Acta Linguistica Academica 73 (2026) 2, 286–326
DOI: 10.1556/2062.2026.01011

1828
AKADÉMIAI KIADÓ

Between Ingrian and Votic: More marginal than core consonants in Lower Luga Ingrian (Finnic)

NATALIA KUZNETSOVA*

Università Cattolica del Sacro Cuore, Milan, Italy

Received: February 1, 2025 • Revised manuscript received: March 22, 2026 • Accepted: March 31, 2026
Published online: June 9, 2026

ABSTRACT

The paper provides a detailed fieldwork-based comparative analysis of the consonant systems of three cognate Finnic varieties: Soikkola Ingrian, Lower Luga Ingrian, and Lower Luga Votic, conducted along unified principles. Consonants are divided into core (appearing in native non-expressive lexicon), robust marginal, and unstable marginal. The consonant inventories of Soikkola Ingrian (North Finnic) and Lower Luga Votic (South Finnic) are different as regards contrasts based on duration and voicing, s/ʃ distinctions, affricates, and palatalised consonants. Still, both varieties contain more core phonemes than marginal phonemes. Lower Luga Ingrian is a convergent variety formed in contact between the Votes and the Ingrians. Its consonant system is closer to that of Lower Luga Votic, but the inventory of core consonants is simplified in comparison with both Soikkola Ingrian and Votic. As an outcome, in Lower Luga Ingrian, marginal consonants outnumber core consonants, which is cross-linguistically rare. However, the system appears unstable and transitory.

KEYWORDS

Ingrian, Votic, marginal consonants, language contact, sound change, phonological rarities

1. INTRODUCTION

Lower Luga Ingrian is one of the four dialects of the vanishing Ingrian language, along with Soikkola and now extinct Hevaha and Oredež (cf. Laanest 1966a; Kuznetsova, Markus & Muslimov 2015). The paper discusses the consonant system of Lower Luga Ingrian, which was formed through contact between two closely related languages, Ingrian (North Finnic) and

* Corresponding author. E-mail: natalia.kuznetsova@unicatt.it

Votic (South Finnic), and is highly different from the systems of other Ingrian dialects. As a result of this contact, the system of core consonants was simplified while the inventory of marginal consonants was significantly widened. By analysing comparable field data from Soikkola Ingrian, Lower Luga Votic, and Lower Luga Ingrian, I show that, in contrast to both donor languages, marginal consonants outnumber core consonants in convergent Lower Luga Ingrian. However, this system appears unstable and transitory.

The structure of the paper is as follows. In the remainder of this introduction, I discuss the nature of Lower Luga Ingrian and its place within the neighbouring Finnic varieties (1.1), describe the data and methods of the study (1.2), and define and classify the types of marginal phonemes analysed in the paper (1.3). Section 2 outlines the consonant inventory of Soikkola Ingrian, one of the core Ingrian dialects, while Section 3 provides a similar description for the Lower Luga Votic dialect. In Section 4, the consonant system of Lower Luga Ingrian is discussed in comparison with both Soikkola Ingrian and Lower Luga Votic. Section 5 is dedicated to a quantitative comparison of core and marginal consonants between the three varieties and to the theoretical and typological implications of this comparison. Section 6 concludes the study.

1.1. Historical background on Lower Luga Ingrian and adjacent varieties

Lower Luga variety is the most idiosyncratic dialect of Ingrian, because it has been heavily influenced by several closely related varieties within the *Sprachbund* in the Lower Luga area: Votic, Ingrian Finnish, and Estonian (Laanest 1966a, 146, 150–151; Muslimov 2005, 5, 13; Kuznetsova 2015). In particular, the dialect is considered a convergent variety between Ingrian and Votic (Rožanskij 2010; Rozhanskiy & Zhivlov 2019), and this is strongly reflected in its consonant system, as discussed below.

Figure 1 shows the main villages where Ingrian used to be spoken in the Lower Luga (lau) area and in the adjacent Soikkola (soi) area. Ingrian villages are marked in grey, and multicoloured circles indicate villages where Ingrian has been traditionally spoken alongside other languages. The massive shift of the entire local Finnic population into Russian after WWII (cf. Kuznetsova, Markus & Muslimov 2015) is not reflected in the map but is mentioned in the paper in relation to loanwords. Lower Luga Votic, often called Vaipooli Votic (Ariste 1967, 78), is marked in black, Ingrian Finnish in white, Russian is marked with dots, and a specific mixed Ingrian-Votic variety spoken in the Kukkuzi village and a couple of neighbouring farmsteads (Posti 1980; Suhonen 1985; Markus & Rozhanskiy 2012; Muslimov 2020) is marked with stripes. The sizes of the sectors in the multicoloured circles roughly represent the share of each ethnic group in the village, if such information were available.[1] The map provides a general idea of a significantly higher proportion of mixed Ingrian/Votic and Ingrian/Finnish villages in the Lower Luga area than on the Soikkola peninsula (especially in its western part, where my field data come from).

