The limits of morphological productivity: L1 acquisition of verbal alternations in Russian

K.Pertsova (UNC-Chapel Hill) pertsova@unc.edu
A. Lopukhina (HSE, Moscow) nastya.lopukhina@gmail.com
O. Vedenina (HSE, Moscow) olgetta0511@gmail.com

FASL 2021

How are morpho-phonological alternations learned?

- Some properties of morpho-phonological alternations:
  - irregularity and idiosyncrasy (lexical and morpheme-specific patterns)
  - phonological opacity
  - semi-productivity
  - paradigmatic and lexical effects (e.g., paradigm uniformity, Steriade (2000))

Russian Verbal Inflection

- Not well studied from the point of view of acquisition in general (but see Svistunova, 2008).
- Several morpho-phonological alternations involving consonants, vowels, stress and suffix-alternations, that break up verbs into many inflectional classes

inf.	non-past.tense sg	CLASS
der**za-**t’	derz**aj**-u, derz**aj**-esh, derz**aj**-et, …	a-aj or ø-j
re**z**-a-t’	re**ʒ**-u, re**ʒ**-esh, re**ʒ**-et, …	a-ø; z-**ʒ** in all pres.
sra**z**-i-t’	sra**ʒ**-u, sraz-ish, sraz-it, …	i-ø; z-**ʒ** in 1sg

- Some verbs in classes ‘i’/‘e’ are famously “defective” in that they have a **gap** in 1sg, the cell in the paradigm that is most prone to morpho-phonological alternations

on pobedit *“he will win”*

*ja *?pobezhu *?pobed’u *?pobezhdu* “I will win”

General Questions

- Specific vs. general debate
  - At what point is **the generalization leap** made?
    - Lexeme-based —> segment-based —> class based (specific to general)
    - Assume the majority pattern as default, gradually relax this assumption
    - General and specific rules co-exist from the beginning

- Is “defectiveness” a natural byproduct of low-productivity or a learned phenomenon? How is it learned by children?
  - Learning from negative evidence that some lexemes are defective (lacking specific forms) and extending defectiveness by analogy (Daland et.al. 2007)
  - Defectiveness = low grammaticality/acceptability of all possible outcomes (Albright 2009, Pertsova, 2016)

Outline

- Background on verbal alternations
- Experiment design
- Results and implications for the specific vs. general question
- Results and implications for the defectiveness

Background: alternations

<u>Consonantal alternations in i-class and e-class verbs</u>

Labial altern.: applies to <u>all labials</u> in the same way, no change of stem

f,p,b,v,m + l Ex. lúbʲ-itʲ lub-lʲ-ú “to love”

Dental alternations: <u>no single rule</u>, change of stem

d ~ ʐ	brédʲ-itʲ	bréʐ-u	“be delirious”
z ~ ʐ	vózʲ-itʲ	voʐ-ú	“drive”
s ~ ʂ	kósʲ-itʲ	koʂ-ú	“to cut grass”
st ~ ç	rastʲ-itʲ	raç -u	“to cultivate”
t ~ tʃ	letʲ-etʲ	letʃ-ú	“fly”
t ~ ç	sokratʲ-ítʲ	sokraç-ú	“to diminish”

Background: gaps

- Paradigm gaps in 1sg of verbs with dental alternations
  - Appear in predominantly low-frequency verbs of -i/-e class (II conjugation) with dental alternations
  - No semantic or apparent phonological reasons for gaps

Non-past tense

sbondit' "to swipe, steal"	sg	pl
1p.	??	sbond^{j}-im
2p.	sbond^{j}-ish	sbond^{j}-ite
3p.	sbond^{j}-it	sbond^{j}-at

Defectiveness as a predictable grammatical phenomenon

- Lexemes that have “remote bases” (Steriade, 2008, Steriade & Stanton, 2020) with the same alternation do not have gaps (Pertsova, 2016; Pertsova & Kuznetzova, 2017)

