

ACQUISITION OF COMPOUNDS IN ESTONIAN AND RUSSIAN: FREQUENCY, PRODUCTIVITY, TRANSPARENCY AND SIMPLICITY EFFECT

Reili Argus, Victoria Kazakovskaya

Abstract. This article discusses the early phases of acquisition of noun compounds in two typologically different languages – Estonian and Russian. Longitudinal data on three Russian and two Estonian typically-developing monolingual children and their caregivers was analysed and compared in relation to first emergence, the development of different compound types, productivity, frequency, simplicity and transparency of compounds, and also to the impact of morphological wealth in the input on the output.

Although the frequency of compounds is somewhat different in the two languages being observed, the first compounds emerge in the speech of Estonian and Russian children almost at the same age and quite early, at the end of the second year of life.

In both Estonian and Russian, children first acquire productive patterns in both languages, Estonian and Russian, respectively. Factors such as simplicity and frequency are intertwined in the acquisition process. Simple structures occur at the same time as frequent ones but it cannot just be assumed that simplicity is a stronger factor influencing the choice of first noun compounds than frequency. Frequency effects are not straightforward: there are differences in the individual usage of compounds by the caregivers, and the frequency of compounds in child-directed speech does not always seem to be reflected in the speech of the children.*

Keywords: language acquisition, compounds, child-directed speech, frequency, productivity, transparency, simplicity, Estonian, Russian

* The research was carried out with the financial support of the RF President's Foundation (grant No 1348.2012.6 'The School of Functional Grammar in St. Petersburg').

1. Introduction

Little or no research has been carried out on the acquisition of compounding in Finno-Ugric languages, such as Estonian, nor in Indo-European languages with rich morphology, such as Russian. In general, the acquisition of compounding has been much less investigated than the acquisition of inflection in many languages (cf. Savickienė, Dressler 2007, Stephany, Voejkova 2009). This is probably mainly due to the fact that the study of compounds is quite a recent development.¹ In Balto-Slavic languages, compounds have not been studied since it was assumed that there are few. However, there are works that analyse occasional compounds in Russian children's speech (Kharchenko, Ozerova 1999) and a study of children's compounds in Lithuanian has been started (Dabašinskienė 2010). The acquisition of skills of compounding in Estonian children with specific language impairment is described by Marika Padrik (2010).

This article is an attempt to add data to the areas as yet not covered in first language acquisition and discusses the early phases of acquisition of compounds in two typologically different languages – Estonian and Russian. Investigation into the acquisition of compounds has so far been concentrated on languages that possess many compounds but not a very rich morphological system (Nicoladis 2006, Berman 2009). Both Estonian and Russian are languages with a rich morphological system, although the system of compounding is, to some extent, different and a comparison with the available data of other languages could offer evidence of how the system of compounding with which children are faced can influence the process of acquisition. Thus, we would like to focus on how the differences in two systems reflect in acquisition and try to answer the question whether children acquire productive patterns first, or whether it is merely frequency exerting an influence.

Compounding is a productive word-formation process in many languages and especially in the Germanic ones (Dressler et al. 2010). According to Natural Morphology, a child acquires the more productive parts of grammar first (cf. Dressler 1999, Kilani-Schoch, Dressler 2005). It can therefore be assumed that the **productivity** of compounding in different languages and the productivity of certain compounding patterns can affect the age of emergence of compounds, the choice of first patterns, and the speed of acquisition of compounding.

The age and speed of the acquisition of morphology has been argued to be dependent on the richness of the system children encounter (Laaha, Gillis 2007). **Frequency** of certain structures, like inflections, derivation suffixes, etc., has been noted to have a great importance on the order of acquisition of different grammatical elements (cf. Argus 2008, 2012). As compounding has a different degree of productivity in different languages and as a result different compounds have different frequency in the input (child-directed speech, CDS), it could be assumed to have an effect on the acquisition of compounding.

Eve Clark and her colleagues have argued that the principle of **simplicity**, amongst others, is important in child word formation (Clark, Berman 1984). Simplicity is measured by the degree of change in a form: the less the word changes, the simpler it is (ibid.: 548). For example, it has been argued that English simple noun+noun (N+N) compounds are easier to acquire than compounds with the

¹ See, in particular, a short description of the acquisition of nominal compounds in Dutch, English, French, German, Greek, Hungarian, Italian and Spanish according to different models compounds (Rivista di Linguistica 1992), as well as English (notably Clark et al. 1985, 1986, Nicoladis 2002), Swedish (Mellenius 1996, 1997), Hebrew (Berman, Clark 1989, Clark, Berman 1984, 1987, Berman 2009), Greek (Stephany 2010) and French (Nicoladis 2002, 2007).

suffix *-er* (Nicoladis 2006: 119). Compounds with differing degrees of simplicity can be found in both Estonian and Russian and an early acquisition of simpler patterns could be supposed, based on work previously conducted.

Simplicity can be considered to be connected with **transparency**. Morphosemantically transparent compounds will be more relevant and may be acquired earlier than opaque ones. According to Libben's basic scale of transparency (1998), the most transparent compounds are those with a structure where both head and non-head (modifier) are transparent. In other words: compounds with the meaning of each constituent morpheme contributing directly to the meaning of a compound (*door+bell* in English, *radio+peredacha* 'radio+transmission=broadcast' in Russian, *seene+supp* 'mushroom+soup' in Estonian) are easier to acquire than others. It also has to be taken into account that interfixed compounds are also morphosemantically less transparent than interfixless ones, because the interfix does not contribute to the meaning of the compound (Dressler et al. 2010: 15). Finally, compounds are semantically more descriptive than derivations, therefore morphosemantically more transparent.

Some prosodic constraints have also been considered to influence the acquisition of morphology (Laalo 1994, Argus 2008). In the early stage of acquisition children use shorter words and prefer disyllabic patterns. For example, in Finnic languages the period of disyllabic structures (among them compounds consisting of two disyllabic words) is often to be considered to last for even five or six months (Laalo 1994: 340, Argus 2008: 32). Since compounds are longer than root-words, it could be that the children do not acquire compounds very early and start to use them only after the disyllabic period, or prefer disyllabic patterns also in compounds. Moreover, it is known that compounds are also harder to read than, for example, phrases (Böer et al. 2012) and on this basis it could be supposed that they could also be harder to comprehend. At the same time there can be some factors which facilitate the acquisition of compounds: the formation of compounds is different from ordinary inflection – it is often merely a juxtaposition of two or several stems. Therefore it could be assumed that it could be easier for the child to acquire than, for example, a derivation.

In this paper we discuss the early phases of first language acquisition of noun compounds. The restriction to nominal compounds, i.e. compounds which function as Nouns (Ns), is justified by the rarity or very different structure of verb compounds in many languages. Compounds are analysed comparatively in Estonian and Russian, concentrating on different factors, like frequency, productivity, simplicity, and transparency, which could influence the age and the speed of their acquisition.

The longitudinal data of three typically-developing, monolingual Russian and two Estonian children and their caregivers has been analysed and compared according to first emergence, development of different compound types, productivity, frequency, simplicity and transparency of compounds, and also the impact of morphological wealth in the input on the output (child's speech, CS).

