THE SOUTH ESTONIAN LANGUAGE ISLANDS IN THE CONTEXT OF THE CENTRAL BALTIC AREA

**Miina Norvik[a,b], Uldis Balodis[c,d], Valts Ernštreits[c],
Gunta Kļava[c], Helle Metslang[a], Karl Pajusalu[a], Eva Saar[a]**
[a]*University of Tartu, EE*
[b]*Uppsala University, SE*
[c]*University of Latvia, LV*
[d]*Western Institute for Endangered Language Documentation, US*

miina.norvik@ut.ee, ubalodis@lu.lv, valts.ernstreits@lu.lv,
gunta.klava@lu.lv, helle.metslang@ut.ee, karl.pajusalu@ut.ee,
eva.saar@ut.ee

Abstract. This article offers a comparative analysis of several morphosyntactic and phonological features in the South Estonian language islands: Leivu, Lutsi, and Kraasna. The objective is to give an overview of the distribution of selected features, their (in)stability over time, and discuss their form and use in a broader areal context. To achieve this goal, comparative information was also included from the closest cognate varieties (Estonian and the South Estonian varieties, Courland Livonian and Salaca Livonian) and the main contact varieties (Latgalian, Latvian, and Russian). The data analysed in this study originated from various sources: text collections, dictionaries, and language corpora. The results reveal a multitude of linguistic patterns and distribution patterns, which means that the studied varieties are similar to / different from one another in various ways and points to multifaceted contact situations and outcomes in this area.

Keywords: morphosyntax, phonology, language contact, borrowing, language areas, Central Baltic area, Southern Finnic languages

1. Introduction

The Circum-Baltic area (CBA) is a meeting point for the languages of the Indo-European and Uralic language families. Based on a number of linguistic features, the CBA can be regarded as a buffer zone between the languages of the Standard Average European (SAE) area and Central Eurasia (Wälchli 2011). Whereas genetic diversity in the CBA is only moderate, continuity of contacts over a long period of time is seen as

the most significant factor characterising the area (Koptjevskaja-Tamm & Wälchli 2001, Wälchli 2011). According to the current view, for instance, Finnic speakers reached the Baltic Sea about 3200 to 2800 years ago but came into contact with Baltic tribes already on their way there (see Lang 2018, Grünthal et al., in press, Nichols 2021).

There have been attempts to establish the CBA as a linguistic area, but no isoglosses that would cover the entire area have been found (Koptjevskaja-Tamm & Wälchli 2001, Seržant (to appear)). Koptjevskaja-Tamm (2006) proposes two isoglosses that could almost unite the entire area: polytonicity and word order in possessive NPs. However, it tends to be more often the case that convergence works on a micro-level mainly involving two or three languages; if more languages are involved this is regarded as an instance of overlapping and superposition of different language contacts (Koptjevskaja-Tamm & Wälchli 2001: 728). Thomas Stolz (1991), one of the many researchers studying the common features of this area, suggested the existence of convergence zones, e.g., he proposed a Latvian, Livonian, and Estonian contact-intensive zone. He, however, looked at the standard varieties.

The present article takes a closer look at a number of morpho-syntactic and phonological features in the southernmost Finnic languages, with a focus on the South Estonian language island varieties: Lutsi, Leivu, and Kraasna. The other varieties included in the study are Võro, Seto, Mulgi, Standard Estonian (represented by North Estonian in Figure 1), two Livonian varieties – Courland Livonian and Salaca Livonian, and the main non-cognate contact varieties – Latgalian, Latvian, and Russian. Here, the respective area where the studied varieties are spoken is called the Central Baltic (see also Figure 1). It should be noted that Võro and Seto are important in terms of tracing the origins of Lutsi, Leivu, and Kraasna. Namely, Kraasna speakers are thought to have migrated to Krasnogorodsk from the Seto areas in the 16th century; the initial migration to the Ludza region (> Lutsi) is thought to have taken place in the 17th and 18th centuries from eastern Võromaa but followed by later waves from different parts of Setomaa and Võromaa; Leivu speakers, in turn, are thought to originate from western Võromaa (see, e.g., Kallas 1903: 46–56, Vaba 1977: 22). By now, the language island varieties have all gone extinct – Lutsi and Leivu became extinct as conversational languages in the 1970s to 1980s, Kraasna already in the first half of the 20th century.
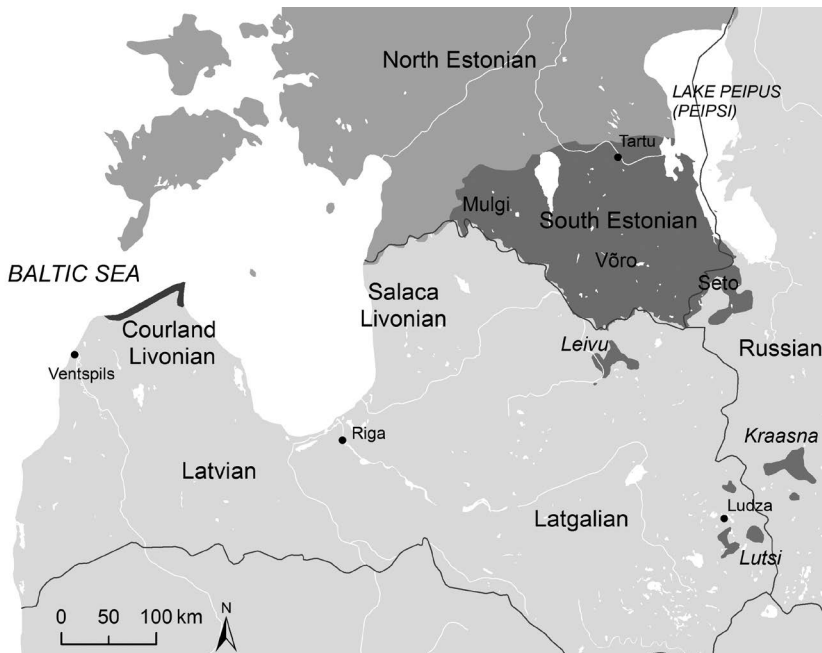
Figure 1. Location of the studied varieties in the Central Baltic area (map by Timo Rantanen, BEDLAN).

This article studies a selection of features, which include (i) four morphosyntactic topics: case-marking and agreement in noun phrases, comparative constructions, person-indexing, and negation; and (ii) ten phonological features: stød, glottal stop, *h*, voiced plosives, short vs. long consonant geminates, short vs. long vowels, central vowels (õ [ɤ]; y [ɯ]), front rounded vowels (ü [y]; ö [ø]), vowel harmony, extensive palatalisation. The selected features are in one way or another characteristic of the South Estonian language islands or form a central part of their language system. Being characteristic of the language islands, however, does not exclude developments that are shared more broadly in the area. For instance, comparative adjective marking with -*b* can be regarded as a joint development in the southernmost South Estonian area (see Pajusalu 2008: 164). Whereas some of the selected features are discussed also in earlier studies (e.g., comparative constructions), there are also topics that previously have found only little attention (e.g., person-indexing). The selection of features described here is connected with the aims of the paper. First, to elaborate on the results of

previous studies that contain information on the respective varieties but to different extents. As far as we know, there is no study that would provide a systematic comparison of all twelve varieties without any gaps. Second, to shed more light on the topics that earlier have found little attention.

Our research questions are as follows:

1. Which of the selected features, if any, show convergence in the southernmost Finnic area?
2. Are there any features in the studied Finnic varieties that have been relatively stable over time?
3. If changes have taken place, then what have they brought along and what might have caused them?

We also have three main hypotheses. First, we hypothesise that there are instances where similarities involve two or three languages spoken in close proximity, rather than the languages of the entire southern-most Finnic area (as also proposed by Koptjevskaja-Tamm & Wälchli 2001; see above). Second, we assume that there are features which have been relatively stable over time regardless of the multifaceted contact situations found in the area. Third, we hypothesise that we can detect two kinds of changes: (i) changes caused by the neighbouring Indo-European languages, (ii) changes that cannot be attributed (at least not directly) to the non-cognate contact varieties. As regards the former, we primarily expect 'PAT(tern)' transfers (for the term and explanation, see Matras & Sakel 2007). To exemplify the latter, the unexpectedly broad use of external local cases in Lutsi and Vana-Laitsna (Latvian: Veclaicene) is regarded as an instance of a development that cannot be explained by the direct influence of Latvian (Pajusalu 2008: 164).

Outcomes of language contacts depend on several factors. In addition to purely linguistic factors, such as genetic similarity, and the role and inherentness of the feature to the language structure, language external factors also play a role, e.g., the sociolinguistic situation and type of language contact (Aikhenvald 2006, Sakel 2007, Seržant (to appear)). The South Estonian language islands also existed in multi-faceted contact situations. For instance, Lutsi speakers lived side by side with Latgalian, Russian, Belarusian, Polish, and Yiddish speakers

(see Ariste 1981: 36); the Central dialect of Standard Latvian may also be considered a contact language to some extent. This complex contact situation may have contributed to Lutsi remaining "pure" longer (see Vaba 2011: 208). Latgalian has been in close contact with other languages for centuries, which is evident in its historical development and due to the multilingualism of the Latgale region. Traces of older and newer contacts can be found in Latgalian, reflecting Finnic, Lithuanian, Polish, Belarusian, Russian, and others (Breidaks 2007). It was only at the end of the 19th century when Russian became dominant in the region (Ariste 1962). Due to its location in western Russia, Kraasna speakers ultimately switched to Russian, but in the 1850s there were still people in the Kraasna villages who did not speak Russian (Kallas 1903: 33). The Leivus, in turn, assimilated into the Latvians. The first documentations of Leivu from the middle of the 19th century already show a strong Latvian superstrate in its phonetics, morphology, syntax, and lexicon (Vaba 1997: 39). The phonetic, grammatical, and lexical innovations of Leivu that are shared with Livonian are due to Latvian influence on both Leivu and the Livonian varieties (Vaba 1997: 39ff.).

This article has the following structure. First, we introduce the materials and methods used in our study. Second, we introduce the main results of our study divided into two sections: morphosyntax and phonology. Finally, we present a discussion of the results and our conclusions.

2. Materials and methods

This article is a comparative study of several morphosyntactic and phonological features in twelve language varieties with a focus on the South Estonian language islands: Leivu (Lei), Lutsi (Lut), and Kraasna (Kra). The other varieties in the study are (i) the South Estonian varieties historically most closely related to them – Võro (Vro), Seto (Set), and Mulgi (Mul); (ii) Latgalian (Ltg) as the main non-cognate local variety; (iii) two Livonian varieties historically spoken in Latvia – Courland Livonian (CLiv) and Salaca Livonian (SLiv); and (iv) Estonian (Est), Latvian (Lav), and Russian (Rus) as the main standard varieties in the area. We also draw some parallels with local Russian varieties if relevant.

The selection of linguistic features depended on several considerations. As the focus is on the South Estonian language islands, the feature had to be in some respect characteristic/relevant to at least one of the language islands or, in general, central to language structure (e.g., noun and verb morphology). For instance, vowel harmony can be found in all three varieties (as well as in other dialects of South Estonian), whereas back negation is common not only to Lutsi and Kraasna (e.g., Mets et al. 2014: 14–20), but also to Seto and Võro. Both domains (morphosyntax and phonology) had to be represented by features showing at least some variation in that area. As one of the goals was to elaborate on the previous studies, the selected features could already be listed as special features of South Estonian included in earlier areal typological studies (e.g., concerning CBA, SAE). The features chosen for the analysis represent only one possible selection of various features that could be given a closer look.

The linguistic examples used in the study come from a multitude of sources that represent somewhat different periods. The following paragraphs explain the considerations used to select the data to compare.

Linguistic examples from **Leivu, Lutsi, and Kraasna** were mainly found in the text collection *Eesti murded IX* ('Estonian dialects', Mets et al. 2014). This collection contains transcriptions of recordings from the 1910s to 1970s and is approx. 270 pages long (a Standard Estonian translation takes up half of each page). For Lutsi, we also used the grammatical overview by Balodis (2020). As there are only 16 pages of Kraasna texts in Mets et al. (2014) and all are recorded from a single speaker in 1911 or 1912, we also gathered data from two additional sources, one compiled by Kallas (1903) and the other by Ojansuu (1938; henceforth AES 202). However, regardless of this, the Kraasna data are the scarcest.

