

ARTICLE

Do contact languages influence the distribution of prepositions in Estonian dialects?

Mirjam Ruutma

Institute of Estonian and General Linguistics, University of Tartu, Jakobi 2, 51014, Tartu, Estonia
Email for correspondence: mirjam.ruutma@ut.ee

(Received 15 November 2018; revised 06 March 2019; accepted 08 April 2019; first published online 12 July 2019)

Abstract

The origins of prepositional phrase structure in Finnic languages shows little evidence of being contact-induced. However, whether language contact has influenced the structure at a later stage is debatable. The current paper provides new findings on the topic of contact-induced change by comparing the distribution of prepositions in Estonian dialects with the respective contact languages. The purpose is to determine whether the usage frequency of prepositions is higher in areas mainly in contact with prepositional Indo-European languages. The topic is approached from a corpus-based, frequency-driven viewpoint. The results show a small, gradual decrease in the use of prepositions from the northeastern to the western dialect areas. Thus, the uneven but regular distribution of prepositions in Estonian dialects cannot be explained with language contact. This evidence supports the general understanding that adpositions are an unlikely class to be influenced by contact.

Keywords: adpositions; dialects; Estonian; language contacts; prepositions

1. Introduction

In Finno-Ugric languages, an adpositional phrase is generally built up of a complement noun followed by an adposition – a postposition. However, in the Finnic and Sami branches of Finno-Ugric languages, an adpositional phrase has two possible structures: the adposition can either precede or follow the complement noun, i.e. both prepositions and postpositions are used (Tauli 1966, Palmeos 1982, Ehala 1995a, Grünthal 2005, Janda, Antosen & Bals Baal 2014). Several researchers (e.g. Ehala 1995a, 1995b, Erelt & Metslang 1998, Grünthal 2003, 2005) have suggested that the development of this atypical prepositional pattern in an otherwise postpositional language family has been enabled by various language-internal changes. The influences of prepositional languages are considered less significant or irrelevant. This is because in Finno-Ugric context, there is little evidence of borrowing adpositions structurally or as lexical units even in the most intensive examples of contact (e.g. Latvian and Livonian, or Russian and Erzya) (Ehala 1995a, 1995b, Grünthal 2003, 2005).

Furthermore, Lass (1997:190) has argued that adpositions are not likely to be borrowed. However, existing but marginal grammatical constructions may acquire a higher frequency of use due to the influence of equivalent structures used in a contact language (e.g. Koptjevskaja-Tamm & Wälchli 2001:627, Heine & Kuteva 2005:4–50). In addition, structural barriers are often outweighed by the intensity of contact and cultural pressure (Thomason & Kaufman 1988, Matras 2007, 2009).

A convincing proof of contact languages having an effect on adpositional system is provided in Janda et al.'s (2014) study of the adpositional phrase structure in North Sami dialects. The western dialects of this Finno-Ugric language are in contact with Norwegian and Swedish. Both languages mainly use prepositions. The major contact language in the Sami eastern area is the predominantly postpositional Finno-Ugric language, Finnish. Janda et al. (2014) have shown how these contacts have an effect on the use of adpositions that can be used as either prepositions or postpositions (hereinafter 'ambipositions'). Dialects in contact with Norwegian and Swedish tend to use ambipositions as prepositions, whereas dialects in contact with Finnish are more likely to use these same ambipositions as postpositions (Janda et al. 2014). The influence of contact language is also present in Romani varieties spoken in Finland and Turkey, where otherwise prepositional Romani has adapted the postpositional structure, and in exclusively postpositional Turkish, where in the dialects spoken in North Macedonia, the Macedonian prepositional phrase structure is replicated (Matras 2009:257).

The contacts between Estonian and neighbouring languages have been somewhat similar to that of North Sami – over the course of several millennia, Estonian has had contact with languages predominantly using either prepositions (Indo-European languages) or postpositions (the Finno-Ugric languages). Thus, although prepositions in Finnic languages show evidence of having developed due to internal changes, rather than as a consequence of external pressure, the question remains of whether the use of prepositions and postpositions in different parts of Estonia might have been influenced by language contact. Since grammatical constructions may acquire a higher frequency of use as a result of equivalent structures used in a contact language (as in Koptjevskaja-Tamm & Wälchli 2001:627; Heine & Kuteva 2005; Janda et al. 2014), we can expect Estonian dialects mainly in contact with prepositional Indo-European languages to use prepositions more frequently.

Preference regarding the placement of adpositions in Estonian dialects has been the focus of research previously – e.g. Ruutma et al. (2016) have examined the position of five common ambipositions in Estonian dialects. They found some dialectal variation in speakers' preferences regarding the placement of ambipositions (Ruutma et al. 2016:102–109). However, the focus of that study was not on the possible influence of language contact. Thus, similarly to the tendencies found by Janda et al. (2014) in North Sami dialects, the dialectal differences in the position of Estonian ambipositions in Ruutma et al. (2016) could be contact-related. In addition, bilingual areas or areas with long-term contacts with predominantly prepositional Indo-European languages could have an unusually high use of prepositions (i.e. also the use of those prepositions that are not ambipositional) as compared to areas with strong Finno-Ugric contacts.

In this study, I adopt a corpus-based, frequency-driven approach to examine whether the Estonian adpositional system has been prone to contact-induced

changes. I use frequency data obtained from the Corpus of Estonian Dialects (<http://www.murre.ut.ee/murdekorpus>) in order to determine whether prepositions are similarly distributed in all Estonian dialects. Furthermore, I examine whether any possible differences in usage frequencies could, at least in part, be explained by language contact. I expect to find that dialects in contact with Indo-European languages use prepositions more than those dialects mostly in contact with Finno-Ugric languages, similar to Janda et al.'s (2014) findings. However, it must be noted that language contacts across Estonia are not geographically as clearly divided into Indo-European and Finno-Ugric contact areas as in the North Sami case. This means that the contact areas may overlap. Moreover, the intensity, scale and duration of contacts differ throughout the area. This may make it difficult to assess the relative importance of a contact on the distribution of prepositions.

In the next section, I present some background on prepositions and postpositions in Estonian. Then, in Section 3, I examine the languages Estonian has been in contact with and describe the adpositional systems of these languages. In Section 4, I explain the methodology and data used in the article, and in Section 5 I present the results of the study.

2. Prepositions vs. postpositions in Estonian

Estonian, similar to other Finnic languages, makes use of both prepositions and postpositions. The main difference between the two structures, besides the difference in word order, is that postpositions govern nouns in genitive [N[GEN]+POST] and prepositions in partitive [PREP+N[PART]] (EKG I 1995:37–38). As lexemes, postpositions are more common than prepositions in Standard Estonian. For example, out of all 266 adpositions included in the online version of the Estonian explanatory dictionary (EED), 193 are exclusively postpositions, 29 exclusively prepositions, and 19 ambipositions. Thus, nearly 81% of adpositions can occur as postpositions while only around 19% can occur as prepositions. In addition to postpositions being lexically more common than prepositions, the usage frequencies of postpositions are also considerably higher than prepositions in written texts. For example, in the Morphologically Disambiguated Corpus of Estonian (MDC), which consists of 513,000 words from various types of texts, the frequency of prepositions and postpositions is 3,038 to 10,116, making the proportion of prepositions out of all adpositions around 23%.