[1]The map was created by Yury Koryakov based on maps in Kuznecova & Sidorkevič (2012) and the author's indications. Note that villages where Ingrian population was not known to be present (i.e. purely Ingrian Finnish, Votic, Estonian, or Russian villages) are not shown on this map. Equal sector sizes indicate either similar proportions of different ethnic groups or a lack of precise information about these proportions. The ratio of 1/4 to 3/4 illustrates a situation in which one population is smaller than the other but still forms a distinct group. The ratio of 1/8 to 7/8 illustrates a scenario where only a few members of a group reside in a village mostly inhabited by another group.

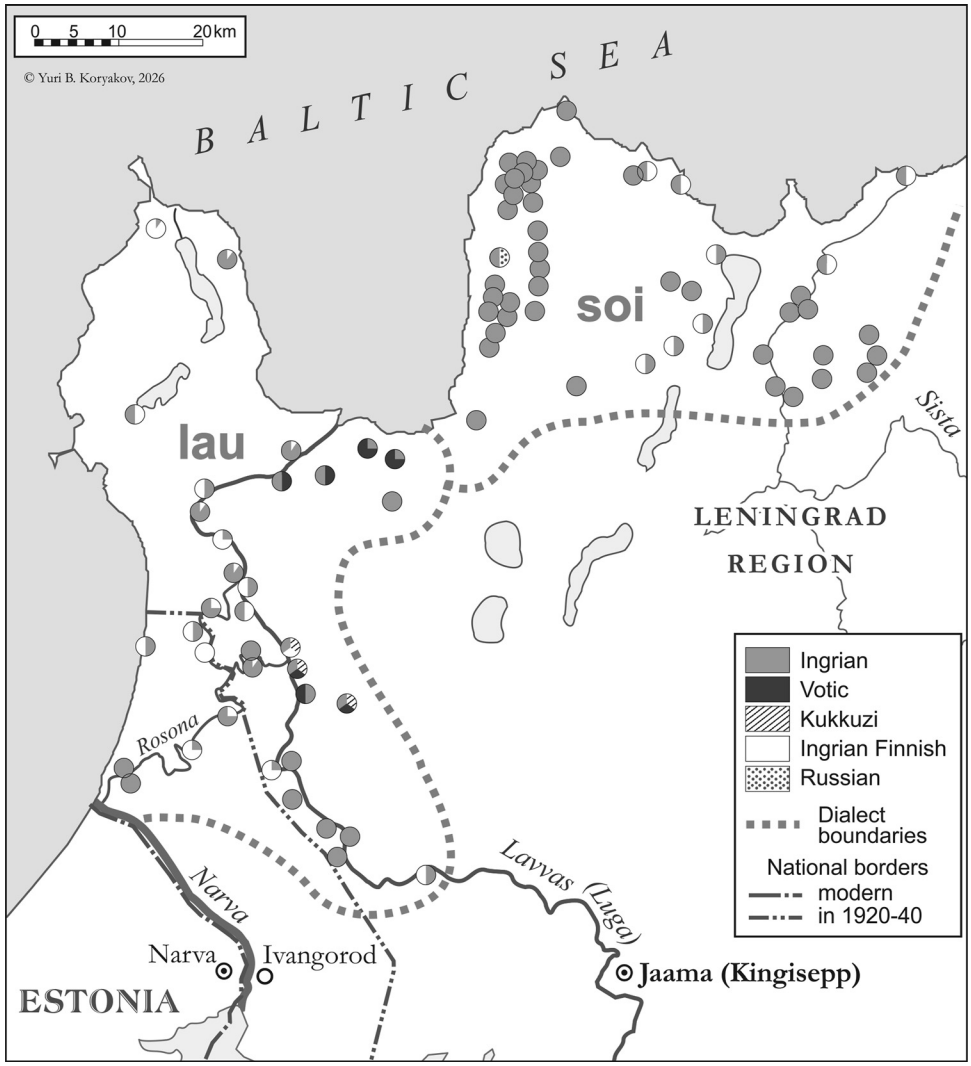

Figure 1. Ingrian-speaking villages in the Soikkola (*soi*) and Lower Luga (*lau*) areas