“to plant” perf.

posadit’	sg	pl
1p.	posaʒ-u	posad’-im
2p.	posad*j*-ish	posad’-ite
3p.	posad*j*-it	posad’-at

“to plant” imperf.

saʒat’	sg	pl
1p.	saʒ-aj-u	saʒ-aj-em
2p.	saʒ-aj-esh	saʒ-aj-ete
3p.	saʒ-aj-et	saʒ-aj-ut

“one that has been planted” (PPP)

posaʒ-enn-ij “ m. …	
Nom.	posaʒ-enn-yj
Gen.	posaʒ-enn-ogo
…	…

(a) Practically all verbs with **known gaps** (based on dictionaries) have no related forms with the same alternation (e.g., intransitive verbs don’t have PPPs, denominal verbs have few other verbal deriv.).

(c) Verbs which have no related forms with the same alterations (including novel borrowings) are treated by speakers in a similar way as those that have gaps = higher degree of variation, less certainty.
- Pertsova & Kuznetsova (2017)
- Here we call such verbs “candidate” verbs

Experiment

Experiment: general information

- **Production study**: subjects produce 1sg. and 3sg. of existing and nonce pseudo-words (roughly based on Svistunova, 2008)

- Subjects
  - Children: 38 (3 - 5.5 years old) recruited in a day-care center in Moscow
  - Adults: 30 adults, mostly college students

- 2 sessions (order balanced across subjects)
  - All real before all pseudo

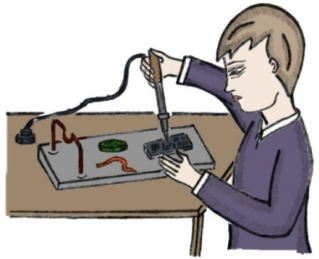
Experiment

- Task:

  E: What is this person doing?
  C: _____ (3sg. target)
  E: “Right, he is rolling (a barrel)”
  E: And if he was talking about himself, he would say: “I am ____” (1sg target)

  For pseudo-verbs: “This is *rafidit’* (inf.)” (the rest is the same as above).

Pictures come from this database http://stimdb.ru/



Experiment: stimuli

Verb type	Example	Num real /Num pseudo
1. no alternations class-i	totʃit' "sharpen" (totʃu) 1sg	8/4
2. dental in 1sg class-i	kra**s**it' "paint" (kraʃu) 1sg	7/15 (after exclusions)
3. labial in 1sg class-i	bombit' "bomb" (bomb**l'**u) 1sg	8/8
4. variation class-a/aj	*pa**x**at' "plow"* *(paʃ- / pax-aj-)*	8/4
5. dental gaps	ubedit' "convince"	8
6. gap candidate	zanozit' "splinter"	8

Groups 1-3 were matched on
- lemma frequency
- segment and syll.length
- aspect

Question 1: specificity

Hypotheses about specificity

- Specific-before-general models
  - Segment-specific patterns are acquired first, before class generalizations are made in the domain of phonotactics (Albright&Hayes, 2003, Adriaans and Kager 2010), in the domain of word learning (Siskind 1996, Pertsova 2007)
- Simultaneous general and specific generalizations
  - Segment- and class-based generalizations are present simultaneously and can affect one another (e.g, constraint-based models in phonology, e.g. MaxEnt (Wilson & Hayes), GMECCS (Moreton et.al.)
- General-before-specific models
  - RULEX (Nosofsky & Palmeri, 2)
    - Begin with the simplest (most general) rules first

Labial vs. dental alternations

Specific rule	# of unique non-defect.roots*		General rule
b → bl	44		
m → ml	38		
v → vl	62		ø → /l/ / [+lab] + 1sg (189)
p → pl	38		
f → fl	7		
d → ʐ	75		? [+dental +voi] → [+postalv +cont] / __ + 1sg (111)
z → ʐ	36		
s → ʂ	38		
t → tʃ	68	* (except t → ç) 16	
st → ç	42		