Unlike the language varieties of the South Estonian language islands, **Mulgi, Seto, and Võro** are local varieties that still remain in active use. According to the 2011 Estonian census, there were 74,512 Võro speakers, 12,532 Seto speakers, and 9,682 Mulgi speakers in Estonia. Still, among the youngest group (0 to 14 years) knowledge of local varieties is scarce, e.g., only 247 children of that age were reported to know Seto (ESA 2011). For the purposes of the present study, we used language examples included in the dictionaries, e.g., the Mulgi dictionary (Laande & Todesk 2013), the Estonian-Võro Dictionary

(Faster et al. 2014), the Seto-Estonian Dictionary (Saar et al. (in preparation)), the Seto Dictionary of Unique Words (Saar et al. 2020). These dictionaries represent the language use of the 20th century, mainly of the first half of the century, which makes it more comparable to the data of the language islands. South Estonian dictionaries were compiled using the word slips from the Institute of the Estonian Language Wiedemann card file (EKI WK). We also included data from various studies (e.g., Iva 2007, Pajusalu 1996, Tanning 1961, 2004).

Latgalian is currently spoken by 164,510 individuals, or approximately 8% of Latvia's population (see Lauze 2017: 50). The Latgalian data are taken from *A short grammar of Latgalian* by Nau (2011); we also used the Corpus of Modern Latgalian (MuLa) to check for the presence or absence of features. Nau's grammar is partly based on the language used in blogs, short stories, short journalistic prose, literary self-portraits of contemporary writers (referred to as "modern texts" in the grammar). Additionally, she used a corpus of traditional narratives, fairy-tales collected from two villages of Central Latgale in the 1890s and 1920. The Eastern-Central variety is also the language that forms the basis of written Latgalian (i.e., the partly standardised written form of Latgalian).

Similarly to the South Estonian language islands, **Salaca Livonian** is also extinct now. As it faded out of use already in the second half of the 19th century, the Salaca Livonian examples represent the oldest data included in this study – they come from the mid-19th century. The examples were collected from the grammar and dictionary compiled by Winkler and Pajusalu (2016, 2018) retaining the orthography used there[1]. **Courland Livonian** was in everyday use in the 20th century (Blumberga 2013: 182), presently, it is actively being developed, has a standardised form, and a handful of L2 speakers (Ernštreits 2012: 159). The linguistic examples of Courland Livonian included in this study date to the 20th century. They are mainly taken from the Livonian-Estonian-Latvian dictionary (Viitso & Ernštreits 2012), which reflects the Livonian language of the second half of the 20th century. Regardless of

---

1   Originally, the Salaca Livonian data were collected by A. J. Sjögren in 1846 and included in the Livonian grammar published in 1861 (see Sjögren & Wiedemann 1861). As Sjögren was working with a translator, and about 60% of example sentences are translations of Bible sources, Hesselberg's grammar, and riddles (Winkler & Pajusalu 2018: 155), Salaca Livonian examples are to be treated with some caution.

the source, the Courland Livonian examples are presented in the modern Livonian orthography.

Latvian and Estonian are national languages that were subject to language standardisation in the 19th–20th centuries. The Estonian examples used in this study follow the present-day standard. Only where the language of the 19th–20th centuries differs has a separate comment been provided. With regard to **Russian**, we take into account, first of all, Standard Russian, and within the limits of available information, also Central Russian dialects spoken in the areas of Russia neighbouring Estonia and Latvia.

Our approach to data collection was to collect as comprehensive of a data set as possible from Lutsi, Leivu, and Kraasna. This mainly meant reading the text collections mentioned above and using the ctrl + F function for PDFs. The other varieties were included for comparative purposes, thus we did not attempt to obtain maximally complete data. We provide a qualitative analysis regarding the uneven amount of source material.

Several sources on South Estonian varieties contain transcribed text that varies somewhat from source to source. As the main part of the paper is concerned with morphosyntax for which phonetic details are not essential, we simplified the Kraasna, Seto, Võro, Mulgi, and Leivu transcriptions following the principles of the South Estonian literary standard (see, e.g., Faster et al. 2014), some Lutsi language examples are also presented in this literary standard and others in the new Lutsi orthography (see Balodis 2015).

3. Morphosyntactic features

In the following, sections 3.1 and 3.2 present a comparative analysis of nominal features, which fall under the topics of case-marking and noun phrases, and also comparative constructions. Sections 3.3 and 3.4, in turn, concentrate on verbal features relating to person-indexing and prodrop as well as negation.

3.1. Case-marking and agreement in noun phrases

This section takes a closer look at the cases that in South Estonian are of more recent origin, or that have become unproductive over time.

Additionally, we comment on general changes in the case system and study agreement between the adjectives and the head noun in the NP. In Latvian, Latgalian, and Russian, an adpositional construction is used in several instances; or the respective meaning represents one of several uses of a case and, therefore, would require more specific treatment. Thus, this section sets the main focus on the Finnic varieties.

The case paradigm of South Estonian varieties and language island varieties differs from the Standard Estonian by one case – the essive that is not listed in the case paradigm (e.g., see Balodis 2020: 76–77 for Lutsi) or among productive cases (Iva 2007: 41, 56–57 for Võro). The remaining shared cases are the nominative, genitive, partitive, illative, inessive, elative, allative, adessive, ablative, translative, terminative, abessive, and comitative. The presence of the essive in Standard Estonian is actually the result of language planning that started at the end of the 19th century. At that time, the essive had fallen out of use in most Estonian dialects and was productive only in the Northwestern Coastal dialect of Estonian, which was used as the source for its reintroduction. To compare, the literary standard for Võro was developed only at the end of the 20th century. Iva (2007: 56–57) sees one of the reasons for its unproductivity as the formal similarity between the essive and inessive cases, but he also mentions that the nominative and translative cases can appear in the function of the essive. Thus, although, on the one hand, we might be dealing with case syncretism, on the other hand, there are also indications of unproductivity, i.e., other cases taking over. In the entire South Estonian area (including in the language islands), the endings are similar to those of the inessive with the two cases being distinguished primarily by their functions (Metslang & Lindström 2017, Prillop et al. 2020: 307–309): Kra, Set *-h*, *-hnA* (1–2); Lut, Vro *-hn(A)*, *-n* (3); Lei, Mul *-n* (cf. Standard Estonian *-na*, e.g., *paadi**na*** 'as a boat' vs. *paadi**s*** 'in a boat'). By contrast in Courland Livonian, the essive is preserved only in some lexicalised forms, e.g., *pivāpāva**n*** 'on Sunday' (Viitso 2008: 328); in Salaca Livonian, there do not seem to be any remaining traces of the essive.

(1)   Kra: *mine terve-**hnä**!* (Kallas 1903: 29)
            go.IMP.2SG healthy-ESS
            'go in health (lit. healthily)!'

(2)   Set: *üts´ tütär´ oll´ koto-h tütriku-**h** viil* (Saar et al. (in preparation))
one daughter be.PST.3SG home-INE girl-ESS still
'one daughter was still unmarried (lit. as a girl) at home'

(3)   Lut: *naka=s inäp rūvli-**n** olyma* (Mets et al. 2014: 152)
begin=NEG.PST more thief-ESS be:SUP
'[s/he] was no longer a thief'

Whereas the status of the essive might be considered somewhat debatable, the following cases are unproductive in all of the Southern Finnic languages: the excessive (e.g., Lut, Mul, Set, Vro *ma**nt*** 'from (near)', Est *kodu**nt*** 'from home', CLiv *tagā**nd*** 'from behind'), prolative (Vro *vesi**lde*** 'by water', *mai**ldõ*** 'by land'), and instructive. Whereas the instructive case has lost its inflectional ending in Estonian and the South Estonian varieties, the historical ending *-*n* has been preserved in Northern Finnic; Livonian -*ņ* also probably has the same origin (see Ross 1988: 21–23). Typically, the instructive forms appear in the plural and are used with numeral phrases (4–6a) but there are also other uses (e.g., 6b).

(4)   Kra: ***kuuzi*** *koi* (Kallas 1903: 113)
six:PL.INS knit:PST.1SG
'[I] knitted six at a time'

(5)   Set: ***katsi pätsi*** *anda-s poodi-h leibä* (Saar et al. (in preparation))
two:PL.INS loaf.of.bread:PL.INS give-IPS.PRS shop-INE bread.PRT
'they give two loaves of bread at a time at the grocery store'

(6)   CLiv: a. *kakš-**ī-ņ*** (Viitso 2008: 329)
two-PL-INS
'two at a time'

b. *īe-**i-ņ*** (Viitso & Ernštreits 2012)
night-PL-INS
'for nights'

In general, shifts in the inflectional system of the South Estonian language island varieties correspond to the general developments of the inflectional systems of Estonian and South Estonian. The Livonian varieties also show further developments. One of the notable differences

is the marginalisation/conflation of the external local cases (for Courland Livonian, see, e.g., Viitso 2008: 328, Blokland & Inaba 2018). An interesting parallel can be observed between the Livonian varieties and Mulgi: they reveal conflation between the adessive and allative cases, e.g., Mul *seinäl* 'onto the wall, on the wall', SLiv *ybil* 'to the horse, on the horse' (Winkler & Pajusalu 2018: 77–79), and CLiv *põrandõl* 'onto the floor, on the floor'. The Salaca Livonian example dates to the mid-19th century and illustrates the situation at that time, the Courland Livonian example represents one of the few instances where the external cases occur in the 20th century (see more in Viitso 2008, Blokland & Inaba 2018). In all other respects, Mulgi follows the system of the South Estonian varieties described above (see also Tanning 2004: 85–97).

The South Estonian varieties, including the language islands, show differences from Standard Estonian and the Livonian varieties with regard to the marking of agreeing adjectives in case and number with the head noun in the NP. In Standard Estonian, agreement is a general rule, but there are four case suffixes – the terminative, essive, abessive, and comitative – that are only attached to the final word of the phrase (7a–d) and can thus be regarded as phrase markers (Metslang & Lindström 2017: 60; for phrase markers, see also Hansen 2000; this phenomenon can also be considered to be suspended affixation, see, e.g., Despić 2017). In the South Estonian varieties, in turn, also essive and terminative agreement turned out to be possible (e.g., 8–9), and a few examples of comitative agreement could be found in Leivu (e.g., 10). According to Iva (2007: 54), using the terminative suffix only on the final word in Võro is an influence from Standard Estonian.

(7) Est: a. *väikse puu-**ni*** (small tree-TERM) 'up to a small tree'

b. *väikse puu-**na*** (small tree-ESS) 'as a small tree'

c. *väikse puu-**ta*** (small tree-ABE) 'without a small tree'

d. *väikse puu-**ga*** (small tree-COM) 'with a small tree'

(8) Lut: a. *sūre-**ni** vanhuze-**ni*** (Mets et al. 2014: 126)
big-TERM age-TERM
'until an old age'

Kra: b. *suure-**ni** lehmä nüssängu-**ni*** (Kallas 1903: 108)
big-TERM cow.GEN milking-TERM
'until the big cow-milking'

(9)   Lei: *vanu-**n** aju-**n** olle varbun aid* (EMS)
            old.PL-ESS time.PL-ESS be.PST.3SG rod fence
            'in the old days there was a fence of rods'

(10) Lei: *madali-de-**ge** lak´u osse-**ge*** (Mets et al. 2014: 24)
            low-PL-COM wide.PL.GEN branch.PL-COM
            'with low wide branches'

Regular agreement of the adjective and the noun in Standard Estonian is the result of language planning, or the so-called congruence reform that started at the end of the 19th century. In the 19th to 20th century, case endings in the North Estonian dialects (including the illative and allative endings) as well as plural suffixes that formed a separate syllable were generally not used on adjective attributes. This was also common in the literary language, which is largely based on the North Estonian Central dialect, and in the common spoken language (11) (see Nurkse 1937, Saari 2004[1995]). Relying on the South Estonian dialects where congruence was more common, Karl August Hermann included in his 1884 grammar book a requirement for agreement in all cases except the aforementioned four cases (see Hermann 1884).

