doi:10.5128/ERYa14.02

EESTI RAKENDUSLINGVISTIKA ÜHINGU AASTARAAMAT 14, 23–39

ACQUISITION OF NOUN DERIVATION IN ESTONIAN AND RUSSIAN L1

Reili Argus, Victoria Kazakovskaya

Abstract. Acquisition of derivation is not a well-studied area in first language research and a comparative approach to the acquisition of derivation in different languages doesn't exist. There is no information on how a child acquires derivation in a language with a rich and regular system of derivational patterns, or in a language where derivation is productive, but the system of derivational patterns is opaque. According to general ideas of complexity in a language, the child should start to use simplex stems first and, only after that, complex ones, that is, complexity should increase in the course of acquisition. Our paper is intended to address these issues, based on longitudinal child data from typologically different languages, Estonian and Russian. The results revealed significant differences in the acquisition of noun derivation in the two languages under observation. The system of noun derivation is acquired at a faster pace in Russian, while Estonian children have far fewer noun derivatives in their speech and they use different derivation suffixes with less regularity. Even so, the so-called building block model may be applied for both languages only partially.*

Keywords: language acquisition, noun derivation, building-block model, acquisition of non-formation patterns, Russian, Estonian

1. Introduction

The study is a first attempt to analyse the early acquisition of noun derivation in typologically different languages, Finno-Ugric Estonian (EST) and Eastern Slavic Russian (RU). In both languages, scientific literature dealing with the acquisition of these phenomena is scarce. However, some work on the acquisition of word formation already exists: in Estonian on the acquisition of causatives (Argus 2012, Suurmäe 2013) and word formation by children with SLI (Padrik 2010), in Russian mainly on the acquisition of occasional words, both derivatives (Ceitlin 2009) and compounds (Xarchenko, Ozerova 1999). Besides these studies based

* The research was carried out with the financial support of the Russian Scientific Foundation (grant 14-18-03668).

on parental diary data, there are some preliminary attempts to describe early Russian diminutives (Protassova, Voeikova 2007) and nominal word formation in L1 and L2 (Russian-German bilingual situation) (Kazakovskaya 2017a) based on spontaneous speech data.

Previous research on acquisition of derivation in L1 was mainly concentrated on the problem of diminutives, which are usually dominant in early 'caregiver-child' dialogues. It does not clearly identify when, exactly, children start to use different noun derivatives and when they acquire knowledge about the internal structure of words. Some information can be found in a comparative cross-linguistic study on the acquisition of diminutives (Savickienė, Dressler 2007). Based on twelve typologically different languages, authors claim that in all the languages investigated, diminutives are the earliest emerging derivational category.

Our paper fills the first gap in this area and provides a description of the acquisition of noun derivation comparatively, presenting an analysis of not only diminutive, but all semantic categories of noun derivatives in the spontaneous speech data of young children.

Being permanent participants in the cross-linguistic project "Pre- and Protomorphology in Language Acquisition", we share theoretical and methodological background with its leader, W. U. Dressler (2005), and consider that the acquisition of derivational morphology is an important step in mastering the complexity of lexical items (see also Schipke, Kauschke 2011). This process is interrelated with the issue of lexical knowledge as well as of lexical categories, since the central functions of derivation are change of meaning and transposition of word classes (Clark, Berman 1984, Ravid 2004). The cross-linguistic differences are expected to be considerable, being dependent on the role of derivation in the different languages (e.g. frequency, transparency, functions etc.). Particularly, Eve Clark (1993) mentions the typologically related sphere of word formation (Dressler et al. 2017: 1).

The main tendencies of Estonian and Russian noun derivation, along with the aims, methodology of investigation and information concerning the data analysed are briefly presented in Chapter 1. An overview of the general course of acquisition of noun derivatives, main patterns and affixes is provided in Chapter 2. Chapter 3 deals with the idea of the validity of the building block model (Zurek 1990, Dziubalska-Kołaczyk 2014) in the acquisition of noun derivation. The data is analysed according to emergence of simplex stems versus derivatives and compound versus simplex basis of derivatives.

1.1. Noun derivation in EST and RUS

The languages under investigation are morphologically and morphemically rich. However, mostly agglutinative Estonian is moving towards being a fusional-type language, whereas Russian is an inflectional one. In these languages, compounding and derivation[1] are very important ways of word formation but not the only ones. For example, the Russian system of noun word formation has the less frequently used abbreviation (*RAS* < Russian Academy of Sciences) and conversion (*stolova-ja*

[1] Here we should mention that in the grammatical tradition of both languages, the term '(word formation) derivation' is used in parallel with the term 'word formation' (i.e. in a wider sense). For instance, noun word formation includes such means as compounding, affixation (which coincides with the term 'derivation' in the European tradition, or derivation in the so-called narrow sense), abbreviation (EST *treening* 'training' > *trenn*), conversion (EST *mõju-ma* 'influence-INF' > mõju 'influence') etc. as well as mixed means (e.g. compounding with simultaneous affixation: RUS *mor+e+plava+tel'* 'seafarer'). Here we will use the term 'derivation' as meaning 'affixation' in order to be comparable with other studies based on the data of European languages.

'dining room-N' > *stolov-aja* 'dining-ADJ') as well. But the main method of word formation concerning nouns is different: compounding is the most productive word formation strategy in Estonian nouns, while derivation is the most productive in Russian.

Within derivation in both languages, the most frequent and productive means of derivation is suffixation (EST *ava-**ja*** 'opener' < *ava-ma* 'open-INF', RUS *beg-**un*** 'runner' < *bega-t'* 'run-INF') including zero suffix (EST *jahti-ma* 'hunt' > *jaht* 'hunting', RUS *beg-Ø* 'running' < *bega-t'* 'run-INF'), in comparison with other means of derivation, namely prefixation (EST ***eba**-piisav* 'insufficient' < *piisav* 'sufficient', RUS ***pra**-vnuk* 'great-grandson' < *vnuk* 'grandson') and simultaneous prefixation and suffixation (EST ***eba**-piisava-**lt*** 'insufficiently' < *piisav* 'sufficient', RUS ***so**-zvezd-**ij**(e)* 'constellation' < *zvezd(a)* 'star').

