# VOICE OR VOT?

FKA: SYSTEMATIC AND INCIDENTAL SOUND ERRORS IN CHILD LANGUAGE PRODUCTIONS

Clara Levelt, Leiden University

Word forms in early child language

- Word forms in child language productions often deviate from their target adult forms.
- Deviations can be systematic or variable
- The big question: what is the source of these deviations in production?

[bat]

[pat]

baard /bart/
(beard)
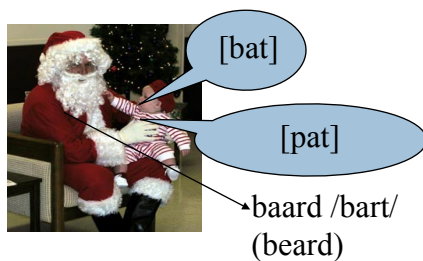
Sources for deviations

- The lexical representation is incomplete
- The grammar interprets the lexical representation in an overly constrained way
- The phonetic encoding is flawed
- There are physical inabilities to execute the phonetic-articulatory plan
- Timing/planning problems can arise at different levels
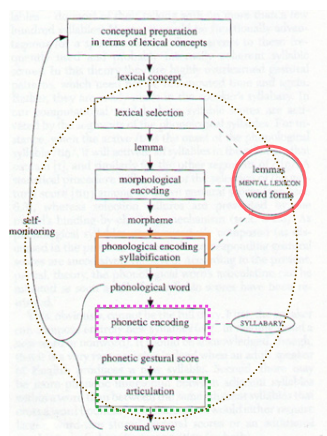
Hypotheses

- [-representation]: default interpretations of representational gaps: *systematic deviations*
- [-grammar]: *systematic deviation patterns*, cross-linguistic variability possible, categorical deviations from target forms
  (also: regular correct productions!)
- [-phonetic]: *non-categorical deviation patterns* (p.e. durational aspects, timing errors)
- [-motor]: *systematic* inability to produce certain sound(combination)s.
- [-timing/planning]: *variable deviations*

Dutch stops

- Dutch is a pre-voicing language
  - voiced stops (b d): voicing lead of -4 ms
  - voiceless stops (p t): short lag VOT between 0-25 ms

- Table below from Kager et al. (2007)

	Voicing Lead	Short Lag VOT	Long Lag VOT
Dutch	-4 ms: b, d	0-25 ms: p, t	
German		16 ms: b, d	51 ms: p, t
English		32 ms: b, d	59 ms: p, t

Table 1: *VOT in Dutch, German and English*

Fikkert-Kager group results

- Predominantly devoicing errors:

Examples of laryngeal errors in Robin's utterances
|				
a.	douche	'shower'	tus	(1;10.21)
b.	dier	'animal'	tiɹ	(1;10.21)
c.	beer	'bear'	pi	(1;7.13)
d.	bal	'ball'	pɑl	(1;7.13)
e.	baby	'baby'	pipi	(1;8.10)
f.	thuis	'home'	dœys	(1;5.10)

- Error pattern is independent of Place of Articulation

Fikkert-Kager group results

☐ Input frequency does not seem to play a role:

	Labials			Alveolars		
	p		b	t		d
types	151 (40.7%)		220 (59.3%)	104 (41.3%)		148 (58.7%)
tokens	1492 (30.9%)		3342 (69.1%)	1481 (13.6%)		9389 (86.4%)

Table 7: *Distribution of voicing in child-directed speech from van de Weijer corpus*
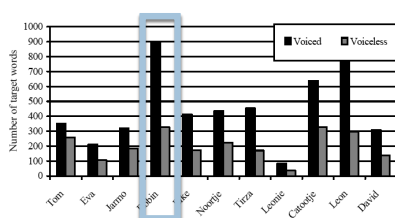
Dutch stops in acquisition