(11) Est: *targa professori-te-**le*** (Nurkse 1937: 51)
            smart professor-PL-ALL
            'to smart professors'

In Courland Livonian, agreement is a characteristic of grammatical cases and internal local cases, whereas the dative (12a) and instrumental (12b) cases show non-agreement. In Salaca Livonian, the extent of non-agreement is even greater, as it also involves external local cases (13). Here again, Mulgi shows interesting parallels with Salaca Livonian as the modifier may lack case marking in external as well as internal local cases (14) (see also the comment about North Estonian dialects above; for further examples from the Karksi subdialect, see Tanning 2004: 85–87).

(12) CLiv: a. *jõvā sõbrā-**n*** (good friend-DAT) 'to/for a good friend',
              b. *ūd veisõ-ks* (new knife-INS) 'with a new knife'

(13) SLiv: *vana vallisnika-l* (Winkler & Pajusalu 2018: 180)
             old cottager-ADE;ALL
             'to/for an old cottager'

(14) Mul: *ma kinksi ubine väikse latse-l* (Laande & Todesk 2013)
             1SG give:PST.1SG apple.GEN small child-ADE
             'I gave a small child an apple'

Considering that Latvian, Latgalian, and Russian nominal dependents show agreement in all cases present in each language, it could be suggested that this is also responsible for a more elaborate agreement system in Võro, Seto, and the language islands. Still, Salaca Livonian, which was under strong Latvian influence, does not seem to offer support for this: non-agreement turned out to be typical even in local cases that show agreement in Latvian.

3.2. Comparative constructions

Typically, comparative constructions consist of two noun phrases, the object of comparison (*the comparee NP*, see *ta* in (15)) and the object to which it is compared (*the standard NP*, see *minust/mina* in (15)). The main differences among languages in forming comparative constructions are shown by the marking of the standard NP. Relying on (Stassen 2013), a distinction can be made between locational comparatives (15a), particle comparatives (15b), conjoined comparatives, and exceed comparatives. Under locational comparatives, he further lists: (i) *from*-comparatives (the standard NP marks the source of a movement associated with the meanings 'from', 'out of'), (ii) *to*-comparatives (the standard NP marks the goal of a movement associated with the meanings 'to, towards', 'over, beyond', 'for'), and (iii) *at*-comparatives (the standard NP marks a location associated with the meanings 'in', 'on', 'at').

(15) Est: a. *Ta on **minu-st** ilusa-m*
             3SG be.3SG 1SG-ELA beautiful-COMP

             b. *Ta on ilusa-**m** **kui mina**.*
             3SG be.3SG beautiful-COMP than 1SG
             'S/he is more beautiful than I'

The presence of the particle comparative type is usually listed as one of the traits of SAE (e.g., Haspelmath 2001, Heine and Kuteva 2006). Moreover, it can be regarded as a joint innovation in these languages as the older type in the Indo-European languages was the locational type (see Haspelmath 1998). A characteristic of European languages is also the occurrence of a comparative suffix (Stassen 2013). Table 1 shows that a comparative suffix can be found in all the studied varieties.

Table 1. Comparative suffixes in the analysed varieties.

Est	Vro	Set	Kra	Lut	Lei	Mul	SLiv	CLiv	Lav	Ltg	Rus
-m	-mb, -mp	-mb	-mb	-mb(i), -mp	-mb	-mb, -mp	-m	-m(i)	-āk-	-uok-	-ee
	-p	-b, -p	-b	-p	-b	-p					

As (15) shows, a single language can include a locational as well as particle comparative (see also Metslang 2009 for Estonian and Finnish). Table 2 reveals that this is true for most of the studied varieties. The particle type (referred to as Ptcl in Table 2) is the only option in Latvian and Latgalian (see also Endzelīns 1951: 478). It is important to note that the source marking type of locational comparative (referred to as Loc in Table 2) involves different kinds of marking for the standard – elative, partitive, genitive – depending on the language. Historically these are all related to source marking and thus are subsumed here under the same type (e.g., see Bernštein 2005: 28 for the Indo-European languages and Prillop et al. 2020 for the Finnic languages; for a comment on the Russian genitive, see Koptjevskaja-Tamm & Wälchli 2001: 683–685).

Table 2. Types of comparative constructions.

Type		Est	Vro	Set	Kra	Lut	Lei	Mul	SLiv	CLiv	Lav	Ltg	Rus
Loc	SOURCE	+	+	+	+	+	+	+	+	+			+
	GOAL			+		+	+		+				
Ptcl	'than'	+	+	+		+	+	+	+	+	+	+	+
	NEG					+	+					+	
	NEG + 'than'					+	+		+	+	+	+	+

In the **locational type**, the standard NP in the Finnic languages is typically marked either with the elative (in Kra, Lei, Est[2] (15a), Mul, Vro, Set, and CLiv) or the partitive (Lut (16), Kra, Vro, Set, SLiv). Thus, Kraasna, Võro, and Seto show examples of two locational types. (17) shows variation even within the same sentence. In our survey of Lutsi data, only examples of the partitive comparative were found.

(16) Lut: *vane-mb **minno*** (Mets et al. 2014: 169)
           old-COMP 1SG.PRT
           'older than me'

(17) Kra: *mis on madala-mb **maa haina** ja pike-mb **mu-i-st pu-i-st*** (Mets
           et al. 2014: 290)
           what be.3SG low-COMP meadow grass.PRT and tall-COMP other-
           PL-ELA tree-PL-ELA
           'what is lower than the meadow grass and taller than other trees'

The Leivu, Lutsi, Seto, and Salaca Livonian data also revealed instances that can be subsumed under goal marking (see also above). Namely, Seto and Leivu texts contained examples of *vasta* 'to, towards' (18a–b), and Salaca Livonian data contained examples of 'over' (19). The Salaca Livonian *yl* 'over' seems to be a PAT-borrowing from Latvian *par* (or *pār*) 'over'. According to Endzelīns (1951: 672), historically the older form is *pār*, which in modern Latvian retains its location meaning and other smaller specifics, but *par* has a broader range of meanings (e.g., of, about, than, for, as, too, etc.). The particle *par*, which originates in Latvian, is also found in Latgalian, where it can be regarded as a late influence from Standard Latvian (Nau 2011: 72).

(18) a. Lei: *tuu om kümme aastagu-t nuorõ-b **vasta** minnu* (EMS)
             that be.3SG ten year-PRT young-COMP towards 1SG.PRT
             '[s]he is ten years younger than me'

---

2 Whereas in present-day Estonian, the partitive is only attested in the comparative correlative construction, e.g., *mida varem, seda parem* (what.PRT soon-COMP that.PRT better.COMP) 'the sooner, the better', the Corpus of Old Written Estonian (VAKK) and the standard language from the beginning of 20th century show wider use, e.g., *wannem mind* 'older than me' (VAKK [1739]), *ausam sind* (VAKK [1766]) 'more honest than you', *selgem vett* (clear.COMP water.PRT) 'clearer than water' (Kallas 1903: 61).

  b. Set: *kuld om **vasta** hõpõ-t õks viil pallo kalli-p* (Saar et al. (in prepa-
             ration))
             gold be.3SG towards silver-PRT still more much expensive-COMP
             'gold is still much more expensive than silver'

(19) SLiv: *Läeli-m **yl** kaks birkau* (Winkler & Pajusalu 2018: 88–89)
              heavy-COMP over two ship_pound.GEN
              'heavier than two ship-pounds'

The **particle type** is present in all of the analysed varieties, although to different extents. In Kraasna, however, it was not attested at all. Although this might be due to the limited amount of data, one should not forget that Kraasna data revealed two kinds of locational comparatives (17). Probably, Russian has also enabled the Kraasna variety to preserve the locational type (cf. Latvian and Latgalian that only contain the particle type).

According to the particle used in the construction, the following types can be distinguished:

(i) 'than'[3] – Est, Set *kui*, Lei *ku(i)*, Kra *ku*, *ko*, Lut, Set *ku*, Mul *ku*, *nagu*, Vro *ku(q[4])*, CLiv *ku*, SLiv *kuj*, Ltg *kai*, Lav *kā*, Ltg, Lav *par*, Rus *čem*

(20) Lut: *pihlappuu-st um parõ-mp **kuu** tammõ-st vil* (Mets et al. 2014: 247)
             rowan-ELA be.3SG good-COMP than oak-ELA moreover
             'rowan [cart] is even better than oak [cart]'

(ii) negative marker – Ltg *na* (only in 19th century texts (see the description of sources used for Nau 2011), in Modern Latgalian this type of use is very rare), Lut *ei*, Lei *ei*, *is* (= negative marker in the past tense, see also section 3.4)

(21) Lut: *to oļļ sūre-mb **ei** şõ* (Vaba 1977: 20)
            that be.PST.3SG big-COMP NEG this
            'that one was bigger than this one'

---

3   Here 'than' stands for what can be regarded as a neutral particle that does not carry any additional meaning.

4   Following the Võro Standard language, *q* is hereinafter used for the laryngeal stop.

(22) Ltg: *sieniok bieja ciszi daudź kieniniu, wajrok **na** tagad* (Nau 2011: 72)
            earlier be.PST.3SG very many king.PL.ACC more NEG now
            'in earlier times there were many kings, more than today'

(iii) negative marker + 'than' – CLiv *äb ku ~ äbku*, SLiv *ap ka*, Ltg *nakai ~ nikai*, Lav *nekā*, Lut *eigu*, *üskui* (= negative marker in the past tense + 'than'), Lei *ei ku*, Rus *neželi*

(23) Lei: *küla laib om magusa-mb **eiq ku** uma laib* (EMS)
            guest.GEN bread be.3SG sweet-COMP NEG than own bread
            'bread (eaten) as a guest is sweeter than one's own'

(24) CLiv: *Jo kovāl⁵ **äb ku** tämā* (EMK [Setälä 1953])
            PTCL smart NEG than 3SG
            'smarter than him/her'

Unlike the constructions based on 'than', other types of particle comparatives are restricted to particular groups of varieties. The negative marker, either with or without 'than', is found in Lutsi and Leivu as well as in the Livonian varieties where it can be regarded as a PAT-borrowing from Latvian/Latgalian. Already Vaba (1977: 20, 24) described Lutsi *ei*, *eigu* (< *ei* 'not' + *kui* 'than') and *üskui* (< past negation marker *is* + *kui*) as translations of Latvian *ne* 'not' and *nekā* ('not' + 'than') and argued that this also applies to their counterparts in Leivu: *eiq ku*, *eikku*, *e ku*, and *isku.* It is noteworthy that both varieties also make use of the past tense marker even though Latvian and Latgalian negative markers are not inflected for tense (for more see Section 3.4). A closer look at examples in Leivu and Lutsi does not, however, enable one to conclude that their usage would be determined by temporal reference.

Although the negative marker can be found in Russian – *neželi* (*ne* 'not'), at least in the standard language it is regarded as old-fashioned (Timberlake 2004: 215). The Kraasna data, which contained no examples of the negative marker included in the particle type (nor the particle type in general), also seem to suggest that there was no (strong) model in the neighbouring Russian varieties. Still, it should be kept in mind that the Kraasna data are scarce. In any case, it can be concluded that the usage of the negative marker unites the varieties (once) spoken in the territory of present-day Latvia.

---

5   For the usage of the comparative degree marker *jo* in Livonian, see Stolz (2013: 107).

3.3. Person-indexing in the indicative mood and pro-drop

This section takes a closer look at person-indexing in the indicative mood (see Table 3), which in previous studies has found less attention than, for instance, the presence/absence of personal endings in the conditional and quotative moods.

Table 3. Person-indexing in the studied Finnic varieties.

