Commentary on Virve-Anneli Vihman et al. (2021). “Many ways to decline a noun: elicitation of children’s novel noun inflection in Estonian.” In: *Language and Cognition* 13.4, pp. 693–733

Olivier Bonami
olivier.bonami@u-paris.fr

Université Paris Cité — Lexical Matters — 20/05/2022

Motivation

- The appeal of dual mechanism accounts of the acquisition (and processing) of inflection largely rests on a coincidence that holds for English but not other languages:
  - One inflection class strongly dominates the system in terms of type frequency
  - That inflection class also is describable in terms of simple suffixation patterns, with one of the forms unsuffixed
- The goal of the paper is explore Estonian declension as an example of a system where neither of these holds.

Basics of Estonian declension

- Widespread overabundance in the plural, which we will ignore.
- Local cases are based on the genitive with regular suffixation.

	SINGULAR	PLURAL	alternate PL
NOMINATIVE	sepp	sepad	
GENITIVE	sepa	seppade	
PARTITIVE	seppa	seppasid	seppi
INESSIVE	sepas	seppades	sepis
ILLATIVE	sepasse	seppadesse	sepisse
ELATIVE	sepast	seppadest	sepist
ADESSIVE	sepal	seppadel	sepil
ALLATIVE	sepale	seppadele	sepile
ABLATIVE	sepalt	seppadelt	sepilt
COMITATIVE	sepaga	seppadega	sepiga
ABESSIVE	sepata	seppadeta	sepita
ESSIVE	sepana	seppadena	sepina
TRANSLATIVE	sepaks	seppadeks	sepiks
TERMINATIVE	sepani	seppadeni	sepini

Basics of Estonian declension

- Most of the action is in the relation between the three core cases.
- Main classes in child directed speech corpus (97.8% of types):

	NOM	GEN	PART	GEN formation	PART formation	Freq.
I	*kuu*	*kuu*	*kuud*	∅	*-d*	2.6%
II	*päike*	*päikese*	*païkest*	*-(s)e*	*-(st)*	9.4%
III	*auto*	*auto*	*autot*	∅	*-t*	7.6%
IV	*raamat*	*raamatu*	*raamatut*	-V	-V*t*	12.5%
V	*maja*	*maja*	*maja*	∅	∅	9.5%
VI	*pilt*	*pildi*	*pilti*	weakening + -V	-V	49.2%
VII	*tigu*	*teo*	*tigu*	weakening	∅	2.4%
VIII	*aken*	*akna*	*akent*	strengthening	*-t*	2.8%

- In the highlighted classes, the theme vowel (*-a*, *-e*, *-i* or *-u*) is unpredictable.
- Interesting situation: the main inflection pattern
  - relies on stem alternations, and
  - is opaque when starting from the citation (and most frequent) form.

State of the art

- Theory:
  - Models of the acquisition of morphology can be ranked on a gradient, with dual mechanism “Words and Rules” models at one extreme, and single route exemplar-based models at the other extreme (Ambridge 2020).
  - Mounting evidence that, even for English, analogy to existing forms plays a role in the processing and acquisition of regular forms (Albright & Hayes 2003; Ramscar & Yarlett 2007; Ambridge 2010).
  - On the other hand no emerging consensus on how much abstraction is necessary/warranted (compare Albright & Hayes 2003 with Keuleers 2008).
- Linguistic diversity
  - Research on languages other than English highlights the absence of a general alignment of (i) cell frequency with morphological simplicity; (ii) different aspects of ‘default’ inflection (high type frequency, formal simplicity, productivity…)

Factors affecting morphological accuracy

1. Wordform frequency: more frequent words are easier to inflect accurately.
   - The present paper uses wugs to neutralize that effect.
2. Neighborhood density: words whose phonological neighbors inflect in a more coherent way are easier to inflect accurately.
3. Age: Older children are more accurate
4. Granlund et al. (2019) found an interaction between age and neighborhood density in Polish and Finnish: neighborhood density became less important with age
   - This was a study with existing words. With nonce words, one expects the opposite interaction, with children getting more proficient at using analogy.
5. System effects: properties of the linguistic system lead to expectations. In Estonian:
   - We expect differential behavior when predicting from the (unsuffixed) nominative or a suffixed form
   - We expect differences between affixal inflection and stem allomorphy
   - We expect differential opacity to play a role: e.g. allative and genitive should be interpredictable.