☐ Contrast that needs to be acquired:
  ☐ phonology: [+voice] vs [-voice]
    ▪ representation: monovalent [voice] or binary [±voice]
  ☐ phonetically: -VOT vs. +VOT
☐ What is the source of voicing errors in Dutch (onset) stops?
  ☐ representation (phonology)
    ▪ initially no [voice] = default –voice
  ☐ phonetic encoding
  ☐ articulatory effort (Kager et al.)
    ▪ initial preference for short lag VOT (in Dutch: voiceless stops)

Variability in stop-voicing

- Representational account *seems* to predict a categorical development:
  - initial devoicing, across the board
  - acquisition of [voice]: voicing, across the board for [voice] segments
- However: variable productions

Session	Robin.1990-03-21
Inventory	
Result format	All
Result	*Count*
b ↔ b	4
b ↔ p	6
b ↔ t	1

Session	Robin.1990-03-21
Inventory	
Result format	All
Result	*Count*
d ↔ d	6
d ↔ d̪	1
d ↔ t	3

Effort today

- Get across the inkling: representation is OK but phonetic encoding needs to be worked out
- Data:
  - VOT measurements of longitudinal data
    - voiced and voiceless stops produced by Robin (1;5 – 2;5)
  - Production experiment
    - 9 two-year old Dutch children

Petit Grand Dessert

Results VOT measurements /b/

Results VOT measurements /d/

Results VOT measurements /p/

Results VOT measurements /b/

Results VOT experiment

VOT measurements: method

- Phonex searches:
- measured/ measurable:
  - 89 d
  - 65 t
  - 125 b
  - 68 p

Aligned Phones

Target phonex b:Onset

Actual phonex {}:Onset

☐ Search actual before target

Example

Phonex Help

Group Position

Word Position

Search by word: ☑

Singleton words: ☑ (groups with only one word)

Multiple words: ☑ Initial ☑ Medial ☑ Final

Syllable Position and Stress
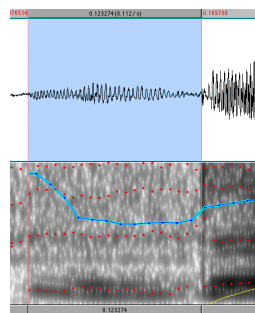
Search by syllable: ☑

Singleton syllables: ☑ (words with only one syllable)

Multiple syllables: ☑ Initial ☑ Medial ☑ Final

Syllable stress: ☑ Primary ☐ Secondary ☐ Unstressed

VOT measurements: method

- VOT measurements in PRAAT:
 - from burst to first zero-crossing of periodicity

---
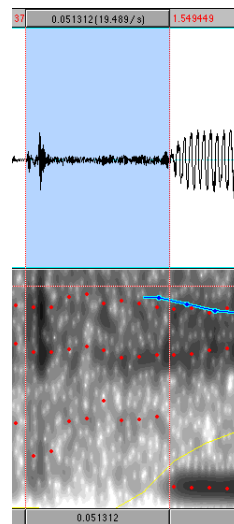
/b/ /p/ targets

Robin		b=b	b=p
81189	baby	-1.06	
	boing	-3.54	
	bah2		1.29
221189	bah	-1.9	
71289	boom	-5.79	
71289	boom	-14.99	
	boom2	-1.26	
	bal	-23.68	
201289	baby	-11.26	
	boom1	-17.53	
	boom2	-2.33	
	bad		0.69
	bal	-2.24	
	boing	-0.35	
100190	boing2	-18.92	
	beer	-13.26	
	beer2		0.1
	bad		0.47
240190	boom	-6.22	
	bal		2.28
	beer		1.54
	beer		0.56
	boom	-7.87	
	boven1	-1.65	
	boven2	-3.91	
	buiten		2.14
	bus		1.75

| Robin p		p=p	p=b
81189	pop	1.68	
100190	pappa48		-1.52
	pappa60	1.21	
240190	paard3	0.61	
	paard29		
	paard35	0.12	
	pap34	0.81	
210290	paard24	0.24	
	paard25		-1.9
	paard31	0.74	
	pet71	1.06	
	pop74	0.59	
70390	paard4	0.76	
	pappa22		-7.04
	paard50		-3.01
	speen76		-1.45
	speen75		-1.92
210390	paard9		-0.26
	plons44		-3.84
	speen59	0.66	
	pijn66	0.97	
	speen76		-1.01
180490	pappa		-4.84
	poppebuggy	0.91	
	plukken	2.47	
	paardebloeme	0.34	
	paardebloeme	2.89	