The most frequent and productive pattern of noun derivatives in Russian is 'noun (N)+suffix'. The other types with a motivating stem (base) are verbs (V) (see *begun* 'runner' above) and adjectives (ADJ) (*chern-**ik**(a)* 'blueberries' < *chern-yj* 'black-ADJ'), but both are less frequently used. Adverbs (ADV) as a base for noun derivatives are rare. Russian nominal patterns are mostly semantically transparent, in spite of many consonant alternations on the boundaries of root and suffix morphemes. The suffix system is extremely rich and diverse.

The Estonian system consists of both productive and frequent patterns 'N+suffix' (*koer-lane* 'canine'< *koer* 'dog') and 'V+suffix' (*joo-mine* 'drinking' < *joo-ma* 'drink-INF') along with the patterns 'ADJ+suffix' (*must-us* 'dirt' < *must* 'black'), 'ADV+suffix' (*taga-tis* 'warranty' < *taga* 'behind') and 'Numeral (NUM)+suffix' (*kolm-ik* 'triplet' < *kolm* 'three'), which are less common. Most Estonian suffixes attach only to nominal or verbal stems, but some suffixes can be used with both stems, however, with different semantics: e.g. *-ur* in *ved-ur* 'locomotive' < *veda-ma* 'pull-INF' (instrument), *kang-ur* 'weaver' < *kangas* 'fabric.N' (actor). Also, suffixes used in the pattern 'ADV+suffix' coincide with suffixes belonging to the pattern 'N+suffix' (Kasik 2013: 98). According to Kerge (2016: 3228), the most regular fields of modern derivation are deadjectival and deverbal noun formation.

The following features might be considered differences in the noun derivation of the languages. The repertoire of Estonian suffixes[2] is not as large or as diverse as in Russian; additionally, some suffixes have overlapping or ambiguous meaning. Moreover, such methods of word formation as prefixation and 'mixed' – simultaneous prefixation and suffixation – are rare in Estonian. In contrast to Russian having quite regular nominal patterns, only some Estonian derivatives can be described as being regular (when considering both form and meaning), and many less productive and semantically opaque patterns are also used.

1.2. Aims and methodology

Our study has two general aims, which can be regarded as being comparative and so-called developmental.

The comparative aim is to describe and compare the acquisition of noun derivation in two languages based on the analysis of a) the first occurrences of noun derivatives, b) their word formation features (including their belonging to a definite

[2] According to Vare (2012), there are 42 noun suffixes (plus some phonologically different variants of the same suffix) in Estonian.

pattern, the presence of affixes, and their semantics), c) frequency in child speech (CS), d) further development, e) similarities and differences between the acquisition of this phenomenon in Estonian and Russian.

The developmental objective in our study is to discover whether a building block model of complexity, according to which complexity increases in the course of acquisition (Zurek 1990, Dziubalska-Kołaczyk 2014), can be applied to the order of emergence of nominal derivatives. More specific research questions are the following. 1) Do simplex stems occur first and, only afterwards, derivatives? 2) Does the child use derivational affixes first with simple stems/roots and then, subsequently, with complex words (compounds and derived stems)? 3) What is more important in the course of development in the two languages, general acquisition strategy (following the building block model) or a different system of derivation?

1.3. Data

Noun derivation is analysed based on longitudinal spontaneous data of 4 monolingual typically developing children (3 boys and 1 girl) from middle SES[3] families from age 1;3–3;0. The total volume of recordings is 62.5 hours which were transcribed and coded in CHILDES (MacWhinney 2000). The total number of noun tokens analysed in CS is 12,301 (see Table 1).

We have analysed the naturalistic observations of two Estonian children: the corpus "Martina", collected by A. Kapanen (Tallinn), and the corpus "Andreas", collected by M. Vija (Tartu). The two Russian boys were from St. Petersburg (the same Baltic region as Estonia). The first corpus, "Filipp", was collected by T. Pranova (in the late 1990s) and the second corpus, "Kirill", was collected by K. Ivanova (in the early 2000s).

Table 1. Data

Corpus	Gender	Age	Length of recordings (hours)	Noun tokens (CS)
Martina (MAR)	F	1;3–3;0	14.5	4301
Andreas (AND)	M	1;7–3;0	15	3333
Filipp (FIL)	M	1;5–2;8	28	3703
Kirill (KIR)	M	1;9–3;0	5	964
Total			62.5	12301

2. Results

2.1. General overview of the number of derivatives and their distribution in child speech

Table 2 shows a general overview of the number of new derivatives, i.e. first-appearing lemmas (including their types and tokens) and their percentage from all noun tokens in the data analysed from the two languages.

[3] SES – socio-economic status, the social standing or class of an individual or group. It is often measured as a combination of education, income and occupation.

When looking at this table, one can see that there are many more noun derivative lemmas (both in types and tokens) in the speech of the Russian children.

Table 2. Number of new noun derivatives (shown cumulatively)

Age	EST				RUS			
	MAR		AND		FIL		KIR	
	Lemmas/ Types/ Tokens	% tokens among all noun tokens	Lemmas/ Types/ Tokens	% tokens among all noun tokens	Lemmas/ Types/ Tokens	% tokens among all noun tokens	Lemmas/ Types/ Tokens	% tokens among all noun tokens
1;3	1/1/1	0.3						
1;5	4/4/30	6.3			3/3/9	5.5		
1;6	1/1/1	0.3			7/7/9	5.1		
1;7			1/1/1	0.4	2/2/3	2.9		
1;8			1/1/3	0.9	16/16/26	7.4		
1;9	3/3/3	0.7	0/0/0	0			0	0
1;10	3/4/16	3.5	3/3/3	0.9	24/28/34	28.3	0	0
1;11	3/3/11	3.0	0/0/0	0	16/17/22	21.8	0	0
2;0			7/9/16	4.2	29/41/54	26.5	1/1/1	1.7
2;1	8/14/25	5.5	1/1/1	0.4	50/61/86	19.1	0	0
2;2					31/32/38	9.4	0	0
2;3	6/7/8	6.1	6/6/10	7.1			5/5/9	22.5
2;4	1/1/1	0.6	1/1/2	1.3	30/35/62	15.5	7/7/11	14
2;5			4/5/6	3.7	23/29/37	9.9	8/8/10	10
2;6			7/7/12	4.3	13/15/20	9.2	6/8/13	10
2;7	9/10/17	4.7	9/9/9	4.7	18/20/24	6.6	2/4/5	8.2
2;8			7/8/15	6.8	17/17/21	4.8	3/3/6	6.5
2;9							5/8/9	7,2
2;10							5/5/6	5
2;11							3/3/22	17
3;0	5/6/6	1.6	5/6/7	4.1			8/8/12	6
Total	44/54/ 119		47/57/ 85		280/325/ 456		53/60/ 104	

