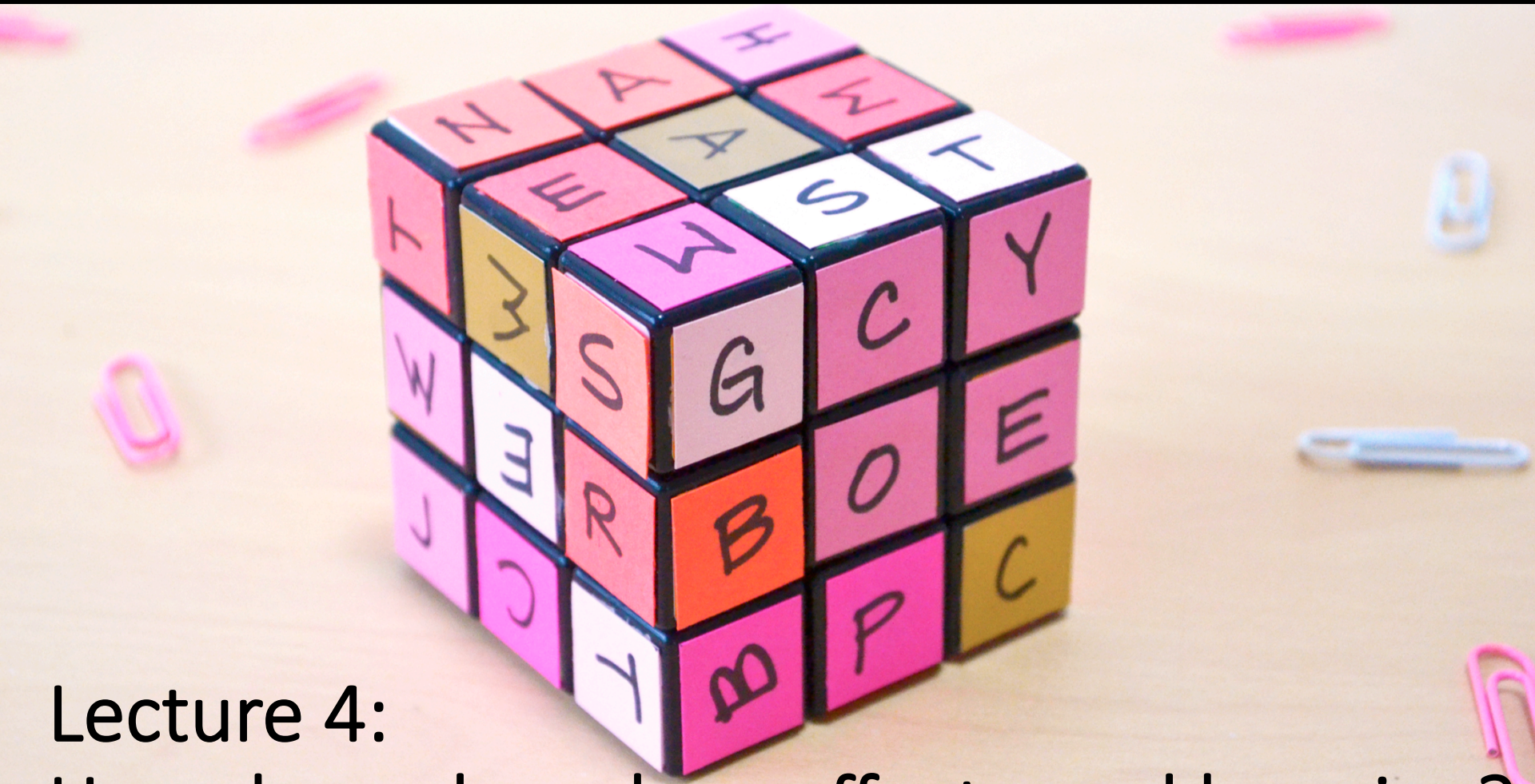
The relationship between early phonological and lexical development



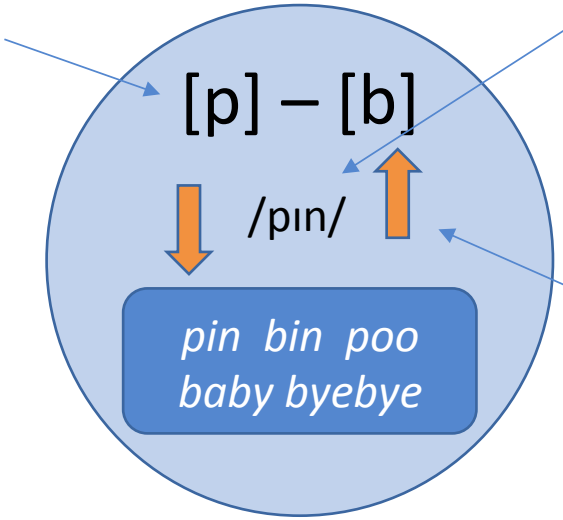
Lecture 4:
How does phonology affect word learning?

Mits Ota
University of Edinburgh

Summer School of Linguistics, 2019
mits-ota.github.io/SSoL2019.html

Last time

Can sounds be learned independent of words?



How do phonological representations develop?

How do lexical factors affect phonological development?

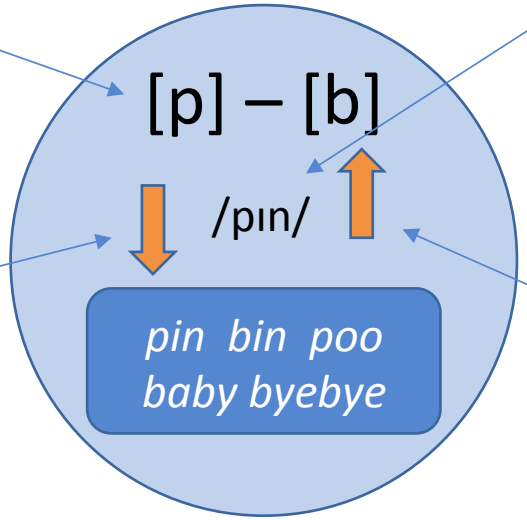
This lecture

Can sounds be learned independent of words?

How does phonology influence word learning?

How do phonological representations develop?

How do lexical factors affect phonological development?



Part 1

Do learners selectively learn words with certain phonological patterns?

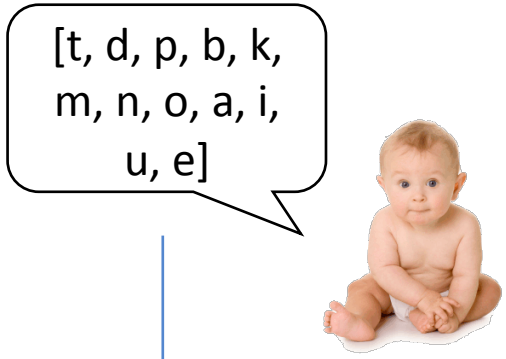
Learners show idiosyncratic phonological biases in lexical selection

Disproportionate number of words with certain phonological profiles found among children's first 50 words

- 'T' (Ferguson & Farwell, 1975): *ice*, *eyes, keys, shoes, juice*
- Daniel (Stoel-Gammon & Cooper, 1984): *quack, rock, clock, sock, whack, milk, frog, yuk, block, walk*

These biases are not a simple reflection of the lexical input (Schwartz & Leonard, 1982)

Initial language sampling



14-15 months

IN vs. OUT sounds

IN [t, d, k, e ...]

OUT [f, g, ʃ, ɔ, ...]