/b/ /p/ targets

- plot of the mean VOT values over time
  - mean VOT values are significantly different: p<.002
  - productions of /b/ seem to always have negative VOTs
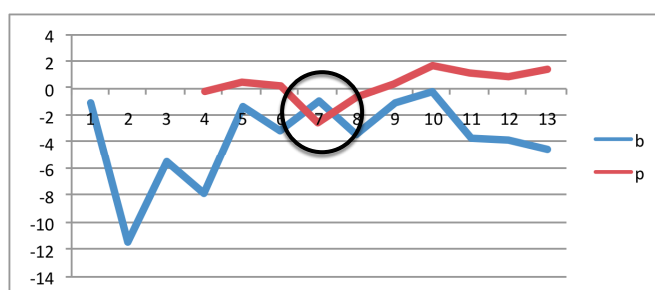
/b/ targets

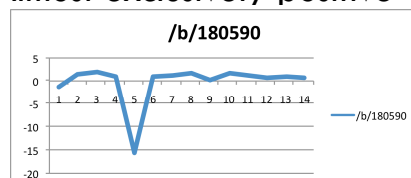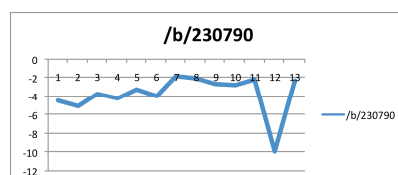- Both positive and negative values in session 070390
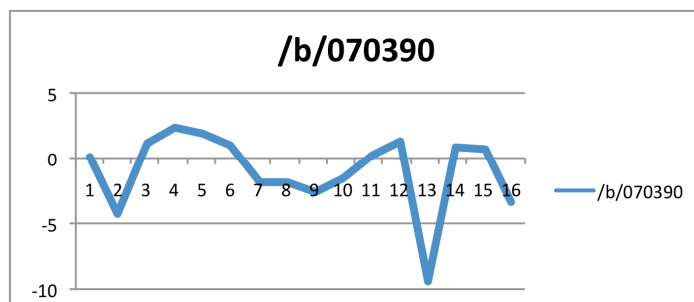
/b/ targets
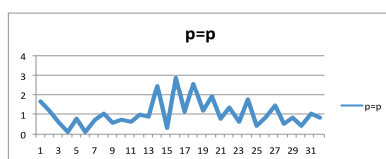
- Almost exclusively positive values in session 180590

/b/180590

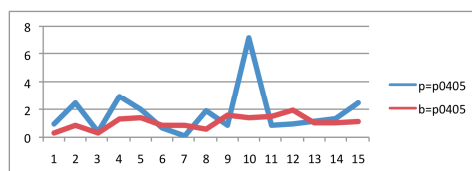- Exclusively negative values from session 230790 onwards

/b/230790

/p/ targets

- +VOT value increases, then decreases, no SD over time

p=p

- Increase occurs at the high point of b=p period. Suggestive, but not significant at this point

/b/ /p/ targets

- Could /b/ and /p/ be represented differently?
  - Are the [p]s from /b/s phonetically different from the [p]s from /p/s?
  - Yes, in the first 8 sessions (13 p=p cases): $p < .05$
    - No in the next two sessions, no more b=p in last three sessions
  - Interestingly/puzzling:
    in first 8 sessions
    - b=p mean VOT: 1.5 ms
    - p=p mean VOT: 0.79 ms
  - At high-point of b=p:
    - b=p mean VOT: 1.05 ms
   - p=p mean VOT: 1.7 ms

