and Dor-back intersyllabic CV combinations
3. intrasyllabic Lab-Cor sequences (than Cor-Lab sequences) in #**CV.CVX** contexts.

Data: cross-linguistic samples

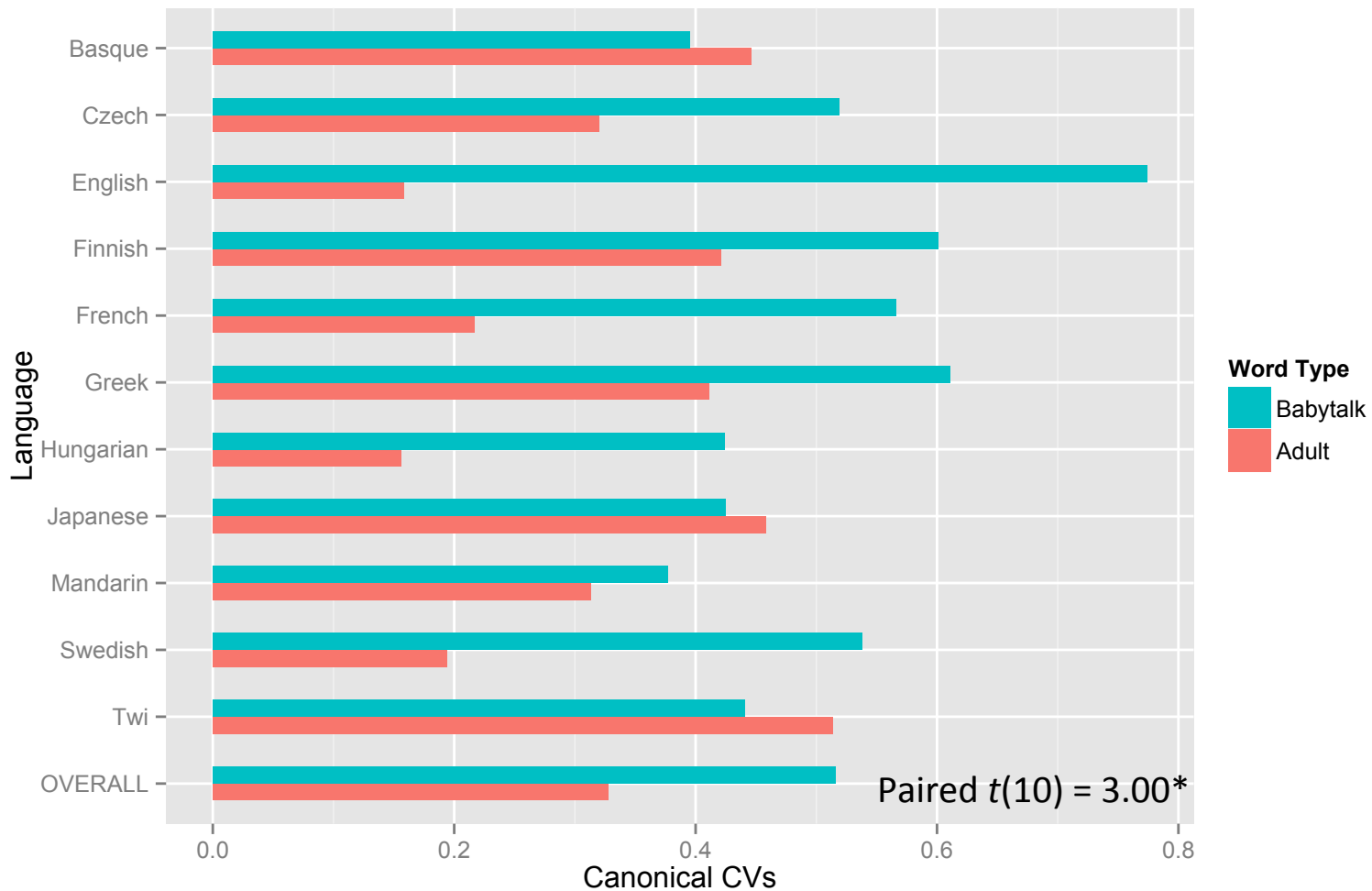
- Baby-talk words + adult equivalents
e.g., bunny – rabbit, choochoo – train, tummy – stomach
- Elicitation template based on Ferguson (1964): 45 meanings in 8 semantic categories + any others:
‘actions’ (e.g., walk, clap), ‘animals’ (e.g., dog, cat), ‘baby objects’ (e.g., bottle, pacifier), ‘body parts’ (e.g., breast, nose), ‘food’ (e.g., water, sweets), ‘quality’ (e.g., dirty, small), ‘people/kinship’ (e.g., baby, mother), ‘vehicles’ (e.g., car, train)
- Cross-checked between two informants