2. Compounds in Estonian and Russian

Compounding has a different degree of productivity in the languages under observation and it is a productive strategy of word formation in Estonian. Although compounding includes the formation of compound verbs, nominals (Ns, ADJs, PRONs) and adverbs, it is very productive only in nouns and adjectives. Noun compounds are quite frequent in colloquial speech and also in child-directed speech (henceforth CDS) (e.g. compounds like *lutipudel* 'bottle with pacifier', *riisipuder* 'rice porridge'). But they are even more frequently used in terminology, for example, in scientific and technical disciplines like biology (most parts of names of plants are compounds) and engineering.

Although compounding is quite often used also in Russian, other means of word formation are used more frequently, namely, suffixation, prefixation, a combination of suffixation and prefixation, substantivation etc. (Russian Grammar 1980). Therefore, compounds used to be less frequent, but are now becoming more frequent in CDS and, consequently, in CS.

Compounding is not always used when building new words in Russian, for example, it is uncommon in verb building. Such rare examples are 'calques' from Greek: *blagodarit'* 'to thank' or *zloslovit'* 'to say spiteful things'. However, compounding is a very useful way of building nouns: *knigoljub* 'book+to love=book lover', *inoplanetjanin* 'different+a planet=alien', and adjectives: *ostroumnyj* 'sharp+mind=witty', *svetlo-goluboj* 'light+blue=light blue', similar to Estonian.

The Corpus of Spoken Russian data has shown that compounds are not frequent in colloquial speech, but with some exceptions like *kofemolka* 'coffee+to mill=machine for milling of coffee', *nervotrepka* 'nerves+to spoil=hassle'. However, they are favoured in scientific and technical disciplines as in Estonian: *mirovozzrenie* 'world+view=world-view', *jazykoved* 'language+to know=linguist'.

Table 1 presents the classification and comparison of compound nouns in Estonian and Russian according to the word class and inflectional form of the head and modifier.

Compounds in both languages can concatenate components with different structures². Table 2 provides a classification and comparison of compounds in Estonian and Russian according to the degree of complexity of constituents forming the compound.

² A compound has been defined here according to structural characteristics – a compound is a word consisting of at least two root morphemes (see also Kerge 1990: 4).

Table 1. Estonian and Russian compound nouns according to the word class, form of constituents and productivity of formation pattern (Vasilevskaja 1962, Russian Grammar 1980, Zemskaja 1992, Erelt et al. 1995, Gudilova 2005)

| Word class | Estonian | Russian |
|---|---|---|
| Head: Verb | Compound nouns consisting of head functioning as a verb are missing in Estonian. | The dominating structure of compounds X+V : the modifier can be a noun (or ADJ, PRON, NUM, V) followed by a verb as the second member, a head of the compound. The compound might be with an interfix (<i>o, e</i> etc.), and with suffixation – zero (\emptyset): <i>vod-o+pad</i> ‘water+fall+ \emptyset =waterfall’ or material one: <i>mor-e+plava-tel</i> ‘sea+sail+suffix tel =seaman’. A complex suffixation is the most productive pattern. Compounds where a verb is the first member, and a head (V+X), are relatively rare: <i>sorv-i+golova</i> ‘tear away+head= madcap’. The verb part with an interfix coincides with IMP.SINGL. The sub-pattern is unproductive. |
| Head: Noun | All Estonian noun compounds belong to this type. | Compounds might be with an interfix: <i>les-o+step</i> ‘forest+steppe=forest-steppe’ or without one: <i>plachsh+palatka</i> ‘raincoat+tent=duckbill’. This pattern is less productive. |
| Modifier: Substantive | Most productive and common pattern in Estonian: <i>riisi+puder</i> ‘rice+porridge=rice porridge’. | This pattern is productive: <i>nos-o+rog</i> (horn-nose) ‘nose+horn=rhinoceros’; <i>jug-o+vostok</i> ‘south+east=southeast’. |
| Modifier: Adjective (or its part in Russian) | The pattern is not very productive: <i>noor+tööline</i> ‘young+worker=young worker’. | The pattern is very productive: <i>sukh-o+phrukt</i> ‘dry+fruit=dried fruit’. Quite often some part of ADJ and N are concatenated (a kind of abbreviation): <i>med(icinskaja)+sestra</i> ‘med(ical)+sister=nurse’. |
| Modifier: Pronoun | The pattern is unproductive: <i>enese+kaitse</i> ‘self+defence=self defence’. | The pattern is not very productive: <i>sam-o+proverka</i> ‘oneself+check=self-test(ing)’. |
| Modifier: Numeral | The pattern is unproductive: <i>kahe+kõne</i> ‘two:GEN+speech= dialogue’. | The pattern is not very productive: <i>perv-o+istochnik</i> ‘first+source=original source’. |
| Modifier: Verb | The pattern is productive: <i>istu-ma+jäämine</i> ‘sit-INF+staying=remain seated’. | The pattern is unproductive: <i>vert-o+let</i> ‘turn+fly=helicopter’. |
| Modifier: Particle | The pattern is not very productive: <i>all+kiri</i> ‘under+letter/writing=signature’. | The pattern is unproductive: <i>pol-u+mrak</i> ‘semi+darkness=shade’. |

Table 2. Constituents of compounds in Estonian and Russian and productivity of formation pattern

| Constituent | Estonian | Russian |
|--|--|---|
| Root | Productive and frequent pattern: <i>paber+raha</i> 'paper:NOM+money=paper money. The most productive pattern is with singular genitive case form which is a base form without any suffixes occurring as a modifier: <i>seene+supp</i> 'mushroom:GEN+soup=mushroom soup'. | May be used as an interfix: <i>sam-o+let</i> (aeroplane) 'oneself+fly= aeroplane' or without one: <i>divan+krovat'</i> 'sofa+bed=sofa bed'. The pattern is productive. |
| Stem | The pattern is productive: <i>kauba+ndus+töötaja</i> 'trade+derivational suffix+worker=salesperson'. | The pattern is very productive: <i>kreslo+kachalka</i> 'armchair+rocking chair=rocking chair'. |
| Inflected form | The pattern is productive: all noun case forms can occur as a modifier: <i>kala-l+käik</i> 'fish-ALL'+going:NOM =fishing'; <i>küla-s+käik</i> 'village-ILL+visit=visit'. | The pattern is productive: <i>zav+kafedroj</i> 'head (shortened from <i>zavedujuchshij</i>)+sub-department:INSTR SINGL=head of sub-department'. Typical case form of inflected modifier is INSTR. |
| Interfixed form (containing a meaningless linking element) | The pattern is not very productive: <i>töötaja-s+pere</i> 'worker-s+family=people who work together'. | The pattern is productive: <i>gr'az-e+vod-o+lechenie</i> 'mud+water+cure =water-and-mud cure'. |
| Part of a word | The pattern is productive: <i>lamamis+tool</i> (<i>lama+mise+tool</i>) 'lay+derivational suffix+chair=reclining chair'; <i>kiir+tee</i> (<i>kiire+tee</i>) 'quick+road=highway' – the last phoneme of a stem disappears. | The pattern is very productive: <i>sten(naja)+gazeta</i> 'wall+newspaper=newspaper on the wall'. |
| Phrase | The pattern is not very productive: <i>kahe+inimese+voodi</i> (<i>kaks inimest+voodi</i>) 'two:GEN person:GEN+bed=bed for two persons'. | The pattern is unproductive: <i>vremj-a+provozhdenie</i> 'time+pass= pastime'. |

The systems of Estonian and Russian compounds differ (see Table 2).