Conclusions /b/ /p/

- Target /b/ shows U-shaped development:
  - sessions 1-4: -VOT
  - sessions 4-9: variable VOT
  - session 10: +VOT
  - sessions 11-13: -VOT
- Means of VOT values for target /b/ and /p/ are significantly different from each other
- VOT values for b=p and p=p significantly different in first 8 sessions, switch in VOT at high-point of b=p.
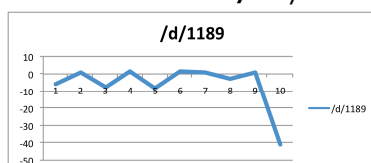
Different representations?

/d/ /t/ targets

- plot of mean VOT values over time
- mean values differ significantly: p< .05
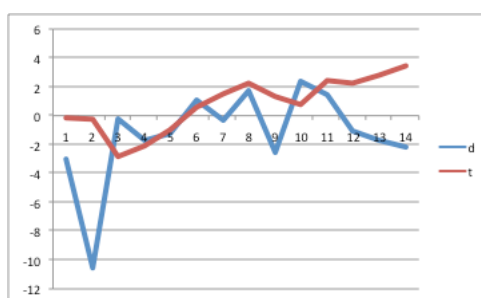
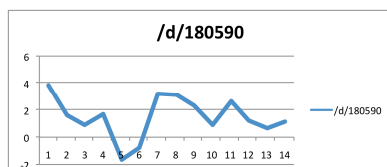

/d/ targets

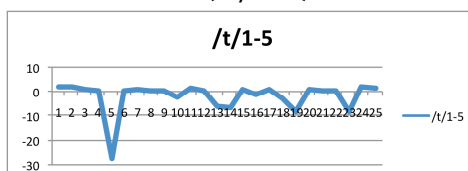- more variability +/- VOT than /b/ in initial sessions

- like for /b/, almost exclusively positive – but highly variable – VOT values in session 180590
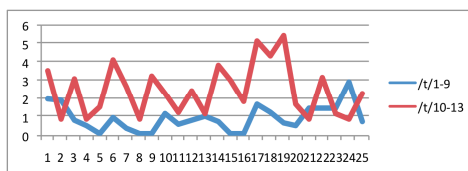
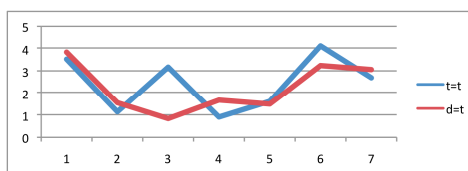# /t/ targets

- Negative VOT values (8/25) in first 5 sessions

/t/1-5

10
0
-10
-20
-30

1 2 3 4 5 6 7 8 9 10111213141516171819202122232425 — /t/1-5

- Increase in VOT value over time (significantly different first vs last sessions $p<.001$)

6
5
4
3
2
1
0

1 2 3 4 5 6 7 8 9 10111213141516171819202122232425

— /t/1-9
— /t/10-13

Conclusions /d/ /t/

- More variability in +/- VOT for /d/ throughout
- No significant difference between d=t and t=t anywhere
- No specific increase in VOT of [t] at high-point of d=t

5
4
3
2
1
0

1   2   3   4   5   6   7

— t=t
— d=t

Conclusions /d/ /t/

- Different representations for /d/ and /t/?
  - less clear, but still: different mean VOT values overall
- Possible account
  - /d/ targets are in 53/89 cases (60%) demonstratives, only 4/89 nouns!
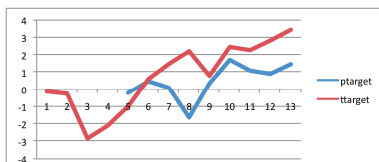 - +/- VOT variability can be tolerated in these items

Comparison /b/ /d/

- Development of /b/ and /d/ (mean VOT values) is pretty parallel!
- /b/ is more "voiced" than /d/ (tolerance account)
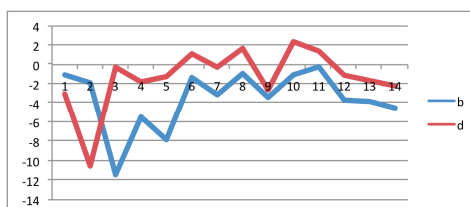
Comparison /p/ /t/

- Again: pretty parallel development!