Word learning session

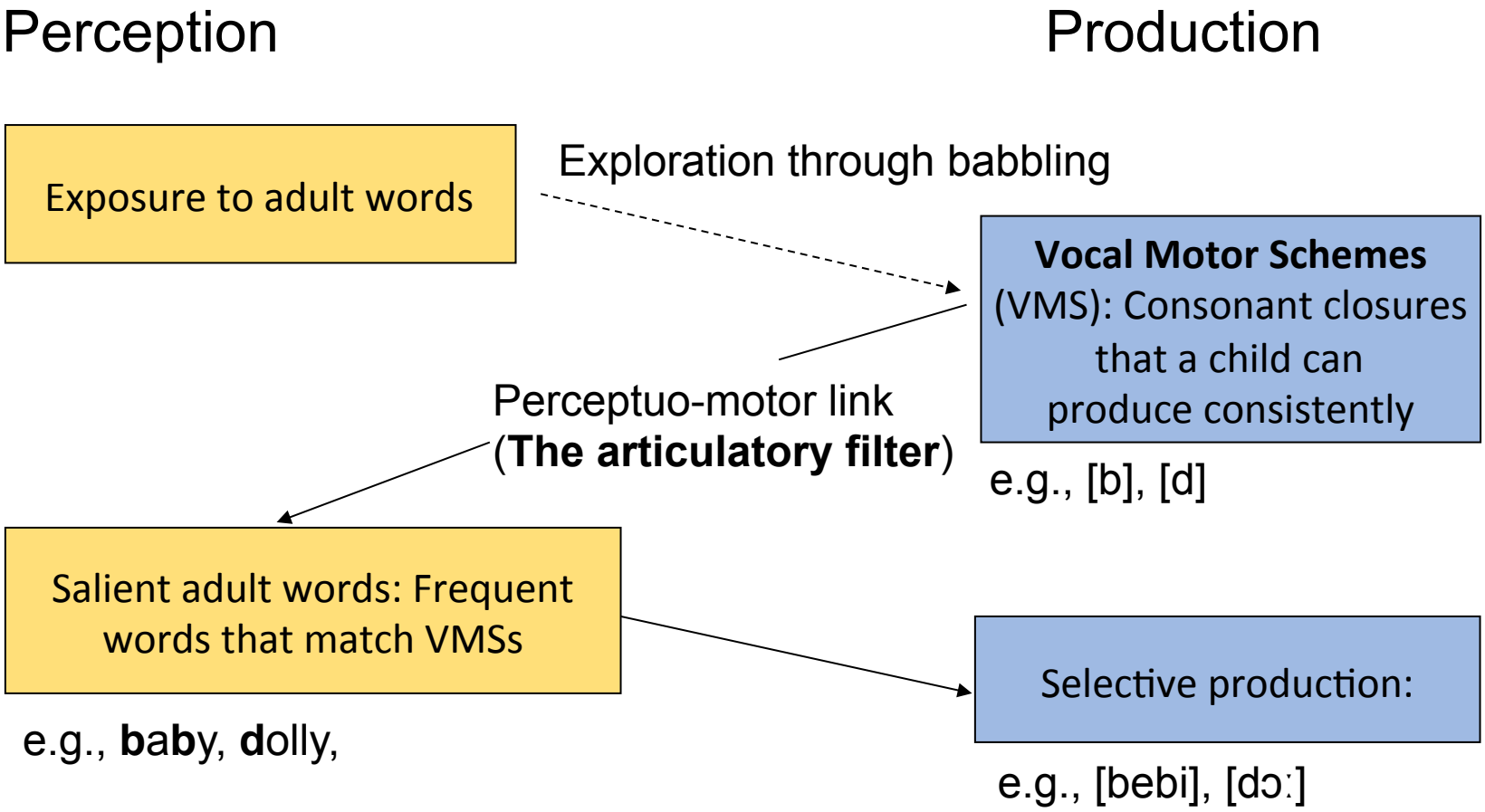
These biases are not a simple reflection of lexical input (Schwartz & Leonard, 1982)

Five month later: Number of concepts named by each child

	Non-imitative		Imitative	
Child	IN	OUT	IN	OUT
LF	6	4	7	6
CB	6	3	6	4
JR	8	4	6	2
BP	5	6	5	4
KS	5	7	4	3
MK	6	4	7	4
RC	2	3	3	2
TH	5	5	6	4
KP	7	4	5	3
KB	6	3	5	4
BG	8	6	6	6
LT	7	4	6	3
Totals	71	53	66	45

But no difference in accuracy of production.

Source of the bias: the articulatory-filter hypothesis (Vihman, 1993, 1996)



Beyond VMSs: Templates – broader idiosyncratic phonological patterns

Select		Adapt	
Target word	Child form	Target word	Child form
GERMAN: Annalena, 10–12 months <σ_1 σ_1>, that is, reduplicated syllables (Elsen, 1996)			
Pipi /ˈpipi/ 'peepee'	[pipi:]	*Annalena* /ˈanalena/	[nana]
wauwau /ˈvauvau/ 'bowwow'	[vava]	*kikeriki* /kikeriˈki::/ 'cock-a-doodle-do'	[ki:ki:]
		Bauch /baux/ 'belly'	[baba]
		Tag /tak/ '(good)day'	[dada]
		Zahn(bürste) /ˈtsa:nbürʃtə/ 'tooth(brush)'	[nana]
ENGLISH: Alice, 14 months <CVCi> (Vihman *et al.*, 1994)			
baby	[bebi]	*bottle*	[baɖi, ba:tʃi, batji]
daddy	[tæɟi]	*hiya*	[ha:ji]
lady	[jɛiji]		
mommy	[maɲi]		
FRENCH: Laurent, 15 months <C_oVIV> (Vihman & Kunnari, 2006)			
allo 'hello' /alo/	[alo]	*canard* 'duck' /kanaʁ/	[kɔla]
dans l'eau, de l'eau 'in/some water' /dɑ̃lo/, /dəlo/	[dəlo]	*chapeau* 'hat' /ʃapo/	[bɔlo]
ballon 'big ball' /balɔ̃/	[palɔ]	*la brosse* 'the brush' /labrʌs/	[bəla]
pas là 'not there' /pala/	[pala]	*la cuillère* 'the spoon' /lakɥijɛʁ/	[kola]
		voilà 'there you are' /vwala/	[lala]

Part 2

Are there any universal phonological biases for word learning?

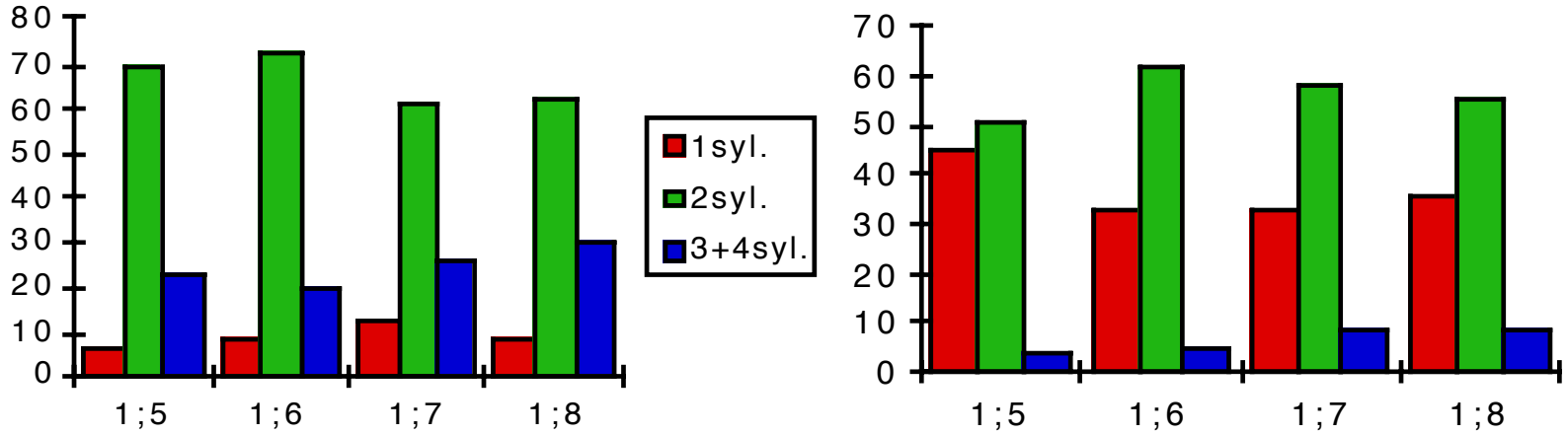
There are some general phonological characteristics in early-acquired words

Words listed in the American-English version of the MacArthur Communicative Development Inventory (Stoel-Gammon, 1998) :

- 60% monosyllables, 35% disyllables
- 81% without initial onset clusters
- 90% of multisyllabic words have initial stress
- 57% with an initial stop
- 22% with an initial /b/

But some of these simply reflect the lexical input

- 95% of lexical words in infant-directed English have initial stress (Kelly & Martin, 1994).
- 59% of words in infant-directed English are monosyllabic (Vihman, 1991). Distribution of target word size is language-dependent.