1. The type of head-word. In Russian it can be a noun or a verb, which mainly follows a noun, while in Estonian it is always a noun, if the noun is a compound itself. However, the similarity is that both languages have compounds with a head that is a noun.
2. The inflected forms. In Estonian compounds the inflected form is first in a compound, whereas in Russian, it is second.
3. The interfixed form. In Russian the interfix (only vowels) is widely used, while in Estonian (where the interfix could be a consonant) it is quite rare. Moreover, in both languages, compounds could concatenate not only the roots of words, but also the stems.
4. The degree of productivity. In Russian language where the modifier is an adjective are productive, while the most productive pattern of compounding in Estonian is one consisting of noun in genitive as modifier following

the patterns where the modifier is an adjective or a verb. However, the compounds in which the modifier is a pronoun, numeral or particle are equally rare in both languages.

3. The data

Table 3 gives a general overview of the data used for the analysis. The longitudinal corpora of three Russian (St. Petersburg) and two Estonian (Tallinn and Tartu) children (~135 hours of recorded speech (119.5 hours for the Russian children, 17 hours for the Estonian ones) were transcribed using CHILDES³. Although the quantity of Estonian data was significantly smaller, the number of compounds is almost the same as in the Russian corpora.

Table 3. Number of compound types and tokens in the children’s speech (Vanja, Liza, Philja, Hendrik, Andreas) and the CDS. MLU – Mean Length of Utterance (here counted in words)

| Russian | Compound Types | Compound Tokens | Types/Tokens |
|---|----------------|-----------------|--------------|
| Vanja 1;9 (MLU 2.008) – 3;6 (MLU 3.210) | 22 | 119 | 5.36 |
| Vanja’s CDS | 34 | 251 | 7.38 |
| Liza 1;7 (MLU 1.008) – 3;6 (MLU 4.170) | 18 | 73 | 4.05 |
| Liza’s CDS | 39 | 95 | 2.43 |
| Philja 1;5 (MLU 1.009) – 2;8 (MLU 2.900) | 11 | 68 | 6.18 |
| Philja’s CDS | 16 | 144 | 9.00 |
| Estonian | | | |
| Hendrik 1;8 (MLU 1.651) – 2;5 (MLU 2.811) | 28 | 70 | 2.50 |
| Hendrik’s CDS | 93 | 200 | 2.15 |
| Andreas 1;7 (MLU 1.093) – 3;0 (MLU 3.044) | 167 | 329 | 1.97 |
| Andreas’s CDS | 184 | 511 | 2.78 |

The total number of compound types and tokens differs in sub-corpora. Among the Russian children, Vanja has 22 types and 119 tokens, Liza has 18 and 73, and Philja has 11 and 68 within their material, while one of the Estonian children, Hendrik, has 28 types and 70 tokens⁴ but the other, Andreas, has 167 types and 329 tokens in his data.

³ Child Language Exchange Data System, <http://chilides.psy.cmu.edu> (accessed 23 October 2012).

⁴ It must be mentioned here that the corpora of Hendrik is somewhat restricted, the quantity of speech samples is much smaller in the early period (only 20 minutes per month).

4. Acquisition of compounds

4.1. Emergence of first compounds

First compounds emerged in the speech of the Estonian children early: *lume+pall* ‘snow+ball’, *notsu+sig* ‘pig:DEM+pig=cavy’ (Andreas 1;9), *kaka+nunnu* ‘shit:DEM+poop:DEM’ (Hendrik 1;10). In the beginning both children had very few compounds: there were four compound types and four tokens in Andreas’s speech and only one compound type with two tokens occurring in the speech of Hendrik.

In the speech of Andreas the N+N compositions emerge at the beginning of the protomorphological period⁵, in the case of Hendrik it emerges three months before this period or, in other words, before noun inflection starts to develop productively. Compounds are used only in one case form at first by both children, the first true mini-paradigm appeared only at 2;7 in Andreas’s speech. Compounds as longer words do not cause any difficulties for children. Andreas starts to use first compounds at the end of his disyllabic period. First compounds emerge in the speech of Hendrik at the period when he has almost only disyllabic words. His first compound clearly demonstrates the preference for using trochaic phonological pattern – he uses the diminutive form from one-syllable form *nunn* (junn) > *nunnu* to adapt the whole compound to the pattern consisting of two disyllabic words *kaka+nunnu*.

Comparing compounding with other word formation possibilities, it must be mentioned that compounding emerges in the speech of Andreas at the same time as other word formation strategy starts to develop – he used first compounds at the same time as first derivated verbs. Hendrik’s compounding emerges much (6 months) earlier than derivation (cf. Argus 2012: 10).

According to Nicoladis (2006: 110), children seem to have little trouble ordering N+N compounds. The same can be stated on the basis of the Estonian data. The ordering of parts in compounds is not absolutely clear for both children, but only at the start of using compounds. Hendrik used both *abi+iuiu* and *iuiu+abi* ‘emergency car’ at 1;11, even later, at 2;4 he used a compound *onu+kass* (should be *kassi+onu*) ‘uncle+cat=male cat’ and from the context we can see that he intended to use the correct version *kassi+onu*. Andreas had only one such ordering error *pika+päka* (should be *päkapikk*) ‘dwarf’ at 2;1.

First compounds emerged in the speech of the two Russian children at the same age 1;9: e.g. *sam-o+let:DIM* (‘oneself+to fly=aeroplane’) (2 types/7 tokens for Philja); *photo+apparat* (‘photo+apparatus=camera’)⁶, *syr-o+ezh-ka* (‘raw+to eat=russula’) (2 types/4 tokens for Liza). Vanja’s earlier compounds emerge at 2;2: e.g. *metall-o+lom+ø* ‘metal+to break=scrap’, *vert-o+let+ø* ‘to turn+to fly=helicopter’, and also a compound which is typical for colloquial speech *chern-o+plod-ka* ‘black+foetus=black chokeberry’ (4 types/12 tokens).

In the speech of all the Russian children the first compounds emerge at the stage of protomorphology, because the number oppositions of nouns (SING vs. PL) are already observed, and the contrast verb forms (the mini-paradigms) are used. In contrast to early Estonian compounds, the erroneous order of parts is not present in the Russian data.

⁵ The protomorphological period is considered to start from the emergence of first mini-paradigms (for criteria, see Bittner et al. 2003: xvi) and signifies the beginning of active acquisition of morphology.

⁶ This concept *photo+apparat* ‘camera’ is named by a specific children’s word (*pack*), but at the same time in the girl’s speech the word *apparat* ‘apparatus’ does exist, which she uses for the ‘camera’.

Although the frequency of compounds is somewhat different in the two languages under observation, the first compounds emerge in the speech of the Estonian and Russian children at almost the same age, at the end of second year of life, which is quite early.

4.2. The structure and productive patterns of early compounds

A productive compound pattern is one that is applied freely and unconsciously to new loanwords (Dressler et al. 2005). In the case of Estonian, the most productive pattern among N+N compounds is a pattern with modifier in the genitive like *riisi+puder* 'rice:GEN+porridge'. Estonian early compounds are of a similar structure. Almost all compounds in the speech of both children were N+N, where the modifier is a singular genitive stem (see for patterns the tables 1 and 2): *une+laulu* 'sleep:GEN+song=lullaby' (Andreas 1;8), *tati+nina* 'snot:GEN+nose' (Hendrik 2;0).

