

## Many ways to decline a noun: elicitation of children's novel noun inflection in Estonian\*

VIRVE-ANNELI VIHMAN 

*Institute of Estonian and General Linguistics, University of Tartu*

FELIX ENGELMANN

*School of Health Sciences, University of Manchester<sup>†</sup>*

ELENA V. M. LIEVEN

*School of Health Sciences, University of Manchester<sup>†</sup>*

AND

ANNA L. THEAKSTON

*School of Health Sciences, University of Manchester*

*(Received 18 December 2020 – Revised 6 August 2021 – Accepted 10 August 2021)*

### ABSTRACT

**Aims:** This study investigated three- to five-year-olds' ability to generalise knowledge of case inflection to novel nouns in Estonian, which has complex morphology and lacks a default declension pattern. We explored whether Estonian-speaking children use similar strategies to adults, and whether they default to a preferred pattern or use analogy to phonological neighbours.

**Method:** We taught children novel nouns in nominative or allative case and elicited partitive and genitive case forms based on pictures of unfamiliar creatures. Participants included 66 children (3;0–6;0) and 21 adults.

---

[\*] We gratefully acknowledge support from the European Union's Seventh Framework Programme (Marie Curie IEF grant no 623742 to first author), which made this research possible, and from the ESRC (UK) for the International Centre for Language and Communicative Development (LuCiD, grant no ES/L008955/1) which supported Elena Lieven's, Anna Theakston's, and Felix Engelmann's participation in this research. Many thanks to Research Assistant Kristiina Vaik for data collection; to artist Tonis Kriisa, who drew the pictures used as stimuli; and to the children and teachers in day-care centres in Estonia who made the study possible. Address for correspondence: e-mail: [virve.vihman@ut.ee](mailto:virve.vihman@ut.ee).

[†] This article was originally published with an error in these authors' affiliations. This has now been corrected and an erratum is available at <https://doi.org/10.1017/langcog.2022.33>.

Because of multiple grammatical inflection patterns, children's responses were compared with those of adults for variability, accuracy, and morphological neighbourhood density. Errors were analysed to reveal how children differed from adults.

Conclusions: Young children make use of varied available patterns, but find generalisation difficult. Children's responses showed much variability, yet even three-year-olds used the same general declension patterns as adults. Accuracy increased with age but responses were not fully adult-like by age five. Neighbourhood density of responses increased with age, indicating that analogy over a larger store of examples underlies proficiency with productive noun inflection. Children did not default to the more transparent, affixal patterns available, preferring instead to use the more frequent, stem-changing patterns.

**KEYWORDS:** morphological acquisition, case-marking, novel nouns, wug method, Estonian

## 1. Introduction

How do children learn to generalise over inflectional forms in a complex morphological system? Mechanisms underlying the acquisition of inflectional morphology have been debated vehemently over the past three decades. Yet the debate has mostly spun around data from the impoverished morphological system of English, which has formed the basis for most of the theoretical discussion contrasting dual-route (e.g., Hartshorne & Ullman, 2006; Pinker & Ullman, 2002; Prasada & Pinker, 1993) and single-route approaches (e.g., Keuleers & Daelemans, 2007; Maratsos, 2000; Maslen, Theakston, Lieven, & Tomasello, 2004; McClelland & Patterson, 2002). This opposition is scaffolded by a simplified conceptualisation of the structure of languages and an idealised model of the mechanisms involved in their acquisition.

The sparse morphology of English past tense verbs or plural nouns can be described using a contrast between prevalent regular forms and a smaller set of irregular lexemes (frequent in tokens but not types). It is not evident, however, that this binary contrast is universally applicable; a more complex description may be required for more highly inflected languages, which have become increasingly visible in research on morphological acquisition. This paper focuses on Estonian nouns, which are inflected for fourteen cases and two numbers. The nominal paradigm includes many classes, none of which accounts for even half of noun types or tokens in child-directed speech (Granlund et al., 2019, p. 187). Hence, it is not straightforward to label the noun classes as regular or irregular. In this study, we asked whether Estonian-speaking children aged 3 to 5 have similar strategies as adults when declining novel nouns, and whether children and adults use a default pattern or derive inflections by analogy to phonological neighbours.

## 2. Theoretical background

'Words and rules' accounts claim that two separate mechanisms are needed to explain the psychological facts (Prasada & Pinker, 1993; Prince & Pinker, 1988). Irregular forms are said to be learned by rote and accessed through analogy, making use of declarative memory, while regular forms are learned and retrieved by way of rules, using procedural memory. This bifurcated view of morpho-syntactic processing relies primarily on evidence from the effect of phonological neighbours, namely the claim that irregulars are affected by phonologically similar words, while regulars are not (Berent et al., 2002; Marcus, 1995).

Yet synchronic, diachronic, and experimental evidence was presented as early as the 1980s, indicating that speakers are sensitive to whole word forms, and that they may use prototypes to classify English verbs into morphological classes (Bybee & Moder, 1983). Bybee and Slobin (1982) proposed that speakers generalise not in the form of rules, but rather in schemas defined by whole word properties of English verbs. Schemas are invoked to describe a morphological formation which does not generate "forms by combination, but rather [...] lexical associations among existing forms" (Bybee & Moder, 1983, p. 255). Bybee and Moder claim that "producing a past form of an irregular verb is not a matter of applying a feature-changing rule to an underlying form, but is rather a matter of lexical selection" (1983, p. 256). The schema-based approach indicates that rules are not necessary if general cognitive mechanisms underlie morphological processing. Even morphological patterns which can be described by rules may emerge from networks of connections between word forms sharing similar phonology and/or semantics. Morphological generalisations based on similar patterns across the network reinforce output-based schemas. When new word forms are fitted into such schemas, speakers derive morphological generalisations from analogy. Productivity then depends on the type frequency of a schema and its openness to accept new forms (Krajewski et al., 2011, p. 835).

Indeed, more recent research has shown that regular forms may be generalised according to analogy and similarity to prototypes or nearest neighbours, just like irregular forms (Ambridge, 2010; Blything et al., 2018; Ramscar & Yarlett, 2007). Even in English, studies have identified problems with positing different cognitive processes for regular and irregular inflection: an approach relying on such a division cannot explain findings that show a good deal of overlap between them. A pattern emerges which is more suggestive of 'quasi-regularity', with multiple patterns described with partially regular rules (see McClelland & Patterson, 2002; see also Mirković et al., 2011 for quasi-regularity in Serbian). Irregular forms demonstrate various sorts of partial regularity, and regular forms can include partial exceptions. Moreover, regular forms have also been found to be influenced by analogy to other – regular and irregular – forms (Albright & Hayes, 2003). This is problematic for a dual

mechanism account, lending support instead to a model based on associative memory and analogical generalisation.

While the theoretical approaches to morphological acquisition are often seen as a battleground between two camps, the above discussion demonstrates that in fact models of acquisition include varying degrees of both rules and analogy. Granlund et al. (2019) set out a continuum of approaches, from generative, structural linguistic approaches which rely on highly abstract formal rules and feature checking (e.g., Deen & Hyams, 2006), to ‘similarity-based’ exemplar theories, which, at their most extreme, abandon abstractions altogether, as in radical exemplar models (see, e.g., Ambridge, 2020b; Chandler, 2010; but also Ambridge, 2020a).

Despite the apparent dichotomy between dual- and single-route models, most approaches fall somewhere between these poles. The continuum includes “essentially rule-based approaches that incorporate some sensitivity to statistical and phonological properties of the system, and essentially similarity-based approaches that incorporate ‘rules’ in the sense of stored representations that abstract to some degree across individual lexical items” (Granlund et al., 2019, p. 170). Cross-linguistic evidence, reviewed below, increasingly lends support to the similarity-based approaches allowing for a certain degree of abstract knowledge, such as connectionist models (e.g., Mirković et al., 2011), which focus on individual input–output mappings, and Usage-Based models using abstract schemas (e.g., Bybee & Moder, 1983; Bybee & Slobin, 1982; Maslen et al., 2004). Two factors affecting acquisition hold the key to differentiating between the theoretical accounts: form frequency and phonological neighbourhood density, discussed below, in [Section 2.2](#).

### 2.1. LANGUAGES WITH MORE COMPLEX MORPHOLOGY

Models of acquisition rely on assumptions about linguistic structure more generally. If regular inflections tap into a more efficient learning mechanism and representational capacity, a rule-based paradigm would be expected to emerge as a default, across languages. For cognitive plausibility, this claim must be tested cross-linguistically, in languages with more complex patterns.

The focus on English in research in this area, and the idealised binary classification of inflectional forms, has had at least two unfortunate consequences. First, linguistic models built around the relatively impoverished morphological patterns of English cannot readily be scaled up to describe the patterns found in morphologically richer languages. Second, the English system also leads to models of acquisition which may mischaracterise the acquisition of morphological forms and paradigms of more complex languages. Data from languages with rich morphology challenge largely unexamined assumptions, including the regular–irregular distinction, and the equation of

regularity (productivity) with rule-based affixation, and irregularity with stem change. As noted by Blevins, "it is neither the case that all productive processes are affixal nor the case that all forms that can be 'decomposed into stem + affix' reflect productive processes" (1999, p. 1015). Stem changes can participate in regular inflectional morphology, as in Danish (see Kjaerbaek et al., 2014; Laaha et al., 2011), or Estonian, as discussed below; and affixes can co-occur with phonemic stem changes, as in Polish, Finnish, or Estonian.

Typological and psycholinguistic data also challenge the assumption that there is one default, regular pattern and other deviant, irregular ones. The notion of a default can refer to the most widely applicable forms (associated with the most types), the most productive forms (most likely to be applied to new words), or the most regular forms (most transparently described by a rule). In English, regular past tense *-ed* or plural *-s* formation are defaults on all three counts, but languages display considerable diversity in this regard. Plurals in German and Hebrew (Berent et al., 1999) are examples of defaults in terms of regularity and productivity: e.g. in German, the *-s* plural is infrequent, but newly coined words and foreign loans take the *-s* plural (Hahn & Nakisa, 2000; Marcus, 1995). Many languages, or paradigms within a language, cannot be said to have any default, as with Dutch plurals (Keuleers & Daelemans, 2007) and Polish genitive case forms. Bermel, Knittl, and Russell have shown that in Czech, both phonological neighbourhood density and corpus frequency matter in adults' selection of morphological forms with multiple possible inflections (Bermel et al., 2015, 2017). Dąbrowska (2001) demonstrates that Polish children use all three genitive forms in overgeneralised ways, rather than any one acting as a default. Dąbrowska and Szczerbiński (2006) show that type frequency and phonological diversity predicted children's accuracy with Polish case inflection better than regularity. Without a default pattern, it is not apparent why differing morphological processes would be learned or stored in different ways, nor how this could be effected. In Estonian, the inflection of novel words and names is often open to disagreement when two equally frequent paradigms are in competition; the data on Estonian is discussed further in Section 2.4.

Cross-linguistically, then, we find differences in the availability of morphological defaults, and diversity with regard to how many productive inflectional paradigms exist side by side. When two (or more) equally large classes of phonological neighbours are in competition, speakers may well disagree on how to resolve the inflection of a novel noun. This disagreement may remain in some languages over time, while in others, something like a default may emerge, based on the strength of cues in competition and the distributional characteristics of word forms. What looks like a default may be an emerging property of some, but by no means all, morphological systems. The lack of a default in many languages suggests that the likely mechanism underlying morphological productivity is analogy.

The Estonian system is useful for examining the above assumptions on three counts: (a) multiple declension classes are productive, and the nominal system is not readily divided into regular and irregular classes; (b) no class accounts for over half of noun types or tokens – hence, it does not have a clear candidate for a default class based on frequency; and (c) stem changes play a crucial role in the nominal inflection system. Further, the questions of how children learn the system and how adults inflect novel words are intriguing, as: (a) the classes are not fully predictable from either phonological form or semantics; and (b) predicting the class of a noun sometimes requires more than a single form (see Blevins, 2008).