SPANISH GERMAN

Number of syllables in early words (Lleo & Demuth, 1999)

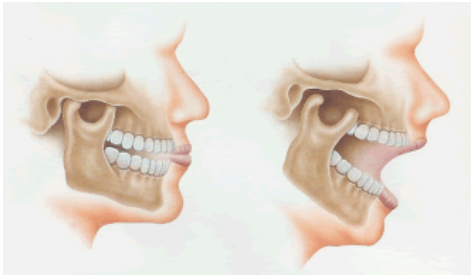
Learner- or input-bias?: The case of LC bias

Rank	Word	% samp
1.	daddy	54.9
2.	mommy	52.9
3.	abye	43.1
4.	ahi	39.3
5.	auh-oh	35.5
6.	baa-baa	31.9
7.	ball	30.9
8.	dog	30.6
9.	ano	28.5
10.	bottle	25.2
11.	woof	24.9
12.	baby	24.6
13.	ayum-yum	24.1
14.	grr	23.5
15.	kitty	21.8
16.	vroom	20.2
17.	book	19.9
18.	bird	19.6
19.	duck	18.8
20.	balloon	18.4
21.	cat	18.2
22.	anight-night	17.1
23.	quack	17.0
24.	shoe	17.0
25.	meow	16.6
26.	banana	16.3
27.	ahot	16.0
28.	juice	15.4
29.	eye	14.8
30.	grandma	14.3
31.	moo	14.2
32.	athank-you	14.0
33.	aup	14.0
34.	cookie	13.5
35.	nose	13.5
36.	aouch	13.4
37.	cracker	12.3
38.	grandpa	12.3
39.	ashh	12.0
40.	bath	11.8
41.	keys	11.8
42.	bubbles	11.4
43.	adown	11.4
44.	car	11.2
45.	ayes	11.0
46.	cheese	10.9
47.	bear	10.7
48.	ahello	10.6
49.	fish	10.4
50.	aallgone	10.3

First 50 English words in production (Caselli et al., 1999)

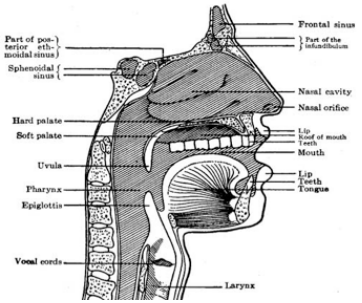
Reason for the LC bias: Frame/Content Theory (MacNeilage, 1998)

‘Frame’



- Mandibular oscillation

‘Content’



- Controlled motor movements of lips, tongue and jaw

Biomechanical complexity

labial → Coronal

timing

Less ‘content’

More ‘content’

‘easy-to-difficult’

Children tend to produce/learn more words with LC sequences (vs. CL)

Subject	LC	CL
Aa	30	18
An	13	9
C	6	2
J	39	0
K	9	0
Ma	21	2
Mi	2	0
N	19	1
P	6	3
R	36	36
Total	181	71

Numbers of labial-coronal (LC) sequences vs. coronal-labial (CL) sequences in the first words of American-English learning infants (MacNeilage et al., 1999). Same pattern reported for German, French, Dutch and Czech-learning infants (MacNeilage & Davies, 2000)

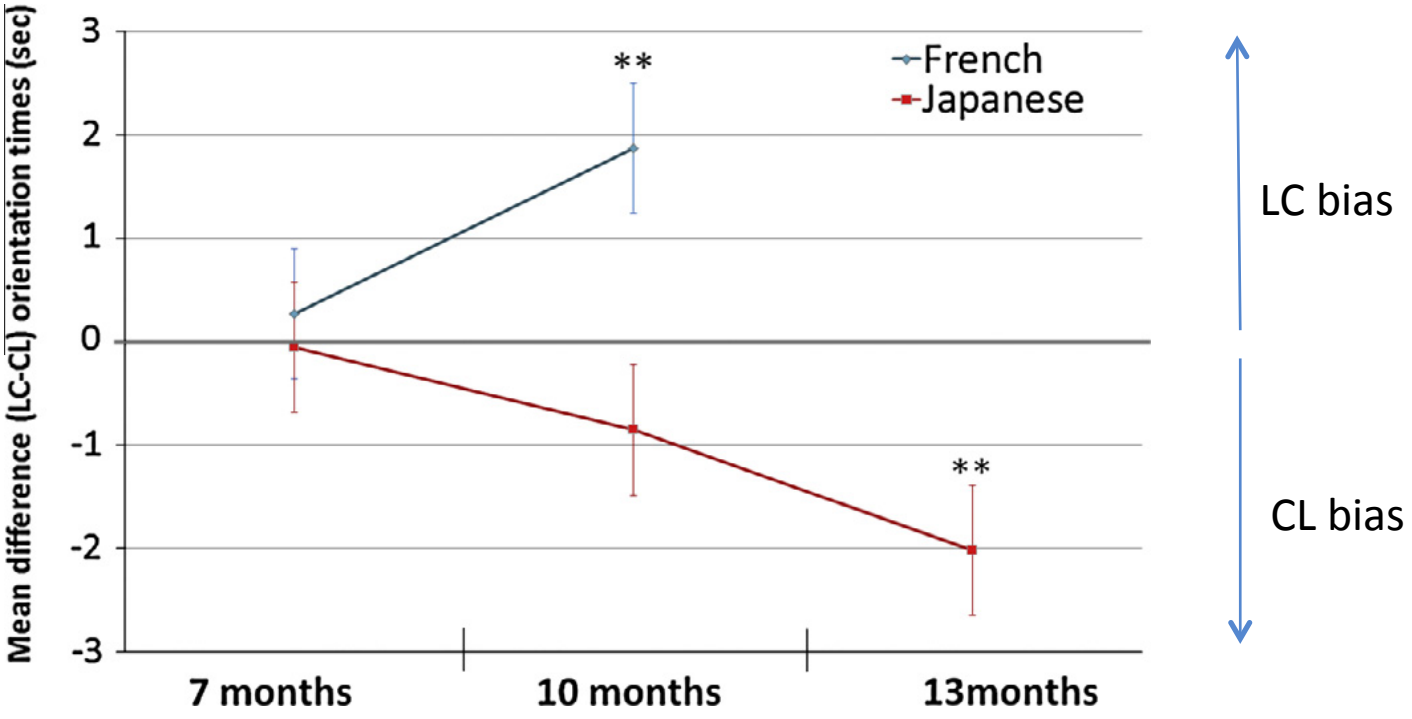
But the input languages also have an LC bias

Language	Number of words sampled	Ratio
English	942	2.55
Estonian	171	2.10
French	1,015	1.48
German	382	2.78
Hebrew	92	2.29
Japanese	68	0.84[a]
Maori	1,063	2.46
Quichua	65	3.33
Spanish	585	3.09
Swahili	164	1.34[a]
Mean ratio		2.23

[a]Not significant.

Ratio of LC to CL sequences in various languages (MacNeilage et al., 1999)

Indeed the LC bias emerges in response to biased input



Orientation difference towards LC (e.g., /pita/) and CL (e.g., /tipa/) stimuli (Gonzalez-Gomez et al., 2014)

Another case: Adjacent syllable or onset consonant repetition ('reduplication')

Rank	Word	% samp
1.	daddy	54.9
2.	mommy	52.9
3.	[a]bye	43.1
4.	[a]hi	39.3
5.	[a]uh-oh	35.5
6.	baa-baa	31.9
7.	ball	30.9
8.	dog	30.6
9.	[a]no	28.5
10.	bottle	25.2
11.	woof	24.9
12.	baby	24.6
13.	[a]yum-yum	24.1
14.	grr	23.5
15.	kitty	21.8
16.	vroom	20.2
17.	book	19.9
18.	bird	19.6
19.	duck	18.8
20.	balloon	18.4
21.	cat	18.2
22.	[a]night-night	17.1
23.	quack	17.0
24.	shoe	17.0
25.	meow	16.6
26.	banana	16.3
27.	[a]hot	16.0
28.	juice	15.4
29.	eye	14.8
30.	grandma	14.3
31.	moo	14.2
32.	[a]thank-you	14.0
33.	[a]up	14.0
34.	cookie	13.5
35.	nose	13.5
36.	[a]ouch	13.4
37.	cracker	12.3
38.	grandpa	12.3
39.	[a]shh	12.0
40.	bath	11.8
41.	keys	11.8
42.	bubbles	11.4
43.	[a]down	11.4
44.	car	11.2
45.	[a]yes	11.0
46.	cheese	10.9
47.	bear	10.7
48.	[a]hello	10.6
49.	fish	10.4
50.	[a]allgone	10.3