Only some compounds with different structure (usually the first part as an onomatopoeic word) were found from the early compounds: *iiuiu+abi* 'sound of the emergency car+help=emergency car' (Hendrik 1;11). First compounds with different structure start to emerge in Andreas's speech at 2;0: *kustu+kumm* 'erase+gum=rubber'. These compounds belong to different formation patterns, the modifier could be a noun in nominative, like *mootor+ratas* 'engine+bicycle=motorcycle', as well as an adjective *vana+isa* 'old+father=grandfather' (2;4), or a shortened genitive form like *hingamis+rohi* 'breathing+medicine=inhalator', or a verb stem like *röst+sai* 'roast+bread=toast' (2;1). The number of compounds with modifier in singular nominative remained small during the whole observation period, and there were only five such compounds in the data. First compounds with a modifier in plural genitive occur at Andreas's age 2;8: *isa-de+päev* 'father-PL.GEN+day'.

The distribution of different structures of compounds remains quite similar from the age of 2;0 until the end of the observation period: Approximately two-thirds of all compounds in child speech belong to the most productive pattern (N+N, modifier in singular genitive) and one-third of compounds consist of compounds with different structure. The pattern where the modifier is an adjective (e.g. *must+sõstar* 'black+currant' (Andreas 3;0)) remains the most frequent pattern among infrequent patterns.

The frequency of compounds formed according to the most productive pattern in CDS changes when the child gets older: it can vary from 90% at the beginning of the observation period down to 67% at 2;8. Hence it can be argued that productivity and frequency are factors which are intertwined in the process of acquisition of compounds because the frequency of compounds formed according to productive pattern can change in CDS. It can be seen as a carer's fine-tuning.

The structure of children's first compounds is similar in modern Russian. The boys' data shows that the predominance of X+V compounds (like *tepl-o+khod* 'heat+to go=motor boat') is very marked, with X+N type (*zoo+park* 'zoological+garden=zoo', *mul't+phil'm* 'animation+film=animated cartoon') ranking behind them, but the predominance of this pattern was found in the girls' corpus.

Both the patterns X+V and X+N are the most numerous types of early compounds in Russian Child Speech. Moreover, in all cases these first compounds are subordinate endocentric right-headed ones, according to the classification of compounds (Scalise, Bisetto 2005).

Then, within the X+V type of compound, such structures as PRON+V (*sam-o+kat* 'oneself+to roll=scooter'), ADJ+V (*syr-o+ezhka* 'raw+to eat=russula (an edible mushroom)') and V+V (*vert-o+let* 'to turn+to fly=helicopter') were found.

In the frames of the X+N type of compound, in second place in the boys' corpora was the structure N+N: *nos-o+rog* 'nose+horn=rhinoceros' (V. 2;7). Nevertheless, there are less numerous compounds in comparison with the prototypical presenters of these models, namely N+V and ADJ+N.

The members of compound are connected by one of the prototypical interfix *o*: e.g. *metall-o+lom+ø* 'scrap', *sam-o+let+ø*=aeroplane'.⁷ In most cases, the formation of compounds is followed by the addition of suffixation (usually zero one, see examples above), which also corresponds to the leading trends in Russian compounding. Zero suffixation occurs in the verb part: *-let* (*samolet* 'aeroplane') > *let(et)* 'to fly'+ø', etc.

In Vanja's and Liza's data X+N compounds where two words (stem+ending, which is material or zero one) are concatenated by a hyphen prevailed: e.g. *sorok-a+belobok-a*⁸ 'magpie+white-sided=magpie' (Vanja 2;10).

As to preferences and more detailed characteristics of Russian children's compounds, it can be observed that Vanja's data displayed the most variation in compounds in both patterns and their types, although earlier compounds emerge quite late (1;9 is the time of appearance of the first compounds in his CDS). This corpus has the middle 'types/tokens ratio' (TTR) of compounds: 5.36 for CS, 7.38 for CDS.

Liza's data shows, as mentioned above, a lot of so-called hyphen (N+N) models of compounds. The fact that there is an abundance of these words, and most of them are heroes of Russian folk tales, can be explained by the significant role of book reading in the life of this family. The girl's mother (who is the main caregiver) works as a teacher. TTR of compounds in this corpus is the lowest: 4.0 for CS, 2.43 for CDS. In Philja's data the list of compounds varied less but they are very frequent. This corpus has the highest TTR of compounds: 6.18 for CS and 9 for CDS.

Thus, the data of Estonian and Russian children's speech reveal a similar tendency in the emergence of first compounds: children start with the most productive patterns of compounding in both languages regardless of their perceived complexity.

4.3. Simplicity and morphosemantic transparency

Formation of compounds is a simple process in almost all cases in Estonian, only in some cases will a meaningless element be interfixed. The modifier can, in some patterns, be an inflected form like allative or illative case form, but these patterns are not frequent. The most frequent and productive pattern is with modifier as a root in singular genitive and it cannot be considered to be more complicated than the pattern with singular nominative. The main reason is that the genitive case form serves like a base form for other (semantic) cases, it has no suffix and

⁷ It does not correspond with some comments on compound acquisition (Dressler et al. 2010: 15), at least on these Russian data.

⁸ Interestingly, this part (*beloboka*) of the compound is itself a compound, formed by concatenated *bel-o+bok-a* 'white+side' with a special rhyming ending.

is often the same form as the nominative. Hence, when Estonian children start with compounds consisting of modifier in genitive case, we cannot claim that they start from more complicated patterns because the compounds with genitive noun as modifier can be as simple for them as those with nominative modifier. These compounds are also transparent, as both members of the compound, head and the modifier are contributing directly to the meaning of a compound (e.g. in the compound *une+laul* ‘sleep+song=lullaby’ both parts are directly contributing to the meaning of the compound).

The first compound with opaque meaning occurs in Andreas’ speech at the age of only 2;0, and it was *meri+sig* ‘sea+pig=cavy’. Hendrik (2;2) used a compound *mesi+mumm* ‘honey+spot=bee’ in which the head is opaque. There were very few compounds in CS where the head or modifier of the compound was opaque: *rohu+tirts* ‘grass+bit=grasshopper’ (Andreas 2;6), *musi+mumm* ‘kiss+spot=dear’. This kind of compounds also seems infrequent in CDS.

In Russian the most frequent compounds are formed according to an X+V pattern which is accompanied by zero suffixation. It seems that the process of suffixation complicates compounding: *molok-o+voz(i-t)* ‘milk+carry=milk tanker’. However, a zero suffix, which is added to the verb stem *voz(i-t)+ø* ‘to carry’, opens this stem, truncating both a thematic vowel (*a, e, i*) and a form-building suffix of an infinitive (*t*). Simultaneously a bare root is ‘uncovering’: *voz*; see also *pyl-e+sos* ‘dust+suck=vacuum cleaner’: *sos > sos(a-t)+ø*. Modifiers are a root: *molok-o* ‘milk’ or *pyl* ‘dust’.

The use of interfixes, in our opinion, is needed for phonetic reasons. Interfixes open syllables (according to ‘the law of open syllable’) and avoid clusters, which are acquired by Russian children with some difficulty (Gvozdev 1995). In other words, we tend to consider the use of interfixes as a bootstrapping element. Consequently, the forming of such compounds is quite a simple process, and as our results have shown, all children start with this pattern. Until 2;5 in Vanja’s and Liza’s data this is the main type of compounds⁹: Liza has 12 such compounds out of 18, Vanja has 12 out of 22, Philja 8 out of 11.