Evidence for indeterminacy of inflection classes

- Kaalep (pc) claims that Estonian speakers tend to show strong consensus on the inflection of novel word.
  - E.g. NOM.SG *äpp* (< En. *app(lication)*) uniformly leads to GEN.SG *äpi*, despite the existence of

NOM	GEN	PART	
näpp	näp-u	näpp-u	'finger'
käpp	käp-a	käpp-a	'paw'
täpp	täp-i	täpp-i	'dot'

- However, strong philological evidence of hesitation when a noun suddenly reaches prominence.
- In addition, children commonly produce declension errors amounting to assigning a noun to the wrong inflection classes (20% of noun tokens at age 1;7 according to Argus 2009)

Research questions

1. Do children's responses show development toward those of adults?

   H1 Children's responses will not be fully adult-like at age 5, but accuracy will increase with age.

   H2 Children's responses will vary more than adults' responses; variability will decrease with age.

2. How do presentation and target case affect children's responses?

   H3 Target and presentation case will both affect accuracy.

   H4 Children will not use the affixal partitives as a generalised default, but rather use both affixal and vowel-final patterns.

3. Do children make greater use of neighbourhood density to inflect novel nouns than adults?

   H5 With novel nouns, akin to very-low-frequency forms, we expect to find a main effect of neighbourhood density and an interaction with age, with a greater effect of neighbourhood density with increasing age.

Methods I

- 70 children in 3 age groups (3, 4, 5) and 21 adults as controls
- Items:

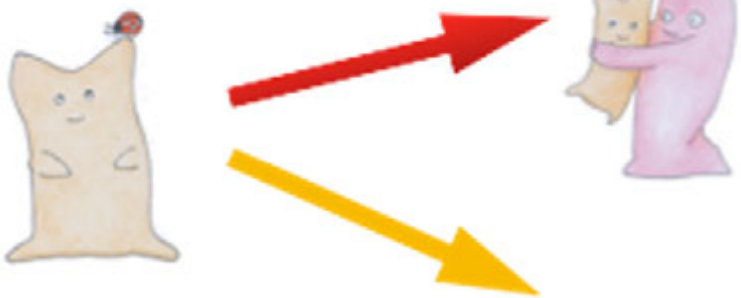
	Presentation forms		
	Nominative	Allative	Declension class
1	esu	esu-le	III, V
2	keenel	keenli-le	IV
3	keesik	keesiku-le	IV
4	kidu	kidu-le	III, V
5	käle	käle-le	III, V
6	lada	lada-le	III, V
7	miga	mea-le	I, III
8	mii	mii-le	I, III
9	mool	mooli-le	IV, VI
10	nuplik	nupliku-le	IV, VI
11	palas	palase-le	II, IV
12	pei	pei-le	I, III
13	rupp	rupi-le	III, VI
14	sebu	sebu-le	III, V
15	sipp	sipu-le	IV, VI

Methods II

- Conditions:

Predictor	Target
Nominative	Genitive
Nominative	Partitive
Allative	Genitive
Allative	Partitive

- Materials and procedure:



Presentation (ALLATIVE):

Exp: *Vaata! Lepatriinu tuli külla* ***keesikule***.
Keesikule *väga meeldivad lepatriinud.*
Kas saad mulle öelda, kellele meeldivad lepatriinud?
Exp: Look! A ladybug came to visit this ***keesik***.
A ***keesik*** really likes ladybugs.
Can you tell me who likes ladybugs?
Child: "**keesikule!**"

Elicitation (a) PARTITIVE:
Exp: *Siin on tema issi. Keda issi kallistab? Issi kallistab…*
Exp: Here's his/her daddy. Who is Daddy hugging? Daddy's hugging…
Child: "**…keesik-ut!**"

Elicitation (b) GENITIVE:
Exp: *Siin on tal lilla mantel seljas. Kelle seljas on mantel? Mantel on…*
Exp: Here s/he has a purple coat on. Who has a coat on? The coat is on… ["whose back"]
Child: "**…keesik-u!**"