- Significantly different VOT values final sessions (p<.002)
  - "Leiden" accent
    - tekenen (to draw)
    - kas'teel (castle)
    - thuis (home)

Tentative conclusion

- There seems to be a representational difference between voiced and voiceless stops
  - voiced stops show parallel development
  - voiceless stops show different parallel development
- Initial negative VOTs for /b/ /d/ targets imply that articulation is not the problem,
- Phonetic encoding of phonological difference has to be figured out

VOT experiment: rationale

- If phonological representation is OK...
  - voiced stops have a [voice] representation
- but phonetic encoding is iffy...
- maybe, better performance with voiced consonants can be provoked in children's productions

Base-Berk & Golderick 2009 study

- Looked at influence of neighborhood density on production of VOT in adults
- Experiment 2: participant A has to tell participant B to click on one of three words that appear on screen:

Condition			
Context Condition	cod	god	yell
No Context Condition	cod	lamp	yell
No Competitor Condition	cop	lamp	yell

- English: longer VOTs were expected

Results experiment 2

- VOT of words with minimal pair neighbors is significantly longer
- VOT increases even more if competitor is presented in context condition:
 context > no context > no competitor
  (82.5)   (77.4) (72.4)
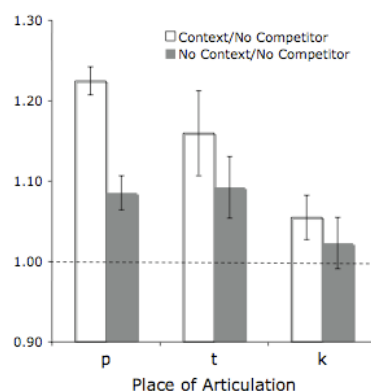
Increase seems very small…

- VOT increases with the same proportion across the three different places of articulation :
  - 68.9 ms versus 65.7 for /p/
  - 84.3 ms versus 80.3 ms for /t/
 - 95.5 ms versus 90.6 ms for /k/