The preference for postpositions in Estonian is to be expected as, firstly, the prepositional phrase structure is relatively recent compared to the postpositional structure. It has been suggested that it emerged after Estonian diverged from the Proto-Finnic language (e.g. Palmeos 1982, Grünthal 2005). The postpositional phrase structure is not considered to be typical to the Finno-Ugric languages either – the general viewpoint in Finno-Ugric linguistics (as in e.g. Rätsep 1979, Palmeos 1982:69–72, Grünthal 2003) is that the adpositional system has developed during the Proto-Finnic era and the functions of postpositions (and prepositions) used to be filled by alternative means such as case endings and/or constructions of case endings and independent nouns.

An often presented example (e.g. in EKG I 1995:38, EKK 2007:195) illustrating the development of an adposition from a noun in Finnic languages is *maja ees*

'in front of the house'. Initially, the postposition *ees* 'in front of' was an inflectional form of a noun in local case **edesnä* 'the fore area of an entity'. *Maja* 'house' was the attributing noun in genitive **majan*. The head noun **edesnä* gained new meanings, then acquired a new grammatical function and became a postposition. As the attributing noun is in genitive, the original phrase **majan edesnä* 'the front area of the house' forms a possessive construction. Grünthal (2003:36) has suggested that in Finno-Ugric languages, the possessive construction consists of a genitive attribute which generally precedes its head noun. Thus, because the adpositional phrase originally followed the structure of the possessive construction, postpositions are more natural to Estonian and therefore more commonly used than prepositions.

The third reason for the preference of postpositions in Estonian lies in the relation the Estonian postpositions most often express – spatiality (Palmeos 1982, Grünthal 2003, Ruutma 2016). The process of grammaticalisation is associated with a tendency toward metaphorical abstraction, where abstract phenomena are expressed by concrete concepts. Spatial expressions, often transparent in their origin, are semantically less abstract than others. Therefore spatiality has been proposed to be the base from which other more abstract expressions, such as temporal ones, may derive (see e.g. Heine, Claudi & Hünemeyer 1991, Traugott & Heine 1991:4–5, Hopper & Traugott 1994, Haspelmath 1997, Hagège 2010:162). Traugott & Heine (1991:4) note that the change towards abstraction in grammaticalisation is, nonetheless, only a pre-theoretical notion. However, in the case of Estonian, adpositions do seem to undergo a metaphorical change. Returning to the phrase *maja ees* 'in front of the house', the relation it was originally used for was spatial. The adposition has, however, gained multiple abstract meanings, e.g. denoting events that are happening or about to happen (1) or when indicating entities in relation to which something is being done (2) (examples from the EED).

(1) Riik seisa-b uue majanduskriis-i ees
state stand-3SG new.GEN economic.crisis-GEN in.front
 'The state faces a new economic crisis.'

(2) Pea-n täitma oma kohus-t isamaa-Ø ees
have-1SG fulfil mine.GEN duty-PAR fatherland.GEN in.front
 'I have to fulfil my duty towards the fatherland.'

Thus, it could be hypothesised that as *ees* – an adposition with a clear spatial background – has acquired abstract meanings; spatiality could indeed have been the original relation expressed in the Estonian adpositional system. As the prepositional structure is a more recent innovation in Estonian than postpositions, it might have emerged after the expansion of meanings. Thus, some prepositions may have 'skipped' the earlier stage of expressing spatial relations and started directly to denote more complex relations. Therefore, postpositions could be hypothesised to be more associated with spatial relations and prepositions with abstract relations.

The association of spatiality with postpositions can also be drawn from Palmeos' (1982) handbook of Estonian adpositions, where around 200 adpositions are listed

along with their expressed relations: the most common expressed relation of postpositions in Estonian is spatiality. This kind of association between the position of an adposition and the relation it expresses is also mentioned in studies by e.g. Grünthal (2005), Huumo (2013:320–321) and Janda et al. (2014:100), where it has been noted that abstract, temporal relations typically associate with prepositions, and concrete, spatial relations with postpositions. This notion, however, concerns the preference regarding the placement of individual ambipositions and does not yield all adpositions. Nevertheless, in a more cognitive-semantic sense, the more concrete relations, such as spatiality, are considered to be primary to the language user and therefore, again, more frequent (e.g. in the Estonian context, Veismann 2009). Thus it is natural that postpositions are more frequent than prepositions in Estonian, as they convey spatial relations – the primary and most frequently expressed relations.

Considering that postpositions are more frequent in usage and diachronically, typologically and semantically more common in Estonian than prepositions, prepositions could be seen as a rather unnatural grammatical structure in Estonian. The development of the structure may consequently be expected to be contact-induced. However, there is little evidence of such influence. Grünthal (2005) has suggested that the development of prepositions is caused mostly by language-internal changes, such as SOV word order becoming SVO and shifts in the grammatical relations of adverbs. The role of language contacts is seen to be minimal, for Finnic languages have few prepositions as loanwords. Nevertheless, the suggestion that the frequency of marginal structures may increase due to the pressures of language contact (see Koptjevskaja-Tamm & Wälchli 2001:627; Heine & Kuteva 2005) leads to the possibility that after initial development, the prepositional structure itself might have become more frequent due to foreign influences. A few examples of contact language having an effect on the pattern of the word order of the adpositional phrase are presented by e.g. Janda et al. (2014) and Matras (2009).

Janda et al. (2014) have shown that North Sami dialects in long-term contact with Swedish and Norwegian prefer using ambipositions as prepositions, whereas dialects in contact with Finnish are found to favour postpositions. A non-Finno-Ugric example presented by Matras (2009:257–258) can be found in otherwise prepositional Romani, where dialects in contact with postpositional Finnish and Turkish have adapted the postpositional phrase structure. Another example described by Matras (2009:258) is Macedonian Turkish, where after the model of prepositional Macedonian, the otherwise postpositional Turkish also uses prepositions.

Interestingly, a rise in the use of prepositions has been detected in 20th century Standard Estonian texts (Hint 1990:1401, Ehala 1994:177). Again, the reasons behind the change are much debated (e.g. in Hint 1990, Ehala 1994, Erelt & Metslang 1998) and the general view is that language contact has not played a crucial role in the rise of prepositions. However, the studies concentrate on the recent, 19th–20th-century, Russian and German influences on Standard Estonian. The old local contacts of Estonian dialects have not been in focus. The duration of the old contacts, as will be established in Section 3, are considerably longer than those of Russian and German. Thus the old contacts might, nonetheless, have had an effect on the frequency of prepositions in Estonian dialects. The present study sets out to

Table 1. Thomason & Kaufman's (1988:74–76) borrowing scale

Casual contact	(1) content words
↓	(2) conjunctions, adverbial particles, minor phonological, syntactic, and lexical semantic features
	(3) adpositions, derivational affixes, pronouns, numerals, phonemes, minor syntax features (such as borrowed postp in prep languages (and vice versa))
	(4) major structural changes that cause little typological change, distinctive features in phonology, word order, inflectional morphology
Very intense contact	(5) significant typological disruption, phonetic changes, word structure rules

cast light on the question of whether language contact might have an effect on the adpositional system, increasingly favouring prepositions.

3. Contact languages of Estonian dialects

Traditionally, grammar was considered to be resistant to foreign influences. In recent times, it is acknowledged that also grammar is open to contact-induced change. The openness to influence is seen to depend on the duration and intensity (i.e. status, attitudes) of the contact, and on the typological barriers the grammatical unit carries (Thomason & Kaufman 1988; Matras 2007, 2009:154). Thomason & Kaufman (1988:74–76) have presented a scale of the likelihood of borrowing (Table 1), where grammatical units are listed according to the pressure by a contact language needed for a change.