In Estonian, children start to use first derivatives at an early age, that is, at the same time as first compounds emerge and noun inflection starts to develop (Argus, Kazakovskaya 2013, Argus 2009). Still, there is only one derived noun lemma occurring in one inflectional (grammatical) form (i.e. type) during the recording session, at the beginning of the observation period. Children start to use more than one form of the same noun only 6–8 months after they have used their first nouns.

A similar situation is observed in the Russian data. Filipp begins to use his first noun derivatives earlier than Kirill at 7 months, and does it more frequently. However, his repertoire of noun derivatives is quite large and diverse: 280 lemmas (325 types and 456 tokens) are documented only for newly appeared derivatives. However, the first grammatical forms of derivatives appear in his speech production in 3 months, whereas in Kirill's data it happens earlier, in a month. Simultaneously

(i.e. during the first recording session) the boys start to use both noun derivatives in an initial form (NOM or ACC) and in any grammatical form 5–6 months later. Filipp has a so-called derivative spurt from 1;10–2;0 (28–26,5%) and after that the number of new derivative lemmas gradually decreases. The peaks in emergence of new derivative lemmas in Kirill's speech are at 2;3 (22,5%) and at 2;11 (17%).

The percentage of derivatives (including their repetitions) amongst all noun tokens can vary in different recording sessions (see Diagram 1) and there is still a higher percentage of derivatives among nouns in the speech of one of the Russian children, namely Filipp. Analysis of the distribution of derivative noun tokens, including their repetitions, in our corpora has shown the following: Estonian derivatives increase moderately during the observation period, with no sudden increase. Martina and Andreas start to increase their derivative vocabulary at a moderate pace: in every recording a couple of new lemmas are added to the 'old' (i.e. ones already used in previous months) derivatives. Russian children, conversely, have several derivational spurts, showing high frequency of derivative usage. So the first notable increase of derivative tokens in Filipp's speech occurs at 1;10 and 'holds' till 2;1 (see also above), and the second one occurs at 2;6 (54,4%). The second Russian boy, Kirill, also has two derivative spurts at 2;3 and 2;11, both coinciding with the spurts of lemmas, excluding repetitions.

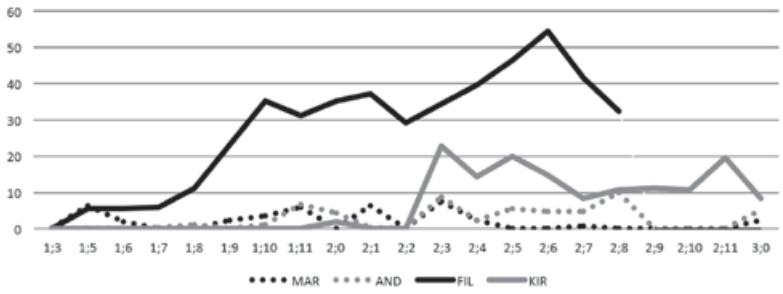


Diagram 1. Percentage of noun derivatives among nouns (in tokens)

Comparing the Estonian and Russian data, one can say that the general number of noun derivatives is larger and the growth of noun derivative vocabulary is faster in the Russian CS. Differences between children are less significant than differences between languages. For example, the percentage of derivatives among noun tokens reaches no higher than 10% in the speech of Estonian children, while at some age points the percentage of derivatives is as high as 54 in the speech of Russian children.

2.2. First occurrences of noun derivatives

As mentioned, both Estonian children were already using derivatives in their first recordings. The first semantic category to emerge in their speech was diminutives: Martina used the diminutive *nuk-u* 'doll' (1;3) and Andreas *jänk-u* 'bunny' (1;7). The next suffix they used was also a diminutive one *-ke*: *linnu-ke* 'bird:GEN-DIM' (Martina 1;5), *kohu-ke* 'rise-DIM=sweet made from cottage cheese' (Martina 1;10). She used the most "intense" diminutive suffix *-kene* (*emme-kene* 'mummy-DIM') at age 1;9.

The first Estonian derivatives not belonging to the semantic class of diminutives (non-diminutives, with very general semantics of an object) emerged some months later. At age 1;10 Andreas used a compound consisting of a derivative as a modifier *puhk-**e**+päev* 'holiday' < *puhka-ma* 'rest-INF'[4], and a compound consisting of a derivative as a head *saja+jalg-**ne*** 'centipede' < *jalg* 'foot'. There was also a similar compound with a derivative as a head in the speech of Martina *merne+irmu-**tis*** [=herne+hirmu-tis] 'scarecrow' < *hirm* 'fear' (1;5).

Within the three initial noun derivatives (3 types / 9 tokens) used by Filipp, produced in his first recording (1;5), there is 1 diminutive with a typical suffix (*mish-k(a)* 'bear-DIM') and 2 lemmas belonging to the class of so-called stylistic modifications (SM) of words (Russian Grammar 2005)[5]: e.g. *mam-k(a)* 'mother-SM'. Within the next month Filipp has already uttered 5 diminutive lemmas (with new diminutive suffixes *-ok*, *-ik*: e.g. *fant-ik* '(candy) wrapper-DIM') as well as 2 non-diminutive derivatives. Except for colloquial variants, the title of a female person appears: *dev-ochk(a)* 'girl'. According to our calculations, in Fillipp's data diminutives are quite numerous, they account for 50 to 70% in different months (see also Protassova, Voeikova 2007). The first motivating base is nouns (see above); verb- and adjective motivated noun derivatives occur later: *kup-k(i)* 'bathing'< *kupa-t'sja*-INF (1;7), *kos'-ak* 'doorjamb' < *kos-oj*-ADJ 'oblique' (1;8). The first compound *magnit+o+fon* 'record player' was documented at 1;11.