Variability I

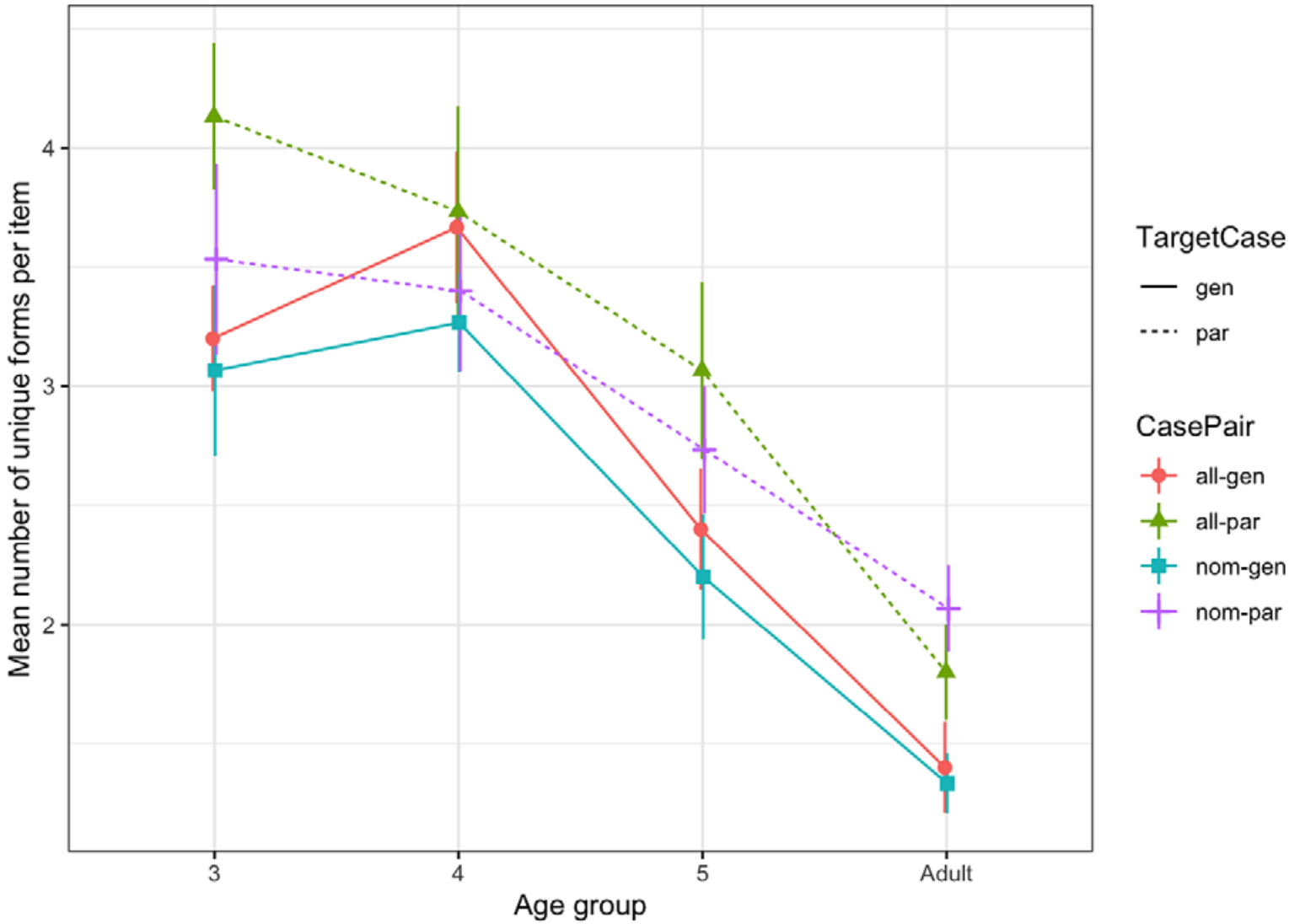


Fig. 2: Variability: mean number of unique responses, by age group and case condition.

Variability II

- Model structure:

$$\texttt{number of forms} \sim \texttt{age} + \texttt{predictor} * \texttt{target} + (1|\texttt{item})$$

- When comparing adults to children (binary coding of age):
  - Adults have less variability
  - Partitive leads to more variability than genitive
- When comparing among children (gradient coding of age in months):
  - Older children have less variability
  - No other documented effect
- Examination of detailed results indicates that variability in adults predicts variability in children, but not the other way around.