## 2.2. FACTORS AFFECTING MORPHOLOGICAL ACCURACY

In this study, we look at three factors which may affect morphological acquisition: (a) statistical properties of the input; (b) the child's age; and (c) the nature of the morphological task.

### 2.2.1. *Properties of the input*

WORD FORM FREQUENCY and NEIGHBOURHOOD DENSITY have both been demonstrated to affect processing in adults (e.g., Lõo, Järvikivi, & Baayen, 2018; Lõo, Järvikivi, Tomaschek, et al., 2018; Vitevitch & Luce, 2016) and in children learning English (e.g., Ambridge et al., 2015; Maslen et al., 2004; Matthews & Theakston, 2006) and languages with complex morphology (e.g., Dąbrowska & Szczerbiński, 2006; Granlund et al., 2019; Kjaerbaek et al., 2014; Räsänen et al., 2015; Savičiūtė et al., 2017). Savičiūtė and colleagues investigated children's abilities to inflect both real and novel nouns in Lithuanian, testing the whole noun paradigm, and found main effects of both token frequency and neighbourhood density, but no interaction. Granlund and colleagues used a rigorous, cross-linguistically comparable noun form elicitation paradigm for testing the effects of input frequency and neighbourhood density in three languages – Polish, Finnish, and Estonian – but their study tested only real nouns, which children may be able to inflect based on stored knowledge of word forms. As noted by Savičiūtė et al., none of the earlier studies included both familiar and novel nouns, which would be “crucial for establishing when and how children generalise their knowledge of individual word forms to inflect noun stems that we can be certain they have not encountered before” (2017, p. 645). We build on the cross-linguistic, familiar noun elicitation study conducted by Granlund et al. (2019) to investigate children's productive knowledge and ability to extend morphological paradigms to novel words in one of those languages: Estonian.

In this study, we use novel nouns which, by definition, do not appear in the input. Hence the frequency of a particular word form cannot affect the child's accuracy. We investigate the effect of morphophonological neighbours (friends

and enemies), which were shown to affect children's accuracy with inflected word forms by Granlund et al. (2019). If speakers make use of analogy, they will draw on more prevalent schemas for inflecting words with similar base forms when deciding on the target form for a new word. Some, but not all, studies have found negative interactions between form frequency and neighbourhood density, indicating that, in order to inflect word forms encountered less frequently, children rely on similar-sounding neighbours. Räsänen et al. (2015) found a greater effect of neighbourhood density with low-frequency verbs in Finnish. Savičiūtė et al. (2017) found no interaction between the same predictors in Lithuanian. Granlund and colleagues tested three languages and found a negative interaction (again, a greater effect of neighbourhood density with low-frequency forms) only in Polish, although a pooled model confirmed the effect across languages, and showed no effect of language when that was included in the model. This suggests that a larger sample size may reveal the same effect in all languages.

For the present study, in which form frequency is kept constant at zero, neighbourhood density should matter: the stimuli are unfamiliar from the input (thus, equivalent to very low-frequency nouns), and the results discussed above show that WHEN an interaction is found, neighbourhood density is more important with lower-frequency forms. Take, for example, the nonce noun *keesik*. Its form is similar to real Estonian nouns like *seelik* 'skirt' and *küülik* 'rabbit', which also have similar partitive forms: *seelik-ut*, *küülik-ut*. The more neighbours a word has with similar base forms, the easier it should be to produce an inflection even when the word is unfamiliar: we expect children as well as adults to reliably produce *keesik-ut* as the partitive form, as all its neighbours provide the same model for analogy.

### 2.2.2. Age

Children's knowledge of the grammatical system and facility with applying that knowledge is expected to improve with age. However, how improved grammatical knowledge interacts with the effect of neighbourhood density is unclear. On the one hand, as children's abstract knowledge of the morphological system grows with age, we might expect them to rely less on analogy to form new inflections, if they can rely instead on morphological rules. In this case, neighbourhood density would decrease in importance with age. On the other hand, as children's vocabulary and exemplar knowledge grows, they might be expected to make more efficient and extensive use of analogy, leading to an increased effect of neighbourhood with age.

In their experiment on children's ability to inflect familiar nouns, Granlund et al. (2019) predicted a positive interaction between neighbourhood density and age, expecting children to rely less on the frequency of individual inflected forms and more on analogy with similar neighbours as their language

developed. However, with one measure of neighbourhood density, they found the opposite interaction in Polish and Finnish, with neighbourhood density becoming less important with age rather than more, and no relationship between age and neighbourhood density in Estonian. With NOVEL nouns, we expected neighbourhood density to play a greater role in their inflection, with more experienced speakers relying on analogy to similar words; younger children have a smaller lexicon to draw on.

### 2.2.3. Linguistic system and morphological task

Finally, the form in which a word is learned and the elicited form may affect children's accuracy. Most Estonian nouns are learned first in the nominative form (Argus, 2009), and most nonce word elicitation studies across languages present novel words in the nominative singular. Krajewski et al. (2011) took exception to this practice, noting that one cannot identify a single, zero-marked form which serves as a base for Polish noun inflection across all morphological classes: hence, "the ability to use the system means the ability to switch endings, rather than add them to a bare stem" (2011, p. 835); that is, competence in Polish means the ability to switch from any noun form to another. Children without full knowledge of the paradigms may be more prone to error when switching between cases other than the nominative. The same holds for Estonian, which also has other frequent cases besides the nominative. Case suffixes are not always added to a zero-marked nominative form. For example, the genitive can be derived from the nominative via consonant omission (*maasikas* 'strawberry.nom' <Please put "nom" in small capitals.> *maasika* 'strawberry.GEN') or stem-internal lengthening (*mõte* 'thought.nom' <Please put "nom" in small capitals.> *mõtte* 'thought.gen'; see Section 2).

What's more, as the case system involves a mixture of stem changes and affixation, target forms may involve various complexities, depending on the phonological form of the lexeme and which form is learned first. Nouns presented in the nominative case may be easier because of higher frequency, but nouns presented in the allative case provide more cues about the genitive case form. Partitive case may be more complex because it has affixal and non-affixal allomorphs.

It is important to test whether children are able to apply general knowledge of nominal paradigms when inflecting nouns, or are only able to derive forms by applying a rule-based change to a nominative base. Hence, we included two different presentation case conditions (nominative and allative) to test for the effect of morphological task. The details of Estonian case morphology are sketched in the following section.



## 2.3. THE NOUN CASE SYSTEM IN ESTONIAN

Estonian is a Finno-Ugric language, known for its complex inflectional morphology. Its noun morphology involves an extensive case-marking system with fourteen distinct cases in two numbers, but no grammatical gender. The complexity in nominal morphology lies in the multiplicity of declension classes (ways of inflecting different groups of nouns), rather than the number of cases (Kaalep, 2012; Peebo, 1997; Viht & Habicht, 2019). The standard grammar distinguishes 26 classes, but, based on just the singular paradigms, we can distinguish at least eight distinct, productive classes (see Table 1, productive classes I–VIII). The most frequent class (VI) represents 49.2% of noun types in a corpus of child-directed speech (CDS) (Granlund et al., 2019, p. 187). Importantly, this class involves not affixation but stem change, contrasting ‘strong’ and ‘weak’ stems (e.g., *kirsi/kirssi* ‘cherry.GEN/PAR’, *kiige/kiike* ‘swing.GEN/PAR’).

We used genitive and partitive as target cases in our study: they are known as the ‘grammatical’ cases, along with the unmarked nominative. These three cases have the most general grammatical functions (marking subjects and objects) and are by far the most frequently used (together accounting for 64.2% of the noun tokens in the CDS corpus). Moreover, these are the diagnostic cases for determining the declension class of a noun, as shown in Table 1. As such, they also capture the variability of the Estonian declension system, demonstrated in Table 2. Apart from the grammatical cases, the eleven ‘semantic’ cases are regular, with mostly invariant affixes attached to a stem identical to the genitive case.

## 2.3.1. Variability in the ‘grammatical’ cases

GENITIVE case forms are always vowel-final, but can be derived in various ways from the nominative, through consonant or syllable DELETION (*maasikas* > *maasika* ‘strawberry.NOM > GEN’), vowel or syllable ADDITION (*raamat* > *raamatu* ‘book.NOM > GEN’), STEM CHANGE (*hobune* > *hobuse* ‘horse.NOM > GEN’ and *pilt* > *pildi* ‘picture.NOM > GEN’), or NO CHANGE (*isa* = *isa* ‘father.NOM = GEN’). When a vowel is added to form the GEN case, this ‘theme vowel’ (-a, -e, -i, -u) must be learned lexically, item by item. The genitive is more informative regarding noun declension than the nominative form, as it contains information about which theme vowel is used in the other cases (-u in the example in Table 2).

The PARTITIVE is formed according to two general patterns, depending on declension class: AFFIXAL (-t, -d) or VOWEL-FINAL, the latter often involving stem change. Partitive formation varies also in whether the affix is added to a nominative (e.g., classes I, VIII), genitive (e.g., IV), or a third stem form. Affixal partitives comprise 36.8% of the noun types in the CDS corpus (shaded cells in Table 1). Vowel-final partitives (with or without stem changes) comprise 61.4% of

TABLE 1. *Noun declension classes in Estonian with examples and type frequency, adapted from Granlund et al. (2019, p. 187) and Kaalep (2012, pp. 434–445). Shaded cells indicate partitives with (-t/d) affixation. Classes IX–XII are unproductive. V = vowel, C = consonant.*

Class	Nominative form: Final phoneme; example	Genitive: Formation from Nom.; example	Partitive: Formation from Nom.; example	Stem changes	Type %, CDS
I	V-final <i>kuu</i> ‘moon’	Ø <i>kuu</i>	+ <i>d</i> <i>kuu-d</i>	None	2.6
II	Ends in <i>-e</i> or <i>-s</i> <i>päike</i> ‘sun’	+ <i>(s)e</i> <i>päikese</i>	+ <i>(s)t</i> <i>päikest</i>	If stem <i>-ne</i> , > <i>-se</i>	9.4
III	V-final <i>auto</i> ‘car’	Ø <i>auto</i>	+ <i>t</i> <i>auto-t</i>	None	7.6
IV	C-final <i>raamat</i> ‘book’	+ <i>V</i> <i>raamat-u</i>	+ <i>Vt</i> <i>raamat-ut</i>	+ <i>V</i>	12.5
V	V-final <i>maja</i> ‘house’	Ø <i>maja</i>	Ø <i>maja</i>	None	9.5
VI	C-final <i>pilt</i> ‘picture’	weak stem + <i>V</i> <i>pild-i</i>	strong stem + <i>V</i> <i>pilt-i</i>	Weakening gradation	49.2
VII	V-final <i>tigu</i> ‘snail’	weak stem <i>teo</i>	Ø (strong stem) <i>tigu</i>	Weakening gradation	2.4
VIII	C-final <i>aken</i> ‘window’	strong stem + <i>V</i> <i>akn-a</i>	weak stem + <i>t</i> <i>aken-t</i>	Strengthening gradation	2.8
IX	V-final <i>süda</i> ‘heart’	+ <i>me</i> <i>süda-me</i>	+ <i>(n)t</i> <i>süda-nt</i>	Phonemic	1.0
X	C-final <i>kaas</i> ‘lid’	+ <i>(n)e</i> <i>kaa-ne</i>	+ <i>(n)t</i> <i>kaa-nt</i>	Phonemic, + <i>e</i>	0.4
XI	V-final <i>lumi</i> ‘snow’	<i>i&gt;e</i> <i>lume</i>	+ <i>(n)d</i> <i>lu-nd</i>	Vowel change	0.5
XII	V-final <i>jõgi</i> ‘river’	weak stem, <i>i&gt;e</i> <i>jõe</i>	strong stem, <i>i&gt;e</i> <i>jõge</i>	Weakening gradation, <i>i&gt;e</i>	0.3