First 50 English words in production (Caselli et al., 1999)

Reasons for the repetition bias

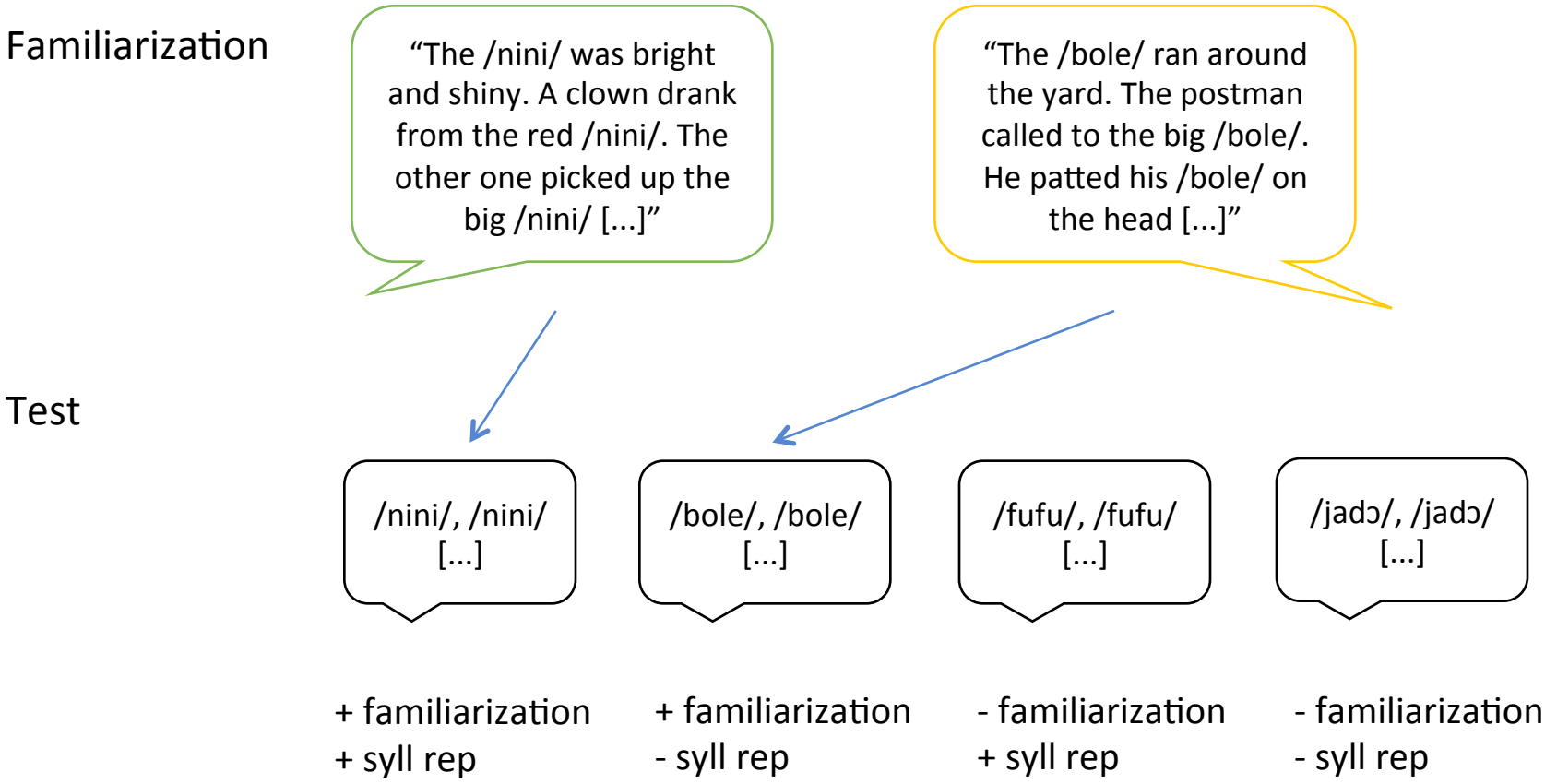
Articulatory filter (Vihman, 1996)

- Babbling favors repetitive articulation
- 'Consonant harmony' prevalent in early production (e.g., [gɔgi] 'doggy'

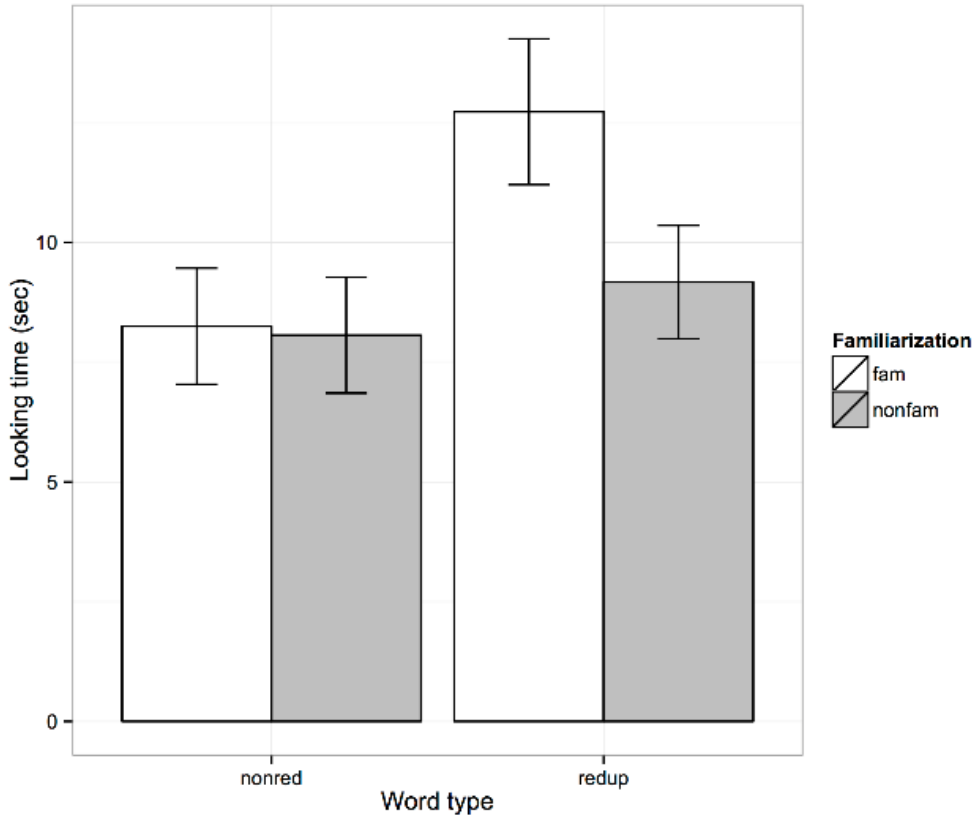
Early perceptual/memory bias for repeated elements (Endress et al., 2009)

- Neurological evidence (/mubaba/ > /mubamu/, /mubage/) (Gervain et al., 2008, 2012)

Infants preferentially segment words with syllable repetition (Ota & Skarabela, 2018)