The pressure needed to influence the structure of an adpositional phrase situates roughly in the middle of the scale, in the 3rd category, meaning that the nature of contact needed for a change has to be rather intense. Thomason & Kaufman (1988) do not explain further, what extra-linguistic features does the 3rd category hold within, but considering that in Estonian dialects, there can be found influences of the local contacts on all linguistic levels (as will be established in this Section), the effect of contact languages on the Estonian adpositional system is likely. In addition, close, long-lasting contacts have been shown to exert pressure on already existing grammatical units, causing them to become primary and more frequent, to take on a new meaning or to find new contexts of use (Heine & Kuteva 2005:1–46).

On the whole, the linguistic landscape of Estonia has been remarkably diverse for thousands of years. On the one hand, Estonian can be divided into two major dialect groups, which are considered to originate from different Proto-Finnic dialect. On the other hand, Estonian has had long-term contacts with neighbouring languages.

In the present paper I follow the dialect division used in the Corpus of Estonian Dialects (CED). The dialect classification used in CED separates three major dialect groups: North, Northeastern-Coastal and South Estonian. The major groups are built up of smaller varieties: the Northern group contains Eastern, Insular, Mid,



Figure 1. Dialect classification according to Corpus of Estonian Dialects. Estonian dialects are divided into three major groups: Northern, Northeastern-Coastal and Southern. The Northern dialect group consists of Eastern, Insular, Mid, and Western dialects; Northeastern-Coastal covers Coastal and Northeastern dialects; and the Southern group comprises of Mulgi, Tartu, Võru, and Seto dialects.

and Western dialects; the Northeastern-Coastal group covers Coastal and Northeastern dialects; and the Southern group consists of the Mulgi, Tartu, Võru, and Seto dialects (see Figure 1). These 10 dialects can be further divided into hundreds of smaller sub-dialects, which roughly correspond to the division of old Estonian parishes. In traditional dialect classifications, instead of three major groups, only a northern and a southern group are distinguished as the phonology, morphology and lexis of the northern groups (North and Northeastern-Coastal) and South Estonian differ the most.

The main local contact languages of Estonian-speaking areas are from three branches of Indo-European languages – Baltic, Germanic and Slavic, and from the western branch of the Finno-Ugric languages – Votic, Finnish, Ingrian, and Livonian. Thus, on the grounds of close, long-term contacts having an effect on grammar (as in Thomanson & Kaufman 1988, Heine & Kuteva 2005, Matras 2009), the diverse language contacts may be mirrored in the usage frequencies of prepositions. Dialects in contact with prepositional Indo-European languages would be expected to use prepositions more than those mostly in contact with postpositional languages. Next I will give an overview of the contact languages of Estonian and briefly introduce the adpositional system of each language. I will not further introduce the adpositional systems of the Finnic languages, as they share a similar background with the Estonian system. All the Estonian language contacts are summarised in Figure 2.



Figure 2. The main contact languages of Estonian. Contacts with Indo-European languages (dashed lines) are most apparent in the southern border areas, the north and north-west coast and Seto. Finno-Ugric contacts (solid lines) are the strongest in the north-east of Estonia.

3.1 Latvian

Contacts between Estonian and pre-Latvian tribes extend back several millennia. As Estonian and Latvian are geographically adjacent, Latvian influences on language, and even on culture, become most apparent on the southern border areas of Estonia in the Võru, Mulgi and Western dialects (Vaba 1997, 2011). Influences have been greatest in the most southwestern part of Võru dialect area (around Hargla parish) – it is notable that around a quarter of all the borrowed Latvian loanwords in Estonian dialects are from Hargla and its surroundings. Numerous Latvian loanwords can also be found in the Mulgi dialect and in the southernmost parishes, Häädemeste and Saarde, in the Western dialect varieties. The border regions of the Mulgi and Western dialect had mixed Estonian-Latvian settlement and even belonged administratively to the neighbouring Latvian parishes, where, for example, church services were held in Latvian. Bilingualism was usual – Estonian and Latvian were used in daily communication by both language communities (Ariste 1981:157, Vaba 1997:479–483). Beside the strong contacts by land, Estonians from Saaremaa and Latvians from Courland were connected by sea, across the Irbe Strait. The contacts were, however, the most intense with Livonian villages and thus the Latvian influences were most likely transmitted through Livonian to the Saaremaa varieties of the Insular dialect (Vaba 1997:478).

As an Indo-European language, the main adpositional phrase structure in Latvian is prepositional. In Hewson & Bubeník (2006), only two Latvian postpositions are mentioned: *dēļ* 'because of' and *apakš* 'under'. Fennell & Gelsen (1980) have also mentioned in their grammar of modern Latvian a pair of ambipositions, *starpā* and *starp* 'between', which differ in the case they assign to the governed term: *starpā* governs genitive and *starp* accusative.

3.2 Swedish

Scandinavian contacts with Estonian are among the oldest, and permanent Swedish settlement was formed in the Western and Insular dialect areas during the 12th–13th century (Viitso 1990:140, Juhkam 1998:28). The settlement reached its largest extent in the 15th century: Swedish villages covered the area from Estonian's southernmost island Ruhnu to the village of Mahu in Coastal dialect area (in Viru-Nigula parish). Thus, at its largest, Swedish settlement covered the coastal areas of the Insular, Western, Mid, and the Coastal dialect. The settlement was at first relatively isolated. The first Estonian-Swedish mixed communities appeared after the 18th century. The long-termed contact and local settlement came to an end in 1944, when Estonian-Swedes fled the Soviet occupation (Juhkam 1992:394–399, 1998:12–14, 27–28). However, despite 700 years of intense contacts, total bilingualism never took root in the area – Swedes were a minority community in Estonia and learned to speak Estonian (Ariste 1965a:72; Juhkam 1998). However, e.g. Juhkam (1998) has shown how both languages have clear signs of mutual influence in the vocabulary, phonetics and syntax. In addition to the linguistic influences of long-term Swedish settlement in western Estonia, Björklöf (2012) has found that a small number of loanwords in the Coastal dialect originate from the easternmost Swedish dialects in Finland (Eastern Nyland dialects).

In Swedish, as well as other Indo-European languages, prepositional phrase structure is predominant. Postpositions occur in Swedish only in certain idiomatic expressions, e.g. *året om/runt* 'whole year round, through the year', *hela natten igenom* 'all night through', *oss emellan* 'between you and me'. Some Swedish prepositions, when emphasised, occur at the end of the clause, e.g. *på* 'on, at, in' and *för* 'for'. Prepositions are postposed also in questions, exclamations, infinitival and relative clauses. In addition, Swedish makes use of so-called circumpositions – adpositions that surround their complement, e.g. *för ... sedan* 'ago', *på ... vägnar* 'on behalf of', *i ... ställe* 'instead of' (Holmes & Hinchliffe 1994:359–459). In the case of Swedish–Estonian contacts, Juhkam (1998) has noted that an increase in prepositional phrases is evident in both Estonian dialects and Estonian Swedish dialects in the islands and North-Western areas.

3.3 Russian

The clearest signs of Russian influences can be found in the areas bordering Russia by land or by Lake Peipus – that is in Seto, Northeastern, Eastern and the northern areas of the Tartu dialect (Must 2000:581). The contacts in the Seto dialect area have been the longest: contacts between Seto and East Slavic tribes started around the middle of the first millennium, when Slavic tribes arrived in the direct neighbourhood of the Finnic people. Contacts intensified when the Seto area went under the rule of Pskov, and the Setos were converted to Orthodox Christianity (Must 2000: 7–9). Contacts with Russian tribes have also been long-term in the Northeastern dialect. There has even been Russian settlement in the southernest parish of Northeastern dialect area (in Iisaku parish), and both Estonian and Russian were used equally as late as in 1965, when it was common in daily communication to mix the languages. For that reason, Russian influence on phonetics, lexicon, morphology and syntax are characteristic to the dialect in Iisaku (Must 1995:107–110).