* Frequency data from Pertsova (2016) based on online list of full paradigms from Zaliznyak's morphological dictionary

Predictions

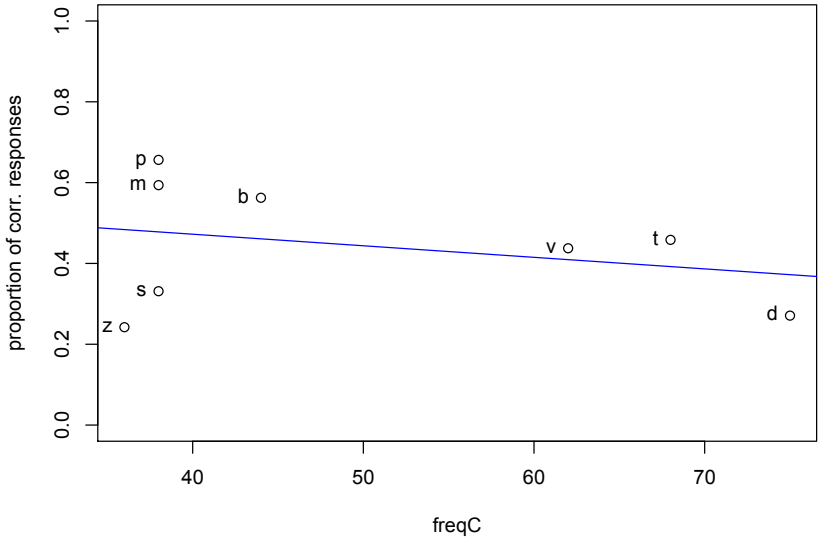
Specific-to-general:

- **Early stage**:
    - performance on each C should be correlated with C type-frequency (within the relevant verb class): assuming more exposure leads to better learning
    - few overgeneralization
- **Later stage:** preference for labial over dental alternations emerges later on

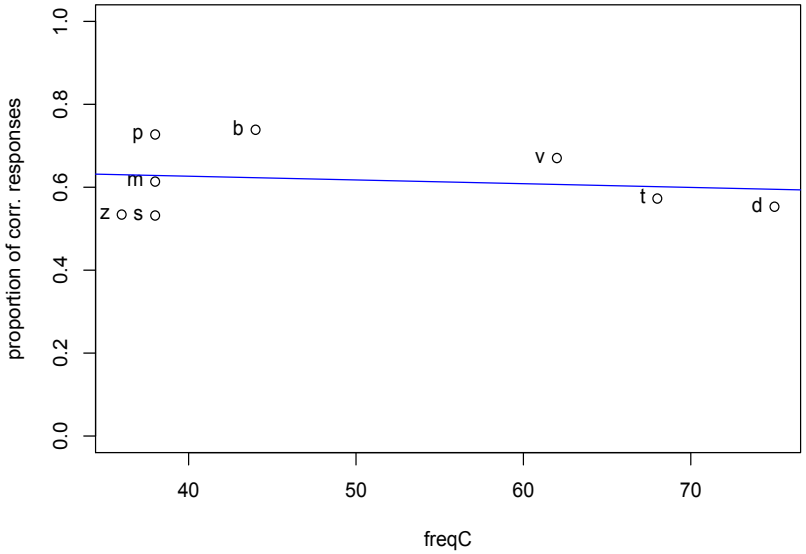
Simultaneous or general-to-specific:

- Advantage of labial over dental alternations from the very beginning
- Within dental alternations:
    - Possibly better performance on “d” and “z” (can be collapsed into one rule)