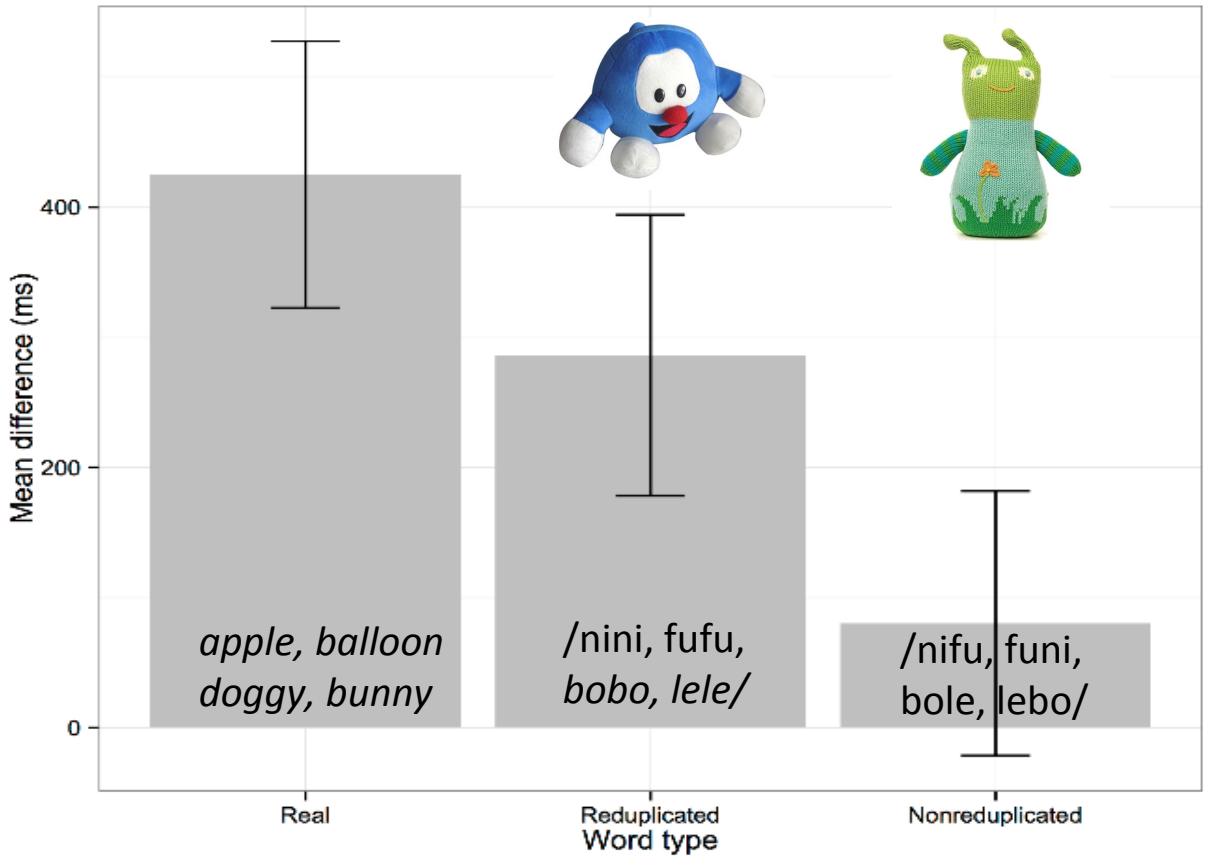


Infants preferentially segment words with syllable repetition (Ota & Skarabela, 2018)



Mean looking time towards visual stimulus (9-month-olds)

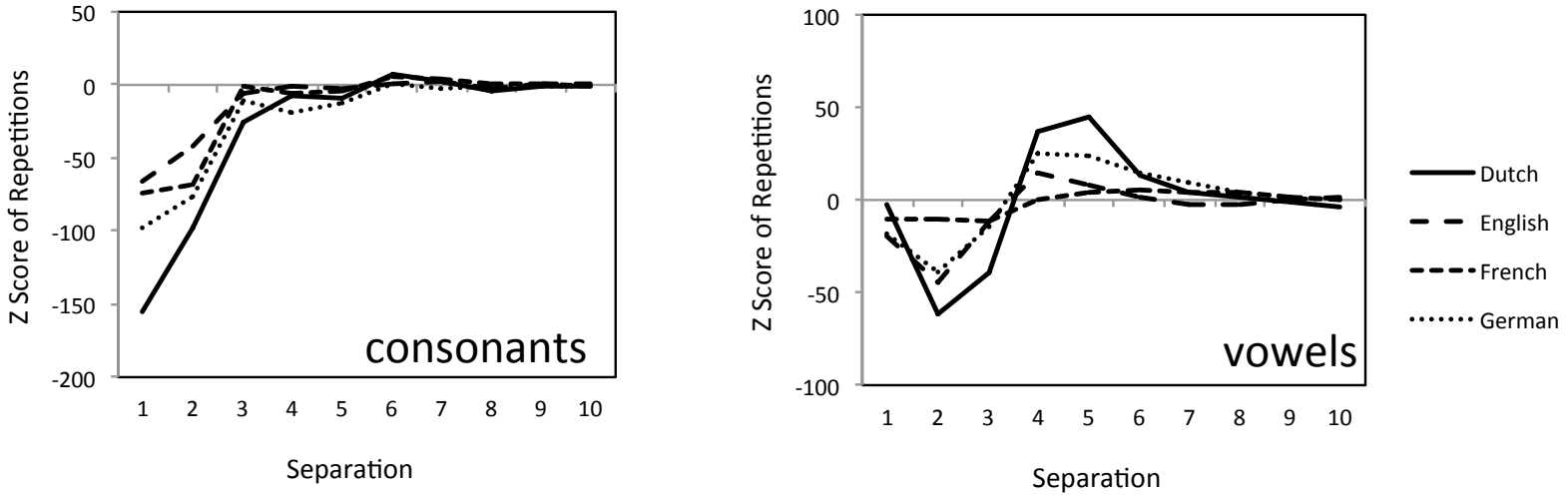
Infants also preferentially learn mapping of words with syllable repetition (Ota & Skarabela, 2016)



Mean increase in looks to correct picture (18-month-olds)

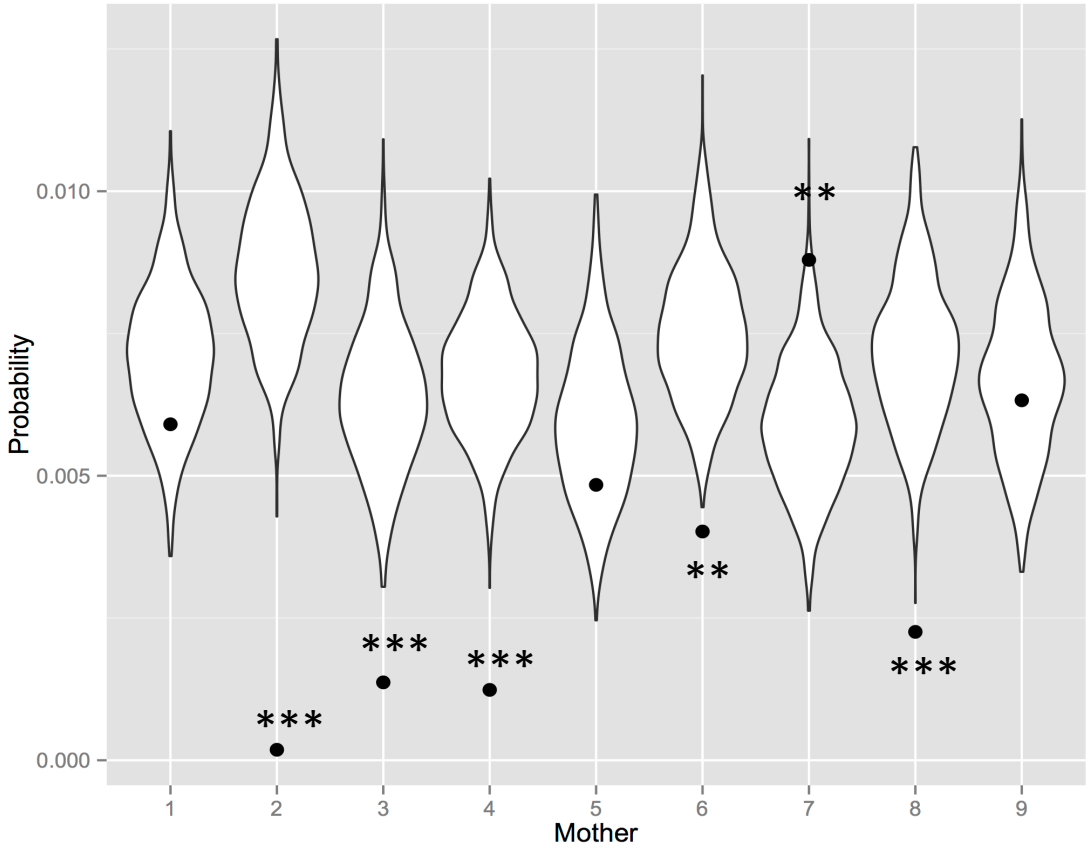
The effect does not mirror the pattern in adult language

(Adult) languages tend to avoid adjacent repetition of identical/similar sounds (Greenberg, 1950; McCarthy, 1986. cf. Obligatory Contour Principle).