There has also been Russian settlement on the shores of Peipus in the Eastern dialect area since the end of the 16th century. By the 18th century, the Russian population increased with the arrival of Old Believers and the Estonian-Russian mixed settlement turned into monolingual Russian villages (Must 2000:8).

Russian adpositional phrases are virtually all prepositional. Hewson & Bubeník (2006:183) emphasise that in Old Slavic, there were only two postpositions: *дѣля* (*dělja*) and *ради* (*radi*), both meaning ‘for the sake of’. Both of the postpositions also exist in modern Russian – *для* (*dlja*) only as a preposition and *ради* (*radi*) as a prepositions or postposition. Additionally, Russian adpositions such as *вопреки* (*vopreki*) ‘contrary to’, *назло* (*nazlo*) ‘to spite (someone)’, *наперекор* (*naperekor*) ‘counter to’, *вслед* (*vsled*) ‘after (someone)’, *навстрѣчу* (*navstreču*) ‘towards’, *спустя* (*spustja*) ‘after’, may be used as postpositions and are therefore ambipositions (Podobryaev 2007:4). Interestingly, none of the grammars (e.g. Borras & Christian 1963, Timberlake 2004, Bailyn 2011) mention there to be other than prepositions in Russian, which may indicate that the postpositional use is, nonetheless, connected to fixed phrases or is of marginal use.

3.4 Finnic languages

Northeastern Estonia has been a meeting point of multiple Finnic languages. The strongest contact in the area has been with Votic: in addition to Votes being a neighbouring nation to Estonians, in the 6th–9th and the 10th–14th centuries there was much migration of Votes to Kodavere parish in the Eastern dialect area (Ariste 1965b:107, Pall 1982:247). Votic influences are also visible in the Northeastern dialect but the linguistic landscape of that area, especially in most eastern parish bordering Russian, is more complex: contacts have been strong also with Ingrian and Finnish (Ariste 1965b, 1981:52–58, Must 1987, Toikka 2003). In addition, the contacts with Finnish have been close in the Coastal and Northeast dialect areas, where various loanwords can be found. Estonians and Finns were involved in a tight barter relationship, known as *sõbrakaubandus* or *seprakauppa* ‘friend trade’, trading fish and cereal grain (Must 1987, Grünthal 1998, Björklöf 2012:14–16). Records of bartering date back to the 19th century although contacts are likely much older – long fishing trips (*rändpüük*) to the Finnish coast were common already in the 16th century (Melander 1937:63, Luts 1960:149, Björklöf 2012:11–16). Must (1987) has suggested that due to contacts with the Finns, the morphophonological characteristics of the region’s dialect remained archaic and untouched by developments taking place elsewhere in Estonia. However, other Finnic contacts have been considered to transmit Russian influences, especially to the North-Eastern areas (Must 2000:577).

While Votic and Ingrian are considered to be transmitters of Russian features, Livonian has transmitted many Latvian features to the south-western variants of Estonian by the Latvian border and to the Saaremaa varieties of Insular dialect (Vaba 1997:478). Similarly to Coastal and Finnish connections, there used to be a close trading relationship between Estonians and Livonians. Furthermore, people from Saaremaa used to work in Livonian farms during summer periods (Ariste 1981:79).

In addition to the local contact languages presented here, since the 13th century, the language of the ruling class in Estonia was Low and High German. Many studies show (e.g. Metslang 1994, Hennoste 1997, Habicht 2000, Laanekask 2004, Raag

2008, Viikberg 2014) that its effect on Standard Estonian is visible at all linguistic levels, and the influence on Old Literary Estonian is particularly noteworthy. However, the direct contacts of Estonian dialects with German and the indirect contacts of dialects with German via Standard Estonian were strongest in the cities. The corpus material used in this paper only covers rural and non-mobile informants born in the late 1800s (before broadcasting mediums using standardised language were widespread), thus it is difficult to assess the strength of German influences on separate dialects. For this reason, I will not focus much on German contacts in the present article.

4. Approach and data

Geographical variation has always been the subject of interest in dialectological work. However, as Szmrecsanyi (2014:3) has concluded, traditional, usually survey- or atlas-based, work has dealt with subjects such as ‘in dialect x, the word y is typically pronounced z’. The attention is on a particular linguistic unit and how it is or is not used in a given dialect. Alternative forms of a linguistic unit in the same dialect are put aside, even though they may provide interesting knowledge. In addition, survey or atlas-based work generally does not yield a realistic linguistic signal, as it relies on dialect speakers’ (usually one-word) answers to questionnaires (Szmrecsanyi & Wolk 2011:564–565; Wolk 2013:3–4, Szmrecsanyi 2014:3). For example, Saareste’s (1955) *Väike eesti murdeatlas* [Estonian dialects atlas] is traditional in essence: it presents which variant of a certain linguistic phenomenon is used in a particular area – the possible variation within a seemingly unified region is not in the focus of interest.

In contrast to the traditional approaches, corpus-based dialectology is frequency-driven. The linguistic variation compared between different dialects or within a dialect is generalised quantitatively (Szmrecsanyi & Wolk 2011:564–565, Wolk 2013:3–4, Szmrecsanyi 2014:3). Szmrecsanyi (2014:3) has summarised the way a typical, corpus-based study examines dialects as follows: ‘in dialect x, feature y is twice as frequent in actual speech than in dialect z’. The text corpora provide a source for natural language data, distinct from elicitation questionnaires. Most importantly, corpus-based dialectology allows researchers to uncover usage frequencies – how often the linguistic variants occur in a particular area (Szmrecsanyi & Wolk 2011:564–565, Szmrecsanyi 2012:4, 2014:2–4, Wolk 2013:3–4). The approach can be thus described as focusing on dialects ‘exactly as they are’. Furthermore, in the light of contact languages, Koptjevskaja-Tamm & Wälchli (2001:627) suggest that frequency data is an important factor which may mirror contact: if a linguistic unit occurs in a particular dialect in much greater quantity than outside that dialect, it may be a sign of a close relationship with a contact language. This view is shared by Heine & Kuteva (2005:4–50), as they state that language contacts may result in an increase in the frequency of use of a linguistic phenomenon.

The corpus-based, frequency-driven approach is also the basis of the present research. I examine the distribution of prepositions and postpositions in Estonian dialects, and investigate whether the variation found could be explained by local language contacts. I have collected data from the Corpus of Estonian Dialects (CED <http://www.murre.ut.ee/mkweb>) and analysed it quantitatively.

The corpus contains transcriptions of unstructured interviews from all the ten Estonian dialect areas (see Figure 1 above). The interviews were mainly recorded in the 1960s–1970s, are often monologues and the topics of the recordings are similar. The non-mobile older rural informants talk about their everyday life, e.g. work, family, past events, customs, and so on. Thus, the dialect material is contextually rather homogeneous and synchronic. However, the research data is not fully representative of the Estonian dialects, as the corpus material is geographically unevenly spread (e.g. Insular, West and Mid dialects are overrepresented). In the present study, the dialect maps present proportions of prepositions and postpositions in each dialect, making the data comparable. The tests used to determine statistical importance do not require the corpus to be balanced.