Data: current status

- 11 languages: Basque, Czech, (American) English, Finnish, (European) French, (Tokyo) Japanese, Hungarian, Greek, (Taiwanese) Mandarin, Swedish, Twi
- Total number of word-pairs: 385
- Average number of word-pairs per language: 35 (range: 25-51)

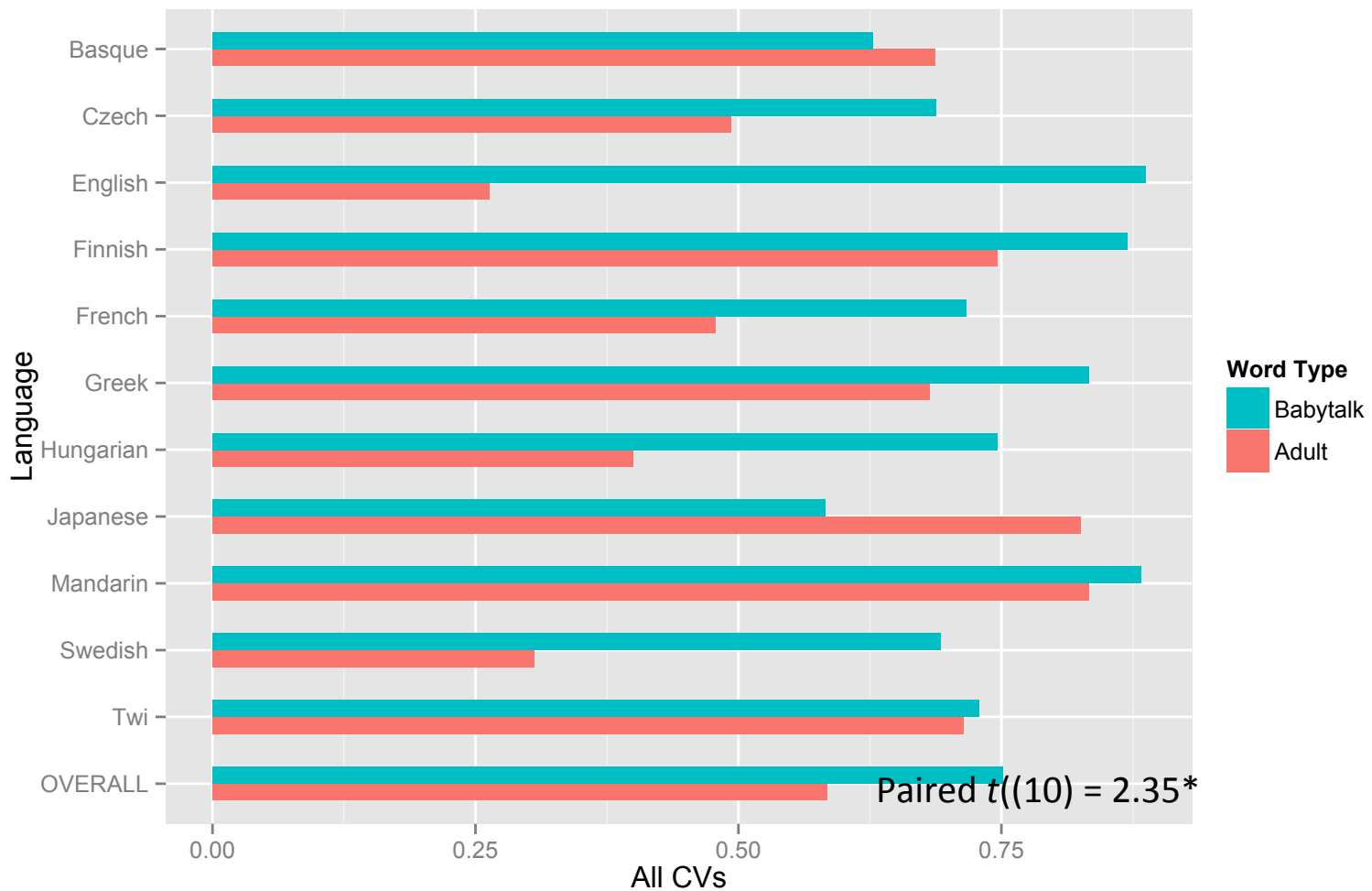
More canonical CVs?