Historically, the Votic population is the most autochthonous in Western Ingria (it was also genetically closer to the local Slavic population than Ingrians and Ingrian Finns in the study by Agdzhoyan et al. 2024). Votes are mentioned under their modern name in Russian chronicles in the middle of the 9th century AD. It has been hypothesised that Votes are direct descendants of the easternmost part of Finnic tribes, represented by the archaeological culture of the so-called Tarand graves, who populated a territory with the centre in modern Estonia already from the first centuries AD (Ariste 1947; R'abinin 1997, 21; Juškova & Kulešov 2011; Hvoš'inskaja 2023; Stas'uk 2023).

Both Ingrians and Ingrian Finns are migratory groups of more recent origin. From the end of the 1st millennium AD (but especially from the 12th to the 13th century), Ingrians branched off from the southern Korela tribes, who lived in the south of the Karelian Isthmus, and began spreading southward. They moved further south along the Oredež river and west along the southern coast of the Gulf of Finland, reaching the westernmost border of their habitat no later than the 17th century (Laanest 1966a, 9–10; R'abinin 1997, 65–66).

Moving west along the coast, Ingrians were increasingly mixing with Votes, and the westmost Lower Luga Ingrian dialect contains the highest number of Votic elements, especially in its southern part. In general, the Votic area apparently used to be larger than attested on the ethnographic map by Köppen (1849), including its extension further to the west and south than attested on that map (see e.g. Ariste 1947, 46). Votic substrate in south Lower Luga Ingrian is traced both in linguistic (Muslimov 2005, 2020) and in populational genetic (Kuznetsova et al. 2022) data. In particular, a linguistic division of the Lower Luga Ingrian area into two major areas, north-east and south-west (Kuznecova 2009, 22), finds parallels in the genetic division of modern Lower Luga Ingrians into more "Ingrian-like" and more "Votic-like".

Historical records show that Ingrian was a more prestigious language than Votic, and an active shift of Votic speakers into Ingrian has been documented since at least the beginning of the 20th century (Kuznetsova, Markus & Muslimov 2015, 154). The structure of the mixed Kukkuzi variety at the border between Ingrian and Votic (cf. Figure 1) confirms this, because it presents a mixture of mostly Votic grammar and mostly Ingrian lexicon (Markus & Rozhanskiy 2012). Kukkuzi phonology is rather part of the Lower Luga Ingrian than of the Votic phonological system (Kuznecova 2009; Markus & Rozhanskiy 2012). Muslimov (2020) reconstructs the rise of the Kukkuzi variety as, first, the formation of Lower Luga Ingrian under Votic influence and, second, as a contact between Lower Luga Ingrian and Votic in the Kukkuzi village. Both the original formation of Lower Luga Ingrian and its subsequent secondary mixing with Votic in Kukkuzi apparently followed the same main pattern of Votes shifting into the Ingrian language.

Ingrian Finns appeared in Ingria later, after the Stolbovo treaty of 1617, when this territory became part of Lutheran Sweden. They were relocated here by Swedish authorities from several Finnish parishes to re-populate the territory after the emigration of the Orthodox population (Laanest 1966a, 11–12; Suni 2003, 469–471). Ingrian Finnish dialects generally influenced Lower Luga Ingrian less than the substrate Votic varieties did. In phonology, the Lower Luga Ingrian vowel system was partly affected by Ingrian Finnish (in the varieties on the left bank of the Luga river), but its consonant system was not (Kuznecova 2009, 292–293).

Estonian influence is also limited — Estonian affected south-west Lower Luga Ingrian varieties along the Rosona river the most, because this territory was part of the Estonian Republic in 1920–1940 (see Figure 1). Finnish and Estonian impact on Lower Luga Ingrian is not discussed below, because the consonant system was not influenced.

All the aforementioned Finnic languages of Ingria also contain the heritage of long-lasting contacts between Uralic and Indo-European languages at different chronological strata: Uralic with Indo-Iranian, Finno-Permic with Balto-Slavic, Finnic with Proto-Slavic and Proto-Germanic, and later between some of their daughter languages (cf. e.g. Laakso 2022, 240–241, 252–253; Holopainen 2025; Kallio, Metsäranta & Honkola 2025 and references therein). Many of the marginal phonemes discussed below occur only or predominantly in different types of loanwords, which is addressed in each particular case separately.