According to an analysis, modifiers might be Ns (6 out of 11 for Vanja, 3 out of 8 for Liza, 4 out of 8 for Philja): *beton-o+meshalka* ‘concrete+mixer=concrete mixer’; PRONs (4 out of 11 for Vanja, 2 out of 8 for Liza and Philja): *sam-o+kat* ‘oneself+to roll=scooter’; Vs (1 out of 11 for Vanja, 2 out of 8 for Philja): *vert-o+let* ‘to turn+to fly=helicopter’; ADJs (3 out of 8 for Liza): *syr-o+ezhka* ‘raw+to eat=russula’.

It should be noted that both parts of X+V compounds are morphosemantically transparent and they are joined by an interfix which is usually a vowel *-o-*: *sam-o+var* ‘oneself+to boil=samovar’ (13 out of 18 for Liza, 17 out of 22 for Vanja, 8 out of 11 for Philja). An interfix *-e-* (*pyl-e+sos* ‘dust+to suck=vacuum cleaner’) was marked only in two compounds (Liza 2;2, Vanja 2;5).

After 2;5 use of the second pattern X+N becomes more active. It is also a typical model of compounds for Russian, but to a lesser extent. The modifiers might be Ns (*nos-o+rog* ‘nose+horn=rhinoceros’ (Vanja 2;7, Liza 2;9)), ADJs (*bos-o+nozhka* ‘bare+foot:DIM=open-toe sandal’ (Liza 2;11)) or just their parts (*plast(icheskaja)+massa* ‘plastic+mass=plastics’ (Vanja 3;0)), and also NUMs (*sorok-o+nozhka* ‘forty+leg:DIM=centipede’ (Vanja 2;11)).

⁹ In the corpus for Philja, no compounds were found after 2;3.

In this model an interfix is used rarely, especially in the situation where the stem is cut and, moreover, concatenation of two words (stems with endings) without an interfix but by means of a hyphen: *ptichka+nevelichka* ‘bird:DIM+tiny creature-a birdy-wirly’ (Liza 2;10), *zajchik+pobegajchik* ‘hare:DIM+to run:N-a running hare’:DIM (Liza 3;10). Both compounds name creatures from Russian folk tales.

Thus, all child compounds are morphosemantically transparent ones, except for a few hyphen compounds such as *akula-+karakula* ‘shark+rhyming meaningless word=a huge shark’ or *kaljaka-+maljaka* (two informal rhyming nouns from ‘to daub’)=scribble’. Such compounds are a juxtaposition of two stems with their endings. These are left-headed compounds with opacity of the modifier which usually does not get used separately. As was mentioned above, the most use of such compounds is a characteristic of Liza’s data, especially her CDS.

4.4. Frequency effects (input vs. output)

The frequency of noun compounds can vary (from 4.8% to 19.8% among all nouns) in different recordings of both Estonian children. The number of compounds could depend very much on the topics of the conversation: for example, in the recording of Andreas (2;1) the mother and the child were looking at pictures in a book and the compounds *rongi+sõidu+raamat* ‘train+ride+book =a book about travelling in a train’ and *lamba+tall* ‘lamb+son’ appeared several times. The percentage of compound Ns still increases from the age 2;5 and does not decrease again during the observation period.

The same tendency can be observed in CDS of Hendrik: At 1;8 there were 5.1% of compounds among all noun tokens and 14.5% of compounds at 2;5. In Andreas’s data the percentage of compounds increases from 4.4 (1;8) to 5.8% (2;5) but remains almost stable (approximately 5% of the nouns were compounds) until the end of observation period.

In the data of Estonian children the frequencies of compound tokens are quite similar to those found in CDS. Usually there are about 30% more compound tokens in CDS than in CS, except in Andreas’s recordings at the end of the observation period, starting from 2;8 where he has even more compound tokens (48) than his mother (28). The difference in the number of compound types is similar, although slightly less significant, there are usually 10–20% more compound types in the CDS than in CS.

Similarly to the Estonian results, the frequency of Russian compounds depends on the topics of conversation and the activity, in particular, games with a particular toy. The data also demonstrates that there is a direct correlation between the number of tokens in CDS and the number of tokens in CS. However, the surprising thing is that CDS does not have many more token and types than CS but more tokens of the same type. For example, there are 22 children types and 34 adult ones in Vanja’s material (that is, half as much again), while Vanja’s tokens are 119, and the adult ones are 251 (that is a 110% increase). Almost the same results were revealed in Philja’s data (1.45 and 2.1). The indexes in Liza’s corpus are 1.2 and 1.3. Figure 1 gives an overview of the volume and increase of compounds in CS.

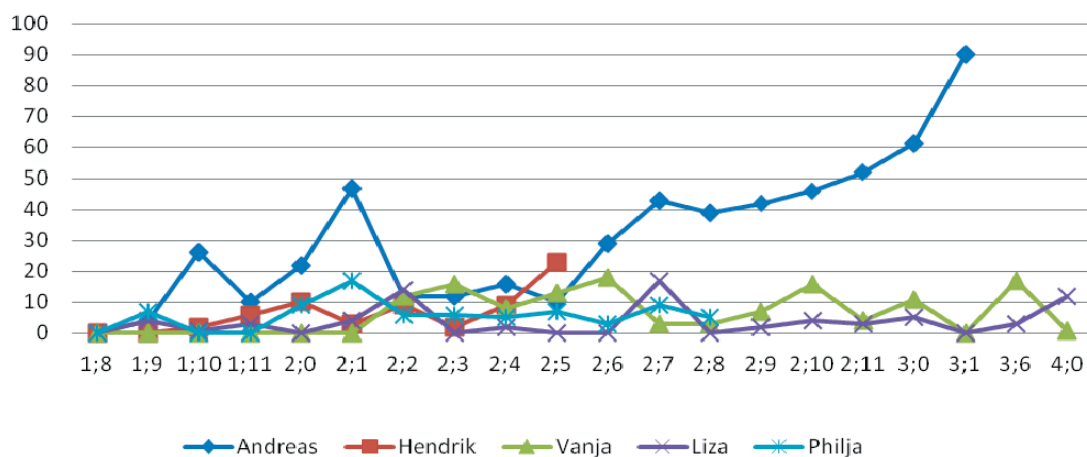


Figure 1. Noun compounds in the speech of Estonian and Russian children

It should be observed that the frequency of compounds in Russian CDS is not as high as in the Estonian data. Compounds consist of approximately 2% or less of all noun tokens: in Liza's CDS it is 0.7% (noun tokens are 12,420), in Vanja's CDS it is slightly higher at 1.1% (noun tokens are 22,786). The highest percentage of compounds was found in Philja's input at 1.63% (noun tokens are 9452). Interestingly, the comparison of both boys' data where the percentage of noun compounds is more than 1% has shown a quite similar dynamic of their development during the same period (1;8–2;9). Figure 2 gives an overview of the amount and increase of compounds in the caregiver's speech.