Results: no correlation btw. C-freq and prop. of correct responses

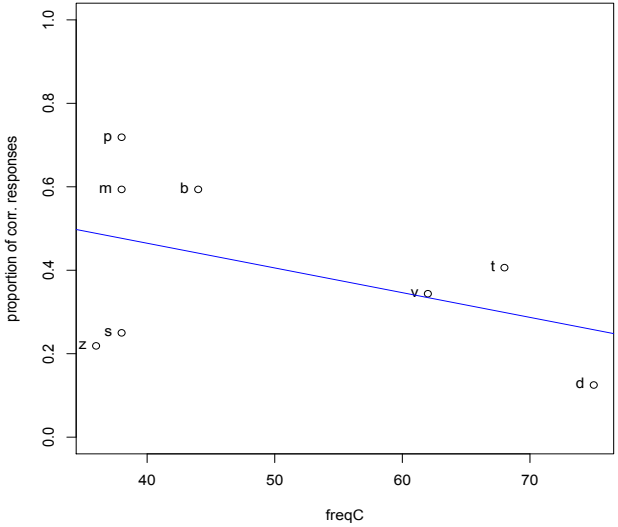


YOUNGER CHILDREN

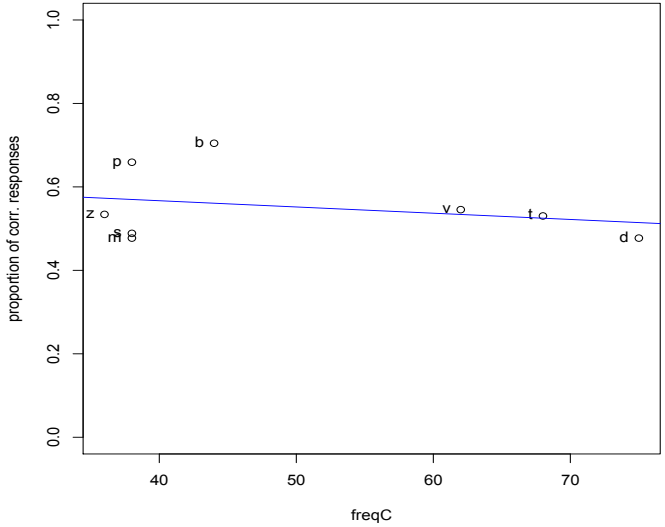


OLDER CHILDREN

Pseudo-words only



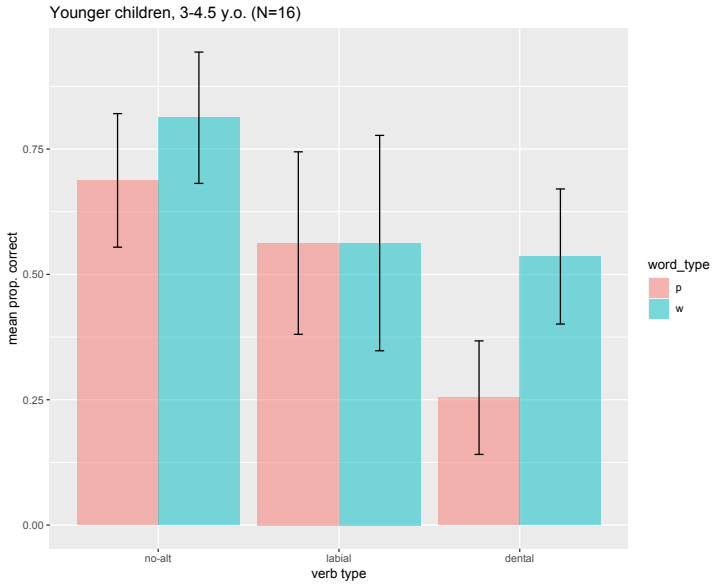
YOUNGER CHILDREN, Pseudo-words only

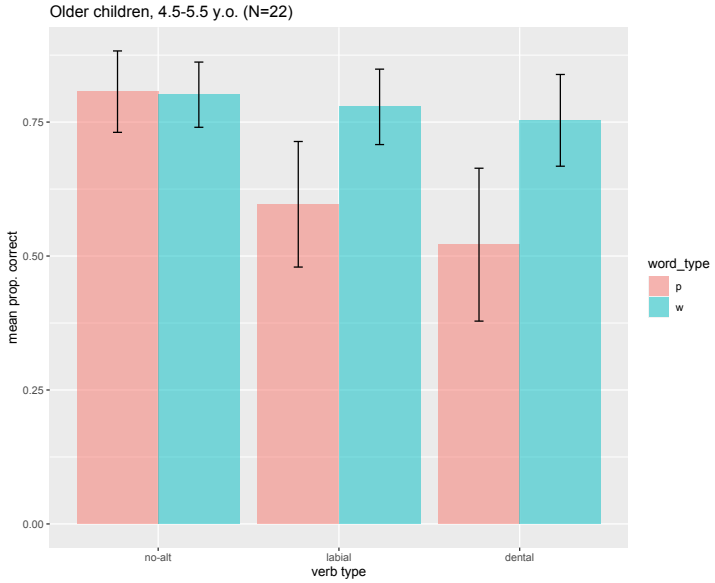


OLDER CHILDREN, Pseudo-words only

Non-significant negative correlation:
R = -0.44 (p=0.26)

Children’s performance on no-alt, labial, dental





No-alt > dental; no-alt > labial; labial > dental