Proportion of CV syllables with oral/nasal stop onsets



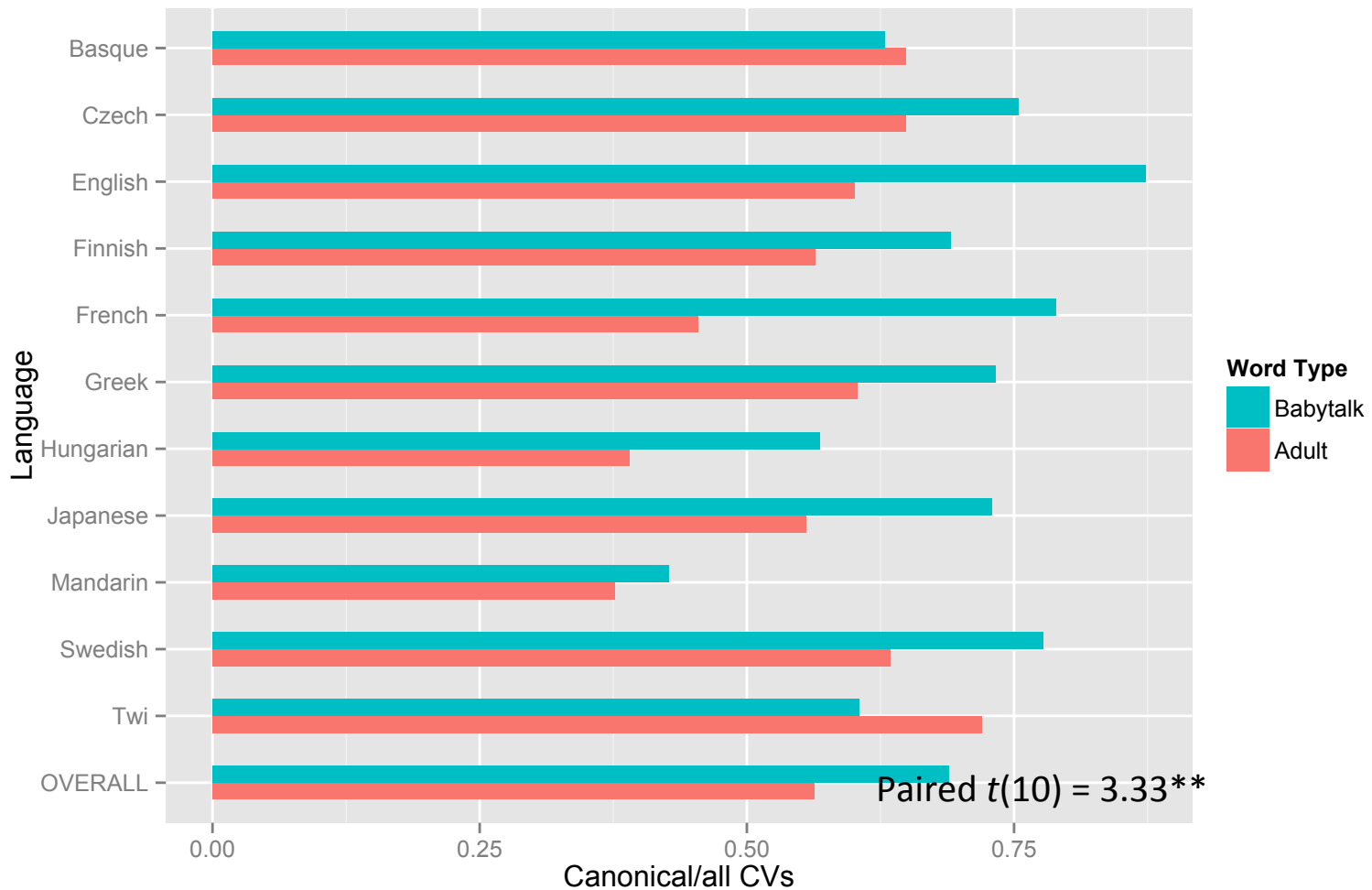
More canonical CVs?