TABLE 2. *Estonian noun cases: form, function, example, and token frequency in Child-Directed Speech*

	Case	Form	Function	Example	% of sing. nouns
Grammatical	<b>NOM (nominative)</b>	Ø	subject, citation form	<i>raamat</i> 'book'	26.2
	<b>GEN (genitive)</b>	Ø/+V	affected object, possessive, complement of postposition	<i>raamat-u</i>	20
Semantic	<b>PAR (partitive)</b>	Ø/+V/+t/+d	partial object, negation	<i>raamatu-t</i>	18
	ILL (illative)	GEN +sse / stem- changing alternative	'into'	<i>raamatu-sse</i>	10.7
	INE (inessive)	GEN +s	'inside'	<i>raamatu-s</i>	4.6
	ELA (elative)	GEN +st	'out of'	<i>raamatu-st</i>	4.2
	ALL (allative)	GEN +le	'to', recipient	<i>raamatu-le</i>	3.8
	ADE (adessive)	GEN +l	'on', 'at', experiencer	<i>raamatu-l</i>	3.5
	ABL (ablative)	GEN +lt	'(away) from'	<i>raamatu-lt</i>	0.3
	TRA (translative)	GEN +ks	change of state	<i>raamatu-ks</i>	1.1
	TER (terminative)	GEN +ni	'until', 'up to'	<i>raamatu-ni</i>	0
	ESS (essive)	GEN +na	'as', temporary state	<i>raamatu-na</i>	0
	ABE (abessive)	GEN +ta	'without'	<i>raamatu-ta</i>	0.3
	COM (comitative)	GEN +ga	'with'	<i>raamatu-ga</i>	7.2

the nouns in CDS. Partitive and genitive are both marked through affixes, stem changes, or a combination of both (classes II, VI, VIII, in Table 1). In some classes (VI, VII, and XII, comprising 51.9% of CDS nouns and adjectives), genitive and partitive are ONLY distinguished through stem change, usually ‘weakening’ or ‘strengthening’ gradation, expressed through contrasts in syllable duration (e.g., *pilt* ‘picture.NOM’, *pildi* ‘picture.GEN’, *pilti* ‘picture.PAR’; Eek & Meister, 2003). Some noun classes have syncretic grammatical cases (e.g., NOM = GEN in class I and III; NOM = PAR in class VII; NOM = GEN = PAR, class V).

While suffixation can be characterised by rules, internal stem changes may apply to phonemes or syllables in various ways and rely more on whole-word shape and prosody than on a specified input-to-output transformation. Stem changes are central to Estonian noun declension in two ways: First, genitive and partitive formation involves stem changes in many productive declension classes. Second, all semantic case forms are based on the genitive form; hence, the genitive stem change is integral to case formation throughout the paradigm.

In summary, while the nominal inflection system is partially predictable, there are declension classes with numerous potential alternatives, not determined by phonological form alone. Crucially for inflecting novel nouns, various options can be considered grammatical in this system, as discussed in the next section.

#### 2.4. PRODUCTIVITY AND OVER-REGULARISATION IN ADULTS AND CHILDRENS USAGE

The Estonian nominal inflection system may pose several challenges for acquisition. The system is not based on semantic principles, nor is the phonology of a single form always sufficient to predict the inflected form. In order to determine how children learn the nominal morphology of Estonian, we need to know how adults generalise, and how much variability they exhibit in inflecting novel nouns. Usually, frequent nouns belong to a single acceptable declension class. Some allow more than one pattern. But for unfamiliar, lower-frequency items or novel nouns, we find variable declension even among adult speakers.

Kaalep (p.c.) claims that with “a rare or previously unseen word, speakers of Estonian immediately exhibit remarkable consensus about what is the generally accepted (i.e. normal, correct) way of forming its inflectional forms”. He gives the example of the widely used word *äpp* (derived from ‘application’, used in technological devices such as smartphones). Phonologically similar words exhibit three possible theme vowels for the genitive: *-u*, *-a*, *-i*, as demonstrated in (1). However, rare words and new loans are inflected almost exclusively with *-i*, as is the case with *äpp* > *äp-i*. This is the presumptive ‘default’ vowel for monosyllabic, consonant-final nouns. The genitive forms

ending in *-u* and *-a* are exceptional, high-frequency words, and Kaalep notes that speakers do not try *äpu* or *äpa* for the new loan *äpp*.

(1)	NOM	GEN	PAR	
(a)	<i>näpp</i>	<i>näp-u</i>	<i>näpp-u</i>	'finger'
(b)	<i>käpp</i>	<i>käp-a</i>	<i>käpp-a</i>	'paw'
(c)	<i>täpp</i>	<i>täp-i</i>	<i>täpp-i</i>	'dot'

Although speakers converge quickly on how to decline new words, there are many examples which counter Kaalep's claim of 'immediate consensus' and which show that the 'defaults' are not categorical: disagreement occurs even in productive classes. To use Kaalep's example, the name *Breivik* came into sudden, frequent use in the news following tragic events in Norway in 2011. Phonologically similar words ending in *-ik* have genitive forms with *-u* (as discussed in Section 2.2); however, *Breivik* may also be classified as a foreign name, and declined with *-i*. According to Kaalep (p.c.), out of 400 examples in the corpus, 75% used *Breivik-i*, and 25% used *Breivik-u*.

Moreover, it is not only foreign names which lead to differences in usage. When Estonian president Kersti Kaljulaid was elected by the Parliament in October 2016, she was not yet publicly well known. Her name, a compound noun meaning 'cliff islet' with a relatively infrequent head noun, could be inflected according to the real noun (NOM, GEN, PAR: *laid, laid-u*) or a more generalised declension for novel nouns or names (*laid, laid-i, laid-ī*). One article published on the Estonian National Broadcasting service included different partitive inflections within the space of four sentences:

- (2) (a) *President Toomas Hendrik Ilves tervita-s Kersti*  
 President T. H. Ilves greet-PAST.3SG K.
- Kaljulaid-i*<sup>1</sup>  
 Kaljulaid-PAR.SG  
 'President Toomas Hendrik Ilves greeted Kersti Kaljulaid.'
- (b) *Ilves lisa-s, et tunne-b Kersti Kaljulaid-u rohkem kui*  
 Ilves add-PAST.3SG that know-PRS.3SG K. Kaljulaid-PAR.SG more than  
 22 aasta-t.  
 22 year-PAR.SG  
 'Ilves added that he has known Kersti Kaljulaid for over 22 years.'

Searches on the Internet that same day turned up results with both names, with 80% in favour of the *-i* inflection. Two years later, after longer national exposure and explicit discussion in the media of the president's own preferences regarding the inflection of her name, searches on Google turned up about 95% responses with the *-i* form preferred by the president herself. The variability in adults' production indicates reliance on analogy with existing forms to inflect new words.

[1] <<http://uudised.err.ee/v/eesti/08e8ef70-ef86-48da-9b94-67a15bd637ee/fotod-president-ilves-tervitas-kadriorus-kersti-kaljulaidu>>

In the face of this level of variability in the system, how do children inflect nouns? Variability in certain parts of the system can lead to children's errors in other parts. Corpus data reveal that error rates with real nouns in early child Estonian are low, but not negligible, as observed by Argus: "at 1;7, there are errors with approximately 12% of noun tokens and 20% of noun types" (2009, p. 129). With older children, we find errors from overgeneralisation (as in (3)), stem selection, and stem change. Both examples (3) and (4) come from the Vija corpus (Vija, 2007), available on the CHILDES database.

- (3) *Adu*                      *kuula-b*                      *tädi\*-t* [: *tädi*]      (2;0.2)  
 Andreas.DIM    listen-to.3SG.PRS    auntie-PAR.SG  
 'Adu is listening to the lady.'

In (3), the partitive *-t* affix is erroneously used with a class V noun, in which the partitive should be syncretic with the nominative. Although the *-t* ending is not the most frequent partitive formation, it is the most transparent and regular.

Errors also occur with vowel-final partitive endings used in place of targeted *-t* forms, as in (4).

- (4) *Õ sööb*                      *\*vahvel-i* [: *vahveli-t*]      *kõmps^kõmps@o*      (2;0.16)  
 eat.3SG.PRS    waffle-PAR.SG                      crunch-crunch  
 '[the guinea pig] is eating a waffle, crunch-crunch'

However, here we cannot be sure that other factors are not involved. Example (4) requires an object in the partitive case, but the error may be morphosyntactic, involving erroneous object marking (see Vihman et al., 2020). These examples show that children do not always default to a single, affixal pattern.

In their elicitation experiment using real nouns, Granlund et al. (2019) found that Estonian-speaking three- to five-year-olds made overt errors in 8.5% of trials, but also used avoidance strategies (e.g., replacing the modelled case form with another form). Both word form frequency and declension class size affected accuracy among the Estonian participants. Stem change was implicated in a majority of errors.

### 3. Research questions and hypotheses

Our elicitation study of case forms using novel nouns was designed to test the following research questions and associated hypotheses:

- RQ1. Do children's responses show *development toward those of adults*?  
 Hypothesis 1. Children's responses will not be fully adult-like at age 5, but accuracy (=similarity to adults' responses) will increase with age.  
 Hypothesis 2. Children's responses will vary more than adults' responses; variability will decrease with age.
- RQ2. How do PRESENTATION AND TARGET CASE affect children's responses?

If children's case formation is based on analogy, then accuracy will vary according to target case; because of differences in frequency of similar forms in the input, presentation case is also expected to affect accuracy (*Hypothesis 3*). We also expect the greater input frequency of vowel-final partitive formation to affect accuracy despite the transparency and distinctiveness of affixal partitives (*Hypothesis 4*).

Hypothesis 3. Target and presentation case will both affect accuracy.

Hypothesis 4. Children will not use the affixal PARTITIVE as a generalised default, but rather use both affixal and vowel-final patterns.

RQ3. Do children make greater use of NEIGHBOURHOOD DENSITY to inflect novel nouns than adults? Although Granlund et al. (2019) did not find an interaction of PND with age in Estonian, they did find an interaction of PND with form frequency in pooled data across three languages.

Hypothesis 5. 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.

## 4. Method

Because children's accurate use of real verbs and nouns may be ascribed to either productive grammatical knowledge or rote memory of individual forms, Berko devised the wug test (Berko, 1958) using novel words. This test demonstrated that children learning English can inflect new words through generalising what they know. It has since been adapted and used for various aims in various languages, but it has not been used widely to test knowledge of morphological variability (though see Hayes & Londe, 2006, who probed speakers' intuitions regarding lexically idiosyncratic vowel harmony in Hungarian). The wug design has been used at least once with Estonian children (Vaik, 2014). Also relevantly, Lyytinen (1987) used a similar method to assess children's morphological productivity in closely related Finnish, and devised a standardised Finnish wug-style test (Lyytinen & Lyytinen, 2004).

Although the present study was not preregistered, all relevant materials have been made accessible for replicability through the Open Science Framework. The experimental materials (experimental protocol, verbal and visual stimuli), corpus-based PND measure, response datasets, and analysis scripts are available on the project site <<https://osf.io/3srw8/>>.

### 4.1. PARTICIPANTS

We recruited 70 children aged three to five through pre-schools in Estonia in the cities of Tartu and Pärnu.<sup>2</sup> None had any known language delays. Four

[2] Before recruitment and testing, we obtained ethics approval from both the University of Manchester's Research Ethics Committee (UREC ref. 15287) and the Research Ethics Committee of the University of Tartu (protocol nr 249/T-14).