The first derivative (1 type / 1 token) used by Kirill emerges later, not only relative to his first recording (cf. 1;9), but also relative to all the children under observation. In contrast to their initial derivations, the base stem for this noun derivative is a verb. The semantics of this derivative is an agent: *beg-un* 'runner' (see above). However, during the next recording containing derivatives (namely at 2;3), 5 'N+suffix' patterns are fixed (5 types / 9 tokens). Amongst them are 2 diminutive lemmas formed by a typical suffix *-k* (*mysh-k(a)* 'mouse-DIM', *bednjazh-k(a)* 'poor fellow-DIM'), 2 lemmas belonging to stylistic modifications (*knizh-k(a)* < *kniga* 'book', *kolen-k(a)* < *koleno* 'knee') and 1 derivative, identifying a female person: *vnuch-k(a)* 'granddaughter' < *vnuk* 'grandson'. In all cases the same formal suffix *-k* is used but, consequently, with different semantics as mentioned above – diminutive, female etc. In comparison to Filipp, diminutives in Kirill's corpus are not represented so widely, at 44%. However, he has a more diverse repertoire of compounds.

[4] The suffix *-e* regularly form compound modifiers of verbal meaning (see Kerge 2016: 3243).
[5] These are colloquial (speech) variants: compare, *Filipp* and *Fil'-k(a)* 'proper name.SM'.

2.3. Noun formation patterns and affixes: further development of derivation

The most frequent pattern of noun derivatives in the speech of all the children under observation was 'N+suffix': e.g. EST *foto-kas* 'photo-SUFF=camera', RUS *nos-ik* 'nose-SUFF=nose-DIM'. Although there were other patterns such as verb, adjective (in Russian), adverb and even numeral stem (in Estonian) + derivational suffixes, derivatives constructed according to this pattern formed approximately 41–55% of all noun derivatives in the Estonian data, and approximately 90% (Filipp) and 70% (Kirill) in Russian (see Table 3).

As mentioned, this pattern is productive and frequent in modern noun word formation in both languages. This is particularly noticeable in the sphere of lemmas.

Table 3. Noun derivational patterns

Patterns and models	EST		RUS	
	MAR	AND	FIL	KIR
	lemma/type/ token	lemma/type/ token	lemma/type/ token	lemma/type/ token
N+affix	22/28/65	28/30/35	244/288/408	34/40/72
a. N+suffix	22/28/65	28/30/35	244/288/408	32/38/70
b. prefix+N	0	0	0	1/1/1
c. prefix+N+suffix	0	0	0	1/1/1
V+suffix	16/17/34	15/16/26	29/30/35	16/17/29
ADJ+affix	3/4/11	3/3/3	7/7/13	0
a. ADJ+suffix	3/4/11	3/3/3	5/5/10	3/3/3
b. prefix+ADJ+suffix	0	0	2/2/3	0
ADV+suffix	2/4/8	1/1/3	0	0
NUM+suffix	1/1/1	0	n/a	n/a
Total	44/54/119	47/57/85	280/325/456	53/60/104
% 'N+affix'	50/52/55	60/53 /41	87/89/89,5	64/67/69

The variety of different suffixes is wider in the most frequent patterns such as 'N+suffix' or 'V+suffix' in both languages (see Table 4).

Table 4. Affixes in early noun derivatives

EST		RUS	
Models	**Number of affixes**	**Models**	**Number of affixes**
N+suffix	11	N+affix	30
		a. N+suffix	28
		b. prefix+N	1
		c. prefix+N+suffix	1, 1
V+suffix	9	V+suffix	18
ADJ+suffix	3	ADJ+affix	8
		a. ADJ+suffix	7
		b. prefix+ADJ+suffix	1, 1
ADV+suffix	3	ADV+suffix	0
NUM+suffix	1		

Table 5 shows the order of emergence of different affixes in CS (cumulatively) in the languages under observation.

Table 5. Models and affixes of noun derivatives

Models	Affixes in order of their emergence
EST	
N+suffix	*u, ke, kene, ik, ne, kas, k, nd, line, lane, stik*
V+suffix	*is, ja, e, k, us, i, mine, ur, kas, is*
ADJ+suffix	*kas, is, us*
ADV+suffix	*ik, kas, is*
NUM+suffix	*ik*
RUS	
N+suffix	*k, an', ok, ik, ush, uh, nik, ut, chik, ek, ic, en(')k, ushk, yshk, ochk, c, echk, ix, ul', ess, n', ink, utk, av, ovn, ich; al'on, ovin*
Prefix+N	*pra*
Prefix+N+suffix	*pod & ik*
V+suffix	*k, ak, enok, ø, (e)n(i)j, d, l, unok, ux, lk, shk, v, nic; un, shchik, ok, tel', nk*
ADJ+suffix	*ak, ysh, ic, ost'; ik, atin*
Prefix+ADJ+suffix	*pod & ik*

2.3.1. Suffixation in Estonian CS

The order of the semantic categories represented by different suffixes is quite similar in the speech of both Estonian children. The most frequent pattern, consisting of a noun stem and suffix was demonstrated with 11 different suffixes by the Estonian children. In cases of the pattern 'N+suffix', aside from three diminutive suffixes (*-u*, *-ke*, *-kene*), the children used several suffixes for deriving nouns with the very broad meaning of an object or person: ***-ik*** (*homm-ik* 'tomorrow-*ik*=*morning*'), ***-k*** (*lennu-k* 'flight:GEN-*k*=plane'), ***-nd*** (*pake-nd* 'pack-*nd*=package'), ***-kas*** (*voto-kas* 'photo-*kas*=camera'), ***-lane*** (*mesi-lane* 'honey-*lane*=bee'), ***-ne*** (*saja-jalg-ne* 'hundred+leg-*ne*=centipede'), ***-line*** (*küla-line* 'village-*line*=visitor'). There was only one suffix for collective nouns, ***-stik*** (*ilu+tule-stik* 'beauty+fire-*stik*=firework') in the speech of Andreas.