Occurrence likelihood of identical segments (Monaghan & Zuidema, 2015)

Or in infant-directed speech



Observed (black dot) vs. simulated (violin plot) probability of adjacent syllable repetition in infant-directed English (Ota & Skarabela, 2016)

Part 3

Does the learners' linguistic environment accommodate phonological biases for lexical learning?

Virtually all speech communities have infant specific lexical items ('baby talk words')

Ferguson (1964, 1977), Skarabela et al. (2015)

bunny

kitty

wee-wee

tummy

night-night

doggy

horsie

choo-choo

dum-dum

Baby-talk words share characteristics aligned with phonological biases in word learning

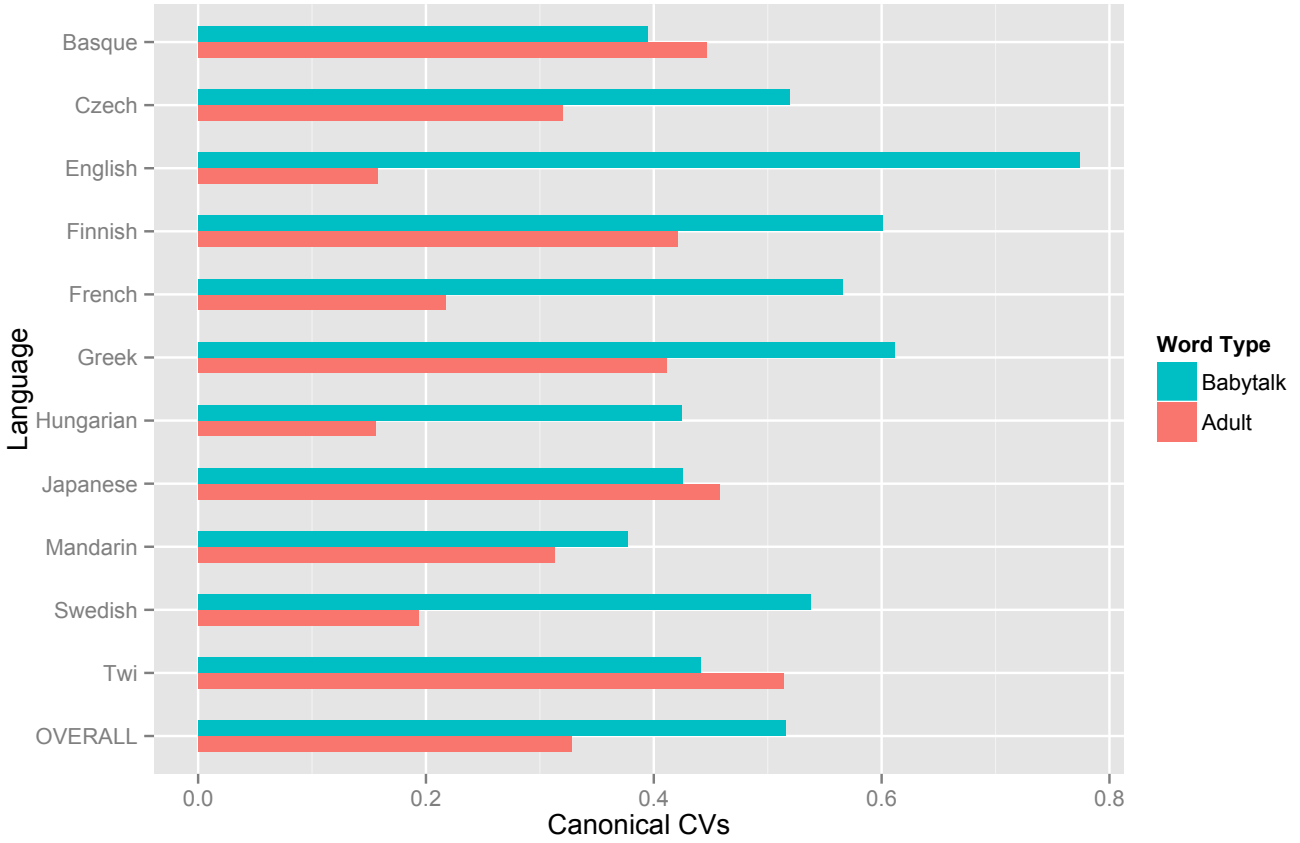
- Reduplication (sound repetition) (Ferguson, 1964, 1977, 1978)
  - Full reduplication: *boo-boo, wee-wee, night-night*
  - Partial reduplication: *daddy, tick-tock, teenie-weenie*
- Canonical CV structure (MacNeilage & Davis, 2000)
  - Stop + Vowel: cf. *tummy* <-> *stomach*, *bunny* <-> *rabbit*
- Predominant prosodic structure (Jusczyk, 1997)
  - 'σσ (English): *DOggy*, *PIGgy*, *BIRDie*, *WEEwee*, *NIGHTnight*
- Diminutives (Berko Gleason et al., 1994; Kempe et al., 2007)
  - *-/i/* (English): *doggy, piggy, birdie, kitty, tummy, bunny*
- Onomatopoeia and sound-symbolism (Laing, 2014)
  - *moo, woof, tick-tock, choo-choo*

Reduplicated words are more likely to be segmented and learned (Ota & Skarabela, 2016, 2018)

'Sleep/nap'

Language	Adult word	Baby-talk word
Arabic (Lebanese)	*noːm; ʇitˈsˤatːˤaħ*	*dodo*
Basque	*lo egin*	*lolo*
Cree	*nipaah*	*miimii; meme*
Czech	*spát*	*nini*
Finnish	*nukkua*	*paapaa*
French	*dormir*	*dodo*
Greek	*kimame*	*nani*
Hungarian	*alvás*	*csicsi*
Japanese	*neru*	*nenne*
Swedish	*sova*	*nanna*

Canonical CV syllables (oral/nasal stop + V) are easier to articulate (MacNeilage & Davies, 2000)



Proportion of canonical CVs in baby-talk words (e.g., *tummy*) and their adult equivalents (e.g., *stomach*) (Skarabela et al., 2015)

Diminutives can facilitate word segmentation

Diminutives in baby-talk words

Arabic: -*o*

Comanche: -ci

Czech: -*ček*, -*ka*

Dutch: -*je*

English: i/y

German: -*chen*

Gilyak: -*k/q*

Japanese: *o*-

Marathi: -*[k]ula*

Russian: -*ik*

Spanish: -*ito/ita*

Turkish: -*cik*

Diminutives can facilitate word segmentation

Marking of word edges (Kempe et al., 2005)

Incidental-learning of L2 words by adults improves with diminutives (Kempe et al., 2005, 2007)

Regularization of prosodic patterns (Echols et al., 1997; Juszyk, 1997)

- *DOG, umBRElla, HANDkerchief* > DOGgy, KITty, BROLly, HANKy
- teLEfono, maMA > *telefoNIto, maMIta*
- English-learning infants' use of Sw pattern in segmentation

Onomatopoeia may aid word learning

Articulatory facilitation

- Admission of 'wild' sounds (e.g., buzzing sound for 'bee') (Laing, 2014)
- Higher proportion of easy-to-articulate sounds (Massaro & Perlman, 2017)

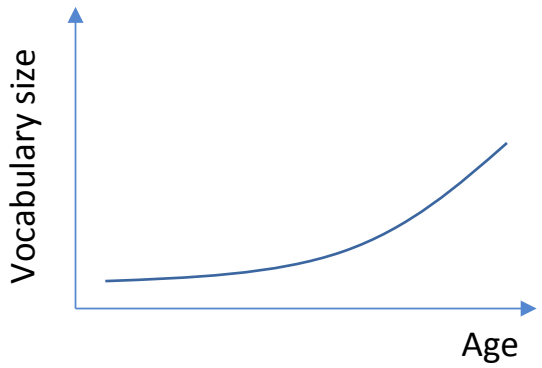
Bootstrapping in form-meaning mapping

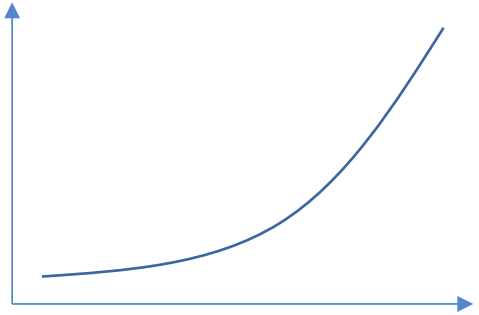
- Referential insight (Imai & Kita, 2014)
- Bias for non-arbitrary sound-meaning mappings (Dingemanse et al, 2015; Maurer et al., 2006)