TABLE 3. *Summary of participants*

Age group	3-year-olds	4-year-olds	5-year-olds	Adults
N (female: male)	23 (15: 8)	23 (14: 9)	20 (10: 10)	21 (12: 9)
Age range in months (years)	36–47 (3;0–3;11)	48–58 (4;0–4;10)	60–72 (5;0–6;0)	(21–53)
Mean age in months ( <i>SD</i> )	42.43 (3.42)	52.61 (3.23)	65.45 (4.07)	396

bilingual participants were excluded from the analysis, resulting in a total of 66 children, ranging in age from 3;0 to 6;0. In order to gain a valid picture of acceptable baseline responses against which to compare children's responses, we tested 21 adults (aged 21 – 53, all native speakers). Table 3 shows a summary of the participants.

#### 4.2. MATERIALS AND DESIGN

Fifteen nonce nouns were devised based on productive declension classes in Estonian. The nonce forms reflected criteria distinguishing nouns in the most frequent declension classes, considering syllable structure, final phoneme, and the medial consonant. We aimed to cover each class with more than 2% type frequency in child-directed speech (as reported in Granlund et al., 2019, p. 187; see Table 1). The fifteen nonce nouns are shown in Table 4, along with their declension classes (see Table 1), according to the responses given by adult participants. A noun form is often compatible with more than one declension class. Most real nouns have one declension pattern, although some allow parallel forms, or what is also known as overabundant morphological paradigms.

Fifteen novel creatures were created and drawn in colour, to represent the fifteen novel nouns. Four pictures depicted each character in various contexts: two for presentation and two for elicitation, resulting in a total of 60 pictures. See Appendix 1 for the pictures.

The nouns were presented in one of two cases, NOMINATIVE or ALLATIVE. Allative was chosen as a presentation case because (a) its form contains the genitive stem plus an invariable affix, and (b) it is used frequently with animate referents. The script included two sentences to describe the picture with the target word, once sentence-finally and once sentence-initially, followed by a prompt for the child to repeat the form. The nominative presentation script read: “Look! This is a ‘*wug*’ [*nonce noun*]. A *wug* is this kind of {funny/friendly/brown/furry} animal. Can you tell me what it's called?” Allative nouns were presented using either the “X is pleasing *to-wug*” construction (“*Wug* likes X”) or “X visits *wug*” (literally “visits *to-wug*”), in order



TABLE 4. *Nonce noun stimuli, with declension classes according to adult responses*

	Presentation forms		Declension class
	Nominative	Allative	
1	esu	esu-le	III, V
2	keemel	keemli-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

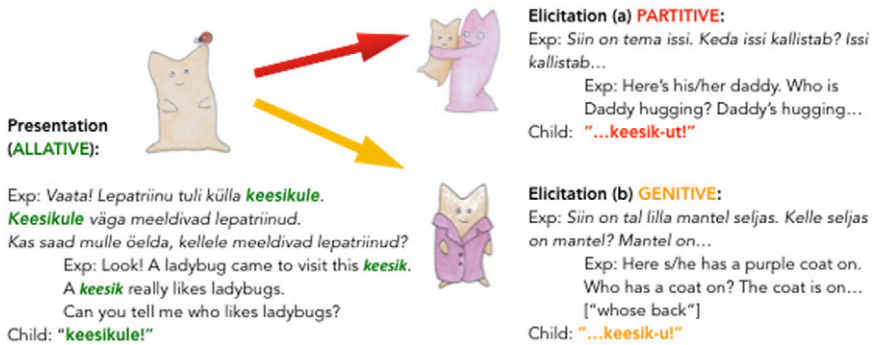


Fig. 1: Example of a trial with keesik in the allative presentation context. The child is shown two subsequent pictures and prompted to produce partitive and genitive forms of keesik.

to alleviate noticeable repetition with this unusual presentation context. An allative example is given in Figure 1; the creatures used can be seen in Appendix 1. Each child saw one presentation picture for each character, and two elicitation pictures.

After introducing the character and novel word in one of the two presentation conditions, the experimenter elicited GENITIVE and PARTITIVE forms, using pictures for each context. Genitive case was elicited using post-positional phrases, e.g., “on wug’s head”, “on wug’s tail”, “on wug’s back [=an idiomatic way to describe wearing clothes]”. Partitive case was elicited as a

direct object with verbs governing partitive complements (e.g., *katsub* ‘touches’, *silitab* ‘pets’, *kallistab* ‘hugs’, *kammib* ‘combs’, *aitab* ‘helps’). Children indicated understanding of the targeted case in responses to training questions with familiar nouns taken from a different experiment. An example of a training trial is: “Look, this is a door [*uks-NOM.SG*]! Here s/he is measuring it [*седа-PAR.SG*]. What is she measuring?”

We used two training items and 12 fillers. Together with two elicited forms for each novel noun, each participant responded to a total of 44 trials. Children of all ages performed nearly at ceiling on fillers with real nouns, for which there was only one correct answer.

#### 4.3. PROCEDURE

Children were tested individually in a quiet corner in their pre-schools. The experimenter (E) described the task to the child as a game and asked for assent to play the game. After presenting real nouns for warm-up, E presented each character as described above and prompted the child to repeat the same form. If the child was unsure or could not produce the form, E repeated the presentation sentences and asked again. If the child was still unsure or unwilling, the response was coded as missing. Children were rewarded with a sticker, regardless of how many responses they had given. Adults were tested with the same procedure. Responses were recorded on audio and also manually.

#### 4.4. MEASURES

Responses were analysed for VARIABILITY (number of different responses given per item), ACCURACY (similarity to adult responses), and the effect of NEIGHBOURHOOD DENSITY (response similarity to analogical familiar nouns). These measures are described in detail below. Adults’ responses were used to determine accuracy, and to compare the variability and neighbourhood density of children’s responses. We used linear mixed effects regression models to test for significant differences between age groups and conditions for each measure.

In addition, children’s non-adult-like responses were analysed separately. We wanted to know whether non-adult-like responses indicated adherence to the morphological system overall. For genitive targets, we looked at whether responses ended in a vowel. For partitives, we looked at whether children used only vowel-final or suffixal (*-t/-d*) forms, or used both productively.

##### 4.4.1. Variability

We examined how many different responses were given per item by participants in each age group. We compared overall means (number of unique

responses) in each age group for each condition (presentation-elicitation context).

#### 4.4.2. Accuracy

Children's responses were coded for matching the adults' responses. In the variability analysis, adults showed significant differences in variability by target case; we did not want this to bias the accuracy score of children's responses. We aimed to capture the variability inherent in the system, but to exclude responses given by very few adults. Hence, we defined a core of acceptable responses using a 'majority vote' rule.

This measure treated as Unique Target Forms any responses which converged in two-thirds of adults for a particular item. Items with no single response achieving a 66% majority were treated as Variable Targets. For instance, the prompt *palas* had adult responses split between two different 'theme vowels' (GEN: 45% *palas-e*, 55% *palas-i*), and so was considered a Variable Target; both responses were considered acceptable forms in the children's data. Variable items were taken to have genuine alternative patterns among adult participants. For Variable Targets, all adult responses were counted as acceptable, and children's responses were marked accurate if they matched any of them.

#### 4.4.3. Neighbourhood density

We used a phonologically based measure of neighbourhood density rather than the declension classes defined in Estonian grammars. As shown in Table 4, adults gave responses which reflected various possible declension classes for nearly all the prompts. Moreover, the classes may not be psychologically valid for children, as they depend on knowledge of the wider morphological paradigm. Young children lack this knowledge when learning vocabulary, and all participants lacked this information for the nonce nouns in our experiment.

Hence, we made use of a form-based measure of Phonological Neighbourhood Density (PND). PND defines neighbourhood with respect to the phonological properties of each individual form, based on the analogical model in Albright and Hayes (2003; for more detail on the measure we used, see Granlund et al., 2019, p. 188). This PND measure of inflectional forms defines neighbourhood based on the transformations (insertions, deletions, substitutions, affixations, or suffix drops) required to derive a target form (e.g., *esugen*) from the base (presentation) form of the same noun (e.g., *esule-ALL*; in this case, one: suffix drop).

Neighbours were defined according to similarity of the transformations: all forms defined as 'neighbours' have the same number of each type of

transformation and differ by a maximum of one bit. The summed similarity of the experimental items was divided by the summed similarity of all corpus items in that case condition, resulting in a summed similarity ratio: our PND measure. PND thus represents both the support of the phonological neighbourhood for successful analogy as well as interference from ‘enemies’.

The PND measure was derived from forms present in a corpus of child-directed speech (CDS, drawn from the Argus (1998), Vija (2007), and Zupping (2016) corpora available on CHILDES), as reported in Granlund et al. (2019). The CDS is addressed to children aged between 1;3 and 4 years of age, and includes over 171,000 word tokens. In this study, to evaluate whether the same measure predicted generalisation of morphological paradigms to unfamiliar words, we used the PND measure in reverse. Children might produce responses using analogies disregarded by adults. In order to determine to what degree children made use of existing analogical forms, we evaluated the PND of all the forms they produced, not just adult-like ones, and compared these across age groups.

## 5. Results

We analysed the responses given according to variability, accuracy, and PND. Models were run in R using the *lmer* or *glmer* function of the *lme4* package. In order to calculate chi-square values and respective *p*-values, the *Anova* function from the *car* package was used. The quantitative results are reported first, followed by an error analysis of non-target-like responses in Section 5.4.

### 5.1. VARIABILITY

Mean numbers of distinct responses for each presentation-elicitation case condition in each age group are shown in Figure 2. Most items received at most two different adult responses, with means between 1.33 (nominative-genitive) and 2.07 (nominative-partitive). Adults showed more agreement (lower variability) in genitive forms, and more disagreement (higher variability) in partitive forms. In the genitive condition, two items generated 3 distinct forms; in partitive, one item was given 3 and two items were given 4 distinct forms.

As Figure 2 shows, the greatest variability in responses was among the younger children and in the allative-partitive condition. To test the significance of the difference in variability between children and adults, we fitted a linear mixed effects regression model with the number of unique forms per item as the dependent measure predicted by group (adults vs. children), presentation case, target case, and an interaction of presentation and target case. Age group was treatment-coded with Children used as the baseline

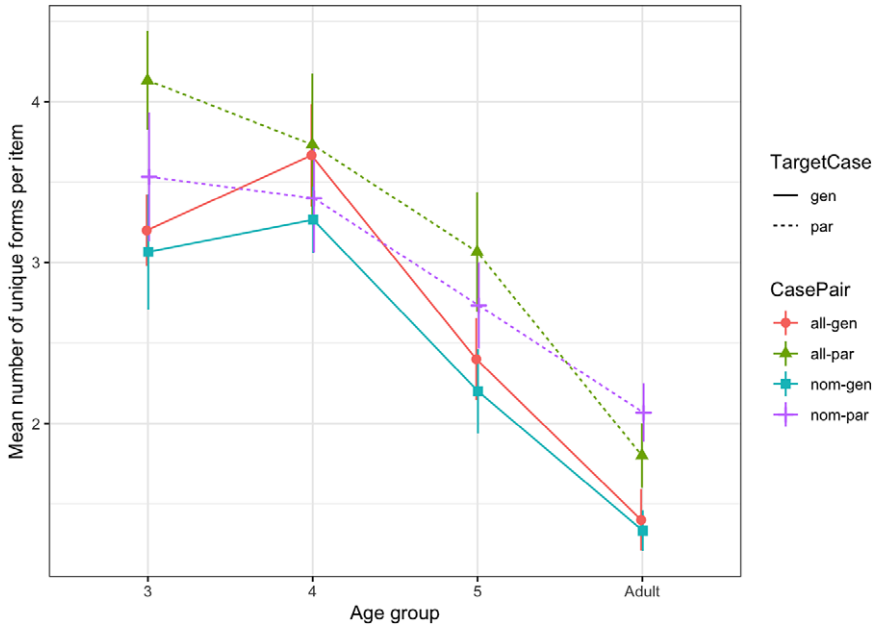


Fig. 2: Variability: mean number of unique responses, by age group and case condition.