The number of suffixes was a little bit smaller (9) in the case of deverbal nouns. The most productive pattern of them all, indicating actions, ***-mine*** (*rääki-mine* 'speak-*mine*=speaking') was not frequent in CS and was used with only three different verbs. For the meaning of a process the suffix ***-us*** was used: e.g. *üllat-us* 'to surprise-*us*=surprise'. The suffix ***-ja*** was used both for indicating an agent (*õpeta-ja* 'teach-*ja*=teacher') or an instrument (*kruvikeera-ja* 'screwdriver'). The verbal base also has the only suffix with pejorative meaning (***-is***, see the example above) used by Martina at age 1;5. The same suffix can also form derivatives having no pejorative meaning but a neutral meaning for an object or a result: *täid-is* 'fill-*is*=filling'. Some patterns with a verbal base can be used with the suffixes also combining with noun stems: the suffix ***-is*** (mentioned above), as well as suffixes ***-k*** (*söö-k* 'eat-*k*=food'), ***-e*** (*kast-e* 'cover-*e*=dressing'), and ***-kas*** (*tulnu-kas* 'come-*kas*=alien'). Two different deverbal suffixes, in addition to the previously mentioned suffix *-ja*, were used for

indicating the meaning of an instrument: e.g. ***-ur*** (*ved*-ur 'pull-*ur*=locomotive'), ***-i*** (*arvut-i* 'calculate-*i*=computer').

Three suffixes used with adjectival stems are the same ones used with either verbal or adverbal stems: ***-is*** (*jäät-is* 'freeze-*is*=ice cream'), ***-us*** (*ulak-us* 'naughty-*us*=naughtiness') or with noun stems ***-kas*** (*sini-kas* blue-*kas*=bruise').

Adverbial stems also combined with three different suffixes, whilst all suffixes were also used with verbal or noun stems: e.g. ***-kas*** (*alu-kas* 'down-*kas*=panties'), ***-ik*** (*ümbr-ik* 'around-*ik*=envelope'), ***-is*** (*täid-is* 'full-*is*=filling'). The pattern with a numeral stem was used only once in one noun derivative: ***-ik*** (*kaks-ik* 'two-*ik*=twin'). Use of these suffixes is quite rare in CS: most suffixes are used only once and with just one stem.

The order of emergence of different semantic categories of derivatives can be summarized as follows: Children start with diminutives. The next suffixes used can be grouped into a very general semantic category of objects (e.g. *-is*, *-k*, *-ik*, *-kas*), or results (often non-distinguishable from objects, e.g. suffix *-nd*) or persons (e.g. *-ne*, *-lane*, *-line*). The category of an instrument is represented by 3 suffixes: *-ja*, *-i*, and *-ur*. These suffixes also occurred with only one stem. The next category to emerge was an agent and the suffix used was the same as already used for instruments (*-ja*). The very general category of processes or results (often with overlapping meaning) is expressed with the suffix *-us* by both children. The suffix *-mine* was used by them for expressing the meaning of an action and this was the only suffix having at least two different stems in CS: *tudu-mine* 'sleeping' (Martina 2;3), *rääki-mine* 'speaking' (Martina 2;7); *pildista-mine* 'photo shooting' (Andreas 2;3), *saagi-mine* 'sawing' (Andreas 2;6).

2.3.2. Affixation in Russian CS

As mentioned, the affix system in Russian is much richer than in Estonian, as is the system of nominal models, which is reflected in the CS (see Tables 4 and 5). Apart from suffixes in the Russian data, a prefix model, along with a mixed model ('prefix+N+suffix') are noted: *pra-ded* 'great-grandfather' (Kirill 2;4), *pod-guzn-ik* 'diaper' (Kirill 2;9). Moreover, simultaneous suffixation and prefixation is used not only with nominal motivating stems, but also with adjectival ones: *pod-osinov-ik* 'orange cap boletus (lit. mushroom, which grows under an aspen)' (Filipp 2;1).

Returning to the process of suffixation, one can observe how 56 different affixes, which are attached to nominal (30), verbal (18) and adjectival (8) stems, gradually increase the repertoire of early semantic categories. For example, at the initial stage of development of noun word formation in Filipp's speech (until 1;8) he is the child with the richest repertoire of derivatives amongst the children under observation. New derivatives are infrequent (8% of all noun tokens, including repetitions). At 1;7 the second pattern (after the initial 'N+suffix') emerges, in which the verb is the motivating stem for noun derivatives. A month later a new stem – an adjective – starts to be used for this purpose. Diminutives predominate in noun models (57% in lemmas during this period), although there are several words with non-diminutive semantics – namely of animal young, persons of female gender, and stylistic modifications of nouns. The first derivatives formed from verbs have the semantics of agent and action. At this stage all derivatives are formed by pure suffixation and

by means of those typical for each semantic category suffix (i.e. prototypical). For instance, *-k*, *-ik*, *-ok*, *-ush,* are used for diminutives (*pitch-k(a)* 1;6, *zub-ok* 1;8, *avtobus-ik* 1;8, *Fil'-ush(a)* 1;8), *-k*, *-ux* are used for stylistic modifications (*dyr-k(a)* 1;5, *Fant'-ux(a)* 1;8) and *-onok* for animal young, 'babies' (*kot'-onok* 1;8).