Variability III

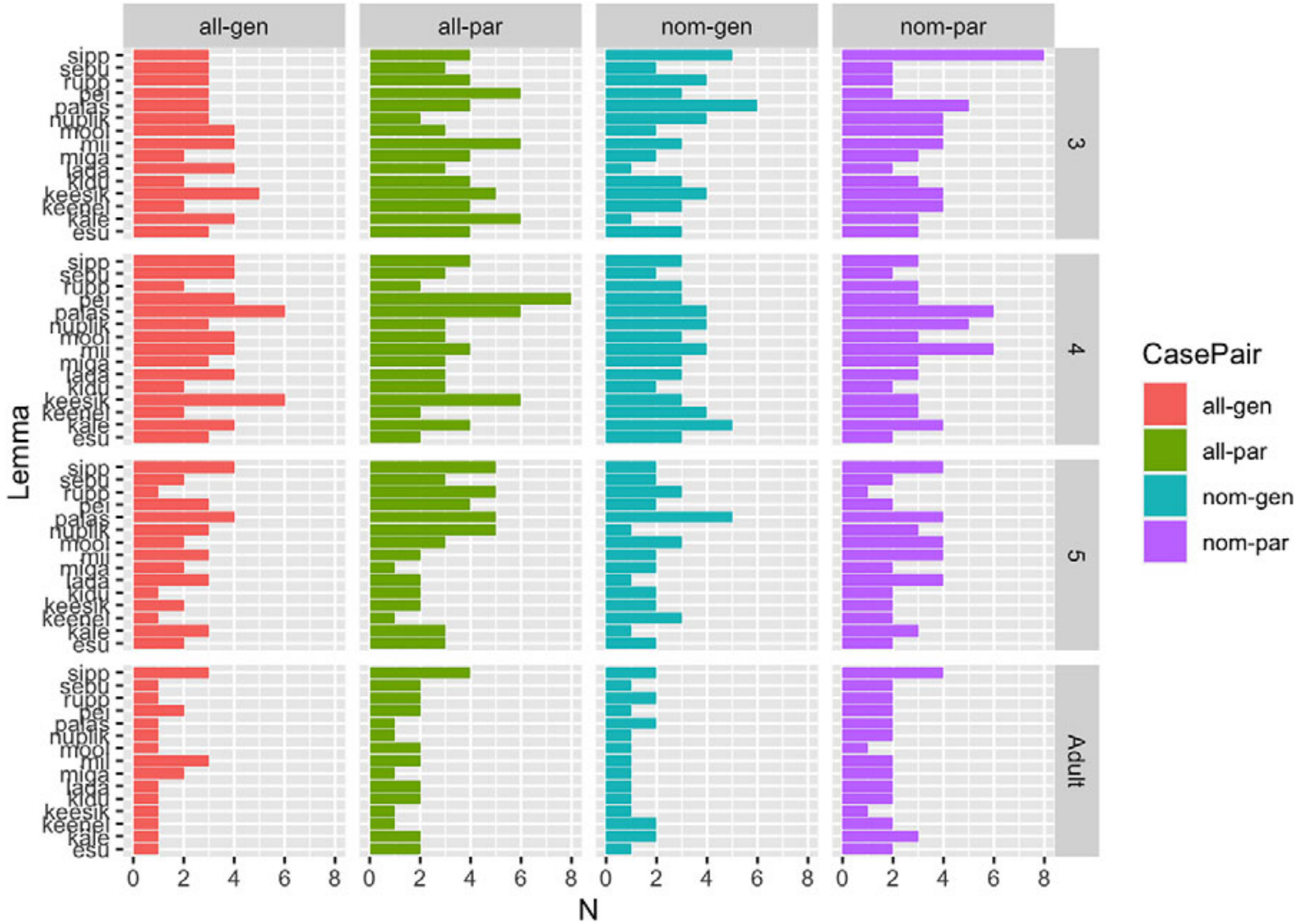