No-alt > dental; no-alt > labial

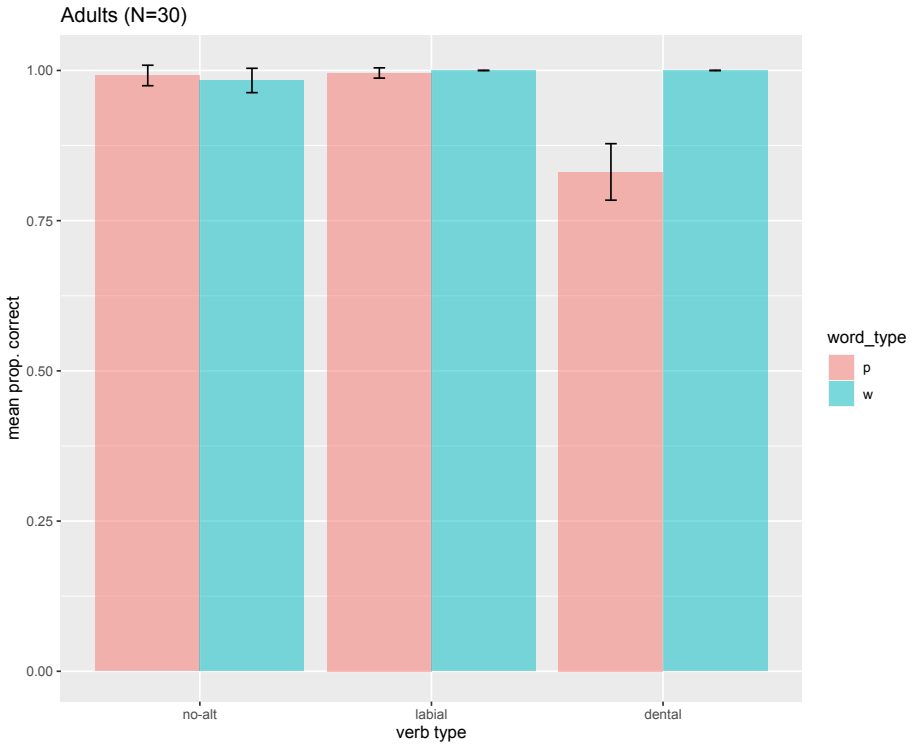
No significant difference for labial vs. dental for real words

By group performance: adults

- Adults are at ceiling on all real verbs
- Lower generalization on dental verbs only

significant effects (post-regression contrast comparisons):

No-alt > dental (1.54, p = 0.004)
Labial > dental (3.85, p = 0.0002)



Overgeneralization

- 2 examples of overapplication of the labial rule:
  - kir-it’ — kirit’l’-u
  - kovyst-at’ — kovyst-l’u