The next period in the development of derivation (until 2;1) is quite important, since the first grammatical forms of derivatives occur. During this time the percentage of derivative tokens is 49% of all noun tokens. The first derivational chains also emerge: e.g. *grib* 'mushroom' – *grib-ok*.DIM – *grib/och-ek*.DIM, which shows the beginning of secondary suffixation. So-called families are developing: cf. *kot* 'cat-MASC' and *kosh-k(a)* 'cat-FEM'; *kot* 'cat' and *kot-ok* 'cat-DIM', *kot-ik* 'cat-DIM'. The number of other patterns (which are formed from verbs and adjectives) increases from 4 to 14, along with the number of non-diminutive patterns, which increases from 8 to 19. The repertoire of suffixes is supplemented by new ones: particularly, zero suffixation within a verb pattern emerges: *pricepø* 'trailer' < *pricepi-t'* 'attach-INF'. At this stage the first and only (in this corpora) occasional diminutives are fixed: *det-ik* 'child-DIM' (1;10), *garmosh-ishk(a)* 'accordian-DIM' (2;1). The first compound is also noted: *magnitofon* 'tape recorder' (1;11), and the first diminutive formed from a compound occurs: *parovoz-ik* 'locomotive-DIM' (2;1). The main pattern of compounding is right-headed 'X+V': *par+o+xodø* 'steamship', *sam+o+letø* 'airplane' (2;0), but there is also one left-headed 'N+N' compound *Kon'ok-Gorbunok* 'The Little Humpbacked Horse' (2;1). Finally, the first grammatical forms of compounds emerge (4/7/10), which start to form the first mini-paradigms: *paroxod* 'steamship' has 4 case forms (Nom, Acc, Gen, Loc). This period ends at 2;1 with the 'derivational spurt' mentioned above. Additionally, we can definitely conclude that compounding occurs during the period of active formation of affixed derivatives (see also Kazakovskaya 2017b, Argus, Kazakovskaya 2013).

The main peak of the third stage (from 2;2 until to the end of the observation) happens at 2;4. Generally, the number of new derivative lemmas decreases. We can observe that the same tendencies and patterns, with diminutive semantics formed by suffixation, prevail. The total number of new affixes is 20, e.g. at 2;2 the suffixes *-ix*, *-chik*, *-d*, *-l*, *-unok* occur; at 2;4 *-ul*, *-ux*, *-lk*, *-shk* appear; at 2;8 *-utk*, *-av*, *-ovn*, *-ic* emerge. Moreover, a new way of derivation (viz. by ending) is fixed at 2;2, and a new method of word formation (namely, non-morphological conversion) is noted: *vann(aja)* 'bathroom.N' (2;6).

2.4. Similarities and differences between languages in the acquisition of noun derivation

Searching for similarities between languages concerning the acquisition of noun derivation, it can be argued that the similarity concerns the occurrence of the first derivatives. In both languages noun derivatives occur early (up to 2;0), at approximately the same age (1;3–1;5 with one exception, Kirill).

They are mainly diminutives, whilst in Russian, the first derivatives include stylistic modifications as markers of colloquial speech. Diminutives form a prominent part among the first noun derivatives at the beginning of acquisition in both languages (as already found in previous research based on the data of other European

languages (e.g. Savickienė, Dressler 2007)). They are from 42% to 65% of all noun derivative tokens in CS (viz. Andreas has 42, Kirill 44, Martina 64, Filipp 65).

The order of emergence of different semantic categories seems to be quite similar, at least for those categories which are represented in both corpora of CS: DIM > person > process, result (see Table 6).

Table 6. Semantic categories in order of emergence

Corpus	The order of emergence of suffix semantics
Martina EST	DIM (1;3) > object and person (1;6) > instrument, actor or agent (1;9) > process or result (2;1) > action (2;3)
Andreas EST	DIM (1;7) > object/person (1;10) > instrument, process/result (2;0) > action (2;3) > actor (3;0)
Filipp RUS	DIM, SM (1;5) > person (females) (1;6) > activity/process (1;7) > animal young, agent, object (inanimate name) (1;8) > object (singulative) (2;1) > instrument (2;5) > person (animate name) (2;6) > abstract name (quality) (2;8)
Kirill RUS	agent (2;0) > DIM, SM, person (females) (2;3) > person (males), instrument, result (2;4) > object (singulative) (2;5) > locative (2;6) > abstract name (quality) (2;9)

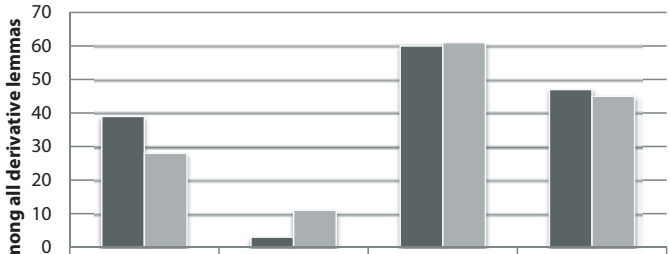
Generally, one can conclude that the main directions of noun word formation in Estonian and Russian L1 reflect the general peculiarities of the relevant domains of each language system and the common tendencies of their modern development.

As regards the differences between languages, the first and most prominent difference concerns the total number of derivatives, both the number and the diversity of derivative affixes, their productive usage and the number of semantic categories to which children's derivatives belong. Derivatives are much more frequent and diverse in Russian CS (where they are used productively within several semantic categories), while Estonian children use only diminutives productively. There are many different suffixes used in both languages, but in Estonian they occur only once or twice within all patterns, whereas in Russian they occur within the less productive ones. Finally, in EST the development of different patterns is clear only in V- and N-stem derivatives (concerning DIM), while in RUS this process is quite clear within all patterns.

2.5. Validity of the building block model

According to the building block model (Zurek 1990, Dziubalska-Kołaczyk 2014), children should start with simplex stems and the complexity increases during development: that is they should start to use derived nouns only after they have already acquired the corresponding simplex stems. To test the validity of this model, we first analysed the data for derivative lexemes and calculated all lexemes emerging first as simplex and only afterwards as derivatives in CS and vice versa (see Diagram 2).

In both languages, 2/3 of all nouns appear for the first time as derivatives and, even more importantly, more than half of them do not have a simplex match in the data. In some cases the derivative and simplex occur simultaneously, within the same recording of CS.