		Est	Vro, Set		Kra	Lut	Lei		Mul	SLiv		CLiv
PRS	**1Sg**	-n	Ø		Ø	Ø	Ø		Ø	-b		-b
	2Sg	-d	-t, -dE		-t, -dE	-t	-t, -dE		-d, -t	-d		-d
	3Sg	-b	Ø		Ø	Ø	Ø		-b, -p | Ø	-b		-b
	1Pl	-me	-miq	Ø	-q	-m	-mE		-m, -me	-mi, -m	-b	-mõ, -m
	2Pl	-te	-tiq, -t, -dE	Ø	-t, -dE	-t	-t, -dE		-t, -de	-ti, -t	-b	-tõ, -t
	3Pl	-vad	-vAq		-vAq	-vAq	-vA, -vAq	Ø	-ve, -va, -v	-bVd	-b	-bõd
PST	**1Sg**	-n	Ø		Ø	Ø	Ø		Ø	Ø		Ø
	2Sg	-d	-t		-q	-q	-t	Ø	-d, -t	-d	Ø	-d, -t
	3Sg	Ø	Ø		Ø	Ø	Ø		Ø	Ø		Ø
	1Pl	-me	-miq | -q	Ø	-q	-m, -mi	-mi	Ø	-m, -me	-mi	Ø	-mõ, -m
	2Pl	-te	-tiq	Ø	-q	-q	-q		-t, -de	-ti	Ø	-tõ, -t
	3Pl	-d	-q		-q	-q	-q	Ø	-ve, -v	-t	Ø	-tõ, -t

As Table 3 illustrates, except for Estonian, the examined Finnic varieties tend to show syncretism between 1Sg and 3Sg. Syncretism depends on the word type, e.g., whereas Lutsi *tulema* 'to come' shows syncretism in the present tense, *tīdmä* 'to know' shows syncretism in the past tense (cf. 25a and 25b, but see the comment below about an additional conjugation type). While in the case of the South Estonian varieties examined here, syncretism between 1Sg and 3Sg forms usually means that there is no personal ending (marked with Ø in Table 3),

in Livonian, the respective forms contain the personal ending *-b* (26). Regarding this feature, Mulgi shares similarities with both types, as in the present tense both Ø and *-b ~ -p* are possible, see (27a–b). Historically, the Proto-Uralic 3rd person forms are thought to have lacked a personal ending (see Janhunen 1982: 35), thus 3rd person with no ending can be argued to represent the earlier stage (Kallio 2014: 156).

(25) Lut: a. *tulema* 'to come': Prs1Sg/3Sg **tule** vs. Pst1Sg **tulli**, Pst3Sg **tuḷḷ** (Balodis 2020: 85–86)

   b. *tīdmä* 'to know': Prs1Sg **tīä**, Prs3Sg **tīd** vs. Pst1Sg/3Sg **tīdze** (Balodis 2020: 88)

(26) CLiv: *tūlda* 'to come': Prs1Sg/3Sg **tulā-b** vs. Pst1Sg/3Sg **tuḷ** (Viitso & Ernštreits 2012)

(27) Mul: a. *laits **aa** tähti müüdä sõnu kokku* (Laande & Todesk 2013)
   child drive.3SG letter.PL.PRT along word.PL.PRT together
   'the child puts the words together based on the letters'

   b. *tule säde **aa-p** maja palame* (Laande & Todesk 2013)
   fire.GEN spark drive-3SG house.GEN burn:SUP
   'a spark of fire sets the house on fire'

Lutsi, Leivu, Kraasna, Võro, and Seto additionally include a conjugation type only available in the present tense that in 3Sg is marked with *-s* and in 3Pl with *-sE(q)*, e.g., Kra *ists* 's/he sits down' vs. *istusõq* 'they sit down' (AES 202: 26). It is important to note that in this conjugation type, there is no syncretism between 1Sg and 3Sg, e.g., Kra *istu* 'I sit down' vs. *ists* 's/he sits down' (*ibid.*), Vro *elä* 'I live', *eläs* 's/he lives'. The forms in *-s* go back to the suffix **-sEn*, which may have expressed a medial or reflexive meaning (see more in Posti 1961). Currently, there are only some pairs of verbs that appear in both conjugation types and could point to such a distinction, e.g., Vro *küdsä* '(someone) bakes sth', *küdsäs* '(something) is baking', cf. words such as Vro *eläs* 's/he lives' (see above), *kirotas* 's/he writes' and many others that only take *-s* in 3Sg (see Iva 2007: 83–84).

By comparison in Latvian and Latgalian, there is no number distinction, so 3Sg and 3Pl regularly overlap regardless of tense (see Table 4). As 3rd person forms with no ending correspond to the proto-language

phase, Vaba (2011: 212) points to an interesting parallel with the Finnic languages, which also lack this ending. Depending on the word type, conflation may also concern the 2Sg form, e.g., in the case of Latvian *runāt* 'to speak' the present tense form for 2Sg, 3Sg, and 3Pl is *runā* (see also Endzelīns 1951: 706). Whereas Standard Russian distinguishes endings for all persons, in Russian dialects spoken nearby, the personal ending may be missing in 3Sg and/or 3Pl forms (see more Zaharova & Orlova 2004: 149, 151).

Table 4. Personal endings in Latvian, Latgalian, and Russian.

Language	Tense	1Sg	2Sg	3Sg	1Pl	2Pl	3Pl
Latvian	PRS	*-u*	*-i*	*-a*	*-am, -ām*	*-at, -āt*	*-a*
			Ø	Ø			Ø
	PST	*-u*	*-i*	*-a*	*-ām*	*-āt*	*-a*
	FUT	*-u*	*-i*	Ø	*-im*	*-iet, -it*	Ø
Latgalian	PRS	*-u*	*-i*	*-a*	*-am, -im, -om*	*-at, -itʲ, -ot*	*-a*
			Ø	Ø			Ø
	PST	*-u*	*-i*	*-a, -e*	*-om, -em*	*-ot, -etʲ*	*-a, -e*
	FUT	*-u*	*-i*	Ø	*-im*	*-itʲ*	Ø
Russian	PRS	*-u, -ju*	*-esh', -ish'*	*-et, -it*	*-em, -im*	*-ete, -ite*	*-ut, -jut, -at, -jat*
	PST		No person distinction, only number and gender distinction				

The present indicative forms in Leivu, Lutsi, and Kraasna show syncretism between 2Sg and 2Pl. While in Lutsi, *-t* appears in both persons, in Kraasna and Leivu, *-t* and *-dE* covary (28a–b), although *-t* is generally more common in 2Sg and *-dE* in 2Pl. A similar pattern is also found in Võro and Seto. The variation between *-t* and *-dE* depends on the subdialect.

(28) Lei: a. **sa** *teija-**t***   ***tii*** *tija-**t***
              (Mets et al. 2014: 42)   (Mets et al. 2014: 86)
              (2sG know-2sG)   (2PL know-2PL)
              'you know'   'you know'

b. ***tii*** *kulle-**de***   ***sa*** *tunne-**de***
(Mets et al. 2014: 54)   (Mets et al. 2014: 56)
(2PL hear-2PL)   (2SG feel-2SG)
'you hear'   'you feel'

In the past indicative, Lutsi shows syncretism in 2Sg, 2Pl, and 3Pl, as in (29) (see Balodis 2020: 84–89). In Kraasna, syncretism additionally involves 1Pl, e.g., *mii anniq* 'we gave' (AES 202: 25–26). The respective forms in Lutsi and Kraasna generally include the marker *-q*, while in Leivu, *-q* is found in 2Pl and 3Pl. Still, according to Table 3, lack of a personal ending is also possible (probably also in 2Pl), e.g., see (30). Although this reveals some similarities with Võro and Seto where the personal ending is commonly dropped in 1Pl and 2Pl, even stronger parallels can be drawn with Salaca Livonian where one and the same form can be used for all persons (see Table 3; see also Winkler & Pajusalu 2018: 115). It appears that this is also true for the Livonian-like subdialects of Latvian in Courland and in northern Vidzeme (Rudzīte 2005: 77). According to Balode & Holvoet (2001: 29), personal endings were first lost in the singular paradigm as a result of the loss of final vowels, which could be the result of a Livonian substrate; homonymy of 3rd person singular and plural forms, in turn, facilitated their further spread in the plural paradigm.

(29) Lut: *andma* 'to give': Pst2Sg/2Pl/3Pl: ***anniq***

(30) Lei: ***sa ütleži***   ***mii elli***
              (Mets et al. 2014: 28)   (Mets et al. 2014: 48)
              (2SG say.PST)   (1PL live.PST)
              'you said'   'we lived'

In addition to Lutsi, conflation of 2Pl and 3Pl in the past tense is regular in Courland Livonian and may additionally involve 2Sg. The choice between *-t* and *-tõ* in 2Pl/3Pl usually depends on word structure (e.g., one-syllable words regularly take *-tõ*, as in (31a)), but there are also words like *ki'zzõ* 'to ask', which permit variation, as in (31b). In Salaca Livonian, the Pst2Pl marker is *-ti* and the Pst3Pl marker is *-t*. However, as it is possible to drop *-i* in the present tense (see Winkler & Pajusalu 2018: 115), it is likewise possible that *-i* could be dropped in the past tense.

(31) CLiv: a. *tõ̬dõ* 'to want': Pst2Sg: ***tõ̬'žt***, Pst2Pl/3Pl: ***tõ̬'žtõ***
              (Viitso & Ernštreits 2012)

              b. *ki'zzõ* 'to ask': Pst2Sg: ***kizīzt,*** Pst2Pl/3Pl: ***kizīzt ~ kizīztõ***
              (Viitso & Ernštreits 2012)

Regardless of the conflating forms in the paradigm, all the Finnic varieties included in the study permit dropping the subject pronoun. This means that in the same way as there are examples where the actual referent becomes clear from verbal morphology (32), there are also instances where morphology is insufficient to establish the referent (33). A further example can be brought from Salaca Livonian (34a) where only the translation suggests that *om* is used for 2Sg. First and foremost, it would be expected to express 3Sg or 1Sg, but as Table 3 illustrates, syncretism is possible in most of the persons (see also Winkler & Pajusalu 2018: 115). In Courland Livonian, the same 3Sg/1Sg forms are additionally used for impersonal reference (e.g., *sīeb* 'I eat, s/he is eating, it is being eaten'). By comparison, in their study on the use of 1st person pronouns in Estonian dialects, Lindström et al. (2009) show that the presence or absence of a pronoun depends on the dialectal area rather than whether the personal ending is used or not. It appears that one such area – where pronouns are commonly dropped – is southern Estonia. In their article, they also list several other factors that play a role.

(32) Lut: ***kääli-mi*** *tuuda škuolla* (Mets et al. 2014: 175)
              go.PST-1PL there school:ILL
              'we went to school there'

(33) Set: ***käve*** *poodih*
              go.PST.1SG/1PL shop:INE
              '(I/we) went to store'

(34) SLiv: a. *Mill om jua miel, ku täru **om**.* (Winkler & Pajusalu 2018: 170)
              1sg:ADE;ALL be.3SG good sense that healthy be.3SG

              b. S: *jag fägnar mig, att du är frisk*
              1SG delight.PRS 1SG.ACC that 2SG be.2SG healthy
              'I am happy that [you] are healthy'

All in all, person-indexing and pro-drop in the Finnic varieties show that interesting developments are broader and do not overlap with syncretic forms in the non-cognate contact varieties, thus, changes in the system reflect language internal developments characteristic of the area rather than contact-induced change.

3.4. Standard negation

Standard negation refers to the negation of declarative verbal main clauses (Miestamo 2007). A distinction can be made based on variations in three properties: 1) symmetricity (whether the form of the lexical verb in the affirmative and negative clauses differs or not), 2) the type of negative marker (a negative auxiliary inflecting for tense and/or person, an invariable negative particle, clitic, or affix), 3) the position of the negative marker relative to the lexical verb (preverbal, postverbal, or double). In the case of asymmetric negation, the form of the lexical verb may vary either appearing as a bare root or also containing an affix. One or more of the following categories are expressed in the negative marker: person, number, or tense (see, e.g., Miestamo et al. 2015, Metslang et al. 2015, Lindström et al. 2021).