(Children = 0, Adults = 1). Treatment coding was also used for presentation case and target case, with allative and genitive as the baseline, respectively. Random intercepts were fitted for items only (variability was aggregated over participants for each item). We first tried running a model with a maximal random structure for items, i.e., random intercepts with slopes for all fixed terms. Because this model had convergence issues, we simplified the random structure by removing random slope terms one by one until the model ran without issues. The detailed steps of this procedure can be found in the analysis R code provided at <<https://osf.io/3srw8/>>. The final model retained only random intercepts for items but no random slopes.

Adults showed significantly less variability ( $\beta = -0.311$ ,  $SE = 0.031$ ,  $p < .001$ ) and the model showed a significant effect of target case, with partitive eliciting more variable responses than genitive. For the model output, see [Appendix 2a](#).

For further analysis of whether variability decreases with age for children, we ran a similar model with adults excluded, using age in months as a continuous measure. As in the previous model, a step-wise simplification of the random structure resulted in a converging model with by-item random intercepts only. Age in months was a significant predictor of variability ( $\beta = -0.013$ ,  $SE = 0.003$ ,  $p < .001$ ). Target case was just below significance in

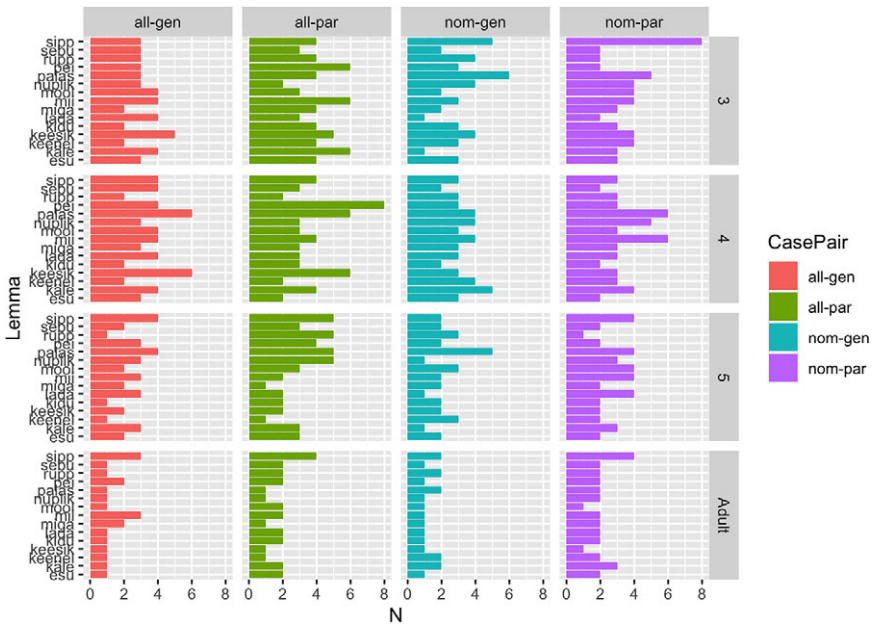


Fig. 3: Variability per item: number of unique responses per lemma, by case pair and age group.

the model with children only ( $\beta = 0.055$ ,  $SE = 0.050$ ,  $p = .058$ ). For the model output, see [Appendix 2b](#).

[Figure 3](#) shows variability per item. Items which elicited variable adult responses usually elicited variability in children's responses. However, many items with variability in children's responses received unanimous responses from adults.

For all presentation-elicitation conditions, we found that the five-year-olds had less variable responses than younger children, and adults had significantly less variability overall than children. Note that adults showed some variability across all conditions, indicating the lack of a default paradigm for any individual form, as discussed in [Sections 2.1](#) and [2.4](#).

## 5.2. ACCURACY

Results for target-like accuracy are shown in [Figure 4](#). These compared children's responses to targets, defined as either a Unique Target, where the adults converged around a preferred response, or Variable Target, when adults did not show a majority preference. For Variable Targets, all adult responses were included.

Five-year-olds responded differently from the younger age groups, moving toward adult patterns of responses, most clearly with genitive targets. With

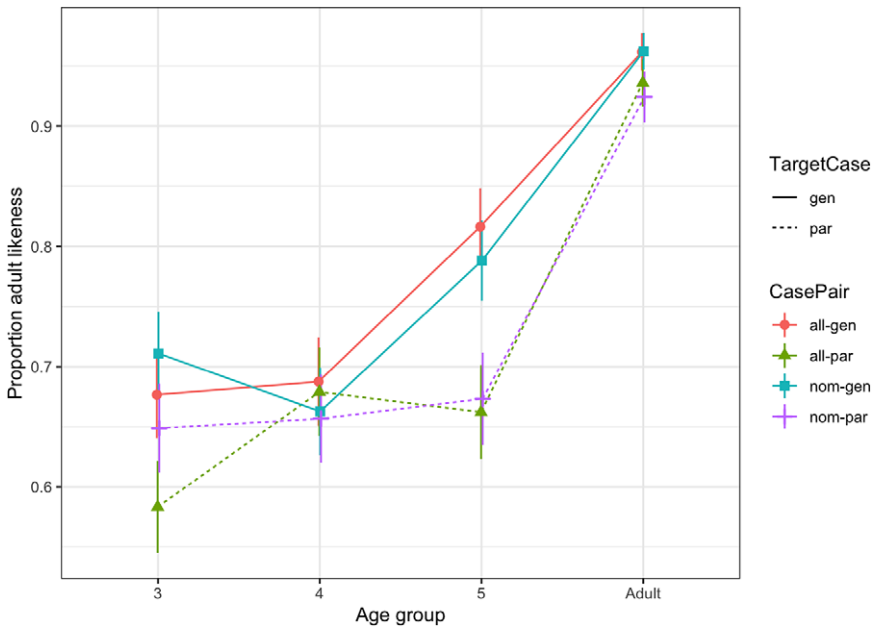


Fig. 4: Accuracy of responses, by case pair and age group.

partitive targets, the five-year-olds can be seen to differ from three-year-olds in the allative condition, while four-year-olds give similar responses to the older children.

We fitted a binomial mixed effects regression model for children only, with a binary measure of accuracy as the dependent variable. As before, we first tried running a model with a maximal random structure for items and participants, i.e., random intercepts for each and slopes for all fixed terms. Predictors were presentation case (NOM, ALL), target case (GEN, PAR), continuous age in months and the two-way interaction between presentation and target case. Age ( $\beta = 0.026$ ,  $SE = 0.006$ ,  $p < .001$ ) and target case condition ( $\beta = -0.424$ ,  $SE = 0.146$ ,  $p < .001$ ) were significant predictors of accuracy. Presentation case was not significant ( $\beta = -0.009$ ,  $SE = 0.150$ ,  $p = .618$ ). As can be seen visually in Figure 4, the three- and four-year-olds did not differ greatly in their accuracy scores. Five-year-olds performed similarly to younger children in the partitive conditions but showed more adult-like responses for genitive. Adults also converged more in the genitive responses. For the model output, see Appendix 2c.

In summary, we found development towards adult-like accuracy at age five, and differences by elicited target case, but not by presentation case, nor any interaction between target and presentation case.

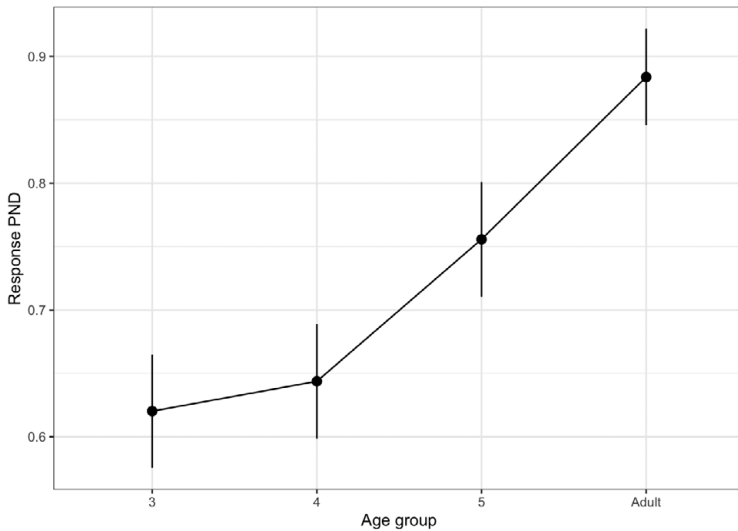


Fig. 5a: PND values of responses for allative-genitive case pairs, by age group.

### 5.3. NEIGHBOURHOOD DENSITY

We implemented a PND measure to investigate the effect of form-based analogy to case forms of real lexical items in the CDS. Given that children may base analogies on different transformations than adults, we evaluated the PND of all responses, accurate (adult-like) and inaccurate, to see how the neighbourhood density of analogical case-form pairs was reflected in responses across ages.

Figures 5a and 5b plot the mean PND values of the responses within each case pair (presentation-elicitation) given by participants for the experimental items. Figure 5b gives the PND values for three case pairs, while Figure 5a shows the allative-genitive PND separately, with mean values much higher than for the other three pairs. Recall that the genitive form is the stem of the allative form, meaning that there is little ambiguity regarding the relation between the allative and genitive forms; this explains the high PND of this particular case pair. Even three- and four-year-olds were able to tap into this regularity. Note that in the nominative-partitive condition, adults do not have recourse to high PND, and do not make more use of analogy than children.

We fitted a linear mixed-effects regression model to test for the effects of age (grouped as 3, 4, 5, adult), presentation case (binary), target case (binary), and the two- and three-way interactions on the PND value of the response (continuous, log-transformed, scaled, and centred within case condition), i.e., whether the extent to which speakers relied on PND in producing their



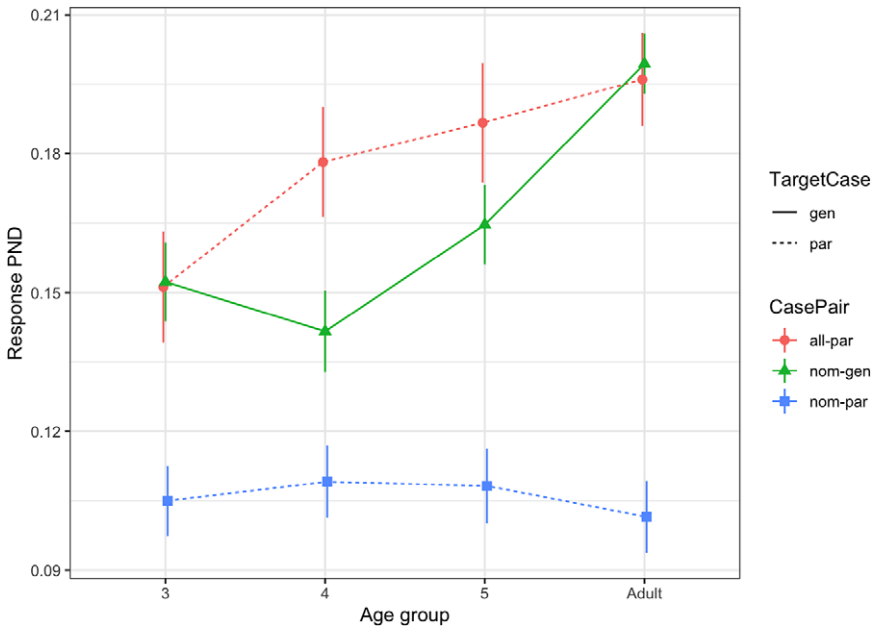


Fig. 5b: PND values of responses by age group and case pair, remaining three pairs.

response varied as a function of these factors. Random intercepts were used for participants and items. We found a significant positive effect of age group ( $p < .001$ , estimates with age 3 as reference: 0.04, 0.26, 0.50) and no main effect of presentation or target case condition ( $p = .97$  and  $.98$ ). The interaction of age and target case, however, was significant ( $p < .001$ , change in age group estimates: 0.14,  $-0.003$ ,  $-0.16$ ), such that the effect of age was smaller when the target case was partitive. For the model output, see [Appendix 2d](#). To test for effects within children, we ran the same model, excluding adults and using age in months as a continuous predictor (scaled and centred). Results showed a significant positive effect of age ( $\beta = 0.1$ ,  $SE = 0.05$ ,  $p = .036$ ) and a significant effect of target case ( $\beta = 0.05$ ,  $SD = 0.06$ ,  $p = .034$ ), such that the response PND was higher when the target case was partitive. For the model output, see [Appendix 2e](#).