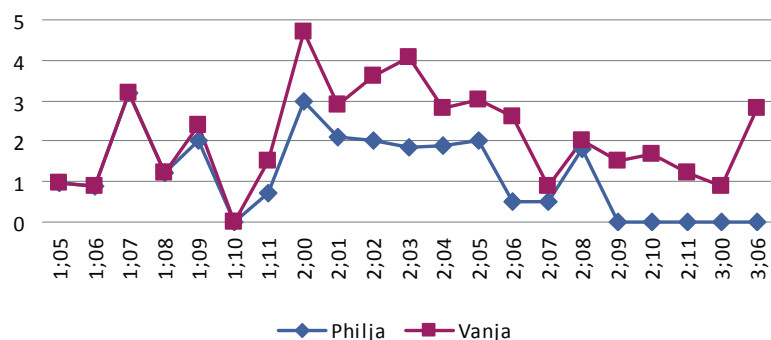


Figure 2. Noun compounds in Russian CDS (% of noun tokens)

Some additional comments are required on the pattern of adult compounds used. A definite correlation between the frequency and order of emergence of some models in CDS and CS was observed in Philja's data, which characterised the biggest percentage of compounds amongst all the Russian corpora. Philja's mother used only 16 prototypical models of noun compounds where the X+V patterns dominated, and this occurred from the beginning of observation period.

In Vanja's CDS the amount of X+V and X+N is almost equal. However, nominal models are slightly more prevalent (19, cf. 15). At the same time some of X+N patterns are represented by hyphen N+N compounds concatenated by two words which are not frequent and consequently they do not occur in the CDS with any stable frequency. Finally, a significant feature of Liza's CDS is an abundance of

such hyphen N+N compounds, which are proper names of fairy-tale folk heroes. These compounds are not frequent (e.g. 1 type / 1 token) in the CDS so they do not enter Liza's speech, which started, similar to the two boys, with the X+V system-prototypical pattern.

4.5. Productive usage of compounds

The emergence of oppositions between compounds and their members demonstrates the child's ability to detect the morphological principle of (de)composing internal form and meaning of words (Dressler et al. 2010: 4). The use of self-invented compounds can also demonstrate the child's ability to use compounding productively.

One of the Estonian children, Andreas, uses compounds productively almost from the beginning of emergence of compounds from age 1;10 when the parts of the compound *kiisu+raamat* 'cat:DEM+book=book about cats' also emerge separately in the recording. Hendrik starts to use compounds productively somewhat later (2;3) when parts of the compound *vanni+tuppa* 'bath+room' (*vann* 'bath' and *tuba* 'room') have also occurred in his speech samples.

First neologisms appeared in the speech of the Estonian children in a different order: Hendrik first starts to use self-invented compounds: *iiuu+abi* 'sound+help=emergency car' (1;10), and thereafter Andreas *kepi+mana* 'stick+old=grandmother with stick' (2;1). Both compounds are formed according to the most productive pattern of N+N compounding in Estonian. It must be mentioned that Estonian children start to use neologisms before their usage of compounds starts to be productive. Similar results were obtained from the materials of Russian diary observations (Kharchenko, Ozerova 1999) but such examples were not revealed in our Russian corpora.

Little evidence was found to support the claim about the productive use of Russian compounds. However, several frequent verbs were revealed in the speech of our children which can be used as members of compounds and separately (Gagarina 2008) at this time: e.g. *sloimat* 'to break' (see *metal-o+lom*), *vezti* 'to carry' (see *par-o+voz*) (Vanja 2;3), *s'jest* 'to eat' (see *syr-o+ezhka*) (L. 1;10).

As for neologisms, only two examples were noted in Liza's data: the type N+N *nos-o+rog-in-a* 'huge rhinoceros' (2;11), and type ADJ+V *syr-o+ezh-ichsh-a* 'huge russula' (4;00). Both words are incidentally-produced ones and are formed from compounds (which were earlier *syr-o+ezhka* (1;9), *nos-o+rog* (2;9)), with the help of augmenting suffixes *-in-* and *-ichsh-* expressing an extraordinarily big size (in contrast to diminutives). So the novelty of these words comes from derivation and this is very different from new compounds found in Estonian, however, both compounds belong to neologisms (Tikhonov 2003).

Moreover, children can form diminutives from compounds, but the diminutives usually appear later than normal compounds: cf. *par-o+voz* (1;10) – *par-o+voz+ik* (Liza 1;11); *sam-o+let* (2;3) – *sam-o+let+ik* (Vanja 3;6). However, in the data of the third Russian child, Philja, his earlier compounds (1;9) were diminutives *parovoz+ik*, *samolet+ik*. The conventional pattern, indicating some productivity of compounds, starts to work in his speech later: *vertolet* (2;1) – *vertolet+ik* (2;3).

5. Results and discussion

The main findings of our cross-linguistic study can be summarised as follows.

1. It has been found that compounds in the analysed data of Russian child language appear infrequently, although compounding is quite a productive pattern of modern noun formation. The very first compounds emerge early (1;9), but probably as lexicalised items because little evidence was found to support the claim of productive use of compounds.

First compounds in Estonian appear at almost the same time, at the end of the child's second year of life. In contrast to Russian child acquisition, Estonians also start to use novel compounds and compounding productively early, one child just a month after the emergence of first compounds at 1;10.

2. The coefficients of compound use by each Russian boy (both in types and tokens) are the same when compared with those that appeared in their CDS: To be more precise, in type, the index is 1.5 in Vanja's corpus and 1.45 in Philja's. Regarding tokens, the index is 2.1 for both boys, and 1.3 for the girl.

Individual differences can be found in the Estonian data: Hendrik's mother produced three times more compounds than Hendrik (index 3.3), while the number of compound types in the speech of Andreas and his mother were almost the same (index 1.1). Regarding tokens, a similar difference can be found: Hendrik produced three times fewer compounds than his mother (2.8) and Andreas approximately as many as his mother (1.6).

3. It is worth noting here that the three Russian caregivers used varying numbers of compounds: Vanja's grandmother and Liza's mother produced more than twice (their indexes are 2.1 and 2.4) as many compounds as Philja's mother. It may be stated, then, that frequency of compounds depends not only on language typology, but, to a great degree, on individual preferences and a dialogue topic or activity type.

The number of compounds in the Estonian CDS is also different due to the different amount of recordings but at the same time the percentage of compounds among all nouns is similar in the two sub-corpora.

4. The effect of simplicity of a compound is not an explicit one. We can assume that in Russian compounds where the first part (N, ADJ, PRON or even V) is followed by a verb, it might be semantically more transparent for a child and may therefore be preferred for productive usage. But at the same time, it is a frequent pattern within Russian compounds. They have a vowel interfix, which makes the stem simpler in a phonetic sense, and quite often a zero suffix, which renders it more transparent in morpheme aspect. After 2;5 the Russian children start to use the second pattern X+N more actively and within that there are compounds formed by concatenation of two words, one of which could be opaque and the other rhyming. Thus, prototypical Russian noun compounds (which are additionally endocentric, subordinate and interfixed) emerge first and are used more frequently than non-interfixed ones. The insertion of an interfix into a cluster of consonants probably makes articulation easier, which may be a reason for a child to prefer the interfixed forms. In addition, practically all the compounds used are morphosemantically transparent ones; and we can make comment on the transparency of both members of the compound.

There are only a few compounds emerging in the speech of the Estonian children with opaque parts and they emerge later than compounds which are

morphosemantically transparent. Compounds consisting of an opaque constituent or constituents are also rare in the input and that may be the other reason why the Estonian children start to use them later and do not produce them very frequently. The Estonian children did not use compounds containing a meaningless linking element and preferred the most transparent pattern with modifier in genitive.