Proportion of CV syllables with any type of onsets



More canonical CVs?

Proportion of CV syllables that are canonical



More preferred CV combinations?

Basque example

Adult words

Baby-talk words

Counts

	Front	Central	Back
Cor	17	8	14
Lab	6	1	6
Dor	5	7	6

	Front	Central	Back
Cor	8	6	5
Lab	7	10	7
Dor	13	10	7

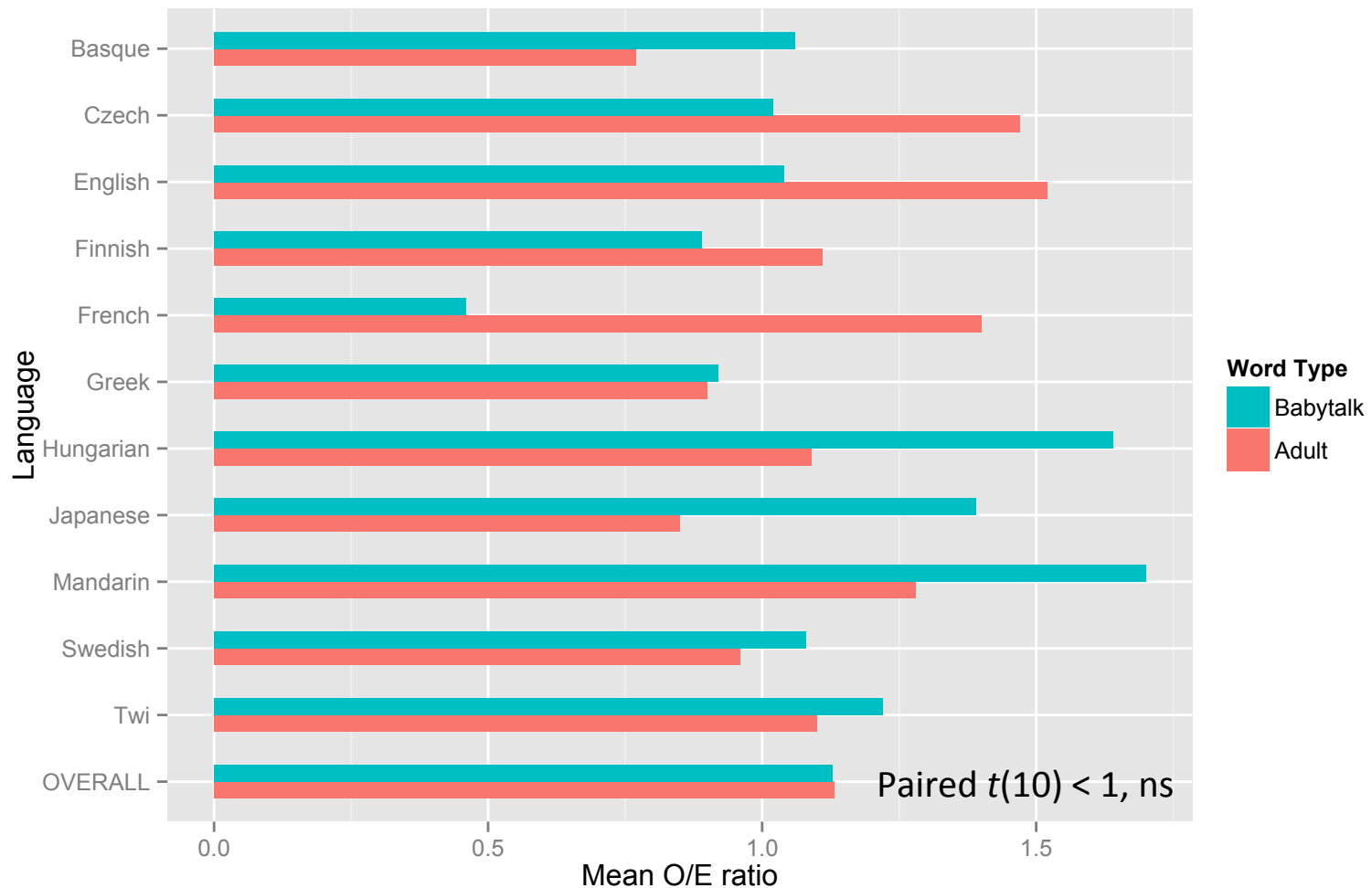
O/E ratios

	Front	Central	Back
Cor	1.09	0.90	0.97
Lab	1.15	0.34	1.24
Dor	0.69	1.70	0.90

	Front	Central	Back
Cor	1.10	0.89	1.01
Lab	0.76	1.17	1.12
Dor	1.13	0.94	0.90

More preferred CV combinations?

Mean observed/expected ratios of Cor-Front, Lab-Central & Dor-Back



A stronger LC bias?