At the time of the retrieval of the data (May 2016) from the morphologically annotated part of CED, the corpus consisted of 834,311 words. In total, the CED data included 14,785 observations of adpositions. In order to exclude erroneous annotations, I reviewed the extracted data manually. The final dataset contains 13,973 utterances containing 31 distinct prepositions and 133 postpositions.

The data in this study is analysed quantitatively. Firstly, I compiled a map of the spread of adpositions in dialects (see Figure 3 below) using R scripts (R Development Core Team 2017) developed by Uihoaed (2016). The map in Figure 3 represents the distribution of prepositions and postpositions in each dialect. The darker a dialect is on the map, the larger the proportion of prepositions to postpositions.

In Section 5.1, the statistical significance of the results is evaluated with Pearson's chi-squared test for independence. The test is applied to categorical data to determine whether the differences in frequency counts are statistically significant and not purely coincidental. In the present paper, it is used to examine whether a dialect conditions how often prepositions or postpositions occur. The chi-squared test does not reveal what kind of relationship holds between individual variables. Pearson residuals (represented in Table 3 below), however, help to define the exact dialects where the correlation between the dialect and the position of an adposition is the most evident. In addition, the residuals tell whether the correlation is negative or positive – whether the postpositions or prepositions occur more or less than expected. In short, Pearson residuals show how much the observed frequency distribution – in this study, the frequencies of prepositions and postpositions in 10 dialect areas – differs from the so-called expected values, that is from the values that would be expected if the counts were at chance. The more the residuals deviate from 0, the greater the effect. The strength of the association between variables is determined with Cramer's V effect size (presented in Section 5.1). Cramer's V result always falls between 0 and 1 – the closer the result is to 1, the stronger the correlation (for more detail about Pearson's chi-square, residuals and Cramer's V effect size, see Agresti 2007, Baayen 2008, Gries 2013).

5. Geographical distribution of prepositions and postpositions

The numbers of prepositions and postpositions in Estonian dialects are presented in Table 2 and mapped out in Figure 3. With reference to Table 2 and the map, it is evident that postpositions are preferred to prepositions in all dialect areas. Even in

Table 2. Distribution of postpositions and prepositions. Presented are the absolute frequencies of both postpositions and prepositions of each dialect and the proportions of prepositions

Dialects	Postpositions	Prepositions	Prepositions (%)
Eastern	694	85	10.9
Mid	1950	268	12.1
Northeastern	597	80	11.8
Western	2609	279	9.7
Mulgi	745	74	9.0
Coastal	737	124	14.4
Insular	2451	198	7.5
Seto	606	81	11.8
Tartu	1046	116	10.0
Võru	1120	113	9.2
Total	12555	1418	10.1

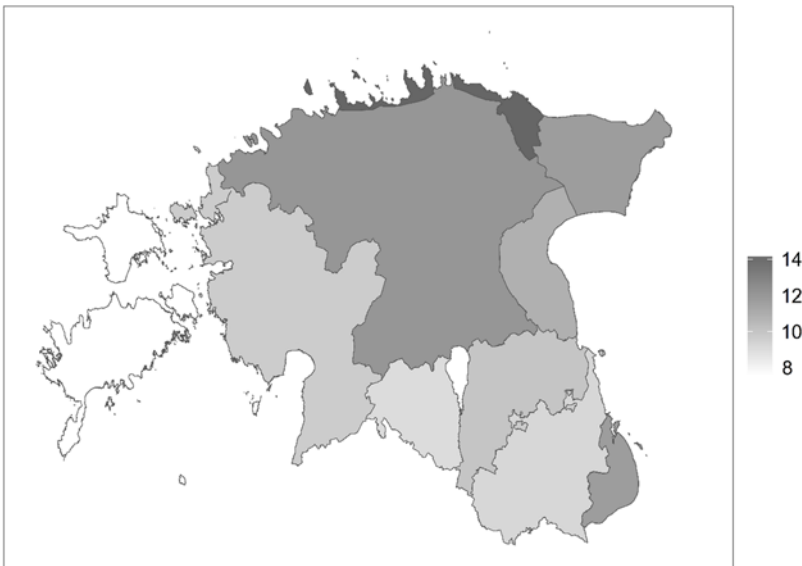


Figure 3. Distribution map of prepositions and postpositions. Darker areas on the map show dialects with larger proportions of prepositions to postpositions. The highest rate of prepositions is found in the Coastal dialect, around 14%. The lowest rate of prepositions is in the Insular dialect, around 8%.

the Coastal dialect, where the use of prepositional word order is the highest compared to the other dialects, the rate of postpositions is around 86%. The preference for postpositions, however, is predictable. As mentioned in Section 2, in Standard Estonian, postpositions are lexically more common than prepositions, as nearly 81%

of all adpositions occur as postpositions. In addition, the postpositional phrase structure is generally considered to be older than the prepositional, and more common not only lexically but also in usage (Palmeos 1982:69–72; Grünthal 2003; Veismann 2009). However, the use of prepositions in dialects is low compared to Standard Estonian, around 10%. Based on Morphologically Disambiguated Corpus of Estonian (MDC), the proportion of prepositions in Standard Estonian is 23% – a rate significantly higher than that of any Estonian dialect (as seen in Table 2). A higher proportion of prepositions in Standard Estonian is probably caused by German influence as its effect on Standard Estonian is attested to be visible on all linguistic levels (e.g. Metslang 1994, 2009, Habicht 2000, Laanekask 2004). In addition, this result could indicate that, at least in the case of adpositions, German influence did not reach dialect speakers as strongly as it reached Standard Estonian users.

Even though postpositions are, as expected, more common in all dialect areas, the varying rates of prepositions across dialects may still reveal the influence of contact languages. I have suggested that dialects mostly in contact with Indo-European languages may use prepositions more than those having closest contacts with Finno-Ugric languages. Thus, the highest rates of prepositions should be found in Western, Insular, Mulgi and Seto areas, as the contact with Indo-European languages has been the most intense in these regions. However, the proportion of prepositions to postpositions in the Western and Insular dialects is the lowest, around 8–10%. In contrast, the highest proportion of prepositions, 14%, is, unexpectedly, in an area mostly in contact with postpositional Finno-Ugric languages – in the Coastal dialect. This fails to support my hypothesis that long-term contacts with Indo-European languages would be reflected in a higher frequency of prepositions. Overall, the use of prepositions seems to decrease gradually from north-east to south-west, with the Coastal, Northeastern, Mid and Seto dialects standing out with a higher use of prepositions than the others. A similar eastern-western distinction, as opposed to a traditional northern-southern dialect division, has also been detected in studies of verbal constructions by Uiboed (2013) and Lindström et al. (2017).

Although the high density of prepositions in the Coastal dialect does not seem to be caused by foreign influences, in the case of Seto, the higher use of prepositions (~12%) may be a result of Russian influence. Russian contacts have also been present in the Northeastern dialect, yet the contacts in that area took place considerably later than those in Seto. Contacts with Finno-Ugric Votic and Ingrian have been stronger and lasted longer in the area. However, both of the languages are considered to transmit Russian influences to Estonian (Must 2000:577). Thus, the higher frequency of prepositions in the Northeastern dialect (~12%) may be mirroring both direct and indirect Russian influences.