1.2. Data and methods used in the study

Most of the data used in the study were collected by the author and colleagues during fieldwork in Lower Luga, Soikkola Ingrian, and Lower Luga Votic villages. Ingrian data used here, were collected in 2006–2008 in the field trips headed by Fedor Rozhanskiy and in total included about 200 h of speech samples of 42 Soikkola Ingrian and 22 Lower Luga Ingrian speakers (see description in Kuznecova 2009, 16). In particular, the author compiled a specific questionnaire regarding marginal consonants based on published resources on Ingrian (most notably Nirvi 1971 and Laanest 1966a) and recorded answers from 21 speakers of Soikkola Ingrian (born between 1924 and 1949) and 11 speakers of Lower Luga Ingrian (born between 1921 and 1943). Some answers were detailed, while others included just a few words (detailed data transcripts and original audio files will appear here: https://osf.io/dxv8t/).

Additional sources on Ingrian included audio recordings on Lower Luga Ingrian from the personal archive by Mehmed Muslimov collected in the beginning of the 2000s (about 35 h from 11 speakers) and archive data from the 1960s (about six hours from nine speakers; courtesy of KOTUS — Institute for the languages of Finland, Helsinki).

All the aforementioned Ingrian data were used in the phonological description of Ingrian by Kuznecova (2009) and are systematically compared in the present paper with the Votic data.

Lower Luga Votic data were collected by the author in 2010 from two speakers residing in Luutsa (or Luuditsa; Rus. *Лу́жицы / Lužicy*), a village combining two earlier settlements, Luutsa (composed of Lower Luutsa and Middle Luutsa) and Liivčülä (or Liivačülä, Rus. *Пески́ / Peski*), and three speakers from Jõgõperä (Rus. *Крако́лье / Krakol'je*), born between 1928 and 1938, in their residences in the Kingisepp district of the St. Petersburg region (Russian Federation). A specific questionnaire on Votic marginal phonemes was recorded, among other things. The questionnaire was based on the data collected by the author on Ingrian marginal phonemes earlier, as well as on published Votic data (most notably, Tsvetkov 1995 and Markus & Rožanskij 2011).

Therefore, the Ingrian and the Votic data are highly comparable, and are transcribed below according to the unified principles. Forms in italic represent phonological transcriptions, while forms in brackets ("[]") show phonetic transcriptions. If not explicitly mentioned otherwise, the stress is always word-initial in all the Finnic material throughout the article. IPA is used in phonetic and phonological transcriptions, except for the vowels *ü* [y], *ä* [æ], *ö* [ø], and *õ* [ɤ]. Examples cited from the literature are either re-transcribed according to these principles or cited in original transcription in double quotation marks (""). All examples from the literature which do not occur in my own field data are referenced.

Lower Luga Votic and Lower Luga Ingrian varieties present several distinct stages of shortening of unstressed vowels in certain prosodic positions. Original long vowels can be realised from long to lengthened to short. Original short vowels can maintain their quality or undergo qualitative reduction, devoicing, and elision (see details in Kuznecova 2012, 2016; Kuznetsova & Verkhodanova 2019). In general, vowel reduction is the weakest in north Lower Luga Ingrian, stronger in central Lower Luga Ingrian and in Lower Luga Votic (with a more advanced stage in Jõgõperä, as compared to Luutsa-Liivčülä), and the strongest in south Lower Luga Ingrian. Different degrees of vowel reduction condition variable status for palatalised consonants before the front vowels that are reduced and elided. The phonetic degree of consonant palatalisation also varies across Lower Luga Ingrian varieties.

Therefore, individual Lower Luga Ingrian and Votic varieties have partially different phonologies as regards reduced vowels and palatalised consonants. These vowels and

consonants are not entirely standardised in phonological transcription and are marked the way they appear in particular tokens. Palatalisation before back vowels or word-finally is always marked. Secondary palatalisation of consonants before non-elided front vowels is not marked in phonological transcription (its degree is variable across speakers and tokens). Primary palatalisation before front vowels, when the primary place of articulation changes (cf. Kochetov 2011; Krämer & Urek 2016), is marked for *tʲ*-types, but not for laterals, which may vary from alveolar to palatal, and sometimes even to velar in this position (cf. 2.2 for Soikkola Ingrian, 3.1.4 for Votic, 4.2.6 for Lower Luga Ingrian).