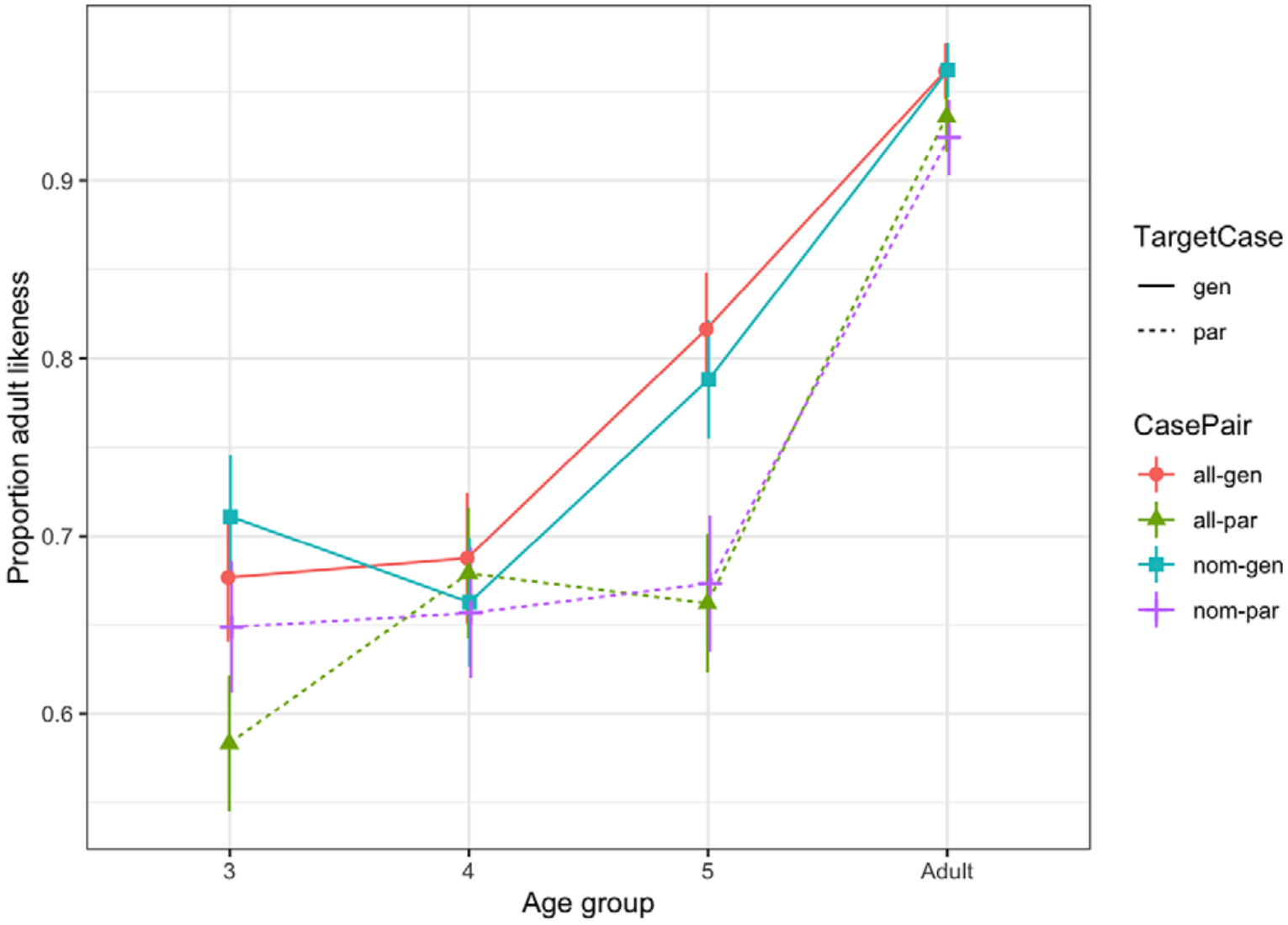