However, the use of prepositions is highest in the Coastal dialect (~14%), which cannot be explained by language contact, as the Coastal dialect has had the closest and most long-term contacts with predominantly postpositional Finnish. The higher use of prepositions could be caused by stronger influence of Standard Estonian on the dialect. However, it would be expected that, if a higher frequency is connected to Standard Estonian influences, the use of prepositions would be highest in the Mid dialect as, out of all Estonian dialects, it is the nearest to Standard Estonian. The use of prepositions in the Mid dialect is nonetheless less frequent than

that of the Coastal dialect (12–14%). Contacts with Swedes living across the Gulf of Finland could not explain the rise in the frequency of prepositions in the Coastal dialect either, because contact with Swedish was not predominant. In addition, if Swedish had had a strong impact on the adpositional system, the same effect should be seen in the western dialects of Estonia, where clear signs of influence in vocabulary, phonetics and syntax are found (e.g. Juhkam 1998). Nevertheless, the use of prepositions is lowest in the western parts of Estonia.

On the whole, with the exception of the Seto and Northeastern dialects, it seems that language contact does not play an important role in the frequency density of prepositions. In fact, although the frequencies seem to follow a clear pattern, the results are the opposite of those hypothesised – the southwestern areas that have had Swedish and Latvian contacts use prepositions the least, and the northeastern areas with the strongest Finnic contacts use prepositions the most. Thus, this should be seen as strong evidence that the adpositional system is not easily influenced by contact languages, at least in one Finnic language. The differences between the frequencies are probably not contact-induced, but associated rather with factors which are not in the scope of the present paper.

5.1 Statistical importance of the distribution

Although the frequency map (Figure 3) shows that the usage of prepositions has a clear pattern of decreasing gradually from north-east to south-west, differences in the proportions of prepositions across dialects are small. For example, the maximal difference between proportions is 6%: while in the Coastal dialect the rate of prepositions to postpositions is around 14%, the proportion of prepositions in the Insular dialect, with the lowest use of prepositions, is 8%. This raises the question of whether the differences in the frequency counts are in fact arbitrary. To calculate the significance of the results, I carried out Pearson's chi-squared test, examined the Pearson residuals, and evaluated the strength of the association with Cramer's V.

The result of Pearson's chi-squared test for independence indicates that the differences in frequency counts of the ten Estonian dialects are not at chance and the effect is highly significant ($\chi^2 = 54.763$, $df = 9$, $p < .001$). However, the chi-squared test only evaluates the statistical significance and does not give an evaluation of answer whether the effect is significant in all of the dialect areas. Pearson residuals seen in Table 3 help to find an answer to this question. The significance of residuals near 0 is small while those smaller than -3.84 or greater than 3.84 are considered particularly noteworthy (Gries 2013:369). Nearly all of the residuals in the present test are around 0, meaning that the effect is not particularly significant. However, the Mid, Coastal and Insular dialects are notable, as they show a greater deviation from 0 than other dialects. The residuals of postpositions in Mid and Coastal dialects are slightly smaller than in the other dialects. In the Insular dialect, the residuals are slightly larger. The same dialects also have a greater significance rate in the use of prepositions: the values of prepositions in the Mid and especially the Coastal dialect are considerably larger than expected (i.e. the values are larger than expected if the counts were at chance levels), and in the Insular dialect the values are remarkably smaller than expected. The small use of prepositions

Table 3. Pearson residuals

Dialects	Postpositions	Prepositions
Eastern	-0.22	0.67
Mid	-0.96	2.86
Northeastern	-0.45	1.36
Western	0.28	-0.82
Mulgi	0.34	-1.00
Coastal	-1.31	3.92
Insular	1.45	-4.32
Seto	-0.45	1.35
Tartu	0.06	-0.18
Võru	0.36	-1.08

in Insular and greater use in the Coastal dialect can be also seen in Table 2 and on the distribution map in Figure 3.

Despite the chi-squared test and the Pearson residuals show statistical significance and, in the case of the Mid, Coastal and Insular dialects, a noteworthy deviation from the expected values, the strength of the association calculated with Cramer's *V* test indicates a very weak association (Cramer's $V = .06$). A comparison between Coastal and Insular dialect – the two dialects that deviated the most from the expected – indicates somewhat larger but still low association (Cramer's $V = .1$). Thus, the correlation between dialects and the position of adpositions is statistically significant and not coincidental, but still weak.

6. Conclusions

The possible effect of Indo-European contact on the adpositional system in Finnic languages has been much discussed in the literature. In the present study, I approached the topic from a corpus-based, frequency-driven perspective. The results show that the usage frequencies of prepositions differ throughout Estonia: there occurs a gradual decrease in the density of prepositions from the northeast to the southwest of Estonia. As the strongest influences with prepositional languages are by the southern border and the northwest coast of Estonia, the final results do not support the prediction that local contacts influence the frequencies of prepositions. The southern and western areas, which have had Latvian, Russian or Swedish contacts, use prepositions the least, and the northeastern areas with strongest Finnic contacts, use prepositions the most. It may thus be concluded that, at least in the case of a Finnic language, the results give support for the view that the adpositional system is not easily influenced by contact languages. A high frequency count of prepositions in Seto and Northeastern dialect might still reflect Russian influences.

The results bring up the question of what lies behind the increase in the number of prepositions in northeastern Estonia. One possibility is that the rise in frequency shows signs of adopting the Standard Estonian adpositional system. It would be

expected, however, that instead of a higher frequency of prepositions in northeastern Estonia, the signs of Standard Estonian would be the most detectable in Mid dialect. It is likely that the increase may be caused by e.g. functional differences (the differences in the use of concrete vs. abstract relations), or there may be certain frequently used fixed idiomatic expressions in the dialects. Whatever the cause may be, the differences in the numbers of prepositions occurring in the dialects under investigation do not appear to be due to local foreign contacts; rather, they appear to be due to factors which were not in the scope of the present paper.

Acknowledgements. This study has been supported by the Centre of Excellence in Estonian Studies (European Union, European Regional Development Fund). I am grateful to my supervisors for their valuable guidance and to the three anonymous *NJL* referees for their useful comments.