Baby-talk words may function as anchor points for further word segmentation

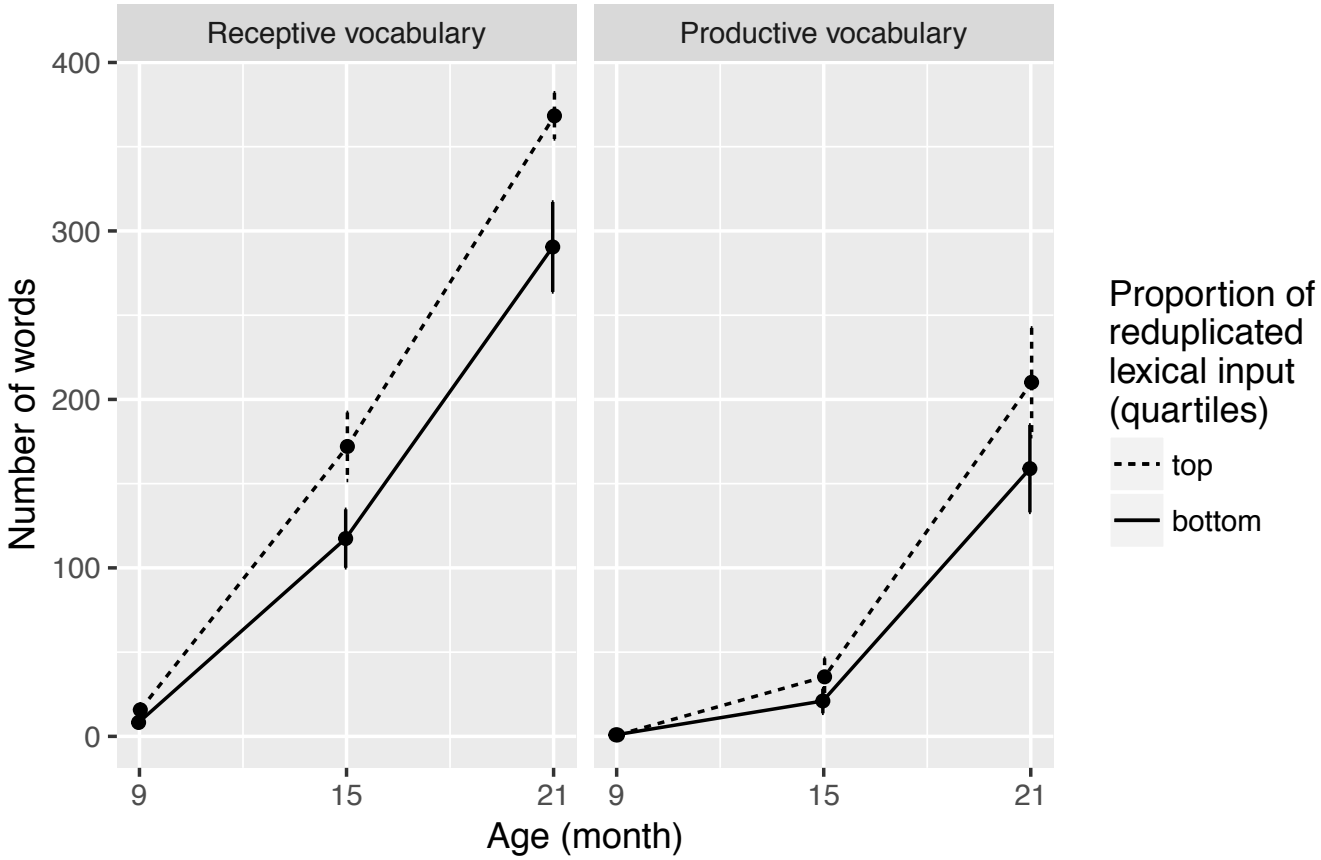
Lorem ipsum dolor sit amet, consectetur adipiscing elit. Praesent neenee nec ipsum at facilisis. Sed et vestibulum dolor.

Lorem ipsum neenee sit amet, consectetur adipiscing elit. Praesent neenee nec ipsum at facilisis. Sed et vestibulum neenee.



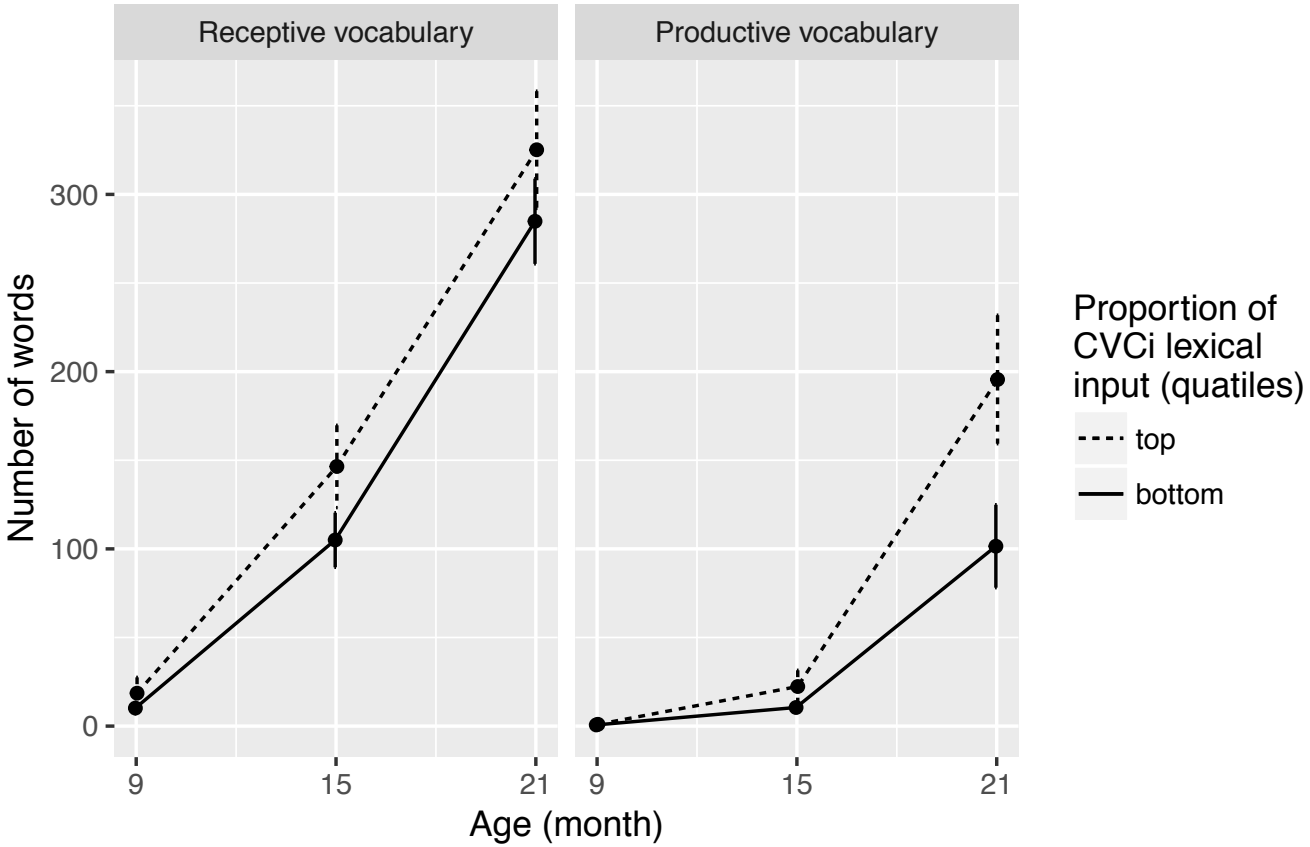


Infants who hear a higher proportion of reduplicated words learn more words



Size of vocabulary at 9, 15 and 21 months depending on the proportion of input words with reduplication (N = 47) (Ota et al., 2018)