The analysed varieties contained examples of all of the types mentioned above (see Table 5). Some features are not shown in the table as they are not found in the analysed languages, e.g., there are only examples of negative prefixes, thus suffixes are not represented in the table; clitics always follow the lexical verb, while particles precede it.

Table 5. Properties of standard negation.

	Est	Vro, Set, Kra, Lut	Lei, Mul, SLiv	CLiv	Lav, Ltg	Rus
symmetric	–	+/–	+/–	+/–	+	+
prevbl neg aux	–	+	+	+	–	–
prevbl ptcl	+	–	–	–	–	+
postvbl clitic	–	+	–	–	–	–
prefix	–	–	–	–	+	–
double neg	–	+	–	–	–	–

Regarding structure, Latvian, Latgalian, and Russian contain symmetric negative constructions as negation differs from affirmation only by the addition of a negative marker, see (35). In Estonian, standard negation is asymmetric, as there are additional differences: unlike in the affirmative, in the negative the lexical verb has no personal ending, see (36). In the other studied varieties, both symmetric and asymmetric negation can be found (indicated by +/–), see (37) and (38). In the South Estonian varieties, including in the language islands, symmetricity regularly involves 1Sg and 3Sg as there is no personal ending; in other persons, symmetricity is possible but there is some variation (see section 3.3). In Courland Livonian, this distinction is made on the basis of number: negation is asymmetric in all singular persons but symmetric in plural, see, e.g., (38a–b). Examples such as (39) indicate that Salaca Livonian also aligns with Courland Livonian.

(35) Rus: *ja **piš-u** : ja ne **piš-u***
1SG write-1SG : 1SG NEG write-1SG
'I write : I don't write'

(36) Est: *sa **tööta-d** : sa ei **tööta***
2SG work-2SG : 2SG NEG work.CNG
'you work : you don't work'

(37) Lut: a. *Ma **kynele** lutsi kīlt : ma **kynele**=**eiq** lutsi kīlt* (Balodis 2020: 83)
1SG speak Lutsi language:PRT : 1SG speak.CNG=NEG Lutsi language:PRT
'I speak Lutsi : I do not speak Lutsi'

b. *Mī **kynele-m** lutsi kīlt : mī **kynele**=**eiq** lutsi kīlt* (Balodis 2020: 83)
1PL speak-1PL Lutsi language:PRT / 1PL speak.CNG=NEG Lutsi language:PRT
'We speak Lutsi : we do not speak Lutsi'

(38) CLiv: a. *ma **nǟ-b** : ma äb **nǟ***
1SG see-1SG : 1SG NEG:1SG see.CNG
'I see : I don't see'

b. *tēg **nǟ-tõ** : tēg ät **nǟ-tõ***
2PL see-2PL : 2PL NEG:2PL see-2PL
'you see : you don't see'

(39) SLiv: *Voj tee ab **uo-ti** korren?* (Winkler & Pajusalu 2018)
            q 2PL NEG be-2PL pick.APP
            'Have you not picked (sth)?'

While Estonian and Russian use a particle, and Latgalian and Latvian use a prefix to mark negation, the other varieties use a negative auxiliary that inflects for tense; in Courland Livonian, the negative marker additionally inflects for person and number (see Table 6). As Table 6 suggests, a distinction in tense is an example of a feature that has been preserved regardless of whether a different model is used in the non-cognate languages spoken in close proximity. However, parallels can be drawn with the use of preverbal non-inflected negative markers and in the Indo-European contact languages (Stolz 1991: 70–73; see also Koptjevskaja-Tamm & Wälchli 2001: 628). Outside the southern-most Finnic area, the negative marker has also inflected for tense in the Insular dialect of Estonian and in the Kodavere subdialect of eastern Estonia. The latter contained a full personal paradigm that went out of use in the 1940s or 1950s (Viikberg 2020: 296).

Table 6. Markers of standard negation.

	Est	Vro, Set, Kra, Lut	Mul	Lei	SLiv	CLiv	Lav	Ltg	Rus
PRS	*ei*	*ei(q)*	*ei*	*ei*	*ab*	*äb* (1Sg, 3Sg, 1Pl, 3Pl), *äd* (2Sg), *ät* (2Pl)	*ne-*	*na-*	*ne*
PST		*es*	*es*	*is*	*is*	*iz* (1Sg, 3Sg, 1Pl), *izt* (2Sg, 2Pl, 3Pl)			

More precisely, Mulgi, Võro, Seto, Lutsi, and Kraasna make a distinction between *ei(q)* and *es* that in Võro, Seto, Lutsi, and Kraasna can be used both preverbally as well as postverbally (see also Table 5) with some variation in the initial vowel of the negative verb (e.g., 40a–b). In Livonian, the general distinction is made between *ab* vs. *is* / *äb* vs. *iz*. As Table 6 illustrates, Leivu shows commonalities with both: *ei* shows similarities with Estonian and South Estonian varieties, including the language islands, whereas the negative past marker *is* shows a parallel with Livonian.

(40) Kra: a. *ma **lää ei*** (Mets et al. 2014: 278)
               1SG go.CNG NEG
               'I don't go'

               b. *timä **lää äs*** (Mets et al. 2014: 278)
               3SG go.CNG NEG.PST
               'S/he didn't go'

With regard to the position of the negative marker in the analysed varieties, the preverbal position prevails (see Table 5). As already noted, both preverbal and postverbal positions are possible in Lutsi, Kraasna, Seto, and Võro (e.g., 40). Still, in Seto and Kraasna, the postverbal clitic turns out to be the primary option (for Seto, see also Lindström et al. 2021). For instance, (42) is one of the few examples containing the preverbal *ei* in Kraasna. However, this may be due to the scarcity of available Kraasna data. Leivu texts, in turn, only contained examples of the negative marker preceding the verb, as in (43).

(41) Vro: ***ei** annaq ~ anna **eiq*** (Iva 2007: 102)
               NEG give.CNG ~ give.CNG NEG
               'doesn't give'

(42) Kra: *ma **tiijä=eiq** / **ei saa** andaq arq* (Weber 2021 in this volume, App. 2)
               1SG know.CNG=NEG / NEG can.CNG give:INF away
               'I don't know, I cannot give [my daughter as a wife]'

(43) Lei: *poig **is** tejja* (Mets et al. 2014: 25)
               son NEG.PST know.CNG
               '[My] son didn't know'

Kraasna, Lutsi, Seto, and Võro also show instances of double negation, as in (44a–b). As example (44b) reveals, double negation can also be attested with negative imperatives (according to Lindström et al. 2021, this is also possible in Seto). The function of double negation in Võro and Seto has included intensifying negation and marking its scope. In present-day Seto, the use of double negation is rare and inconsistent (Lindström et al. 2021).

(44) Set: a. *läts´ arq kõrdsi kotsele ne ináb edese **es saa-ke eiq*** (Saar et al.
              (in preparation))
              go.PST.3SG off tavern.GEN near and more further NEG.PST get.
              CNG-PTCL NEG
              '[S/he] went near the tavern and did not even get any further'

      Kra: b. *är võttu=i setä* (AES 202)
              NEG.IMP take=NEG this.PRT
              'do not take this'

Regarding negation, it can be observed that the distinction between present and past tense forms has been preserved in all of the analysed Finnic varieties except Estonian. Thus, the lack of such a distinction in the non-cognate contact varieties has not had an effect. Furthermore, neither Võro nor Seto are under threat of losing this distinction as a result of influence from Standard Estonian. Also noteworthy is that Kraasna, Võro, Seto, and Lutsi pattern together both in terms of structure and formal properties, while Leivu is closer to Mulgi and Salaca Livonian.

4. Phonological features

South Estonian dialects have a number of characteristic phonological features, which also have broader areal connections (see Pajusalu 2012). Below we consider ten features, all of which occur in Leivu and Lutsi, although several are more marginal in Leivu, and, with one exception are all also characteristic of Kraasna. For this group we chose unique features of word prosody, vowels, and consonants; see Table 7.

Like other Southern Finnic languages and dialects, the South Estonian language island varieties are characterised by complex quantity alternations, which can also be combined with tonal contrasts (see also Koptjevskaja-Tamm & Wälchli 2001: 640−644). Additionally, the realisation of these prosodic phenomena can be associated with sound quality changes.

Table 7. Phonological features (+ – occurrence of a feature; (+) – limited occurrence of a feature; − – absence of a feature).

Feature	Est	Võr	Set	Kra	Lut	Lei	Mul	SLiv	CLiv	Lav	Ltg	Rus
stød	−	−	−	−	(+)	(+)	−	+	+	+	+	−
glottal stop	−	+	+	+	+	+	−	−	−	−	−	−
h	(+)	+	+	+	+	(+)	(+)	−	−	−	−	+
voiced plosives	−	−	(+)	+	+	+	−	+	+	+	+	+
short vs. long cons. geminates	+	+	+	+	+	+	+	+	+	−	−	−
short vs. long vowels	+	+	+	+	+	+	+	+	+	+	+	−
õ [ɤ] / y [ɯ]	+	+	+	+	+	+	+	(+)	+	−	+	+
ü [y] / ö [ø]	+	+	+	+	+	+	+	+	(+)	−	−	−
vowel harmony	−	+	+	+	+	(+)	(+)	−	−	−	(+)	−
extensive palatalisation	−	+	+	+	+	(+)	(+)	−	−	−	+	+

Tonal variation is not typical of the Finno-Ugric languages. **Broken tone or stød** is characteristic, however, of Latvian and Latgalian, and is also found in Livonian (e.g., *lēʼḍ* [leːˀdʲ] 'leaf; page'). In addition to Livonian, broken tone is also found in Leivu and Lutsi, e.g., Leivu *vähämb* [ˈvæ.hæmb] > *väʼämb* [væːˀmb] 'less', *naha* [ˈna.ha] > *naʼa* [naːˀ] 'skin, gen.sg.', *rahaga* [ˈra.ha.ga] > *raʼaga* [ˈraːˀ.ga] 'money. com'; Lutsi *hiʼir* [hiːˀr] 'mouse' (in which broken tone is connected with the third quantity degree), *rehe* [ˈre.he] > *reʼe* [reːˀ] 'threshing barn' (see Balodis et al. 2016), but it has not been observed in other South Estonian dialects or in Kraasna.

Glottal stop is a frequent phoneme in the South Estonian phonological system distinguishing grammatical meanings, cf., e.g., *kala* [ˈka.la] 'fish' and *kalaq* [ˈka.laʔ] 'fish, pl.', *(ma) anna* [ˈan.na] '(I) give' and *annaq* [ˈan.naʔ] 'give.imp'. Typically, glottal stop occurs word-finally in South Estonian; however, it can sometimes also be found word-internally or elsewhere (see Iva 2005), e.g., Seto *iqe latś* [ˈi.ʔe latʲsʲ] 'good child' (Saar et al. 2020). Glottal stop can also occur following

liquids, e.g., Lutsi, Leivu, Kraasna *kolq* [kolʔ] 'three', *ärq* [ærʔ] 'away; off'. Glottal stop does not exist as a phoneme in Latvian or Latgalian but can occur in careful or emphatic pronunciations of broken tone (see Grigorjevs 2011), e.g., Latvian *nē* [neːʔ] 'no' realised as [ˈne.ʔe]. Markus (2012) also notes that pronunciation of broken tone fully or partially as a glottal stop is especially characteristic of the Latgalian subdialects of northeastern Vidzeme. This region, also known as Malēnija, is located near the historical Leivu region in Latvia as well as South Estonian-speaking areas in Estonia. A unique feature of the Lutsi and Leivu language islands is the presence of both broken tone and glottal stop.