	Simplex before derivative	Simplex simultaneous with derivative	Derivative before simplex	Derivative not having a simplex match in the data
EST	39	3	60	47
RUS	28	11	61	45

Diagram 2. Derivatives and simplex stems (in lemmas)

The next question was if children use derivational suffixes first with simple stems/roots and only afterwards with complex words (i.e. with compound or derived stems). In the speech of the Estonian children, first suffixes (within the semantic category of diminutive) were added to simplex stems. Shortly after this, the young Estonians started to use derivational suffixes inside compounds, i.e. the derivative can first occur as a modifier: *puhk-**e**+päev* (Andreas 1;11), while the simplex stem with the same suffix emerged only afterwards (*kast-e* 'dressing' 2;0). Some suffixes (especially deverbals) occurred first, not only with derived verb stems as modifiers of compounds: ***ära-t-us**+kell* 'wakening clock' < '*ära-ta-ma* 'wake-CAUS-INF' (Martina 2;1), but also as heads of a compound: *piima+**vahu-sta-ja*** 'milk+foam-CAUS-ja=milk foamer' (Martina 1;9). Thus, there is still no clear evidence that the same derivational suffix occurs first with simplex and only afterwards with complex stems in EST.

This tendency is comparable with the Russian CS data. Russian children also use simplex before complex derivations, including synthetic compounds which often have zero suffixation. Particularly, they firstly use suffixes (diminutive and so-called stylistic ones) with simplexes *dyr-k(a)* 'hole-SM' (Filipp 1;5), *mysh-k(a)* 'mouse-DIM' (Kirill 2;3), later with derivatives *grib/och-ek* 'mushroom-DIM' (Filipp 1;8), *po/dar-ok* 'gift, present' (Kirill 2;4). After that they begin to produce compounds, e.g. *magnit+o+fon* 'tape recorder' (Filipp 1;11), *foto+apparat* (Kirill 1;9). Slightly later, compounding is accompanied by suffixation, including the zerosuffix: *sam+o+let-Ø* 'airplane' (Filipp 2;0), *beton+o+mesha-lk(a)* 'concrete mixer' (Kirill 2;5).

3. Discussion and conclusions

The acquisition of noun derivation starts early and begins with diminutives in both languages. Even so, similarities in the acquisition of noun derivation are limited. After the emergence of diminutive suffixes, the growth in the number of noun derivatives as well as different patterns and suffixes in Russian CS is fast,

while the Estonian children use noun derivatives more moderately. The system of derivation is rich in both languages, but more regular in Russian and more opaque in Estonian. This difference is also reflected in acquisition. The young Estonians have noun derivatives with less frequency and variety. They use the same suffixes for conveying different meanings, as permitted by the Estonian derivation system. Estonian-speaking children use only diminutive suffixes productively. In contrast to Estonian, Russian is a more 'derivative language' and has more productive patterns and a greater frequency of derivatives. These circumstances have an influence on the acquisition of noun derivatives. Specifically, Russian children start to use nominal derivatives not only with 'N+suffix' diminutives, but also with stylistic modifications and even 'X+suffix' patterns. Their model repertoire has more models and methods, specifically prefixation and simultaneous prefixation and suffixation. The abundance of suffixes provides a opportunity to quickly increase the vocabulary and the number of early semantic categories (e.g. females, singulatives, young animals, agents, activities, results, instruments).

What are the main factors influencing the acquisition of derivation in the languages studied and do the patterns observed support the building block model of acquisition? Also, do the different systems of derivation in two languages influence this process more? We have to favor the latter idea. Following the model mentioned, we can see only usage of suffixes (at least, diminutive ones in both Estonian children) first with simplex stems and after that with complex ones. As mentioned, in both languages about 60% of nouns appear in CS as derivatives. There is no clear evidence that derivational suffixes emerge only after the corresponding simplex stems have been acquired. Furthermore, suffixes do not emerge first with simplex stems; they can occur with derivated stems and as a part in compounds (even being added to one component which is already a derived stem). The only exception is diminutives: diminutive suffixes occur first with simplex stems and only afterwards with compounds in Estonian.

Other suffixes could also emerge first in complex stems (e.g. in Kirill's deverbals). Consequently, one can conclude that the building block model might be applied only partially. Particularly, it was revealed that in both languages more than half of all noun lemmas appear in CS for the first time as derivatives, and later so-called 'decomposition' takes place. The process of acquisition of noun derivatives does not resemble the process of acquisition of morphology. We cannot say that children always start with simple structures, and then continue with structures of increasing complexity, as it's rather similar to the acquisition of vocabulary.

Despite the results presented above, we have to admit that our study is not exhaustive in every aspect and there could be factors having a possible influence on the acquisition process. We would suggest that some other factors (primarily, child-directed speech, or input) could play a greater role in acquisition, but this needs additional investigation.

References

Argus, Reili 2012. Kausatiivsuse omandamisest eesti keeles ['Acquisition of causatives in Estonian']. – Eesti Rakenduslingvistika Ühingu aastaraamat, 8, 5–20. https://doi.org/10.5128/ERYa8.01

Argus, Reili; Kazakovskaya, Victoria V. 2013. Acquisition of compounds in Estonian and Russian: frequency, productivity, transparency and simplicity effect. – Eesti Rakenduslingvistika Ühingu aastaraamat, 9, 23–42. https://doi.org/10.5128/ERYa9.02

Ceitlin, Stella N. 2009. Ocherki po slovoobrazovaniju i formoobrazovaniju v detskoj rechi ['Essays on word formation and formation of word forms in child speech']. Moscow: Znak.

Clark, Eve; Berman, Ruth 1984. Structure and use in the acquisition of word-formation. – Language, 60 (3), 542–590. https://doi.org/10.2307/413991

Dressler, Wolfgang U. 2005. Word-formation in natural morphology. – Pavol Štekauer, Rochelle Lieber (Eds.), Handbook of Word-Formation. Dordrecht: Springer, 267–284. https://doi.org/10.1007/1-4020-3596-9_11

Dressler Wolfgang U.; Mattes, Veronika; Korecky-Kröll, Katharina; Sommer-Lolei, Sabine 2017. Acquisition of derivational morphology. – Paper presented at the Annual workshop on "Pre- and Protomorphology in Language Acquisition", University of Vienna, 2–3 February.

Dziubalska-Kołaczyk, Katarzyna (Ed.) 2014. Special volume on linguistic complexity. Poznan Studies in Contemporary Linguistics, 50 (2).

Kasik, Reet 2013. Komplekssete sõnade struktuur ['The structure of complex words']. Tartu Ülikooli eesti keele osakonna preprindid 3. Tartu: Bookmill.