References

- Agresti, Alan.** 2007. *An Introduction to Categorical Data Analysis*, 2nd edn. New York: John Wiley & Sons.
- Ariste, Paul.** 1965a. Ülevaade Hiiu murretest [An overview about Hiiu dialect]. In Arnold Kask (ed.), *Sõna sõna kõrvale. Paul Ariste teaduslikust tegevusest* [A word next to a word: About the academic work of Paul Ariste], 63–74. Tallinn: Eesti NSV teaduste akadeemia Emakeele Selts.
- Ariste, Paul.** 1965b. Vadja keele ja eesti murrete suhetest [On the relations between Estonian dialects and Votic]. In Arnold Kask (ed.), *Sõna sõna kõrvale. Paul Ariste teaduslikust tegevusest* [A word next to a word: About the academic work of Paul Ariste], 106–110. Tallinn: Eesti NSV teaduste akadeemia Emakeele Selts.
- Ariste, Paul.** 1981. *Keelekontaktid: Eesti keele kontakte teiste keeltega* [Language Contacts: Contacts Between Estonian and other Languages]. Tallinn: Valgus.
- Baayen, R. Harald.** 2008. *Analyzing Linguistic Data: A Practical Introduction to Statistics Using R*. Cambridge: Cambridge University Press.
- Bailyn, John Frederick.** 2011. *The Syntax of Russian*. Cambridge: Cambridge University Press.
- Björklöf, Sofia.** 2012. Viron rantamurteen länsiryhmän sanaston alkuperä suomalaislainojen valossa [The origins of the vocabulary in Eastern Coastal dialect in the light of Finnish loanwords]. Master's thesis, University of Helsinki.
- Borras, F. M. & Reginal Frank Christian.** 1963. *Russian Syntax: Aspects of Modern Russian Syntax and Vocabulary*. Oxford: Clarendon Press.
- EED = Eesti keele seletav sõnaraamat [Estonian Explanatory Dictionary].** 2009. <https://www.eki.ee/dict/ekss/> (accessed 10 April 2014).
- Ehala, Martin.** 1994. Russian influence and the change in progress in the Estonian adpositional system. *Linguistica Uralica* 3, 177–199.
- Ehala, Martin.** 1995a. Explaining the bipositional head–complement order in adpositional systems. In Bohumil Palek, František Dane, Osamu Fujimura & Jiri V. Neustupný (eds.), *Proceedings of LP'94: Item Order in Natural Languages*, 325–339. Prague: Charles University Press.
- Ehala, Martin.** 1995b. *Self-organisation and Language Change: The Theory of Linguistic Bifurcations*. Ph.D. dissertation, University of Cambridge.
- EKG I = Erelt, Mati, Reet Kasik, Helle Metslang, Henno Rajandi, Kristiina Ross, Henn Saari, Kaja Tael & Silvi Vare** (eds.). 1995. *Eesti keele grammatika I* [Estonian Reference Grammar I]. Tallinn: Eesti Teaduste Akadeemia Eesti Keele Instituut.
- EKK = Erelt, Mati, Tiit Erelt & Kristiina Ross** (eds.). 2007. *Eesti keele käsiraamat* [Handbook of Estonian], 2nd edn. Tallinn: Eesti Keele Sihtasutus.
- Erelt, Mati & Helle Metslang.** 1998. Oma või võõras? [Ours or foreign?]. *Keel ja Kirjandus* 10, 657–668.
- Fennell, Trevor G. & Henry Gelsen.** 1980. *A Grammar of Modern Latvian*. Berlin & Boston: De Gruyter Mouton.
- Gries, Stefan Th.** 2013. *Statistics for Linguistics with R: A Practical Introduction*. Berlin & Boston, MA: De Gruyter Mouton.

- Grünthal, Riho.** 1998. Suomalais-virolaisia meriyhteyksiä ja seprakaupan kielellisiä sivutuotteita [Finnish–Estonian contacts by sea and linguistic by-products of the barter relationship]. In Riho Grünthal & Johanna Laakso (eds.), *Oekeeta asijoo: Commentationes Fenno-Ugricae in honorem Seppo Suhonen sexagenarii* 16.V.1998 [Perspectives on the Fenno-Ugric studies in the honour of Seppo Suhonen's 60th anniversary 16 May 1998] (Suomalais-ugrilaisen Seuran Touimituksia 228), 106–121. Helsinki: Finno-Ugrian Society.
- Grünthal, Riho.** 2003. *Finnic Adpositions and Cases in Change*. Helsinki: Finno-Ugrian Society.
- Grünthal, Riho.** 2005. Miksi itämerensuomessa on prepositioita? [Why do the Finnic languages have prepositions?]. *Virittäjä* 1, 28–51.
- Habicht, Külli.** 2000. Grammaticalization of adpositions in Old Literary Estonian. In Mati Erelt (ed.), *Estonian: Typological Studies IV*, 19–58. Tartu: Tartu University Press.
- Hagège, Claude.** 2010. *Adpositions*. Oxford: Oxford University Press.
- Haspelmath, Martin.** 1997. *From Space to Time: Temporal Adverbials in the World's Languages* (LINCOM Studies in Theoretical Linguistics 3). München/Newcastle: LINCOM Europa.
- Heine, Bernd, Ulrike Claudi & Friederike Hünemeyer.** 1991. *Grammaticalization: A Conceptual Framework*. Chicago, IL: University of Chicago Press.
- Heine, Bernd & Tania Kuteva.** 2005. *Language Contact and Grammatical Change* (Cambridge Approaches to Language Contact). Cambridge: Cambridge University Press.
- Hennoste, Tiit.** 1997. Eesti keele sotsiopersioodid. Üldpilt [The social periods of Estonian: An overview]. In Mati Erelt, Meeli Sedrik & Ellen Uuspõld (eds.), *Pühendusteos Huno Rätsepale, 28.12.1997* [Festschrift for Huno Rätsep, 28 December 1997] (Tartu Ülikooli Eesti Keele Õppetooli Toimetised 7), 45–66. Tartu: Tartu Ülikool.
- Hewson, John & Vít Bubeník.** 2006. *From Case to Adposition: The Development of Configurational Syntax in Indo-European Languages* (Amsterdam Studies in the Theory and History of Linguistic Science 280). Amsterdam/Philadelphia: John Benjamins.
- Hint, Mati.** 1990. Vene keele mõjud eesti keelele [Russian influences on Estonian]. *Akadeemia* 2, 1383–1404.
- Holmes, Philip & Ian Hinchliffe.** 1994. *Swedish: A Comprehensive Grammar* (Routledge Grammars). London/New York: Routledge.
- Hopper, Paul J. & Elizabeth Closs Traugott.** 1994. *Grammaticalization*. Cambridge: Cambridge University Press.
- Huomo, Tuomas.** 2013. Many ways of moving along a path: What distinguishes prepositional and postpositional uses of Finnish path adpositions? *Lingua* 133, 319–335.
- Janda, Laura A., Lene Antonsen & Berit Anne Bals Baal.** 2014. A radial category profiling analysis of North Sámi ambipositions. *High Desert Linguistics Society Proceedings* 10, 91–102.
- Juhkam, Evi.** 1992. Eestirootslastest [About the Estonian Swedes]. *Keel ja Kirjandus* 7, 394–402.
- Juhkam, Evi.** 1998. Eesti-rootsi murdekontaktid [Estonian–Swedish contacts]. Master's thesis, University of Tallinn.
- Kask, Arnold** (ed.). 1965. *Sõna sõna kõrvale. Paul Ariste teaduslikust tegevusest* [A word next to a word: About the academic work of Paul Ariste]. Tallinn: Eesti NSV teaduste akadeemia Emakeele Selts.
- Koptjevskaja-Tamm, Maria & Bernhard Wälchli.** 2001. The Circum-Baltic languages: An areal-typological approach. In Östen Dahl & Maria Koptjevskaja-Tamm (eds.), *Circum-Baltic Languages*, vol. 2: *Grammar and Typology* (Studies in Language Companion Series 54–55), 615–750. Amsterdam/Philadelphia: John Benjamins.
- Laanekask, Heli.** 2004. *Eesti kirjakeele kujunemine ja kujundamine 16.–19. sajandil* [Formation and shaping of the Estonian literary language in the 16th–19th centuries]. Ph.D. dissertation, University of Tartu.
- Lass, Roger.** 1997. *Historical Linguistics and Language Change*. Cambridge: Cambridge University Press.
- Lindström, Liina, Maarja-Liisa Pilvik, Mirjam Ruutma & Kristel Uibo.** 2017. On the use of perfect and pluperfect in Estonian dialects: Frequency and language contacts. In Sofia Björklöf & Santra Jantunen (eds.), *Plurilingual Finnic: Change of Finnic Languages in a Multilingual Environment*, 51–89. Helsinki: Finno-Ugrian Society.
- Luts, Arved.** 1960. Rändpüügist Eesti merekalastuses [About the fishing trips in Estonian sea fishing]. *Etnograafia Muuseumi Aastaraamat* 17, 129–168. Tartu: Eesti NSV Teaduste Akadeemia.