Method

- Picture naming experiment, presented in PPT
- 6 different minimal pairs /b/-/p/ = b+ condition
- 3 different minimal pairs /d/- /t/ (x2) = d+ condition
- 6 /b/ + 3 /d/ (x2) "no competitor" words
- context condition (b+)
- no-context condition (b-)
- practice session with individual pictures
- 9 two-year old children (4F, 5M)
- digital recording, files stored in PHON
- VOT analysis in PRAAT
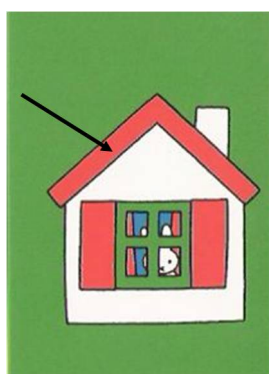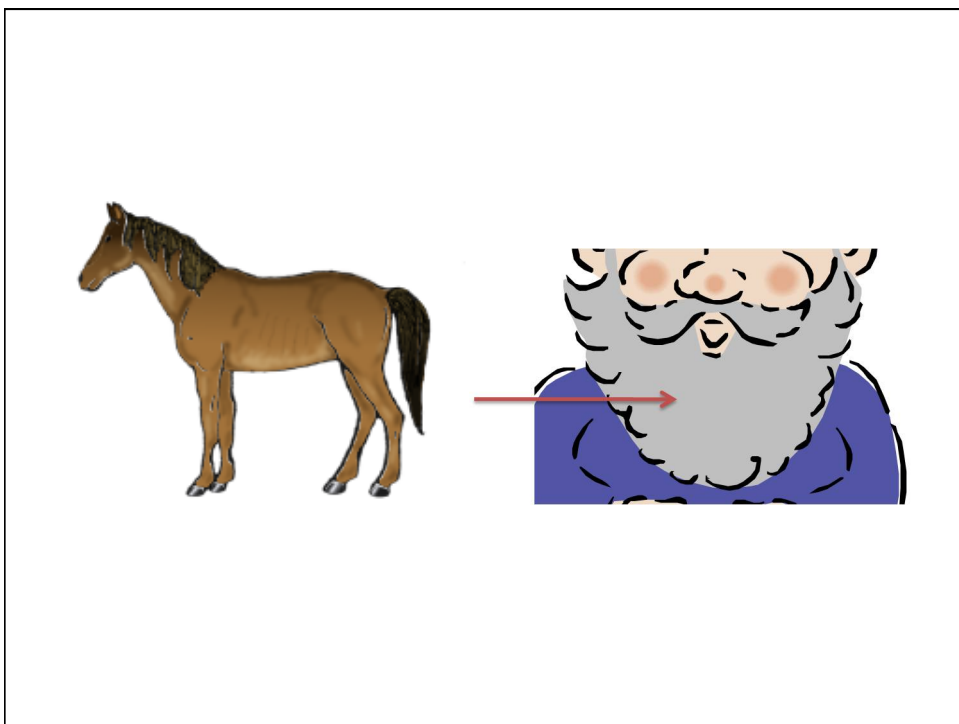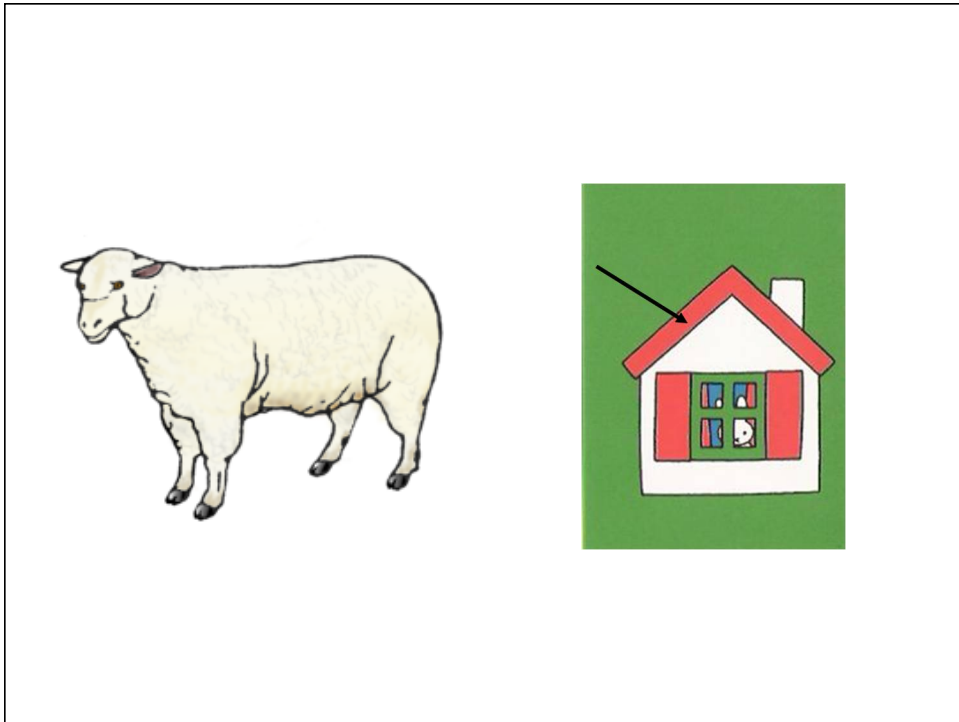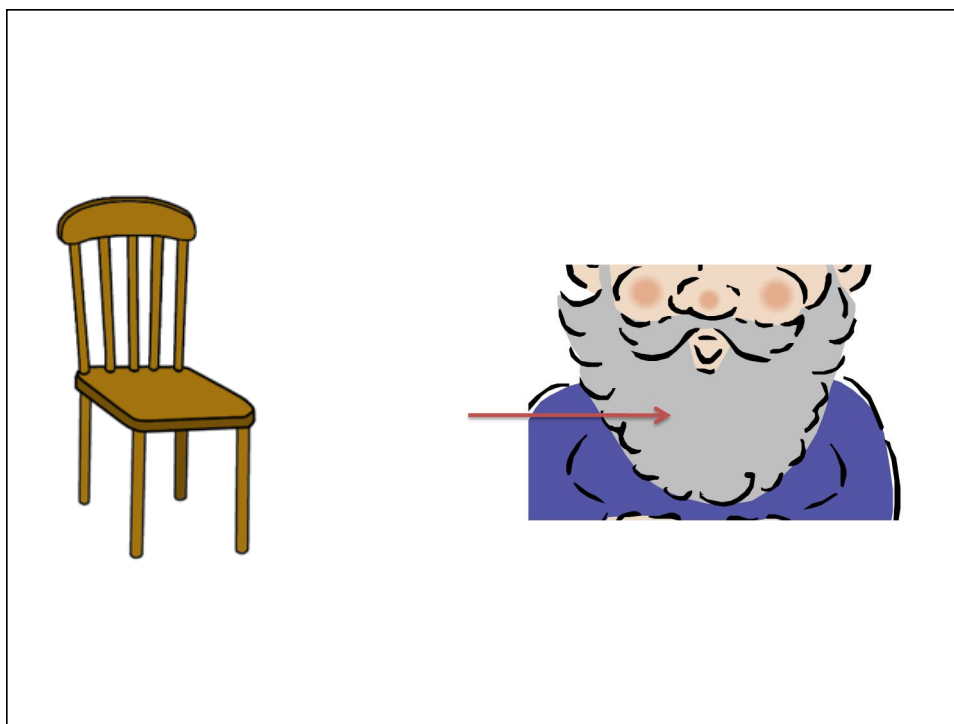
Results