As is historically the case in Livonian, broken tone is also partially related to the **loss of *h*** in Leivu and Lutsi, e.g., *ra'a* [raːʔ] < *raha* [ˈra.ha] 'money' (see also Teras 2010, Viitso 2009: 277–279). In Latvian and Latgalian, *h* is not found as a native phoneme. In Estonian and in eastern South Estonian subdialects as well as in Kraasna, pronunciation of *h* has, as a rule, remained (e.g., Võro and Seto *hõbõhhõnõ* [ˈhɤ.bəh.hɤ.nə] 'silvery', Kraasna and Seto *luhits* [ˈlu.hʲits] 'spoon', *raha* 'money'), which may have been aided by the neighbouring Russian language. As the occurrence of broken tone appears to be sporadic in Lutsi, inter-vocalic *h* also remains common in this variety (e.g., Lutsi *ähäq* [ˈæ.hæʔ] 'wedding'). In North Estonian subdialects and also in everyday spoken Estonian, word-initial *h* is often not pronounced. Loss of *h* is also wide-spread in the western South Estonian subdialects as well as in Mulgi. On the other hand, along with loss of *h*, *h* hypercorrection is also found in Mulgi, for example in the words *kähen* [ˈkæ.hen] 'in hand; at hand' (cf. *käen*), *pähän* [ˈpæ.hæn] 'on/in the head' (cf. *pään*), *pääle* [ˈpæː.le] ~ *pähle* [ˈpæh.le] 'onto, over' (Tanning 1961: 21, 43).

A phonological distinction between long and short sounds is characteristic of the languages of the Baltic region. Unlike in Russian, this distinction is found in all Finnic and Baltic languages in the area under study. In North and South Estonian subdialects (incl. the language islands), there is an additional contrast between long and over-long duration, which is why we speak of three quantity degrees in these languages, cf., e.g., South Estonian Q1 *külä* [ˈky.læ] 'village', Q2 *kül̆lä* [ˈkyl.læ] 'village.part', and Q3 *kùllä* [ˈkylː.læ] 'village.ill'. A three-way contrast for consonants is also found in Livonian; the realisation of this length contrast in a disyllabic trochaic foot is also characteristic of these languages (Markus et al. 2013). The trochaic foot system is

characteristic of Estonian runic songs as well as the Latvian dainas. All of the languages of the language area under study have initial syllable primary stress, including Latvian and Latgalian (but unlike, for example, Russian and Lithuanian); most of the Circum-Baltic languages show basic initial stress (Koptjevskaja-Tamm & Wälchli 2001: 638−640). In contrast to other Russian dialects, stress is also moved to the beginning of the word in the Pskov dialect, which is spoken alongside Finnic speakers, e.g., *rúčej* (cf. *ručéj*) 'stream', *Bóris* (cf. *Borís*) 'Boris (given name)' (Kostjučuk 2018: 119). In the case of stops, the characteristic feature connecting Livonian and the South Estonian language island varieties is intervocalic voicing of single consonants (e.g., Kraasna *hõbõhhõnõ* 'silvery', Lutsi *regi* [ˈre.gi] 'sled', Livonian *tubā* [ˈtu.baː] 'room; house'); voiceless stops occurring intervocalically are always pronounced as geminates: single voiceless stops as short geminates (Kraasna *hõpõ* [ˈhɤp.pə] 'silver', Lutsi *reke* [ˈrek.ke] 'sled.part', Livonian *liepā* [ˈliep.paː] 'alder') and geminate stops as long geminates (Kraasna *tappa* [ˈtapː.pa] 'to kill.inf', Lutsi *rekke* [ˈrekː.ke] 'sled.ill', Livonian *lieppõ* [ˈliepː.pə] 'alder.part'). Similarly in Latvian, voiceless stops are pronounced as half-long when located between a short stressed and short unstressed vowel (Laua 1997, Kalnača 2004) (e.g., *lapa* [ˈlap.pa] 'leaf'). In Livonian, gemination of voiced stops is also present, e.g., *tuʼbbõ* [ˈtuˀb.bə] 'room.part; house.part', *vieʼddõn* [ˈvieˀd.dən] 'water.dat' (see Viitso 2008: 296), which is not characteristic of any of the other languages discussed in this study.

The Southern Finnic languages, Latgalian, and Russian are connected by the presence of the **back or central unrounded vowels** *õ* [ɤ] and/or *y* [ɯ], [ɨ]. The close-mid vowel *õ* [ɤ] is characteristic of Estonian and Mulgi (e.g., Estonian *õde* [ˈɤ.te] 'sister', Mulgi *sõsar* [ˈsɤ.sar] 'sister'), the close vowel *y* [ɯ] of Livonian and Lutsi (e.g., Livonian *sõzār* [ˈsɯ.zaːr] 'sister', Lutsi *syzaŗ* [ˈsɯ.zarʲ] 'sister'); close-mid *õ* [ɤ] and close *y* [ɯ] are found in Seto and Kraasna as well as marginally in Leivu and Võro, e.g., Seto and Võro *sysar* [ˈsɯ.sarʲ] 'sister'. In Latgalian and Russian, *y* is usually pronounced as [ɨ]; only in Latvian is this vowel not found. Despite the strong influence of Latvian, the back unrounded vowels have nevertheless been maintained in all of the Finnic languages and dialects spoken in Latvia.

The **front rounded vowels** *ü* [y] and *ö* [ø] are not found in Latvian, Latgalian, or Russian. At the same time, they have remained in the South

Estonian language islands much as in other Estonian dialect areas, e.g., Kraasna *sügüzelt* [ˈsy.gy.zelt] 'in autumn', *söögi* [ˈsøː.gi] 'food.gen'; Lutsi *üözeq* [ˈyø.zeʔ] ~ *ȫzeq* [ˈøː.zeʔ] 'at night', *sūmä* [ˈsyːː.mæ] 'to eat.inf'; Leivu *püäbä* [ˈpyæ.bæ] 'Sunday'. In the 19th century, *ü* [y] and *ö* [ø] were still found in Livonian, but in the early 20th century *ü* was replaced by *i* and *ö* by *e* in Courland Livonian (see Viitso 2011).

Vowel harmony is characteristic of the Finno-Ugric languages and is also found in South Estonian as front and back *ä*- and *ü*-harmonies, cf. *valla* [ˈvalː.la] 'open' and *vällä* [ˈvælː.læ] 'out', *tulu* [ˈtu.lu] 'profit' and *tülü* [ˈty.ly] 'quarrel', and also as velar *õ*-harmony, cf. *tege* [ˈte.ge] 'makes' and *tulõ* [ˈtu.lə] 'comes'. Atypically for the Finnic languages, vowel harmony is not found in the North Estonian subdialects, Standard Estonian, or Livonian. It also does not occur in most Baltic or Slavic languages, however, Latgalian has a morphophonological harmony which superficially resembles vowel harmony (cf. Nau 2011). Palatal vowel harmony has been maintained in the South Estonian language islands, but it is not entirely consistent. *ü*-harmony fluctuates more, while *ä*-harmony and especially *õ*-harmony are better preserved. Much as in the Western dialect of South Estonian, vowel harmony fluctuates more in Leivu than it does in Lutsi or Kraasna (cf. Wiik 1988).

South Estonian pronunciation differs from that of North Estonian dialects, Standard Estonian, and Latvian due to its extensive **palatalisation** of consonant phonemes, also coarticulatorily near front vowels. As in Russian, and also Latgalian with a few exceptions (for more see Breidaks 2006), all consonants can, in principle, be palatalised in South Estonian, as long as it is articulatorily possible, e.g., Võro, Seto *kapp* [kapʲː] 'cupboard', *pikk* [pikʲː] 'long; tall', *kamm* [kamʲː] 'comb', Seto *ruuhh* [ruːhʲː] 'drug, medicine'. This extensive palatalisation of consonants is also found in all of the South Estonian language islands (e.g., Kraasna *śäŕḱḱi* [ˈsʲærʲkʲ.kʲi] 'coat.part'; Lutsi *tsirguq* [ˈtʲsʲir.guʔ] 'birds', *koṛaṣ* [ˈko.rʲasʲ] '(s/he) gathered'; Leivu *dʹalg* [dʲalːg] 'foot', *tulʹlʹe* [ˈtulʲː.lʲe] '(s/he) came'); however, similar to Mulgi, it is more limited in Leivu.

The above overview of phonological developments in the South Estonian language islands shows that the language islands have maintained the main features of Finnic and, more narrowly, of South Estonian, while also acquiring features characteristic of the Baltic languages such as broken tone and voiced consonants. Latvian influence

is strongest in the westernmost Leivu language island where several phonetic developments are also similar to the western Mulgi dialect. The more eastern position of Kraasna and Lutsi encouraged preservation of characteristics similar to Russian such as the preservation of *h* and extensive consonant palatalisation. At the same time, Finno-Ugric features such as vowel harmony have been better preserved in more eastern South Estonian varieties. Development of tonal contrasts while still preserving the three-way quantity alternation characterises both of the South Estonian varieties spoken in Latvia – Leivu and Lutsi, and, in fact, the importance of these tonal contrasts also increases for quantity alternation (see Balodis et al. 2016).

5. Conclusions

This article examined various features in the domains of morpho-syntax and phonology that were relevant for consideration in an areal perspective. Its main focus was on the South Estonian language islands – Leivu, Lutsi, and Kraasna – but also made comparisons with the other main language varieties of the Central Baltic area: Estonian and the South Estonian varieties – Mulgi, Võro, and Seto; Latgalian and Latvian; Salaca Livonian and Courland Livonian; Russian and its local varieties.

The results of our analysis of the selected features further support the hypothesis that in the Circum-Baltic area, convergence mainly occurs at the micro-level often involving only two to three languages. As this paper makes a more fine-grained distinction at a more specific level, we could see a multitude of patterns among the analysed varieties that point to multifaceted contact situations and their outcomes in the area. In several instances, Lutsi and Kraasna patterned together with Võro and Seto, while Leivu showed greater similarities with Salaca Livonian and Mulgi. This division is evident in the properties of standard negation, occurrence of certain phonological features, e.g., consonant palatalisation, the occurrence of *h*. With regard to comparative constructions, only Lutsi, Kraasna, Võro, and Seto revealed instances of partitive marking (similar to the use of the genitive in Russian) to express the standard, although they all contained additional methods for creating comparative constructions (e.g., using the elative case and/or a particle construction). It is possible that over time, as the partitive lost its

separative meaning the elative took over as an explicitly separative case. This, however, did not happen in the varieties that had a supporting model in the neighbouring languages.

As regards contact induced-changes caused by neighbouring Indo-European languages, we could find cases of PAT-borrowing as hypothesised. The clearest instances of PAT-borrowing were the particle comparatives that make use of the negative marker – only the Finnic varieties that have had close contacts with Latvian and/or Latgalian (Lutsi, Leivu, Courland Livonian, Salaca Livonian) contained such examples. This shows that a pattern is likely to be borrowed if it has spread over a wide territory in a language that has a dominant position in society.

At the same time, there were developments that could not be considered a direct influence of the main non-cognate contact varieties. For instance, changes in the person-indexing system, which have led to various types of syncretism in the Finnic varieties, probably result from language internal developments. Although syncretism is characteristic of the Livonian-like subdialects of Latvian (the contact variety of Courland Livonian and Salaca Livonian), Latvian and Latgalian in general, but also the Russian dialects. As was shown, syncretism is much more widespread in the studied Finnic varieties (except Standard Estonian).

As was also hypothesised, certain structural features have persisted despite the presence of differing models in the main contact varieties. One such example is the distinction between past and present tense in negative markers. This distinction is found in all of the analysed varieties except the main standard varieties in the area (Estonian, Latvian, Latgalian, Russian), which instead use invariable negative markers. In the case of Estonian, the reasons for the simplification of the system could also be sought in the work done by the Germans in developing the written standard, as German also does not have an inflective negative marker. Another example of a characteristic Finnic feature that has been stable is vowel harmony, which has been better preserved in the eastern periphery of the studied area, i.e., closer to predominantly Russian-speaking regions.

This study also contained examples of features or bundles of features, which reflect the impact of conscious language planning. For example, two particular phenomena in Standard Estonian are the result of language planning at the end of the 19th century. Thus, Standard

Estonian shows a different pattern for agreement within noun phrases than is found in the South Estonian language islands or in other non-standard varieties including ones for which a literary standard is of relatively recent origin. Likewise, the essive case was on the verge of disappearing in Estonian, but it was revived in the standard language, while in the other Finnic varieties examined here it has become unproductive. This shows that when favourable conditions exist (not too large of a language community, the standard language is only just developing) language planning can lead to changes in the language system that speakers are ready to accept.