- Several examples of the misapplication of dental alternations within dental stems
  - s ~ ç (instead of s ~ ʂ) kar**s**-et’ — kar**ç**-u
  - t ~ ʂ (instead of t ~ tʃ) pru**t**-it’ — pru**ʂ**-u
  - z ~ ʂ (instead of z ~ ʐ) peren’u**z**-at’ — peren’u**ʂ**-u
  - t ~ ç (instead of t ~ tʃ) prijuʊ**t**-it’ — prijuʊ**ç**-u

Interim Summary

1. No evidence for segment-specific frequency effects for any learners
    - Not consistent with the specific-before general approach (also see Linzen & Gallagher 2014, 2017 for an AGL study that goes against specific-to-general model)
    - But we could be looking at the wrong frequency (only class-internal frequency was considered)

2. Evidence for early effect of class-specific rules: labial alternation is learned better and earlier
    - Significant difference for young children and adults for pseudo-words
    - Dental alternations remain not fully productive for adults: effect of scope or defectiveness?

3. Evidence for lexeme-specific acquisition
    - No differences between labial and dental verbs for real verbs
    - Significant difference between labial and dental pseudo-verbs (labial > dental)

Question 2: defectiveness

Defectiveness as a learned phenomenon

- Daland et.al. (2007) learners learn to match probability distributions of forms within a paradigm. Simulation of a language-change scenario with a Bayesian model.

jest'	**1sg**	**2sg**	**3sg**	**1pl**	**2pl**	**3pl**	**SUM**
D	15	5	45	5	5	25	100
d	0.15	0.05	0.45	0.05	0.05	0.25	1

Table 2. Hypothetical probability distribution

Defective verbs are just verbs whose probability of 1sg is 0.

—> learned as a prob. distribution over different forms

—> analogical force (bias) for forms ending in similar consonants to have similar distributions: effect is centered around ‘d’: strongest analogy to ’t’ and ‘z'

<u>Defectiveness “is orthogonal to productivity”</u>

Alternative: competition with a threshold

Pertsova (2016). Harmonic grammars (similar to MaxEnt)

	Ident(C)-IO	d —> ʐ	Ident(C)-OO$_{rem}$	Harmony
weight	(**5**)	**(10)**	**(5)**	T <10 or below = gap
/gald/-e-t’				
galʐ-u	-1		-1	-10
gald'-u		-1		-10
po-/sad/-i-t’				
☑ posaʐ-u	-1		(saʐ-at’)	-5
posad’-u		-1		-10

Stimuli: Group 2

Verb type	Example	Remote Bases	Num real /Num pseudo
“regular” dental in 1sg class-i	kra**s**it’ “paint” (kraʃu)	kraʃennyj, raskraʃivat’	7/15 (after exclusions)
dental gaps (Ozegov & Shvedova 1992)	ubedit’ “convince” (?) 1sg	—	8
gap candidate	*lazit’ “climb”* (?) 1sg	—	8