In summary, children more often produce target forms with larger neighbourhood density as their linguistic knowledge grows, and adults draw on larger neighbourhoods to produce case forms of novel nouns. As both vocabulary and grammatical repertoire expand, speakers are able to draw on a larger bank of real noun forms to generalise to novel nouns.

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%

#### 5.4. ERROR ANALYSIS

In the analyses above, we examined WHETHER children produced the same forms as adults. We now look at HOW children's responses diverged from those of adults, first in genitive, then in partitive responses.

For genitive targets, adults nearly unanimously gave vowel-final forms. Variability in adult genitive forms was in the selection of the word-final vowel. Table 5 shows proportions of vowel-final responses by age group. Note that only vowel-final genitive forms are acceptable responses, though not all vowel-final responses follow existing paradigms. Children gave vowel-final genitive forms in 45–94% (by item) of nominative presentation contexts and 76–94% of allative contexts. The non-target-like consonant-final genitive forms were mostly unchanged consonant-final stems, such as *keesik*, *mool*, *palas*, or forms retaining part of the allative (*-le*) affix, such as *sebu-l*, *rupi-l*, as well as some examples of case substitution, e.g., partitive-like *t*-final forms like *sipi-t*.

Despite the unfamiliar context of learning a word in allative case and the unusual allative-to-genitive task, three-year-olds were more target-like and showed less variability in allative than nominative presentation contexts. This indicates that they were able to use their knowledge of case forms productively in an unusual task, and they were able to make use of the allative STEM, which is identical to the genitive case form. Three-year-olds produced more vowel-final forms in the allative-genitive task than the older children. The four- and five-year-olds' consonant-final responses resulted from either stripping off more than the allative affix and using a consonant-final stem, or else failing to strip off enough of the affix; both patterns suggest that they may have attended more to the unusual task of transforming an allative form.

Partitive targets elicited more variability among adults and children than genitive: adult responses differed not only in the choice of 'theme vowel' for vowel-final responses, but also in whether they ended in a vowel or consonant. We look here only at the distribution of consonant-final (*-t/d*) and vowel-final forms, which are found in different declension classes (Table 6). Overall, children produced vowel-final responses more than adults. Adults preferred affixal partitives, but not across the board: in the nominative presentation context, adults gave consonant-final responses in 70.5% of trials, with 29.5% vowel-final responses. Children gave consonant-final forms in only half of nominative contexts overall. Allative presentation case led to more affixal

TABLE 6. *Partitive targets: proportions of affixal (consonant-final) responses*

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

responses in all groups, with children still using vowel-final forms more than adults.

Children's elicited partitive forms involved great variability, even among consonant-final responses, suggesting that the production of even affixal responses did not draw on a simple affixation rule, but rather targeted forms analogous to existing forms, sometimes with added syllables. The nonce stimulus *lada*, for instance, elicited the *t*-final forms *lada-t*, *lada-set*, and *lada-st*, while *käle* was inflected as *käle-t*, *käle-set*, and *käle-lit*. *Sipp* inspired *sipe-t*, *sipi-t*, *sipu-t*, and *sipi-kut*: theme vowel selection is the primary point of difference, but here too we find added syllables.

We coded the types of errors children made in deviations from adult responses and found that the greatest proportion of errors consisted of failure to remove the allative affix (*-le*). This was indicated by the presence of the allative affix (*-le*), or part of the affix (*-l*), in 47% of errors. This was followed by unchanged lemmas (17% of errors). Unchanged forms (where change was required) were given only when the stimulus was presented in the nominative case. This indicates a subtle comprehension of the different case forms and the relations between paradigm cells. In allative presentation contexts, a similar error type was the use of an unchanged stem with only the suffix removed (6% of allative contexts). In other words, children did not always inflect the stem correctly, but they showed mastery of the difference between a word learned in allative and nominative form.

Other errors included the omission of relevant phonetic material (e.g., *keesi* < *keesik* or *pala* < *palas*; 13% of errors), the production of the wrong case (e.g., partitive in place of genitive; 4% of errors), and different choice of theme vowel (e.g., *keesik-at* for *keesik-ut* or *sipe* for adult *sipi* or *sipu*; 4%). The selection of theme vowel was a more prevalent error in nominative contexts (6% vs. 2% for allative). Finally, an additional syllable was provided in 8% of nominative contexts, but in no allative contexts.

From this descriptive summary, we see that the errors children made were generally in keeping with the morphological system. They recognised that allative forms need to be changed when switching cases, while nominative forms can be syncretic with the genitive and partitive. They successfully provided vowel-final genitive forms in over 80% of trials overall. They were less adept at selecting the appropriate theme vowel for inflected

forms: unsurprisingly, considering the extent of lexical idiosyncrasy. Generalisation requires familiarity with the lexical system. Finally, children were least adult-like with partitive forms, not because of defaulting to an affixal rule, but the contrary: they preferred vowel-final forms, which involve stem changes, gradation, and theme vowel selection. Analogy to prevalent patterns in the input outweighed any possible preference for a rule-based transformation.

## 6. Discussion

In our elicitation study, we investigated three- to five-year-olds' ability to inflect unfamiliar nonce nouns according to the morphological system of Estonian. Because the system involves variability, we compared children's responses with adult responses to the same stimuli: importantly, adult participants lack consensus in roughly half the conditions. Despite low overall accuracy and much variability in responses, even three-year-olds used the same predominant declension patterns as adults. Young children make use of the available patterns, but find generalisation difficult. We first review the results of our study below with reference to each of our hypotheses.

Our first hypothesis, *Children's responses will not be fully adult-like at age 5, but accuracy will increase with age*, was confirmed. Five-year-olds differed in their responses from both three-year-olds and adults. Accuracy improved with age but responses were not fully adult-like by age five. Although Granlund et al. (2019) demonstrated high proficiency with the case morphology of familiar nouns among Estonian-speaking five-year-olds, the results from the experiment with novel nouns suggests that their productivity with the system is limited.

*Hypothesis 2, Children's responses will vary more than adults' responses, and variability will decrease with age*, was confirmed. Adult responses varied significantly less than children's responses, and age significantly predicted the variability of children's responses. Five-year-olds have knowledge of the system but are not able to productively apply this knowledge equally to all forms in an adult-like way.

*Hypothesis 3, Target and presentation case will both affect accuracy*, was only partially confirmed. Presentation case was not found to significantly affect accuracy, unlike the findings of Krajewski and colleagues for novel nouns in Polish (2011). This finding lends further support to an analogy-based model in which children derive schema-based output from diverse input, rather than applying rules operating on specific input forms to derive output forms.

Target case, however, was a significant predictor of accuracy (and the other measures), with partitive showing lower accuracy and more variable responses than genitive across all ages. The role of stem changes contributes to the challenge of learning the system. Knowing which stem changes apply and when to apply them is much more challenging in partitive case formation than

genitive case, as discussed in Section 2.3. We suspect that this explains why five-year-olds do not produce adult-like partitive forms, despite approaching adult-like inflection in the genitive contexts.

The effect of target case may have also been partly influenced, as a reviewer suggested, by other differences in elicitation contexts, as the contexts for partitive case (object of verbs like 'helping') were arguably more abstract than genitive contexts (object of spatial postpositions like 'on'). However, note that adults were also much more variable in partitive than genitive formation, reflecting the inherent difference in predictability within the paradigms of the two cases. The main difficulty with partitive case formation was not in comprehending THAT partitive nouns were required, but rather HOW to form them. This is more likely due to the parallel presence in the morphological system of different possible partitive case formation paradigms (as outlined in Table 1). Predicting which declension is appropriate for a novel noun is not straightforward, as attested by the variability in adult responses.

The presentation cases in our study placed differing task demands on the participants. Words are usually encountered and learned first in the nominative form, which makes up 26% of singular noun tokens in the CDS, while allative case makes up less than 4% and is not a typical case for first encountering and learning words. The task required in the two presentation contexts was unequal in terms of frequency, typicality, and complexity. Yet, despite the inequality of the two contexts, differences in processing difficulty between conditions did not affect response accuracy overall. Instead, what mattered was the elicited target case. Even the youngest children were able to identify the genitive stem used in allative case forms. Elicited case was a significant predictor of accuracy, variability, and PND, while the presentation case was not significant for any of these. This is compatible with a reliance on output schemas which characterise the shape of the whole word form. Schema-based production shapes the output according to whole-word properties and analogy with known forms.

Hypothesis 4, *Children will not use the affixal PARTITIVE as a default*, despite its transparency, was confirmed. Rather than using an affixal (-t/d) pattern, children were more likely than adults to provide vowel-final forms. Their responses confirm that no single declensional rule underlies generalisation in Estonian (cf. Mirkovic et al., 2011). Even when the affixal option is available, children show no preference for using it, instead making use of vowel-final patterns which are more frequent in the input.

Hypothesis 5, *Children's responses will be associated with greater neighbourhood density with increasing age*, was confirmed, but modulated by target case. Age was a significant predictor of the PND of the responses, with greater PND for adult responses and older children, and no main effect of condition in the model including all ages, but an interaction of age and target case, such that the

effect of age was smaller for the partitive. In the children-only model, both age and target case were significant.

Recall that Granlund et al. (2019) predicted a positive interaction between neighbourhood density and age, but found the opposite interaction in Polish and Finnish, and no relationship between age and PND in Estonian. Our study, however, using novel rather than familiar nouns, confirms Granlund and colleagues' prediction: older children and adults were shown to rely on PND more than younger children. This suggests that, while analogy provides the most reliable mechanism for inflecting novel or low-frequency words, it requires substantial experience with the language. Five-year-olds may not be as productive with the morphological system as adults, but they have a greater store of experience with the lexical items and their inflectional patterns to draw on than three-year-olds. The younger children show no tendency to resort to a default, even when they are less proficient at using PND-based analogy.

This study with Estonian novel nouns adds to the mounting evidence for analogical, similarity-based acquisition of inflectional systems. Analogy is a general cognitive mechanism for pattern-forming and pattern-recognition: it may constitute the "core of human cognition" (Blevins & Blevins, 2009, p. 1). Evidence of the centrality of analogical processing in human cognition draws on non-linguistic and linguistic domains. For language, word-level analogy is most frequently discussed, but when phonological, morphosyntactic, and semantic information align, grammatical analogy is a powerful tool for generating a productive system out of the confluence of multi-level linguistic information. Based on the significance of target forms rather than presentation forms in the children's responses, as well as their non-reliance on an affixal declension and readiness to use stem-changing forms, we understand output schema-based analogy to underlie the children's responses in our study. They relied on whole-word morphological similarity rather than morpheme concatenation. Moreover, the robust variability present in adults' responses underlines the importance of probabilistic information, the reliance on analogical patterns, and the equal availability of multiple paradigms rather than a default, even for novel words.

The original wug test was interpreted as showing the productivity of rule-based knowledge. In our study, children gave many responses that, despite being non-target-like in not exactly matching adults' responses, were compatible with the system in general. They produced forms drawing on various declension classes, rather than defaulting to one. Although they often produced different forms than adults, even three-year-olds displayed a basic understanding of the function of the different cases and the relations between forms. Children's responses demonstrate a sensitivity to whole-word forms and the importance of output schemas derived from the system. The use of forms with support from more similar neighbours – or the increasing

importance of PND with age – forcefully demonstrates the growing reliance on analogical patterns with age and experience.