5. The order of emergence of the different types of Russian compounds reveals a quite clear pattern. According to the data, both X+V and X+N types emerge simultaneously and first in two corpora (Vanja's, Liza's), and subsequently in Philja's data within one month, but model X+V is more frequent. Initially this pattern is represented by four structures of compounds: N+V, PRON+V, ADJ+V, V+V. Conversely, X+N type is not so widely represented, and only two types were found: ADJ+N, N+N in the data of Philja and Liza. In Vanja's corpus NUM+N structure was added.

At the same time, the order of emergence of different types of compounds in Estonian shows a very clear pattern. The first pattern emerging in both CS was N+N type with modifier in genitive case form. This pattern was followed by ADJ+N, after which a pattern with a shortened noun (N+N) emerged. The first acquired pattern was also widely used at the end of the observation period.

6. The main types of compounds and their frequency differ in the three Russian corpora; therefore we cannot assume that the composition of compounds is related to the frequency of occurrence in CS. It seems that compounding serves as a means of enriching the children's lexicon. In the Estonian corpora the frequency has a much stronger effect on the choice of first patterns of compounding.

7. These results may supplement previous cross-linguistic findings about the acquisition of compounds. It can be supposed that Russian is not a compound-poor language, as has been suggested earlier, and it should operate in a similar way to the Hebrew, Baltic (namely Lithuanian) and Romance languages, because compounds in these languages emerge around 2;0. However, children do not use them productively until a later time, in contrast to compound-rich Estonian, where it does not take them very long to start to use them productively. It is also obvious that Russian (an east Slavic language) favours compounding more than west Slavic languages (e.g. Czech (Pačesová 1968)). Thus Russian, like Lithuanian, seems to be placed somewhere between compound-rich languages (Estonian, Germanic, etc.) and compound-poor ones.

In addition to this, a brief comparative analysis has shown that the Russian data is quite similar to that of Lithuanian (Dabašinskienė 2010). The number of Lithuanian child's types is slightly higher than Vanja's, but his tokens are much less than all the Russian children. The same main results were revealed in CDS. The number of types of compounds used by Lithuanian caregivers is higher, whereas they use fewer tokens than in all the Russian data. We can therefore assume that Russian CDS differs from Lithuanian CDS in respect of tokens.

6. Conclusion

The results of our comparative investigation led to a conclusion that the system of compounds reflects the process of their acquisition in the following way. The fact that compounding is just one of the productive word-building methods in the Russian derivation system sets a condition for their lower frequency in CDS and consequently in CS, as compared with compound-rich Estonian. Moreover, in Estonian compounds start to function productively earlier than in the speech of Russian children and Estonian children use more innovative compounds than Russian children.

At the same time, in both languages children start to acquire the compounds with those models which are transparent morphosemantically and are productive. However, the factor of frequency of the caregiver's compounds depends not only on the productivity of a definite pattern in language system and some adult preferences, but also to a large extent on the topic of conversation.

References

- Argus, Reili 2008. Eesti keele muutemorfoloogia omandamine. Tallinna Ülikooli humanitaarteaduste dissertatsioonid 19. Tallinn: TLÜ Kirjastus.
- Argus, Reili 2012. Kausatiivsuse omandamisest eesti keeles. – Eesti Rakenduslingvistika Ühingu aastaraamat, 8, 5–20. <http://dx.doi.org/10.5128/ERYa8.01>
- Berman, Ruth A. 1987. A developmental route: Learning about the form and use of complex nominals in Hebrew. – *Linguistics*, 25 (6), 1057–1086. <http://dx.doi.org/10.1515/ling.1987.25.6.1057>
- Berman, Ruth; Clark, Eve V. 1989. Learning to use compounds for contrast: Data from Hebrew. – *First Language*, 9, 247–270. <http://dx.doi.org/10.1177/014272378900902701>
- Berman, Ruth 2009. Children's acquisition of compounds. – Rochell Lieber, Pavol Štekauer (Eds.). *The Handbook of Compounds*. Oxford: Oxford University Press.
- Bittner, Dagmar; Dressler, Wolfgang U.; Kilani-Schoch, Marianne 2003. Introduction. – Dagmar Bittner, Wolfgang U. Dressler, Marianne Kilani-Schoch (Eds.). *Development of Verb Inflection in First Language Acquisition: A Cross-Linguistic Perspective*. *Studies on Language Acquisition* 21. Berlin/New York: Mouton de Gruyter, vii–xxxvii.
- Böer, Katja; Kotowski, Sven; Härtl, Holden 2012. Compounds vs. phrases: The cognitive status of morphological products. – Presentation on 15th International Morphology Meeting in Vienna. <http://www.wu.ac.at/inst/roman/imm15/abstract-boeer-kotowski-haertl.pdf> (13.02.2013).
- Clark, Eve V.; Gelman, Susan A.; Lane, Nancy M. 1985. Noun compounds and category structure in young children. – *Child Development*, 56 (1), 84–94.
- Clark, Eve V.; Hecht, Barbara F.; Mulford, Randa C. 1986. Coining complex compounds in English: Affixes and word order in acquisition. – *Linguistics*, 24 (1), 7–29. <http://dx.doi.org/10.1515/ling.1986.24.1.7>
- Clark, Eve V.; Berman, Ruth A. 1984. Structure and use in acquisition of word-formation. – *Language*, 60, 542–590.
- Dabašinskienė, Ineta 2010. Acquisition of compounds Lithuanian. – *Baltic Linguistics*, 1, 51–86.
- Dressler, Wolfgang U. 1999. What is natural in natural Morphology? – *Travaux du cercle Linguistique de Prague*, 3, 135–144.
- Dressler, Wolfgang U.; Ladányi, Mária 2000. Productivity in Word Formation: a morphological approach. – *Acta Linguistica Hungarica*, 47, 103–144.