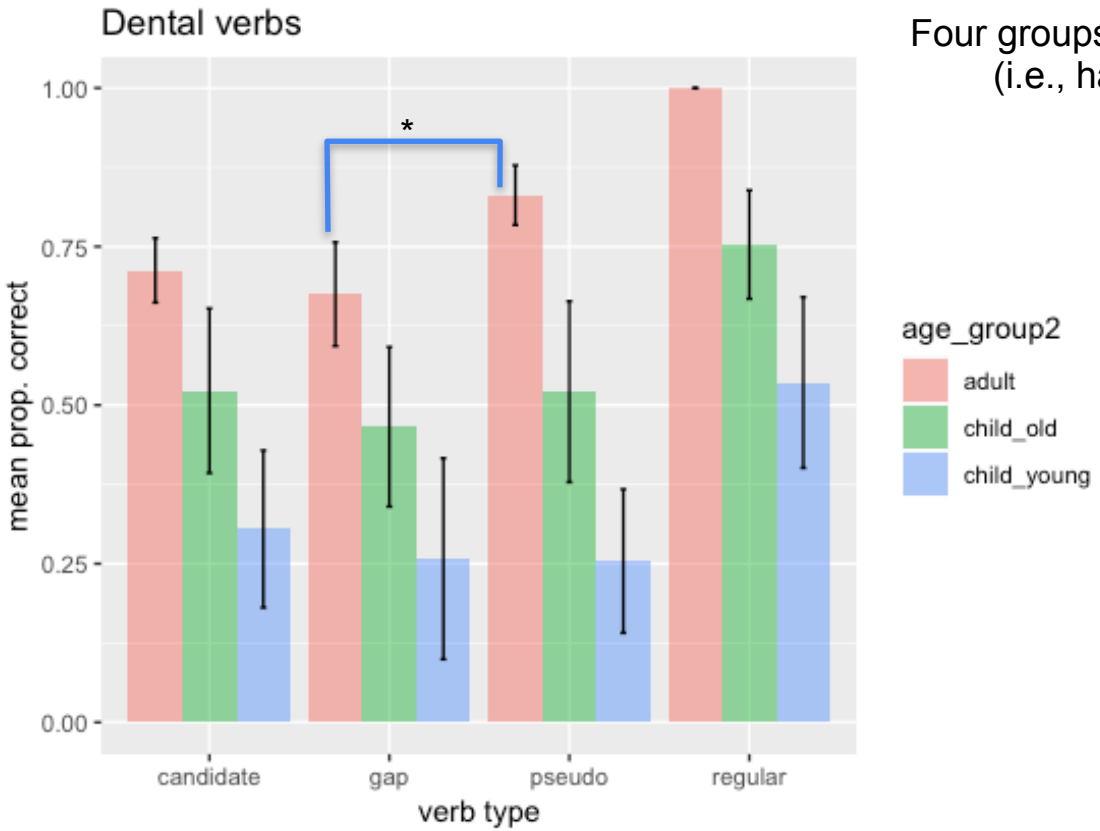
Majority of lexemes in the groups “gaps” and “candidates” are low-frequency, except for a few:
pobedit’ “to win”, ubedit’ “to convince”, shelestet’ “to rustle”
gudet’ “to beep”, osh’utit’ “to sense”, lazit’ “to climb”

Predictions

- Daland et.al. (2007)
  - No role of remote bases or other grammatical factors
  - Most common response in 1sg should be: expected alternation or no response
  - High-frequency defectives should be learned earlier and be different from low-frequency defectives and pseudo-words
  - Pseudo-verbs: may be a little bit gappy (effect of analogy)?

- Pertsova (2016) + OO-Corr bias in L1 (Hayes 2004, Do 2013)
  - Defective verbs = candidate verbs
  - Most common response in 1sg: alternating form, non-alternating form, avoidance response — defectiveness is manifested as variability (with low confidence)
  - Pseudo-verbs should be gappy for all groups (they have no remote bases), but particularly for children who have not learned the dental alternations as well.

Correct = expected dental alternation



Four groups: gap, candidate, pseudo, “regular”
(i.e., have remote bases)

Significant effects:

adults > children
pseudo > gaps (for adults)
regular > all others (children)

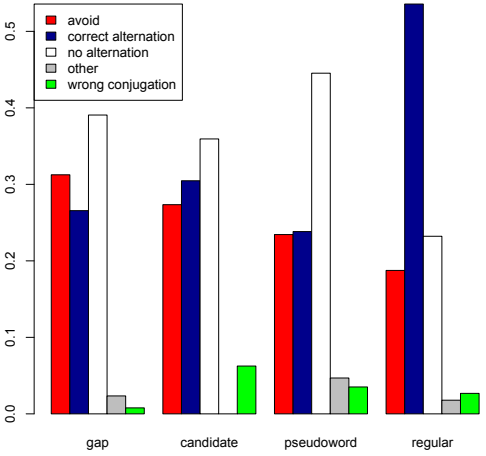
No significant difference found between gaps and candidates