Acknowledgements

Abbreviations

1,2,3 – person, ABE – abessive, ACC – accusative, ADE – adessive, ALL – allative, APP – active past participle, CL – clitic, CLiv – Courland Livonian, CNG – connegative, COM – comitative, COMP – comparative, COND – conditional, CONJ – conjunction, ELA – elative, ESS – essive, Est – Estonian, F – feminine, FUT – future, GEN – genitive, ILL – illative, IMP – imperative, INE – inessive, INF – infinitive, INS – instructive, IPS – impersonal, Kra – Kraasna, Lav – Latvian, Lei – Leivu, Ltg – Latgalian, Lut – Lutsi, M – masculine, Mul – Mulgi, NEG – negative, PL – plural, PRS – present, PRT – partitive, PST – past, PTCL – particle, Q – question particle, REFL – reflexive, Rus – Russian, Set – Seto, SG – singular, SLiv – Salaca Livonian, SUP – supine, TERM – terminative, Vro – Võro

References

AES 202 = Ojansuu, Heikki. 1938. *Akadeemilise Emakeele Seltsi ülevaated. Häälikuloolisi andmeid ja tekste Kraasna murdest.* http://emsuka.eki.ee/view-book-collection/2519 (Tallinn, Eesti Keele Instituut (Institute of the Estonian Language) / Emakeele Selts (Mother Tongue Society))

Aikhenvald, Alexandra Y. 2006. Grammars in contact. A cross-linguistic perspective. In Alexandra. Y. Aikhenvald & R.M.W. Dixon (eds.), *Grammars in contact. A cross-linguistic typology*, 1–66. Oxford: Oxford University Press.

Ariste, Paul. 1962. Ühest keelekontakti juhust. *Keel ja Kirjandus* 5(9). 550–556.

Ariste, Paul. 1981. *Keelekontaktid. Eesti keele kontakte teiste keeltega* (Eesti NSV Teaduste Akadeemia Emakeele seltsi toimetised 14). Tallinn: Valgus.

Balode, Laimute & Axel Holvoet. 2001. The Latvian language and its dialects. In Östen Dahl & Maria Koptjevskaja-Tamm (eds.), *The Circum-Baltic languages: typology and contact. Volume 1: Past and Present* (Studies in Language Companion Series 55), 3–40. Amsterdam, Philadelphia: John Benjamins Publishing Company. https://doi.org/10.1075/slcs.54.04bal.

Balodis, Uldis. 2015. Writing down Lutsi: Creating an orthography for a South Estonian variety of Latgale. In Andra Kalnača (ed.), *Valoda: nozīme un forma 6. Valodas sistēma un lietojums* (= Language: Meaning and Form 6. Language System and Language Use), 55–67. Rīga: Latvijas Universitātes Akadēmiskais apgāds.

Balodis, Uldis. 2020. *Lutsi kiele lementar = Ludzas igauņu valodas ābece.* Ludza: LU Lībiešu institūts, Ludzas pilsētas galvenā bibliotēka.

Balodis, Uldis, Karl Pajusalu & Pire Teras. 2016. Broken tone in South Estonian dialects in Latvia. *Linguistica Lettica* 24. 98–114.

Bernštein, Samuil. 2005. *Sravnitel'naja grammatika slavjanskih jazykov.* Moskva: MGU, Nauka.

Blokland, Rogier & Nobufumi Inaba. 2018. The l-cases in Courland Livonian. *Eesti ja soome-ugri keeleteaduse ajakiri. Journal of Estonian and Finno-Ugric Linguistics* 9(2). 147–164. https://doi.org/10.12697/jeful.2018.9.2.07.

Blumberga, Renāte. 2013. Lībieši 19.–21. gadsimtā. In Renāte Blumberga, Tapio Mäkeläinen & Karl Pajusalu (eds.), *Lībieši. Vēsture, valoda un kultūra*, 169–204. Rīga: Līvõ Kultūr sidām.

Breidaks, Antons. 2006. Dialektnaja leksika latgalskih govorov verhnelatishskogo dialekta i ee istoricheskije svjazi. In Breidaks, Antons. *Darbu izlase. 1*, 225–482. Rīga: LU Latviešu valodas institūts..

Breidaks, Antons. 2007. *Darbu izlase. 2.* Rīga: LU Latviešu valodas institūts, Daugavpils Universitāte.

Despić, Miloje. 2017. Suspended morphology in Serbian: Clitics vs. affixes. *Glossa: a journal of general linguistics* 2(1): 12. 1–43. https://doi.org/10.5334/gjgl.130.

EMK = *Eesti murrete korpus* [Estonian Dialect Corpus]. http://www.murre.ut.ee/mkweb (3 May, 2021).

EKI WK = Eesti Keele Instituudi Wiedemanni kartoteek. Tallinn: Eesti Keele Instituut [Institute of the Estonian Language].

EMS = *Eesti murrete sõnaraamat*. Tallinn: Eesti Keele Instituut. https://www.eki.ee/dict/ems/ (10 May, 2021).

Endzelīns, Jānis. 1951. *Latviešu valodas gramatika.* Rīga: Latvijas Valsts izdevniecība.

Ernštreits, Valts. 2012. Development of the Livonian language situation in Latvia. In Ina Druviete (ed.), *Language situation in Latvia: 2004–2010*, 150–173. Riga: Latvian Language Agency.

ESA 2011 = Eesti Statistikaamet 2011: Rahva ja eluruumide loendus 2011. *Statistika andmebaas.* https://andmed.stat.ee/et/stat (3 May, 2021).

Faster, Mariko, Laivi Org, Urmas Kalla, Sulev Iva & Triin Iva. 2014. *Eesti-võru sõnaraamat. Eesti-võro sõnaraamat*. Võru: Võro Instituut.

Grigorjevs, Juris. 2011. Kopīgais un atšķirīgais fonētisko parādību izpratnē Latvijā un pasaulē. In Ojars Bušs (ed.), *Valodas prakse: Vērojumi un ieteikumi. Populārzinātnisku rakstu krājums 6*, 80–96. Rīga: Latviešu valodas aģentūra.

Grünthal, Riho, Volker Heyd, Sampsa Holopainen, Juha Janhunen, Olesya Khanina, Matti Miestamo, Johanna Nichols, Janne Saarikivi & Kaius Sinnemäki. In press. Drastic demographic events triggered the Uralic spread. *Diachronica*. John Benjamins, https://researchportal.helsinki.fi/en/publications/drastic-demographic-events-triggered-the-uralic-spread. (31 August, 2021).

Hansen, Klaus. 2000. Genitive case or determinative phrase? The status of 's forms in Modern English. In Dalton-Puffer, Christiane & Nikolaus Ritt (eds.), *Words: Structure, Meaning, Function. A Festschrift for Dieter Kastovsky* (Trends in Linguistics 130), 111–124. Berlin, New York: Mouton de Gruyter. https://doi.org/10.1515/9783110809169.111.

Haspelmath, Martin. 1998. How young is Standard Average European? *Language Sciences* 20(3). 272–287. https://doi.org/10.1016/S0388-0001(98)00004-7.

Haspelmath, Martin. 2001. The European linguistic area: Standard Average European. In Martin Haspelmath, Ekkehard König, Wulf Oesterreicher & Wolfgang Raible (eds.), *Language typology and language universals: An international handbook.* Vol. 2 (Handbücher zur Sprach- und Kommunikationswissenschaft, 20.2), 1492–1510. New York: Walter de Gruyter. https://doi.org/10.1515/9783110194265-044.

Heine, Bernd & Tania Kuteva. 2006. *The changing languages of Europe* (Oxford Linguistics). Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199297337.001.0001.

Hermann, Karl August. 1884. *Eesti keele Grammatik*. Tartu: Wilhelm Just.

Iva, Sulev. 2005. Glottal Stop in Võro South Estonian. *Linguistica Uralica* 41(2). 123−133.

Iva, Sulev. 2007. *Võru kirjakeele sõnamuutmissüsteem* (Dissertationes philologiae estonicae Universitatis Tartuensis 20). Tartu: Tartu Ülikooli Kirjastus.

Janhunen, Juha. 1982. On the Structure of Proto-Uralic. *Finnisch-Ugrische Forschungen* 44. 23–42.

Kallas, Oskar. 1903. *Kraasna maarahvas*. Helsingi: Soome Kirjanduse Selts.

Kallio, Petri. 2014. The Diversification of Proto-Finnic. In Joonas Ahola, Frog & Clive Tolley (eds.), *Fibula, fabula, fact: The Viking Age in Finland*, 155–170. Helsinki: Suomalaisen Kirjallisuuden Seura.

Kalnača, Andra. 2004. *Morfēmika un morfonoloģija*. Rīga: LU Akadēmiskais apgāds.

Koptjevskaja-Tamm, Maria. 2006. The circle that won't come full: Two Potential Isoglosses in the Circum-Baltic Area. In Yaron Matras, April McMahon & Nigel Vincent (eds.), *Linguistic Areas: Convergence in Historical and Typological Perspective*, 182–226. London: Palgrave Macmillan UK. https://doi.org/10.1057/9780230287617_8.

Koptjevskaja-Tamm, Maria & Bernhard Wälchli. 2001. The Circum-Baltic languages: An areal-typological approach. In Östen Dahl & Maria Koptjevskaja-Tamm (eds.), *The Circum-Baltic languages: typology and contact. Volume 2: Grammar and Typology* (Studies in Language Companion Series 55), 615–750. Amsterdam, Philadelphia: John Benjamins Publishing Company. https://doi.org/10.1075/slcs.55.

Kostjučuk, Larisa Jakovlevna. 2018. Pskovskie govory v ih sinhronno-diahronnyh otnošenijah s sosednimi jazykami (k fiksacij i issledovaniju). *Slavistica Vilnensis*, 62 (2017). 117–138. https://doi.org/10.15388/SlavViln.2017.62.11686.

Laande, Alli & Triin Todesk. 2013. *Mulgi sõnastik*. (Eds.) Karl Pajusalu & Urmas Sutrop. Abja-Paluoja, Tallinn, Tartu: Eesti Keele Sihtasutus, Mulgi Kultuuri Instituut, Eesti Keele Instituut, Tartu Ülikooli eesti ja üldkeeleteaduse instituut. http://www.eki.ee/dict/mulgi/ (3 May, 2021).

Lang, Valter. 2018. *Läänemeresoome tulemised* (Muinasaja teadus 28). Tartu: Tartu Ülikooli Kirjastus.

Laua, Alīse. 1997. *Latviešu literārās valodas fonētika*. 4. Rīga: Zvaigzne ABC.

Lauze, Linda (ed.). 2017. *The Language Situation in Latvia: 2010–2015. A sociolinguistic study*. Rīga: Latvian Language Agency.

Lindström, Liina, Maarja-Liisa Pilvik, Helen Plado. 2021. Variation in negation in Seto. *Studies in Language* 45(3). 557−597. https://doi.org/10.1075/sl.19063.lin.

Lindström, Liina, Mervi Kalmus, Anne-Liis Klaus, Liisi Bakhoff, Karl Pajusalu. 2009. Ainsuse 1. isikule viitamine eesti murretes. *Emakeele Seltsi aastaraamat* 54, 159–185.

Markus, Dace. 2012. Augšzemnieku dialekta dziļās latgaliskās izloksnes Ziemeļaustrumvidzemē (t.s. Malēnijā). In Bonifacas Stundžia (ed.), *Baltistica. VIII Priedas*, 99–110. Vilnius: Vilniaus universitetas. https://doi.org/10.15388/baltistica.0.8.2114.

Markus, Elena, Pärtel Lippus, Karl Pajusalu & Pire Teras. 2013. Three-way opposition of consonant quantity in Finnic and Saamic languages. In Eva Liina Asu & Pärtel Lippus (eds.), *Nordic prosody: proceedings of the 11th conference. Tartu 2012*, 225–234. Frankfurt am Main: Peter Lang Verlag.