Basque example

Adult words

Baby-talk words

Counts

	Cor	Lab	Dor
Cor	3	0	3
Lab	3	0	0
Dor	5	2	1

	Cor	Lab	Dor
Cor	3	1	0
Lab	3	4	0
Dor	5	2	1

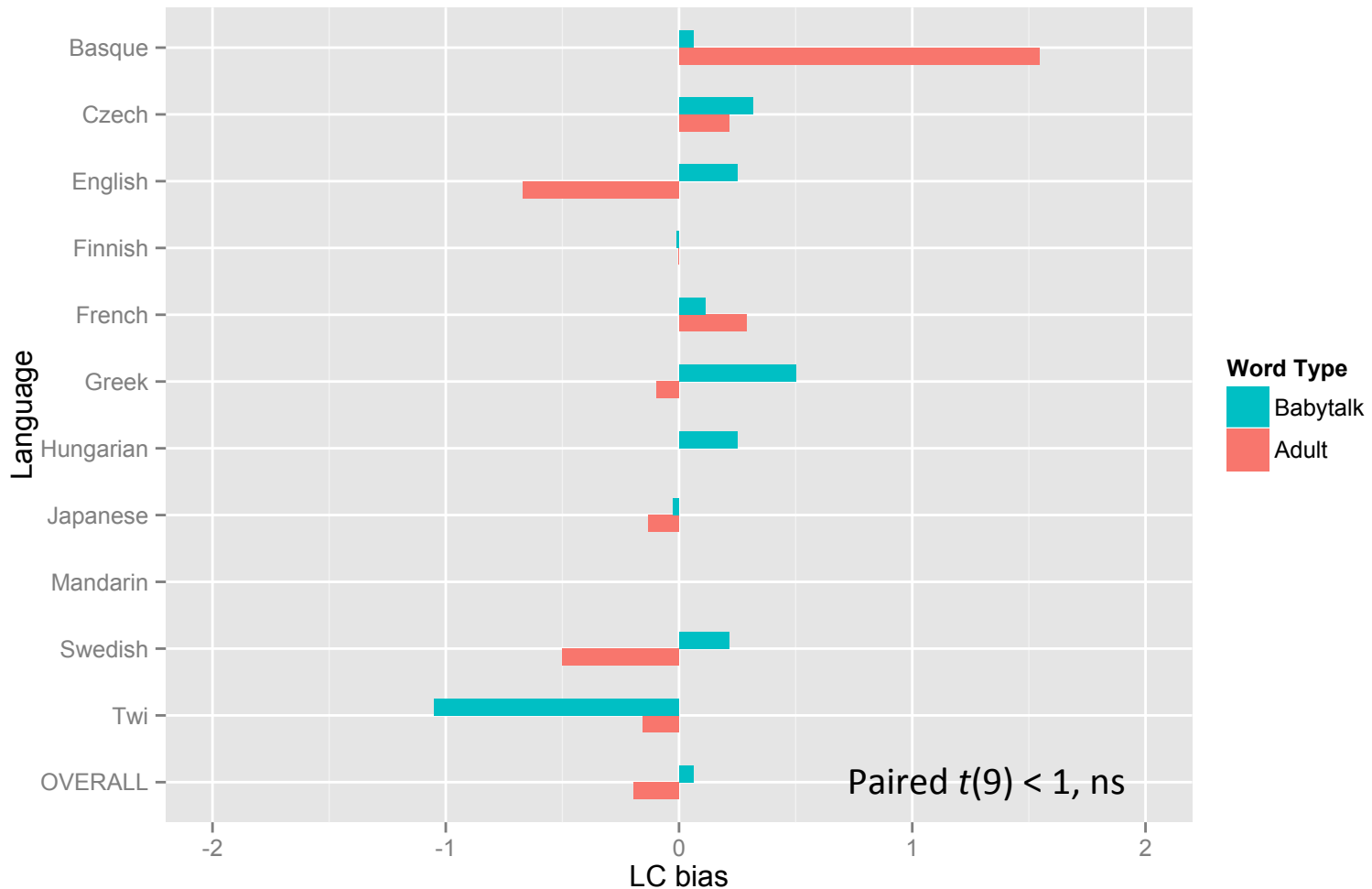
O/E ratios

	Cor	Lab	Dor
Cor	0.77	0.00	0.00
Lab	1.55	0.00	0.00
Dor	0.97	2.13	0.53

	Cor	Lab	Dor
Cor	1.30	0.68	0.00
Lab	0.74	1.55	0.00
Dor	1.08	0.68	2.38

A stronger LC bias?

OE ratios for Lab-Cor sequence minus Cor-Lab sequence



Predictions confirmed?