The results of this study do not provide support for a radical exemplar model of language acquisition, as they show reliance on some level of abstraction even among the younger children, but they strongly support the similarity-based side of the theoretical continuum discussed in Section 2. Further work is needed to distinguish between connectionist (e.g., Mirković et al., 2011) and Usage-Based models which give priority to abstract schemas (e.g., Bybee & Moder, 1983; Bybee & Slobin, 1982; Maslen et al., 2004). The former are more heavily invested in specific input–output mappings, but may involve greater levels of abstraction in hidden layers (see Engelmann et al., 2019), while the latter emphasise the abstract schemas derived from exposure to individual exemplars. These may, in effect, describe the same cognitive representation via different methods and metaphors. Because our study found greater use of PND with age to produce more uniform responses, it indicates a reliance on abstract representations which are strongly linked to morphological, output-based schemas. At the younger ages, the preference for stem-changing forms over the affixal partitive indicates a whole-word approach to nominal morphology. Both the interaction between PND and age, and the lack of a default, lead us to conclude that analogy underlies productivity with nominal inflection in Estonian.

The data from the Estonian noun system provide additional counter-evidence to the rule-based view on three counts. First, in a system with so many viable declension classes, we would need strong evidence of one class providing a default declension pattern in order to know how to identify the 'regular' rule which a child would learn. Neither the frequency distribution of noun classes in the corpus nor the results of our study support the idea of a privileged status for any declension class. Second, a rule-based view would have children learn the rule and apply it across the board, causing errors of over-regularisation before the lexical irregularities of the system are learned. Yet we found continuous development through to age five. Finally, the most transparent way to formulate a transformation from input to output forms is to define a concatenative rule using an affix. In the Estonian system, this is best tested with the partitive case, which has both stem-changing and affixal alternatives. In our study, affixal partitives were preferred by the adults, but not by the children, who provided far fewer affixal responses than adults. The children did not apply the most transparent available affix to form partitives, but used both affixation and vowel-final forms. Participants based their responses on analogical similarity to real nouns, and did so more effectively at older ages.

Children followed the basic principles of Estonian nominal morphology. Although Vaik (2014) found that roughly half of children's errors involved

stem changes, in our study children did not avoid stem-changing declensions. For the genitive case, children's responses were systematically vowel-final, even if the selected vowel did not match the adult targets. This suggests that the children's emergent morphological knowledge is productive before it is complete, and can be flexibly applied.

Children show flexibility regarding the word-learning context. They are able to strip the allative case-marker off a word heard only in the allative case and transform the novel stem to a new case. Children have more limited knowledge than adults of the paradigms overall, as well as lacking access to much of the information adults use in assigning declension classes to new nouns or names. Yet newly acquired words must eventually be inflected when speaking. Although younger children were less proficient at using analogy to phonological neighbours to produce target cases, they produced forms which were in line with the morphological system. Children's ability to productively inflect new words was shown to reflect a reliance on analogy, early sensitivity to systematic morphosyntactic properties of the language, and gradual development.

## REFERENCES

- Albright, A., & Hayes, B. (2003). Rules vs. analogy in English past tenses: A computational/experimental study. *Cognition*, **90**(2), 119–161.
- Ambridge, B. (2010). Children's judgments of regular and irregular novel past-tense forms: New data on the English past-tense debate. *Developmental Psychology*, **46**(6), 1497–1504.
- Ambridge, B. (2020a). Abstractions made of exemplars or 'You're all right, and I've changed my mind': Response to commentators. *First Language*, **40**(5–6), 640–659.
- Ambridge, B. (2020b). Against stored abstractions: A radical exemplar model of language acquisition. *First Language*, **40**(5–6), 509–559.
- Ambridge, B., Kidd, E., Rowland, C. F., & Theakston, A. L. (2015). The ubiquity of frequency effects in first language acquisition. *Journal of Child Language*, **42**(2), 239–273.
- Argus, R. (1998). *CHILDES'i eesti andmepank ja selle suhtluskeskne analüüs* [The Estonian databank on CHILDES and its conversational analysis] (*Hendrik, 1.6-2.6*) [Unpublished master's thesis]. Tallinn University.
- Argus, R. (2009). The early development of case and number in Estonian. In M. D. Voeikova & U. Stephany (Eds.), *Development of nominal inflection in first language acquisition: A cross-linguistic perspective* (pp. 111–152). Mouton de Gruyter.
- Berent, I., Pinker, S., & Shimron, J. (1999). Default nominal inflection in Hebrew: Evidence for mental variables. *Cognition*, **72**(1), 1–44.
- Berent, I., Pinker, S., & Shimron, J. (2002). The nature of regularity and irregularity: Evidence from Hebrew nominal inflection. *Journal of Psycholinguistic Research*, **31**(5), 459–502.
- Berko, J. (1958). The child's learning of English morphology. *Word*, **14**(2–3), 150–177.
- Bermel, N., Knittl, L., & Russell, J. (2015). Morphological variation and sensitivity to frequency of forms among native speakers of Czech. *Russian Linguistics*, **39**(3), 283–308.
- Bermel, N., Knittl, L., & Russell, J. (2017). Still on the brink: The productivity of Czech nominative plural desinences. In A. Makarova, S. M. Dickey, & D. Divjak (Eds.), *Thoughts on language: Studies in cognitive linguistics in honor of Laura A. Janda* (pp. 1–17). Slavica Publishers.










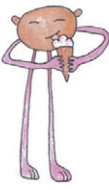



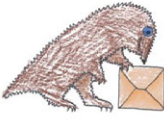




- Blevins, J. P. (1999). Productivity and exponence. *Behavioral and Brain Sciences*, **22**(6), 1015–1016.
- Blevins, J. P. (2008). Declension classes in Estonian. *Linguistica Uralica*, **43**(4), 241–267.
- Blevins, J. P., & Blevins, J. (Eds.). (2009). *Analogy in grammar: Form and acquisition*. Oxford University Press.
- Blything, R. P., Ambridge, B., & Lieven, E. V. M. (2018). Children's acquisition of the English past-tense: Evidence for a single-route account from novel verb production data. *Cognitive Science*, **42**, 621–639.
- Bybee, J. L., & Moder, C. L. (1983). Morphological classes as natural categories. *Language*, **59**(2), 251–270.
- Bybee, J. L., & Slobin, D. I. (1982). Rules and schemas in the development and use of the English past tense. *Language*, **58**(2), 265–289.
- Chandler, S. (2010). The English past tense: Analogy redux. *Cognitive Linguistics*, **21**(3), 371–417.
- Dąbrowska, E. (2001). Learning a morphological system without a default: The Polish genitive. *Journal of Child Language*, **28**(3), 545–574.
- Dąbrowska, E., & Szczerbinski, M. (2006). Polish children's productivity with case marking: The role of regularity, type frequency, and phonological diversity. *Journal of Child Language*, **33**(3), 559–597.
- Deen, K., & Hyams, N. (2006). The morphosyntax of mood in early grammar with special reference to Swahili. *First Language*, **26**(1), 67–102.
- Eek, A. & Meister, E. (2003). Foneetilisi katseid ja arutlusi kvantiteedi alalt (I-II): häälikukeskusi muutvad kontekstid ja välde [Some phonetic experiments and speculations on quantity (I-II), Contexts affecting sound durations, and the quantity degree]. *Keel ja Kirjandus*, **46**(11), 815–837.
- Engelmann, F., Granlund, S., Kolak, J., Szreder, M., Ambridge, B., Pine, J., Theakston, A., & Lieven, E. (2019). How the input shapes the acquisition of verb morphology: Elicited production and computational modelling in two highly inflected languages. *Cognitive Psychology*, **110**, 30–69.
- Granlund, S., Kolak, J., Vihman, V., Engelmann, F., Lieven, E. V. M., Pine, J. M., Theakston, A. L., & Ambridge, B. (2019). Language-general and language-specific phenomena in the acquisition of inflectional noun morphology: A cross-linguistic elicited-production study of Polish, Finnish and Estonian. *Journal of Memory and Language*, **107**, 169–194.
- Hahn, U., & Nakisa, R. C. (2000). German inflection: Single route or dual route? *Cognitive Psychology*, **41**(4), 313–360.
- Hayes, B., & Londe, Z. C. (2006). Stochastic phonological knowledge: The case of Hungarian vowel harmony. *Phonology*, **23**(1), 59–104.
- Hartshorne, J. K. & Ullman, M. T. (2006). Why girls say 'holded' more than boys. *Developmental Science*, **9**(1), 21–32.
- Kaalep, H. J. (2012). Eesti käänamissüsteemi seaduspärasused [Patterns in the declension system of Estonian]. *Keel ja Kirjandus*, **6**, 418–449.
- Keuleers, E., & Daelemans, W. (2007). Memory-based learning models of inflectional morphology: A methodological case study. *Lingue e Linguaggio*, **6**(2), 151–174.
- Kjaerbaek, L., dePont Christensen, R., & Basbøll, H. (2014). Sound structure and input frequency impact on noun plural acquisition: Hypotheses tested on Danish children across different data types. *Nordic Journal of Linguistics*, **37**(1), 47–86.
- Krajewski, G., Theakston, A. L., Lieven, E. V. M., & Tomasello, M. (2011). How Polish children switch from one case to another when using novel nouns: Challenges for models of inflectional morphology. *Language and Cognitive Processes*, **26**(4–6), 830–861.
- Laaha, S., Kjaerbaek, L., Basbøll, H., & Dressler, W. U. (2011). The impact of sound structure on morphology: An experimental study on children's acquisition of German and Danish noun plurals focusing on stem change. *Acta Linguistica Hafniensia*, **43**(2), 106–126.
- Lõo, K., Järvikivi, J., & Baayen, R. H. (2018). Whole-word frequency and inflectional paradigm size facilitate Estonian case-inflected noun processing. *Cognition*, **175**, 20–25.