Matras, Yaron & Jeanette Sakel. 2007. Investigating the mechanisms of pattern replication in language convergence. *Studies in Language* 31(4). 829–865. https://doi.org/10.1075/sl.31.4.05mat.

Mets, Mari, Anu Haak, Triin Iva, Grethe Juhkason, Mervi Kalmus, Miina Norvik, Karl Pajusalu, Pire Teras, Tuuli Tuisk & Lembit Vaba. 2014. *Lõunaeesti keelesaarte tekstid* (Eesti murded 9). Tallinn, Tartu: Eesti Keele Instituut, Tartu Ülikool.

Metslang, Helle. 2009. Estonian grammar between Finnic and SAE: some comparisons. *STUF. Language Typology and Universals* 62(1/2). 49–71. https://doi.org/10.1524/stuf.2009.0004.

Metslang, Helle & Liina Lindström. 2017. Essive in Estonian. In Casper de Groot (ed.), *Uralic Essive and the Expression of Impermanent State* (Typological Studies in Language 119), 57−90. Amsterdam & Philadelphia: John Benjamins. https://doi.org/10.1075/tsl.119.03met.

Metslang, Helle, Karl Pajusalu & Tiit-Rein Viitso. 2015. Negation in Livonian. In Matti Miestamo, Anne Tamm & Beáta Wagner-Nagy (eds.), *Negation in Uralic languages*, 433–456. Amsterdam: Benjamins. https://doi.org/10.1075/tsl.108.16met.

Miestamo, Matti. 2007. Negation – An Overview of Typological Research. *Language and Linguistics Compass* 1(5). 552–570. https://doi.org/10.1111/j.1749-818X.2007.00026.x.

Miestamo, Matti, Anne Tamm & Beáta Wagner-Nagy. 2015. Negation in Uralic languages – Introduction. In Matti Miestamo, Anne Tamm & Beáta Wagner-Nagy (eds.), *Negation in Uralic languages*, 1–41. Amsterdam: Benjamins. https://doi.org/10.1075/tsl.108.01int.

MuLa = *Corpus of Modern Latgalian texts*. http://www.korpuss.lv/id/MuLa (3 May, 2021).

Nau, Nicole. 2011. *A short grammar of Latgalian* (Languages of the World/Materials 482). München: Lincom Europa.

Nichols, Johanna. 2021. The Origin and Dispersal of Uralic: Distributional Typological View. *The Annual Review of Linguistics* 7. 351–369. https://doi.org/10.1146/annurev-linguistics-011619-030405.

Nurkse, Rein. 1937. *Adjektiiv-atribuudi kongruentsist eesti keeles* (Akadeemilise Emakeele Seltsi toimetised XXX). Tartu: Akadeemilise Emakeele Seltsi kirjastus.

Pajusalu, Karl. 1996. *Multiple linguistic contacts in South Estonian: Variation of verb inflection in Karksi* (Turun yliopiston suomalaisen ja yleisen kielitieteen laitoksen julkaisuja 54). Turku: University of Turku.

Pajusalu, Karl. 2008. Lõunaeesti keel Lätis. In Péter Pomozi (ed.), *Navigare humanum est... Pusztay Janos hatvanadik születésnapjara*, 159–167. Budapest: Finnugor Népek Vilagkongresszusa Magyar Nemzeti Szervezete.

Pajusalu, Karl. 2012. Phonological Innovations of the Southern Finnic Languages. In Riho Grünthal & Petri Kallio (eds.), *Linguistic Map of Prehistoric North Europe* (Mémoires de la Société Finno-Ougrienne 266). Helsinki, 201–224.

Posti, Lauri. 1961. Itämerensuomalaisen verbitaivutuksen kysymyksiä. *Virittäjä* 65(4). 351–366. https://journal.fi/virittaja/article/view/34327 (3 May, 2021).

Prillop, Külli, Karl Pajusalu, Eva Saar, Sven-Erik Soosaar & Tiit-Rein Viitso. 2020. *Eesti keele ajalugu* (Eesti Keele Varamu 6). Tartu: Tartu Ülikooli Kirjastus.

Ross, Kristiina. 1988. *Instruktiiv läänemeresoome keeltes.* Tallinn: Valgus.

Rudzīte, Marta. 2005. *Darbi latviešu dialektoloģijā.* (eds.). L. Leikuma, A. Andronovs. Rīga: LU.

Saar, Eva, Paul Hagu, Inge Käsi, Maeve Leivo, Mari Mets, Hanna Pook & Karl Pajusalu. (in preparation). *Seto-eesti sõnaraamat.* Tartu, Tallinn.

Saar, Eva, Paul Hagu, Inge Käsi, Maeve Leivo, Hanna Pook & Karl Pajusalu. 2020. *Seto eripäraste sõnade sõnaraamat.* (eds.) Eva Saar & Jüri Viikberg. Tartu, Tallinn: EKSA.

Saari, Henn. 2004 (1995). Tuulemaa keel. In Henn Saari, *Keelehääling. Eesti Raadio keeleminutid 1975–1999*, 738–742. Tallinn: Eesti Keele Sihtasutus.

Sakel, Jeanette. 2007. Types of loan: Matter and pattern. In Yaron Matras & Jeanette Sakel (eds.), *Grammatical Borrowing in Cross-Linguistic Perspective*, 15–29. Berlin, New York: Mouton de Gruyter. https://doi.org/10.1515/9783110199192.15.

Seržant, Ilja. (to appear). The Circum-Baltic area. An Overview. In Jan Fellerer & Neil Bermel (eds.), *Oxford Guides to the World's Languages: The Slavonic Languages.*

Setälä, Eemil Nestor. 1953. *Näytteitä liivin kielestä* (Mémoires de La Société Finno-Ougrienne 106). Helsinki: Suomalais-ugrilainen seura.

Sjögren, Andreas Johan & Ferdinand Johann Wiedemann. 1861. *Joh. Andreas Sjögren's Gesammelte Schriften. Band II. Theil I, Joh. Andreas Sjögren's Livische Grammatik nebst Sprachproben.* St. Petersburg.

Stassen, Leon. 2013. Comparative Constructions. In Matthew S. Dryer & Martin Haspelmath (eds.), *The World Atlas of Language Structures Online.* Leipzig: Max Planck Institute for Evolutionary Anthropology. http://wals.info/chapter/121 (2 May, 2021).

Stolz, Thomas. 1991. *Sprachbund im Baltikum? Estnisch und Lettisch im Zentrum einer sprachlichen Konvergenzlandschaft.* (Bochum-Essener Beiträge zur Sprachwandel-forschung 13). Bochum: Brockmeyer.

Stolz, Thomas. 2013. *Competing Comparative Constructions in Europe*. Berlin: Akademie Verlag. https://doi.org/10.1524/9783050064994.

Tanning, Salme. 1961. *Mulgi murdetekstid* (Eesti murded 1). Tallinn: Eesti riiklik kirjastus.

Tanning, Salme. 2004. *Karksi murrak.* Tallinn: Eesti Keele Sihtasutus.

Teras, Pire. 2010. Quantity in Leivu. *Linguistica Uralica* 46(1). 1−16. https://doi.org/10.3176/lu.2010.1.01.

Timberlake, Alan. 2004. *A Reference Grammar of Russian*. Cambridge: Cambridge University Press.

Vaba, Lembit. 1977. *Läti laensõnad eesti keeles.* Tallinn: Valgus.

Vaba, Lembit. 1997. *Uurimusi läti-eesti keelesuhetest.* Tallinn, Tampere: Eesti Keele Instituut, Tampereen Yliopiston suomen kielen ja yleisen kielitieteen laitos.

Vaba, Lembit. 2011. Kuidas läti-eesti keelekontakt on mõjutanud eesti murdekeele grammatikat ja sõnamoodustust. *Emakeele Seltsi aastaraamat* 56(1). 204–246. https://doi.org/10.3176/esa56.11.

VAKK = *Vana kirjakeele korpus* [Corpus of Old Literary Estonian]. https://vakk.ut.ee/ (10 May, 2021). https://doi.org/10.15155/TY.0005.

Viikberg, Jüri. 2020. *Eesti murrete grammatika* (Eesti keele varamu 8). Tartu: Tartu Ülikooli Kirjastus.

Viitso, Tiit-Rein. 2008. *Liivi keel ja läänemeresoome keelemaastikud*. Tartu, Tallinn: Eesti Keele Sihtasutus.

Viitso, Tiit-Rein. 2009. Livonian and Leivu: Shared Innovations and Problems. *Linguistica Uralica* 45(4). 269–282. https://doi.org/10.3176/lu.2009.4.03.

Viitso, Tiit-Rein. 2011. Liivi keele põhijooned. In Renāte Blumberga, Tapio Mäkeläinen & Karl Pajusalu, *Liivlased: Ajalugu, keel ja kultuur*, 203–217. Tallinn: Eesti Keele Sihtasutus.

Viitso, Tiit-Rein & Valts Ernštreits. 2012. *Līvõkīel-ēstikīel-leţkīel sõnārōntõz. Liivi-eesti-läti sõnaraamat. Lībiešu-igauņu-latviešu vārdnīca*. Tartu, Rīga: Tartu Ülikool, Latviešu valodas aģentūra.

Wälchli, Bernhard. 2011. The Circum-Baltic languages. In Bernd Kortmann & Johan van der Auwera (eds.), *The Languages and Linguistics of Europe: A Comprehensive Guide*, 325–340. Berlin, Boston: De Gruyter Mouton. https://doi.org/10.1515/9783110220261.325.

Weber, Tobias. 2021. A linguistic analysis of Heikki Ojansuu's phonograph recordings of Kraasna. *Eesti ja soome-ugri keeleteaduse ajakiri. Journal of Estonian and Finno-Ugric Linguistics* 12(2). 313–341. https://doi.org/10.12697/jeful.2021.12.2.13.

Wiik, Kalevi. 1988. *Viron vokaalisointu*. (Suomi 140). Helsinki: Suomalaisen Kirjallisuuden Seura.

Winkler, Eberhard & Karl Pajusalu. 2016. *Salis-Livisch I. J. A. Sjögrens Manuskript.* (Veröffentlichungen der Societas Uralo-Altaica 88). Wiesbaden: Harrassowitz Verlag.

Winkler, Eberhard & Karl Pajusalu. 2018. *Salis-Livisch II: Grammatik und Wörterverzeichnis*. Mit einem Anhang zu den salis-livischen Sprichwörtern. (Veröffentlichungen der Societas Uralo-Altaica 89). Wiesbaden: Harrassowitz Verlag in Kommission.

Zaharova, K. F. & Orlova, V. G. 2004. *Dialektnoe členenie russkogo jazyka*. Izd. 2. Moskva: Editorial URSS.

Kokkuvõte. Miina Norvik, Uldis Balodis, Valts Ernštreits, Gunta Kļava, Helle Metslang, Karl Pajusalu, Eva Saar: Lõunaeesti keelesaared Kesk-Balti mõjuväljas. Artikkel esitab lõunaeesti keelesaarte – Leivu, Lutsi ja Kraasna – mitme morfosüntaktilise ja fonoloogilise joone võrdleva analüüsi. Uurimuse eesmärgiks on anda ülevaade valitud joonte levikust ja püsivusest ajas ning arutleda nende vormide ja kasutuse üle laiemas areaalses kontekstis. Selleks võetakse arvesse lähimate sugulaskeelte (eesti ja lõunaeesti, Kuramaa ja Salatsi liivi) ja -murrete ning tähtsamate kontaktkeelte (latgali, läti, vene) esinemusi. Analüüsitakse erinevatest allikatest, mh tekstikogudest, sõnaraamatutest ja keelekorpustest pärit ainest. Uurimistulemused toovad esile mitmesuguseid vormiseoseid ja muutuste levikuviise, osutades uuritud keelte ja murrete omavaheliste kontaktide mitmelaadsusele ning sellest tingitud erinevatele keelesüsteemi arengutele.

Märksõnad: morfosüntaks, fonoloogia, keelekontakt, laenamine, keele-areaalid, Kesk-Balti areaal, lõuna-läänemeresoome keeled