Infants who hear a higher proportion of diminutivized words learn more words



Size of vocabulary at 9, 15 and 21 months depending on the proportion of input words with ‘diminutives’ (N = 47) (Ota et al., 2018)

Summary

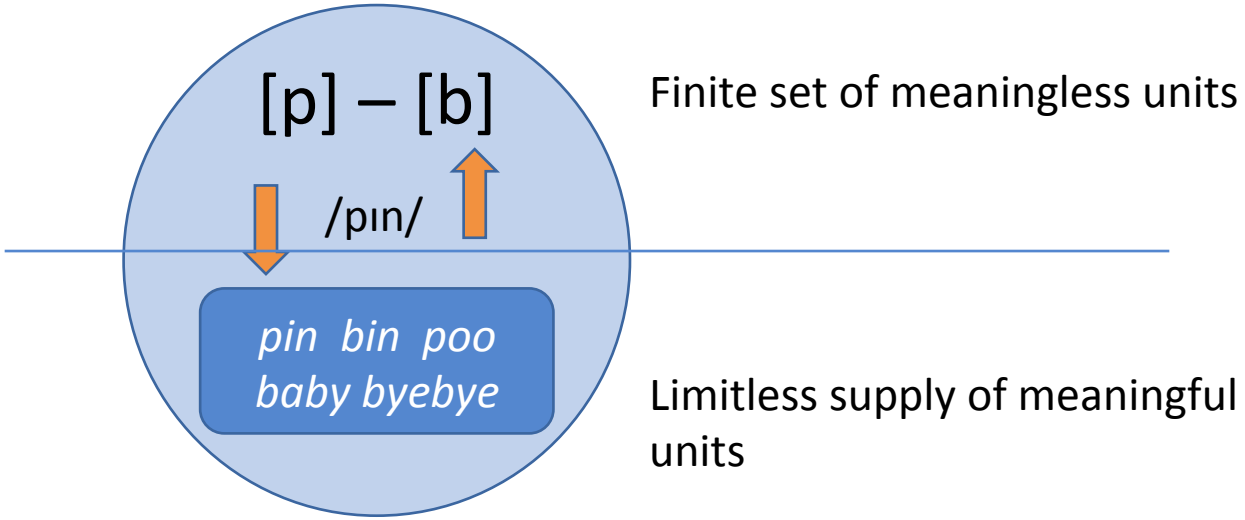
- There are individual phonological biases that lead to selective learning of words. This may be due to an articulatory filter.
- There are also general phonological biases in word learning, such as sound repetition and canonical CV structure.
- Some features of baby-talk words can be characterized as accommodation of phonological biases in word learning. These features may bootstrap early word learning.

General summary

What is the relationship between early phonological and lexical learning?

Our starting point

Duality of patterning



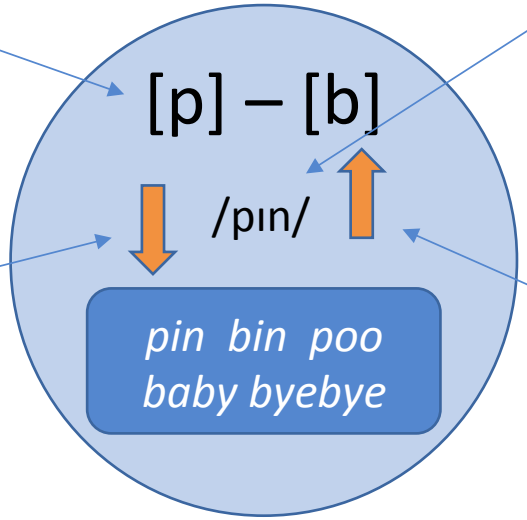
Finite set of meaningless units

Limitless supply of meaningful units

The questions

Can sounds be learned independent of words?

How do phonological representations develop?

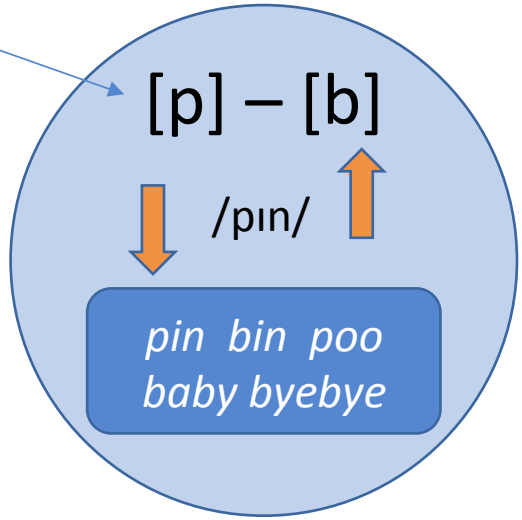


How does phonology influence word learning?

How do lexical factors affect phonological development?

The questions

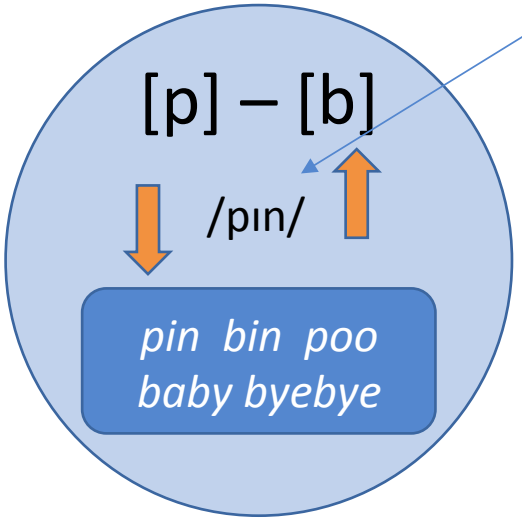
Can sounds be learned independent of words?



Phonetic learning needs to be augmented by lexical information (of *non-minimal* pairs) to successfully derive sound categories.

The questions

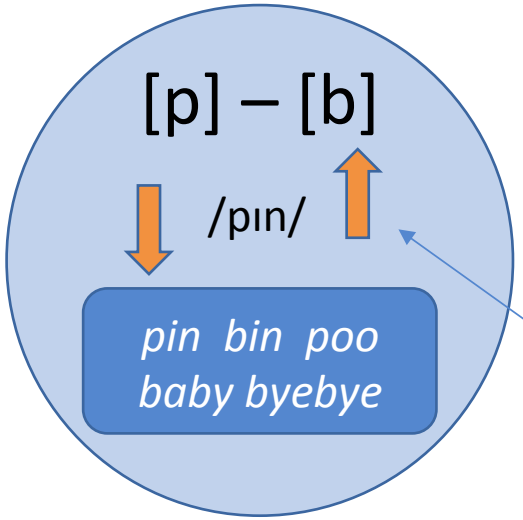
Acquisition of phonological representations is not driven by the need to maintain lexical contrasts, but by encoding detailed phonetic information through repeated exposure.



How do phonological representations develop?

The questions

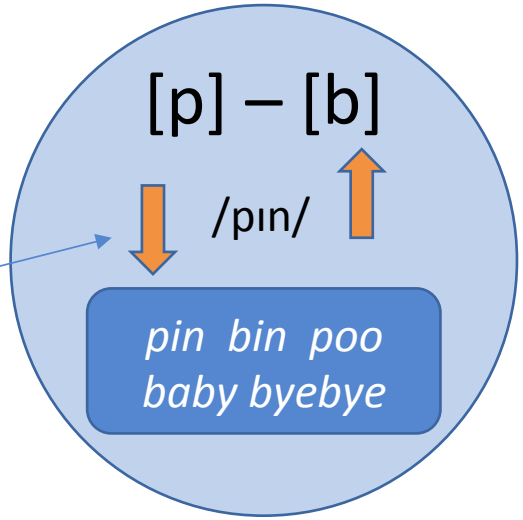
Certain aspects of phonological development are sensitive to the composition of the lexicon and also lexically variable. The sources of the variability include lexical frequency, age of acquisition (of words) and phonological neighborhood density.



How do lexical factors affect phonological development?

The questions

How does phonology influence word learning?



There are individual and general phonological biases in lexical learning. The environment of the learner appears to accommodate these biases.