Kazakovskaya, Victoria V. 2017a. Imennoe slovoobrazovanie v rechi mono- i bilingval'nyx detej rannego vozrasta ['Nominal word formation in early L1 and L2']. – Problemy ontolongvistiki 2017. Osvoenie i funkcionirovanie jazyka v situacii mnogojazychija ['The problems of ontolinguistics']. Ivanovo: Listos, 93–96.

Kazakovskaya, Victoria V. 2017b. Acquisition of nominal compounds in Russian. – W. U. Dressler, M. Kilani-Schoch, N. Ketrez (Eds.), Nominal Compound Acquisition. Language Acquisition and Language Disorders 61. Amsterdam: John Benjamins, 63–90. https://doi.org/10.1075/lald.61.04kaz

Kerge, Krista 2016. Word-formation in the individual European languages: Estonian. – Peter O. Müller, Ingeborg Ohnheiser, Susan Olsen, Franz Rainer (Eds.), Word-Formation. An International Handbook of the Languages of Europe. Vol. 5. Handbooks of Linguistics and Communication Science 40. Berlin, New York: De Gruyter, 3228–3259. https://doi.org/10.1515/9783110424942-009

Laaha, Sabine; Gillis, Steven 2007. Diminutives as pioneers of derivational and inflectional development – a cross-linguistic perspective. Antwerp Papers in Linguistics 112. University of Antwerp

MacWhinney, Brian 2000. The CHILDES Project: Tools for Analyzing Talk. 3rd Edition. Mahwah, NJ: Lawrence Erlbaum Associates.

Padrik, Marika 2010. Word-Formation Skills in Estonian Children with Specific Language Impairment. Dissertationes pedagogicae Universitatis Tartuensis 8. Tartu: Tartu Ülikooli Kirjastus.

Protassova, Ekaterina; Voeikova, Maria 2007. Diminutives in Russian at the early stages of asquisition – Ineta Savickienė, Wolfgang U. Dressler (Eds.), The Acquisition of Diminutives. A Cross-Linguistic Perspective. Language Acquisition and Language Disorders 43. Amsterdam: John Benjamins, 43–72. https://doi.org/10.1075/lald.43.03pro

Ravid, Dorit 2004. Derivational morphology revisited. Later lexical development in Hebrew. – Ruth A. Berman (Ed.), Language Development across Childhood and Adolescence. Trends in Language Acquisition Research 3. Amsterdam, Philadelphia: Benjamins, 53–81. https://doi.org/10.1075/tilar.3.07rav

Shvedova, Natalia J. (Ed.) 2005 [1980]. Russkaja grammatika. T. 1: Fonetika, fonologija, udarenie, slovoobrazovanie, morphologija ['Russian Grammar. Vol. 1: Phonetics, phonology, stress, intonation, derivation, morphology']. Moscow: Nauka.
Savickienė, Ineta; Dressler, Wolfgang U. (Eds.) 2007. The Acquisition of Diminutives. A Cross-Linguistic Perspective. Language Acquisition and Language Disorders 43. Amsterdam: John Benjamins. https://doi.org/10.1075/lald.43
Schipke, Christine; Kauschke, Christina 2011. Early word formation in German language acquisition: A study on word formation growth during the second and third years. – First Language, 31 (1), 67–82. https://doi.org/10.1177/0142723709359240
Suurmäe, Kadri 2013. Eesti verbide leksikaalgrammatiliste kategooriate omandamine: transitiivsus ja kausatiivsus. Master thesis in Tallinn University.
Zurek, Wojciech H. (Ed.) 1990. Complexity, Entropy and the Physics of Information. Redwood City: Addison Wesly.
Vare, Silvi 2012. Eesti keele sõnapered I ['The Estonian Word Families']. Tallinn: Eesti Keele Sihtasutus.
Xarchenko, Vera; Ozerova, Elena G. 1999. Slozhnye slova v detskoj rechi ['Compounds in child language']. Belgorod: Belgorod University Press.

Reili Argus (Tallinn University) research interests are connected mainly with first language acquisition, with acquisition of morphology, aspectuality, modality and evidentiality.
Narva mnt 25, 10120 Tallinn, Estonia
reili.argus@tlu.ee

Victoria V. Kazakovskaya (Russian Academy of Sciences) is a leading researcher at the Department of Theoretical Grammar at the Institute for Linguistic Studies. Her research interests include philology, linguistics, first language acquisition.
Tuchkov per. 9, 199053 St. Petersburg, Russia
victory805@mail.ru

NIMISÕNATULETUSE OMANDAMINE VARASES EESTI JA VENE KEELES

Reili Argus, Victoria Kazakovskaya
Tallinna Ülikool, Vene Teaduste Akadeemia

Nimisõnatuletuse omandamine ei ole esimese keele omandamise uuringute kuigi põhjalikult uuritud valdkond ning nimisõnatuletuse omandamise eri keeli kõrvutavat käsitlust veel ei ole. Seetõttu puudub informatsioon ka selle kohta, kuidas laps omandab erinevad tuletusmallid ja -liited keeles, kus tuletussüsteem on rikas ja reeglipärane, või sellises keeles, kus see süsteem on küll iseenesest rikas, kuid mitte kuigi läbipaistev. Üldiste keelelise kompleksusega seotud põhimõtete järgi peaks laps alustama omandamist kõigepealt lihtsõnadest ja alles pärast seda komplekssetest sõnadest, st komplekssus peaks kasvama arengu käigus. Siinse artikli aluseks on kahe tüpoloogiliselt erineva keele, eesti ja vene keele pikiuuringu andmed. Tulemused osutavad kahe tuletussüsteemi omandamise märkimisväärsetele erinevustele. Tuletus omandatakse vene keeles kiiremini. Eesti laste kõnes on tunduvalt vähem nimisõnatuletisi ning nad kasutavad väiksemat arvu tuletusliiteid. Komplekssete struktuuride omandamise üldine põhimõte ehk nn "lihtsamalt keerulisemale" (ingl *building block model*) kehtib mõlema keele nimisõnatuletuse omandamise kohta ainult osaliselt.

Võtmesõnad: keeleomandamine, nimisõnatuletus, komplekssete struktuuride omandamine, tuletusmallide omandamine, vene keel, eesti keel