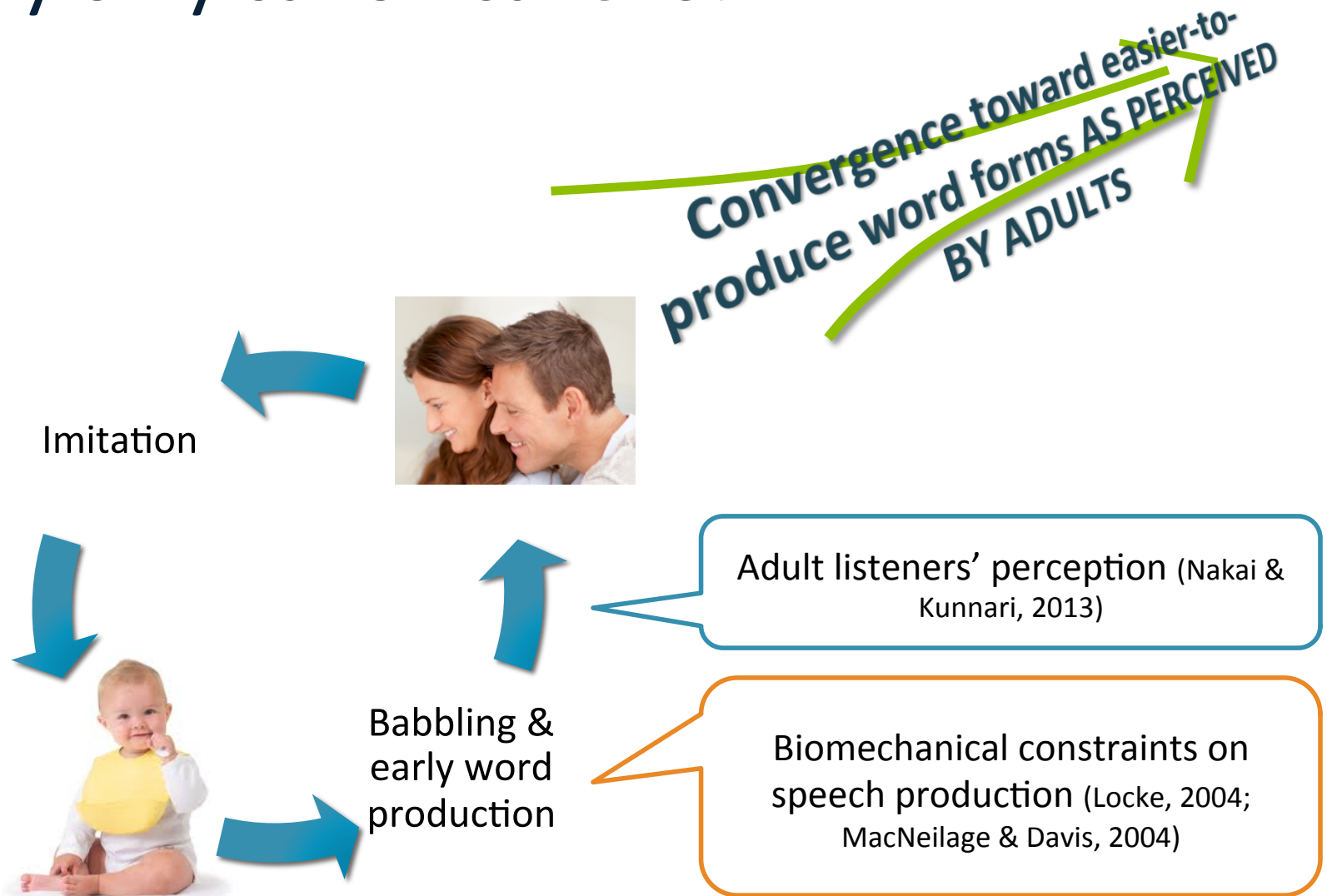
Compared to their adult word equivalents, baby-talk words should have higher proportions of:

1. canonical CVs → YES
2. Cor-front, Lab-central, and Dor-back intrasyllabic CV combinations → NO
3. intersyllabic Lab-Cor sequences (than Cor-Lab sequences) in #CV.CVX contexts → NO

Caveats

- Possible estimation errors
- Lack of power
- Sampling bias

Why only canonical CVs?



Conclusions

1. Effects of biomechanical constraints on speech production on baby-talk words (but only in %canonical CV so far)
2. The role of the adult listeners
3. Effects of child phonology (phonetics) on adult language

Thank you



Please help us
collect more data.

E·S·R·C
ECONOMIC
& SOCIAL
RESEARCH
COUNCIL

This work has been funded by the UK Economic and Social Research Council