Fig. 3: Variability per item: number of unique responses per lemma, by case pair and age group.

Accuracy I

- Accuracy was operationalized as follows:
  - If at least 66% of adults converged on the same response, this was considered the correct form.
  - If no answer reached 66% of adult responses, then all forms produced by some adult were considered correct.

Accuracy II


Proportion adult likeness
0.9
0.8
0.7
0.6
3
4
5
Adult
Age group
TargetCase
gen
par
CasePair
all-gen
all-par
nom-gen
nom-par

Accuracy III

- Model structure (looking at children only):

$$\texttt{correct} \sim \texttt{age} + \texttt{predictor} * \texttt{target} + (1|\texttt{item})$$

- Significant effects of age and target case (partitive less accurate than genitive)

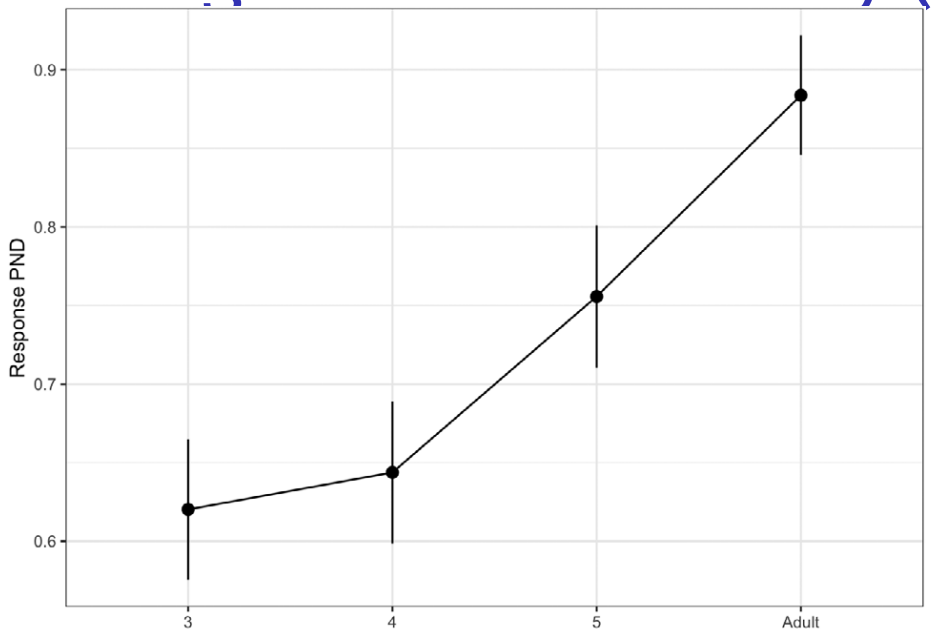
Phonological Neighborhood Density (PND) I

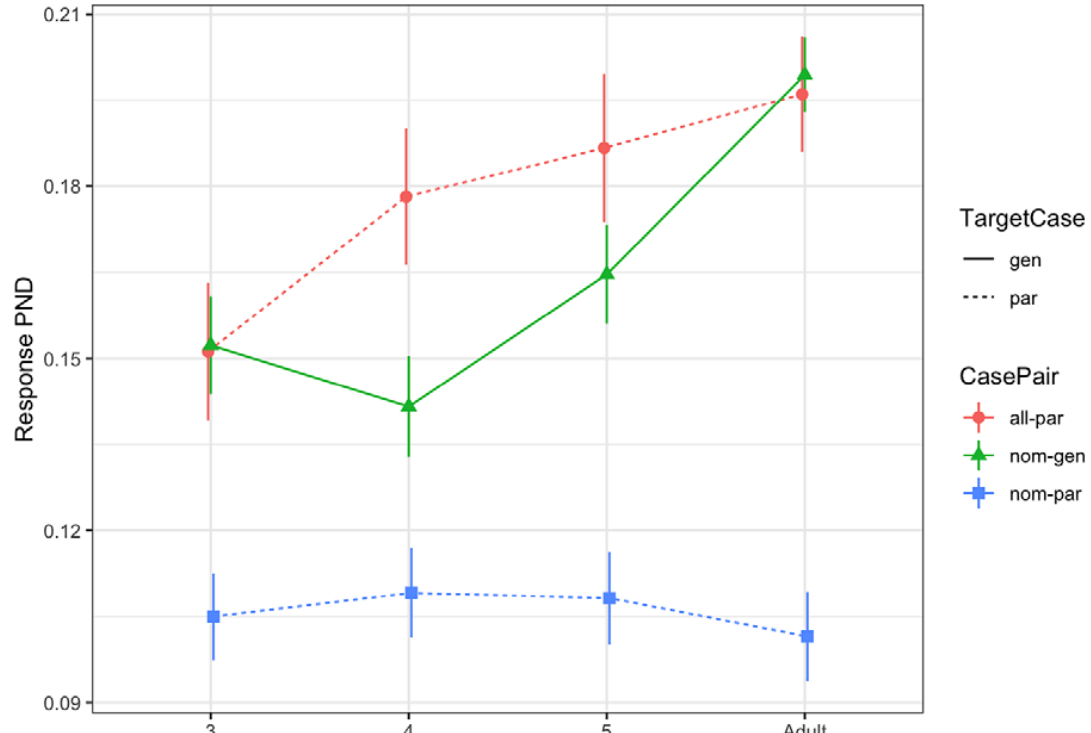
- Misnomer: what they use is a measure of predictability of inflection pattern based on how phonological neighbors behave.
- Taken over from Granslund et al. (2019) and loosely based on Albright & Hayes's 'purely analogical' model.
- The description is quite opaque. From what I gather, the algorithm:
  - Identifies neighbors for each (predictor,target) pair by identifying attested pairs that minimally differ in the edits necessary to go from predictor to target.
  - Final score:

$$\frac{\text{Summed similarity of the target to its neighbors}}{\text{Summed similarity of the predictor to all possible predictors}}$$

Items for comparison were taken from a child directed speech corpus.

Phonological Neighborhood Density (PND) II





Phonological Neighborhood Density (PND) III

- Model structure:

$$\texttt{PND} \sim \texttt{age} * \texttt{predictor} * \texttt{target} + (1|\texttt{item}) + +(1|\texttt{item})$$

- When comparing all four age groups:
  - Significant effect of age
  - No main effect of predictor or target case
  - Interaction of age and target case: smaller effect of age for partitive targets.
- When comparing among children with age in months:
  - Significant effect of age
  - Main effect of target case: higher PND for partitive targets (????)

Error analysis

- For genitive targets, strong preference for vowel final forms (which are the only possible correct responses).

TABLE 5. *Genitive targets: proportions of vowel-final responses*

Presentation Case	3-year-olds	4-year-olds	5-year-olds	Adults
Nominative	83%	78%	85%	100%
Allative	91%	80%	85%	99%

- For partitive targets, more variability.

Presentation case	3-year-olds	4-year-olds	5-year-olds	Adults
Nominative	56%	47%	54%	70.5%
Allative	66%	72%	64%	83.1%

- Interestingly, children preferred non-affixal partitive formation which relies on stem allomorphy and theme vowel selection.

Discussion

- There is true variability: adult participants lack consensus in roughly half the conditions.
- Evidence for awareness of the statistical distribution of inflection patterns from age 3

H1 *Children's responses will not be fully adult-like at age 5, but accuracy will increase with age.* confirmed
  - Monotonous raise in accuracy
  - Far from adult performance at age 5

H2 *Children's responses will vary more than adults' responses; variability will decrease with age.* confirmed

H3 *Target and presentation case will both affect accuracy.* partially confirmed
  - Only target case ended up mattering
  - Speculation: the two presentation cases entail different task demands, which might cancel each other out.

H4 *Children will not use the affixal partitives as a generalised default, but rather use both affixal and vowel-final patterns.* confirmed

H5 *With novel nouns, akin to very-low-frequency forms, we expect to find a main effect of neighbourhood density and an interaction with age, with a greater effect of neighbourhood density with increasing age.* confirmed

More discussion

- Interesting contrast with the results of Granlund et al. (2019), which found an negative interaction between PND and age when testing with real words:
  - With real words, children progressively memorize the lexicon, which entails memorizing patterns that may run counter the predictable distribution of forms. Hence the negative interaction.
  - With wugs, there is no such memorization, so the authors are tapping into the separate process of accuracy of statistical prediction improving.
- Broad support for the strong role of analogy in the acquisition of the inflection system.
- Results do not distinguish between radical exemplar models and other usage-based models relying on abstract representations.
- Children exhibit proficiency at the PCFP in general: for instance they're good at predicting the simpler genitive form from the more elaborate allative form, or at applying stem alternations.

Evaluation

- Generally very informative on the PCFP in acquisition.
- Not clear that the measure of accuracy makes sense.
  - It would make more sense to track how the distribution of answers in children match the distribution in adults overall, e.g. using relative entropy.
- The PND measure is poorly labeled (this is about predictability based on the neighborhood, not density of the neighborhood) is poorly documented, so that it is hard to evaluate what is going on.
- All the data and scripts are available on OSF, so one could easily do better!

References I

Vihman, Virve-Anneli et al. (2021). “Many ways to decline a noun: elicitation of children’s novel noun inflection in Estonian.” In: *Language and Cognition* 13.4, pp. 693–733 (cit. on p. 1).