- Dressler, Wolfgang U.; Lettner, Laura E.; Korecky-Kröll, Katharina 2010. First language acquisition of compounds – with special emphasis on early German child language. – S. Scalise, I. Vogel (Eds.). *Cross-Disciplinary Issues in Compounding*. Amsterdam: Benjamins, 323–344.
- Erelt, Mati; Kasik, Reet; Metslang, Helle; Rajandi, Henno; Ross, Kristiina; Saari, Henn; Tael, Kaja; Vare, Silvi 1995. *Eesti keele grammatika I*. Tallinn: Eesti Teaduste Akadeemia Eesti Keele Instituut.
- Gagarina, Natalia 2008. Становление грамматических категорий русского глагола в детской речи [Stanovlenie grammaticheskikh kategorij russkogo glagola v detskoj rechi / First language acquisition of verb categories in Russian]. St. Petersburg: Nauka.
- Gudilova, Svetlana 2005. Продуктивные типы образования сложных слов в современном русском языке: На материале неологизмов второй половины XX века [Produktivnye tipy obrazovanija slozhnykh slov v sovremennom russkom jazyke: Na materiale neologizmov vtoroj poloviny XX veka / Productive ways of compounding in modern Russian: On the material of neologisms (the second half of the XX century)]. Moskva: Moskovskij Gosudarstvennyj Universitet.
- Gvozdev, Alexandr N. 1995. Усвоение детьми звуковой стороны русского языка [Usvoenie det'mi zvukovoj storony russkogo jazyka / The Acquisition of Russian phonetics by a child]. St. Petersburg: AKCIDENT.
- Kerge, Krista 1990. Liitsõna. Mõisteid ja seoseid. ENSV Teaduste Akadeemia. Preprint KK1-51.
- Kharchenko, Vera K.; Ozerova, Elena 1999. Сложные слова в детской речи [Slozhnye slova v detskoj rechi / Compounds in child language]. Belgorod: Belgorodskij Universitetskij Press.
- Kilani-Schoch, Marianne; Wolfgang U. Dressler 2005. *Morphologie naturelle et flexion du verbe français*. Tübingen: Gunter Narr.
- Laaha, Sabine; Gillis, Steven 2007. Typological Perspectives on the Acquisition of Noun and Verb Morphology. *Antwerp Papers in Linguistics* 112. Antwerp: University of Antwerp.
- Laalo, Klaus 1994. Kaksitavuvaihe lapsen kielen kehityksessa. – *Virittäjä*, 3, 430–448.
- Libben, Gary 1998. Semantic transparency in the processing of compounds: Consequences for representation, processing, and impairment. – *Brain and Language*, 61, 30–44.
- Mellenius, Ingmarie 1996. Children's comprehension of Swedish nominal compounds. – Carolyn E. Johnson, John H.V. Gilbert (Eds.). *Children's Language*. Mahwah, N.J.: Lawrence Erlbaum. Vol. 9, 167–182.
- Mellenius, Ingmarie 1997. *The Acquisition of Nominal Compounding in Swedish*. Lund: Lund University Press.
- Nicoladis, Elena 2002. What's the difference between 'toilet paper' and 'paper toilet'? French-English bilingual children's crosslinguistic transfer in compound nouns. – *Journal of Child Language*, 29 (4), 843–863. <http://dx.doi.org/10.1017/S0305000902005366>
- Nicoladis, Elena 2006. Pre-school Children's Acquisition of Compounds. – Gary Libben, Gonia Jarema (Eds.). *The Representation and Processing of Compound Words*. Oxford: Oxford University Press.
- Nicoladis, Elena 2007. Acquisition of deverbal compounds by French-speaking preschoolers. – *The Mental Lexicon*, 2 (1), 79–102. <http://dx.doi.org/10.1075/ml.2.1.06nic>
- Pačesová, Jaroslava 1968. *The Development of Vocabulary in the Child*. Brno: Universita Purkyně.
- Padrik, Marika 2010. *Word-Formation Skills in Estonian Children with Specific Language Impairment*. *Disserationes pedagogicae Universitatis Tartuensis* 8. Tartu: Tartu Ülikooli Kirjastus.
- Rivista di Linguistica, 4 (1), 1992. Special Issue: The Morphology of Compounding. Sergio Scalise (Guest-ed.).
- Русская грамматика 1980 [Russkaja grammatika 1980 / Russian Grammar 1980]. Shvedova N. (Ed.). Vol. 1: Фонетика, фонология, ударение, интонация,

- словообразование, морфология [Phonetika, fonologija, udarenie, intonacija, slovoobrazovanie, morfologija / Phonetics, phonology, stress, intonation, derivation, morphology]. Moskva: 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.
- Scalise, Sergio; Bisetto, Antonietta 2005. The classification of compounds. – *Lingue e Linguaggio*, IV, 2, 319–332.
- Stephany, Ursula 2010. Preliminary classification of Greek compounds. – Paper, presented at the International Pre- and Protomorphology meeting. 12–14 February 2010.
- Stephany, Ursula; Voejkova, Maria D. (Eds.) 2009. Development of Nominal Inflection in First Language Acquisition: A Cross-Linguistic Perspective. *Studies on Language Acquisition* 30. Berlin: Mouton de Gruyter. <http://dx.doi.org/10.1515/9783110217117>
- Tikhonov, Alexandr N. 2003. Словообразовательный словарь русского языка [Slovoobrazovatel'nyj slovar' russkogo jazuka / Word-formation dictionary of the Russian language]. Vol. 1-2. Moskva: ACT.
- Vasilevskaja, Evgenia A. 1962. Словосложение в русском языке [Slovoslozhenie v russkom jazyke / Compounding in Russian]. Moskva: Vysshaja Shkola.
- Zemskaja, Elena A. 1992. Словообразование как деятельность [Slovoobrazovanie kak dejatel'nost' / Word-formation as an activity]. Moskva: Nauka.

Reili Argus (Tallinn University). Her academic research interests include language acquisition (acquisition of morphology and lexicogrammatical categories).
reili.argus@tlu.ee

Victoria Kazakovskaya (Institute for Linguistic Studies at Russian Academy of Sciences). Her academic research interests include psycholinguistics, functional grammar, conversation analysis, language acquisition (development of communicative competence, acquisition of semantic categories, child-directed speech, negative evidence, fine-tuning).
victory805@mail.ru

LIITSÕNADE OMANDAMINE EESTI JA VENE KEELES: SAGEDUSE, PRODUKTIIVSUSE, LIHTSUSE JA LÄBIPAISTVUSE MÕJU

Reili Argus, Victoria Kazakovskaya

Tallinna Ülikool, Vene Teaduste Akadeemia

Artikkel keskendub varasele liitsõnamoodustuse arengule kahes tüpoloogiliselt erinevas keeles – eesti ja vene keeles. Kui eesti keeles on liitsõnamoodustus produktiivne ja liitsõnu esineb keeles palju, siis vene keele puhul on liitsõnad enamasti kinnistunud leksikaalsed üksused ning neid leidub igapäevases keelekasutuses vähem kui eesti keeles.

Uurimuse aluseks on pikiuuringuga saadud keeleandmed kahelt eesti ja kolmelt vene keelt omandavalt lapselt. Analüüsitud on nii laste enda kui ka neile suunatud kõnet ning analüüsi keskmes on liitsõnade ilmumine laste keelekasutusse, eri liitmismallide areng ning sellised omandamist mõjutavad tegurid nagu liitmismalli produktiivsus, liitsõnade või mallide sagedus, moodustusmalli lihtsus ja liitsõna morfosemantiline läbipaistvus.

Olgugi et liitsõnade sagedus kahes vaadeldud keeles on erinev, ilmuvad esimesed liitsõnad laste keelde samas vanuses, teise eluaasta lõpus. Mõlemas keeles omandavad lapsed produktiivsed liitsõna moodustamise mallid varakult. Kui eesti lapsed hakkavad liitsõnu kasutama produktiivselt juba mõni kuu pärast esimeste liitsõnade ilmumist, ei saa sama väita vene keelt omandavate laste kohta. Samuti ei kasuta vene keelt omandavad lapsed erinevalt eesti lastest omaloomingulisi liitsõnu. Seega paistab kahe vaadeldud keele liitsõnasüsteemi eripära väga selgelt ka laste keelekasutusest.

Sellised tegurid nagu lihtsus ja sagedus on omandamisprotsessis põimunud: lihtsama ehitusega liitsõnad ilmuvad samal ajal kui sagedased liitsõnad. Sageduse mõju ei ole aga alati ühesuunaline: lapsele suunatud kõne puhul võis täheldada, et mõned vanemad kasutavad lapsega kõneldes enam liitsõnu kui teised ning mitte alati ei pruugi lapsele suunatud kõne liitsõnarikkus peegelduda lapse kõnes.

Võtmesõnad: keeleomandamine, liitsõnad, lapsele suunatud kõne, sagedus, produktiivsus, läbipaistvus, lihtsus, eesti keel, vene keel