Response types

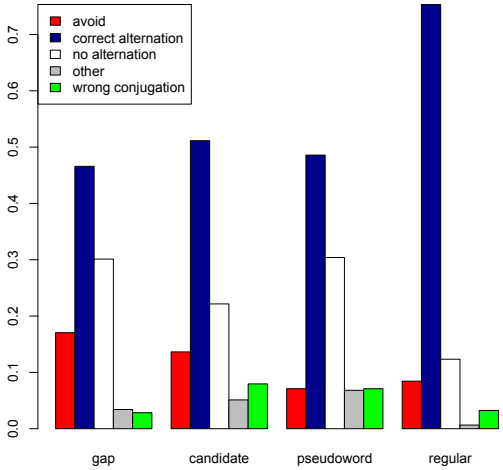
- Example: ***lazit'*** "to climb" (candidate into defective)

  - Expected alternation: laʒ-u
  - Non-alternation: laz'-u
  - Avoid: no response or circumlocution (e.g, *zalezaju, lezu, budu lazit'*)
  - Wrong class: laz-**aj**-u, laz**i-j**-u
  - Other (wrong alternation, stem-truncation, cons.substitution…)

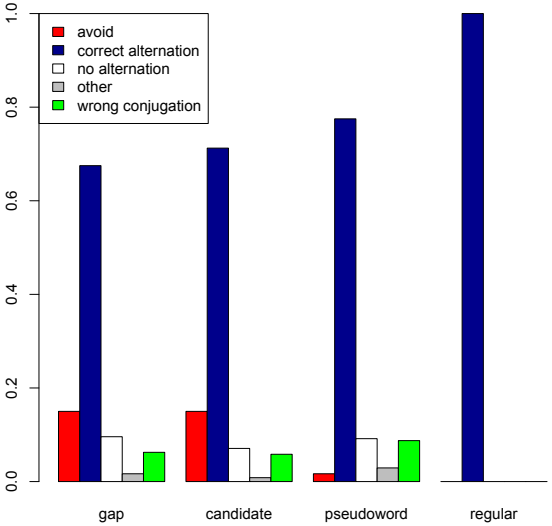
Response types



Younger children



Older children



Adults

- Most common “error” for children is non-alternation
- Adults use avoidance strategies for defective and candidates, but not (as much) for pseudo-verbs
- No effect of lemma frequency for any of the groups

Defectiveness summary

- Children treat gap and candidate verbs in the same way as pseudo-verbs:
  - Non-alternation bias early on
- No significant effect of frequency + significant effect of presence of remote bases with alternations
  - —> low productivity of dental alternations is associated with absence of lexical support from remote bases
- Gap/candidate verbs become distinguished from pseudo-verbs <u>only for adults</u>
  - adults are more likely to apply the dental alternation to a non-word
  - This can be due to
    - memorizing which verbs are defective
    - Or not having the same expectations about non-words

Overall Summary

- Labial alternations are learned early (by 4 years) and via a general rule
- Dental alternations are learned on a lexeme-by-lexeme basis at first
  - there's little evidence for segment specific or class-specific rules early on
- By 5.5 children show an improvement on dental alternations, <u>mainly for those lexemes that have the same alternation in other related forms ("regular")</u>
  - Verbs that lack support from remote bases (including novel psedo-verbs) show high variability in responses, high-number of non-alternating forms, and avoidance strategies which continue into adulthood
- Adults show a somewhat better performance on pseudo-verbs compared to gap/ candidate verbs, suggesting that lexical factors do not play the same role for pseudo-verbs

Remaining questions

- Why are -d, -z, -s stems worse than -t stems (despite the fact that the latter alternation is irregular)?
- Why no frequency effect?
- Can we test performance on ‘f’-final stems? (very infrequent in the lexicon)