- 35 b+/d+ productions
- 29 b-/d- productions
- No significant difference between b+ and b- VOT values (b+ mean VOT -0.59, b- mean VOT -0.38)
- Significant difference between d+ and d- VOT values (p<.05, d+ mean VOT -1.33, d- mean VOT 3.2)

Discussion

- □ Why different outcomes for /b/ and /d/?
  - ■ measured all productions of /b/ and /p/ targets:
    - ■ significant VOT difference
  - ■ measured all productions of /d/ and /t/ targets
    - ■ no significant VOT difference!

Conclusions experiment

- □ If significant VOT difference between target voiced and voiceless plosives is present:
  - □ no significant VOT improvement can be provoked for target voiced plosives
- □ If no significant VOT difference is present:
  - □ VOT improvement can be provoked
- □ Implication: [voice] must be present in representation (in both situations)
  - □ otherwise no improvement/non significant difference would be expected

General conclusions

- Both sets of data seem to show the same thing:
  - different behavior of labial vs coronal plosives
    - possibly because initial /d/ words are more tolerant of VOT variability
  - an inkling that the phonological representation is ok: [voice] is present in the representation
    - significant VOT differences throughout developmental period, between (target) voiced and voiceless plosives, despite their being "incorrect" and/or variable, in Robin's case
    - significant VOT difference present OR significant VOT improvement possible, in experiment
- Source of voicing errors = phonetic encoding

THANK YOU!!

- and
  - Yvan!
  - The rest of the PHON team
  - the participating kids and their parents
  - the students who did the production experiment:
    - Andrea Spruijt, Conny van Paridon, Thijs Nielen
  - my PhD students, for discussions, coffee & tea
    - Marijn van 't Veer, Margarita Gulian, Kathrin Linke, Sita ter Haar
  - the funding agency

NWO
Nederlandse Organisatie voor Wetenschappelijk Onderzoek