- Lõo, K., Järvikivi, J., Tomaschek, F., Tucker, B. V., & Baayen, R. H. (2018). Production of Estonian case-inflected nouns shows whole-word frequency and paradigmatic effects. *Morphology*, *28*(1), 71–97.
- Lyytinen, P. (1987). Cognitive skills and Finnish language inflection. *Scandinavian Journal of Psychology*, *28*(4), 304–312.
- Lyytinen, P., & Lyytinen, H. (2004). Growth and predictive relations of vocabulary and inflectional morphology in children with and without familial risk for dyslexia. *Applied Psycholinguistics*, *25*(3), 397–411.
- Maratsos, M. (2000). More overregularizations after all: New data and discussion on Marcus, Pinker, Ullman, Hollander, Rosen, & Xu. *Journal of Child Language*, *27*(1), 183–212.
- Marcus, G. F. (1995). Children's overregularization of English plurals: A quantitative analysis. *Journal of Child Language*, *22*(2), 447–459.
- Maslen, R. J. C., Theakston, A. L., Lieven, E. V. M., & Tomasello, M. (2004). A dense corpus study of past tense and plural overregularization in English. *Journal of Speech, Language, and Hearing Research*, *47*(4), 1319–1333.
- Matthews, D. E., & Theakston, A. L. (2006). Errors of omission in English-speaking children's production of plurals and the past tense: The effects of frequency, phonology, and competition. *Cognitive Science*, *30*(6), 1027–1052.
- McClelland, J. L., & Patterson, K. (2002). Rules or connections in past-tense inflections: What does the evidence rule out? *TRENDS in Cognitive Sciences*, *6*(11), 465–472.
- Mirković, J., Seidenberg, M. S., & Joanisse, M. F. (2011). Rules versus statistics: Insights from a highly inflected language. *Cognitive Science*, *35*(4), 638–681.
- Pinker, S. & Ullman, M.T. (2002). The past and future of the past tense. *Trends in Cognitive Science*, *1*(6), 456–463.
- Peebo, J. (1997). *Eesti keele muutkonnad* [Estonian declension classes]. University of Tartu Press.
- Prasada, S., & Pinker, S. (1993). Generalisation of regular and irregular morphological patterns. *Language and Cognitive Processes*, *8*(1), 1–56.
- Prince, A., & Pinker, S. (1988). Rules and connections in human language. *Trends in Neurosciences*, *11*(5), 195–202.
- Ramscar, M., & Yarlett, D. (2007). Linguistic self-correction in the absence of feedback: A new approach to the logical problem of language acquisition. *Cognitive Science*, *31*(6), 927–960.
- Räsänen, S. H. M., Ambridge, B., & Pine, J. M. (2015). An elicited-production study of inflectional verb morphology in child Finnish. *Cognitive Science*, *40*(7), 1704–1738.
- Savičiūtė, E., Ambridge, B., & Pine, J. M. (2017). The roles of word-form frequency and phonological neighbourhood density in the acquisition of Lithuanian noun morphology. *Journal of Child Language*, *45*(3), 641–672.
- Vihman, V.-A., Lieven, E. & Theakston, A. (2020). Acquisition of symmetrical and asymmetrical Differential Object Marking subsystems in Estonian. In S. Montrul & A. Mardale (eds), *Acquisition of Differential Object Marking* (Trends in Language Acquisition Research) (pp. 21–49). Amsterdam: John Benjamins.
- Vaik, K. (2014). *Eesti keele muutemorfoloogia omandamisest WUG katse põhjal* [On the acquisition of Estonian inflectional morphology on the basis of a wug test] [Unpublished bachelor's thesis, University of Tartu]. <http://hdl.handle.net/10062/43997%09>
- Viht, A., & Habicht, K. (2019). *Eesti keele sõnamuutmine* [Inflection in Estonian]. Tartu University Press.
- Vija, M. (2007). Pronoomenid lapsekeeles: Mõnda mina ja sina omandamisest eesti laste näitel [Pronouns in child language: Some thoughts on the acquisition of *I* and *you*, based on Estonian]. *Estonian Papers in Applied Linguistics*, *3*, 373–384.
- Vitevitch, M. S., & Luce, P. A. (2016). Phonological neighborhood effects in spoken word perception and production. *Annual Review of Linguistics*, *2*(1), 75–94.
- Zupping, S. (2016). *Zupping corpus*. <https://doi.org/10.21415/T5K89H>

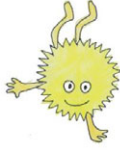
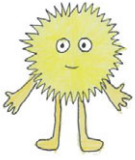
**Appendix 1**

Pictures used as stimuli

	Presentation contexts		Elicitation contexts	
	Nominative	Allative	Genitive	Partitive
	See on [wug] 'This is a ... [wug]'	[wug]-ile meeldib ... '... is pleasing to-[wug]' OR: ... tuli [wug]-ile külla. '... came to visit [wug]'	kelle käes/seljas ... ? '... on whose hands/back, etc.?'	patsutab/aitab/söödab/kammib keda? '... patting/helping/feeding/chasing, etc., whom?'
1				
2				
3				
4				

*Continued*

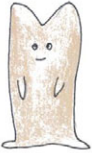
5



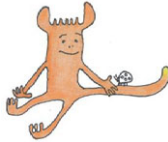
6



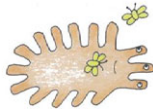
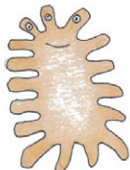
7



8



9

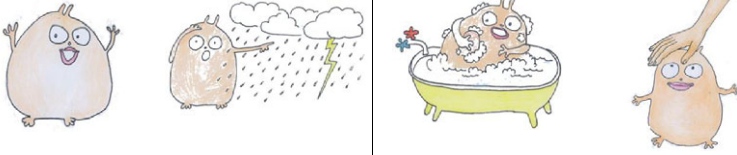


10

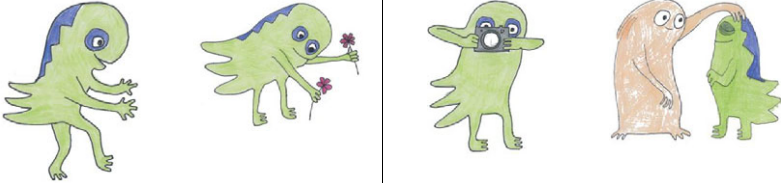


*Continued*

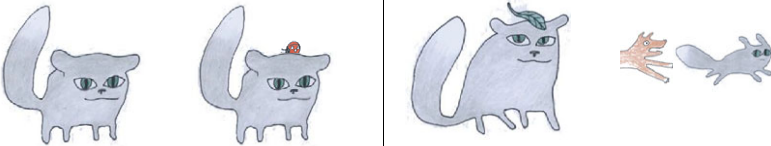
11



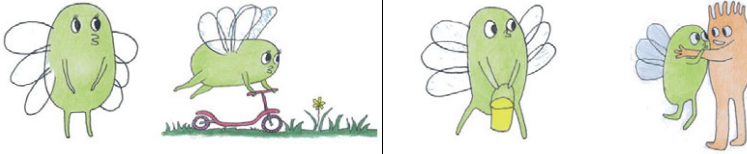
12



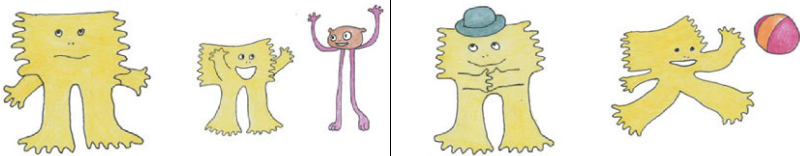
13



14



15



**Appendix 2**

Output of models

2a. *Variability*, model 1 (predictors: age group=Children vs. Adults, Presentation case, Target case)

Term	Estimate	Std. Error	t-value	Chisq.	df	Pr(>Chisq.)
(Intercept)	1.367	0.032	43.299			
GroupAdults	-0.311	0.031	-9.977	99.540	1	<0.001
PresCasenom	-0.021	0.039	-0.550	0.109	1	0.742
TargetCasepar	0.057	0.039	1.466	6.350	1	0.012
PresCasenom: TargetCasepar	0.025	0.055	0.448	0.201	1	0.654

Scaled residuals:

Min	1Q	Median	3Q	Max
-0.9454	-0.7192	-0.3038	1.0380	5.0086

Random effects:

Groups	Name	Variance	Std.Dev.
Lemma	(Intercept)	0.002593	0.05093
Residual		0.278657	0.52788

Number of observations: 1470, groups: Lemma, 15

ESTONIAN CHILDREN'S INFLECTION OF NOVEL NOUNS

2b. *Variability*, model 2 (Children only, Age = 3, 4, 5, Presentation case, Target case)

Term	Estimate	Std. Error	t-value	Chisq.	df	Pr(>Chisq.)
(Intercept)	2.045	0.094	21.825			
AgeMonths	-0.013	0.002	-7.834	61.379	1	<0.001
PresCasenom	-0.038	0.050	-0.761	0.555	1	0.456
TargetCasepar	0.055	0.050	1.113	3.592	1	0.058
PresCasenom: TargetCasepar	0.023	0.070	0.330	0.109	1	0.741

Scaled residuals:

Min	1Q	Median	3Q	Max
-1.2322	-0.7338	-0.4191	0.8164	4.5063

Random effects:

Groups	Name	Variance	Std.Dev.
Lemma	(Intercept)	0.004088	0.06393
Residual		0.334044	0.57797

Number of observations: 1080, groups: Lemma, 15

2c. *Accuracy* (Children only, Age = continuous age in months, Presentation case, Target case)

Term	Estimate	Std. Error	z-value	Chisq.	df	Pr(>Chisq.)
(Intercept)	-0.289	0.365	-0.792			
AgeMonths	0.026	0.006	3.963	15.703	1	<0.001
PresCasenom	-0.009	0.150	-0.059	0.249	1	0.618
TargetCasepar	-0.424	0.146	-2.895	12.623	1	0.000
PresCasenom: TargetCasepar	0.114	0.206	0.554	0.306	1	0.580

Scaled residuals:

Min	1Q	Median	3Q	Max
-2.9692	-1.0146	0.5014	0.6692	1.5729

Random effects:

Groups	Name	Variance	Std.Dev.
PartNum	(Intercept)	0.4264	0.653
Lemma	(Intercept)	0.1115	0.334

Number of observations: 1932, groups: PartNum, 39; Lemma, 15

2d. *Normalised PND*, model 1 (Age group=3, 4, 5, adult, Presentation case, Target case)

Term	Estimate	Std. Error	t-value	Chisq.	df	Pr(>Chisq.)
(Intercept)	-0.196	0.141	-1.393			
PresCasenom	0.082	0.094	0.869	0.001	1	0.972
TargetCasepar	0.005	0.095	0.049	0.001	1	0.982
Agegroup4	0.038	0.113	0.336	19.758	3	<0.001
Agegroup5	0.256	0.116	2.215			
AgegroupAdult	0.497	0.114	4.358			
PresCasenom: TargetCasepar	0.090	0.133	0.675	0.005	1	0.944
PresCasenom:Agegroup4	-0.142	0.134	-1.054	5.452	3	0.142
PresCasenom:Agegroup5	-0.142	0.138	-1.030			
PresCasenom: AgegroupAdult	-0.045	0.136	-0.332			
TargetCasepar: Agegroup4	0.143	0.135	1.059	24.620	3	<0.001
TargetCasepar: Agegroup5	-0.003	0.139	-0.022			
TargetCasepar: AgegroupAdult	-0.161	0.137	-1.180			
PresCasenom: TargetCasepar: Agegroup4	0.007	0.190	0.034	3.740	3	0.291
PresCasenom: TargetCasepar: Agegroup5	-0.077	0.195	-0.394			
PresCasenom: TargetCasepar: AgegroupAdult	-0.320	0.192	-1.666			

Scaled residuals:

Min	1Q	Median	3Q	Max
-2.6445	-0.7743	0.0121	0.7094	3.4067

Random effects:

Groups	Name	Variance	Std.Dev.
Subj	(Intercept)	0.04022	0.2006
Lemma	(Intercept)	0.20261	0.4501
Residual		0.75083	0.8665

Number of observations: 2558, groups: Subj, 87; Lemma, 15



ESTONIAN CHILDREN'S INFLECTION OF NOVEL NOUNS

2e. *Normalised PND*, model 2 (Children only, Age = continuous, Presentation case, Target case)

Term	Estimate	Std. Error	t-value	Chisq.	df	Pr(>Chisq.)
(Intercept)	-0.103	0.126	-0.822			
PresCasenom	-0.011	0.058	-0.191	0.328	1	0.567
TargetCasepar	0.052	0.059	0.888	4.490	1	0.034
scale(AgeMonths)	0.102	0.050	2.019	4.409	1	0.036
PresCasenom:	0.070	0.083	0.844	0.712	1	0.399
TargetCasepar						
PresCasenom:scale	-0.049	0.058	-0.838	3.113	1	0.078
(AgeMonths)						
TargetCasepar:scale	0.017	0.059	0.296	0.030	1	0.863
(AgeMonths)						
PresCasenom:	-0.048	0.083	-0.584	0.341	1	0.559
TargetCasepar:scale						
(AgeMonths)						

Scaled residuals:

Min	1Q	Median	3Q	Max
-2.4114	-0.7871	0.0358	0.7169	3.3089

Random effects:

Groups	Name	Variance	Std.Dev.
Subj	(Intercept)	0.05166	0.2273
Lemma	(Intercept)	0.19999	0.4472
Residual		0.82056	0.9058

Number of observations: 1930, groups: Subj, 66; Lemma, 15