- Matras, Yaron.** 2007. The borrowability of grammatical categories. In Yaron Matras & Jeanette Sakel (eds.), *Grammatical borrowing in cross-linguistic perspective*, 31–74. Berlin: Mouton de Gruyter.
- Matras, Yaron.** 2009. *Language Contact*. Cambridge: Cambridge University Press.
- MDC = Morphologically Disambiguated Corpus of Estonian.** 2017. <http://www.cl.ut.ee/korpused/morkorpus/index.php?lang=en> (accessed 4 January 2018).
- Melander, K. R.** 1937. Virolaista asutusta Suomen etelärannikolla ja sen saarilla [Estonian settlement in the south coast and islands of Finland]. *Kotiseutu* 1937, 63–74.
- Metslang, Helle.** 1994. Grammatikaliseerumise eesti ja soome keeles [On the grammaticalization in Estonian and Finnish]. *Lähivertailuja* 7, 9–32.
- Metslang, Helle.** 2009. Estonian grammar between Finnic and SAE: Some comparisons. *Language Typology and Universals* 62, 49–71.
- Must, Mari.** 1987. *Kirderannikumurre. Häälikuline ja grammatiline ülevaade* [Northeastern Coastal dialects: An overview of phonetics and grammar]. Tallinn: Valgus.
- Must, Mari.** 1995. Vene-eeesti kakskeelsus Kirde-Eestis [Russian–Estonian bilingualism in the northeastern Estonia]. In Harri Moora & Lembit Jaanits (eds.), *Slaavi-läänemeresoome suhete ajaloo* [On the history of Slavic–Finnic relations], 107–132. Tallinn: Kirjastus “Eesti raamat”.
- Must, Mari.** 2000. *Vene laensõnad eesti murretes* [Russian loanwords in Estonian]. Tallinn: Eesti Keele Sihtasutus.
- Pall, Valdek.** 1982. Idamurde vahekorras naabermurretega ja vadja keelega [On the relation of Eastern dialect with neighbouring dialects and Votic]. *Keel ja Kirjandus* 5, 246–251.
- Palmeos, Paula.** 1982. *Eesti keele grammatika: II osa, neljas vihik, kaassõna* [Estonian reference grammar, Part II, Fourth Book: Adposition], 3rd edn. Tartu: Tartu Riiklik Ülikool.
- Podobryaev, Alexander.** 2007. “Postposition stranding” and related phenomena in Russian. In Gerhild Zybatow, Uwe Junghanns, Denisa Lenertová & Petr Biskup (eds.), *Studies in Formal Slavic Phonology, Morphology, Syntax, Semantics and Information Structure*, 197–208. Frankfurt am Main: Peter Lang.
- R Development Core Team.** 2017. *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing. <http://www.R-project.org/> (accessed 3 June 2019).
- Raag, Raimo.** 2008. *Talurahva keelest riigikeeleks* [From a peasantry language to a national language]. Tartu: Atlex.
- Rätsep, Huno.** 1979. *Eesti keele ajalooline morfoloogia II* [Historical morphology of Estonian]. Tartu: Tartu Riiklik Ülikool.
- Ruutma, Mirjam.** 2016. *Kaassõnad eesti murretes* [Adpositions in Estonian dialects]. Master’s thesis, University of Tartu.
- Ruutma, Mirjam, Aki-Juhani Kyröläinen, Maarja-Liisa Pilvik & Kristel Uiboed.** 2016. Ambipositsioonide morfosüntaktilise varieerumise kirjeldusi kvantitatiivsete profiilide abil [Descriptions of the morphosyntactic variation of ambipositions by means of quantitative profiles]. *Keel ja Kirjandus* 2, 92–113.
- Saareste, Andrus.** 1955. *Väike eesti murdeatlas: Atlas Des Parlers Estoniens* [A small atlas of Estonian dialects]. Tartu: Eesti Kirjanduse Selts.
- Szmrecsanyi, Benedikt.** 2012. *Grammatical Variation in British English Dialects: A Study in Corpus-based Dialectometry* (Studies in English Language). Cambridge: Cambridge University Press.
- Szmrecsanyi, Benedikt.** 2014. Methods and objectives in contemporary dialectology. In Ilja A. Seržant & Björn Wiemer (eds.), *Contemporary Approaches to Dialectology: The Area of North, Northwest Russian and Belarusian Vernaculars/Современные Методы В Дialeктологии. Ареал Северных, Северо-Западных Русских И Белорусских Говоров*, 81–92. Bergen: Department of Foreign Languages, University of Bergen.
- Szmrecsanyi, Benedikt & Christoph Benedikt Sebastian Wolk.** 2011. Holistic corpus-based dialectology. *Brazilian Journal of Applied Linguistics/Revista Brasileira de Linguística Aplicada* 11(2), 561–592.
- Tauli, Valter.** 1966. *The Structural Tendencies in Uralic Languages* (Indiana University Publications: Uralic and Altaic Series 17). The Hague: Mouton.
- Thomason, Sarah Grey & Terrence Kaufman.** 1988. *Language Contact, Creolization, and Genetic Linguistics*. Berkeley, CA: University of California Press.
- Timberlake, Alan.** 2004. *A Reference Grammar of Russian*. Cambridge: Cambridge University Press.
- Toikka, Kaja.** 2003. *Kirderannikumurde idaosa murde- ja keelekontaktid* [Dialect and language contacts of the eastern part of Northeastern Coastal dialect]. Tallinn: TPÜ Kirjastus.

- Traugott, Elizabeth Closs & Bernd Heine.** 1991. Introduction. In Elizabeth Closs Traugott & Bernd Heine (eds.), *Approaches to Grammaticalization*, vol. 1: *Focus on Theoretical and Methodological Issues* (Typological Studies in Language 19.1), 1–15. Amsterdam: John Benjamins.
- Uiboed, Kristel.** 2013. *Verbiühendid eesti murretes* [Verb constructions in Estonian dialects]. Ph.D. dissertation, University of Tartu.
- Uiboed, Kristel.** 2016. *Spatial Visualization with R*. Spatial-visualization-with-r 1.0. <https://doi.org/10.5281/zenodo.51473>
- Vaba, Lembit.** 1997. *Uurimusi läti-eesti keelesuhetest* [Studies about the Latvian–Estonian relations]. 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 [How Latvian–Estonian contacts have influenced the grammar and the word formation of Estonian dialects]. *Emakeele Seltsi aastaraamat* 56, 204–246.
- Veismann, Ann.** 2009. *Eesti keele kaas- ja määrsõnade semantika võimalusi* [Semantics of Estonian adpositions and adverbs]. Ph.D. dissertation, University of Tartu.
- Viikberg, Jüri.** 2014. Alamsaksa laensõnadest nende tulekuajas [Arrival of Low German loanwords in Estonian]. *Keel ja Kirjandus* 10, 749–762.
- Viitso, Tiit-Rein.** 1990. On the earliest Finnic and Balto–Slavic contacts. *Itämerensuomalaiset kontaktit: Itämerensuomalainen symposium 7. kansainvälisessä fenno-ugristikongressissa Debrecenissä 28.8–1.9.1990* [Finnic contacts: Finnic Symposium on the 7th International Finno-Ugric Conference in Debrecen, 28 August – 1 September 1990]. Helsinki: Valtion Painatuskeskus.
- Wolk, Christoph Benedikt Sebastian.** 2013. *Integrating aggregational and probabilistic approaches to dialectology and language variation*. Ph.D. dissertation, University of Freiburg.