As regards reduced vowels, Lower Luga Votic tokens are transcribed the way they were pronounced: either with slightly reduced vowels maintaining their quality (*ă ă̆, ĕ*), or with two types of reduced vowels (see Rozhanskiy & Markus 2026), front and non-front, marked here as *ɛ* and *ə*, or without a final vowel at all, if not present in an actual pronunciation of a token. In Lower Luga Ingrian, where the degree of vowel reduction and shortening varies even more than in Lower Luga Votic, the phonological notation reflects the phonology of the particular subdialect (based on Kuznecova 2012) in which a token was attested. The vowel *ə* in parentheses in the phonological transcription refers to the reduced vowel which often elides in speech in south and partially central Lower Luga Ingrian (cf. the transcription principles outlined for south Lower Luga Ingrian in Kuznecova 2020).

Transliterations of Russian words render the orthography rather than the phonology or phonetics of Russian,[2] because the pronunciation could have been partially different in the local Russian dialects that served as sources of borrowing into Ingrian, as compared to modern Standard Russian pronunciation.

Regarding current consensus on Proto-Finnic reconstructions or the origins of certain older borrowings (especially Germanic loanwords), the author consulted etymological data in Wiktionary and the sources cited there.[3] Additionally, the kind advice of Viacheslav Kuleshov on the history of particular lexemes is acknowledged as "p.c.".

1.3. Object of study

A distinction between core/nuclear and marginal/peripheral phonemes is not easy to define, because it can be based on a wide range of criteria (see e.g. the 19 possible dimensions listed by Kuznecova 2014, in the spirit of Bazell 1952, 1954). The criteria may cover, for example, a phoneme's distribution as regards other phonemes and the boundaries of various prosodic and morphosyntactic domains, its relations to prosodic units of different hierarchic levels, its allophonic structure, or distribution across morphemes of various types, and more structure-external criteria, such as diachronic, sociolinguistic, and biological dimensions, as well as orthographic influences.

It is not feasible to provide here a detailed analysis of all consonants of the three varieties in question along all such dimensions. As this work concerns language contact in the transfer of consonant system properties, I primarily focus on the criteria that highlight this aspect of

[2]The transliteration follows the principles outlined in https://en.wikipedia.org/wiki/Scientific_transliteration_of_Cyrillic.

[3]Current Uralic historical databases are being collected from various sources and further developed at wiki-resources, supported by *Kielipankki*, cf. https://sanat.csc.fi/wiki/Sanat:Etymologiawiki.

phoneme marginality. These criteria are naturally related to the distribution of phonemes across different strata of morphemes in the lexicon. First of all, the following two dimensions are important in this work:

(a) native vs non-native lexicon;
(b) non-expressive vs expressive lexicon.

As contacts with Indo-European languages (especially Baltic, Slavic, and Germanic) have occurred for many centuries in Eastern Baltics (cf. 1.2), it is sometimes challenging to define the type (native/non-native) and the exact source (e.g. Proto-Germanic, Proto-Norse, Middle Low German, Swedish etc.; Balto-Slavic, Proto-Slavic, Old Russian, Modern Russian etc.) of certain consonants discussed below. Loanwords of the oldest strata (Indo-Iranian, Balto-Slavic, Proto-Slavic, Proto-Germanic) are generally equated in this work to native words, because these words do not reflect the contacts between Finnic daughter languages, which interest us here the most. Additionally, in the case of closely related languages, such as Votic, Ingrian, Finnish, and Estonian, distinguishing common inherited lexicon from the result of their contact may be challenging. Outstanding issues are discussed below to the extent possible in application to individual consonants.

It is also not easy to draw a line between non-expressive and expressive lexicon. Under expressive lexemes, I generally consider onomatopoeias (natural-sound imitations), vocatives (including words for calling domestic animals and pejorative or diminutive nicknames for people, as well as child-directed speech), and interjections. Note that words for calling animals are normally repeated three or four times in natural speech, while only two iterations are cited below.

Phonemes appearing in at least one non-expressive native (or native-like, cf. above) word are defined in this work as core phonemes and are marked in BOLD. Core phonemes contrast with other consonants in the same positions and are phonetically stable across speakers, i.e. they are not a result of an ongoing sound change. There are, however, units which have all the aforementioned properties of core phonemes but are restricted only to some subdialects of the language, due to inter-dialectal differences in earlier sound changes. These units are still considered as core phonemes but are marked in BOLD IN PARENTHESES.

Phonemes appearing only in expressive lexicon and/or in the entrenched loanwords of post-Proto-Finnic strata are written in PLAIN and considered robust marginal consonants. Phonemes of these two types behave in the languages in question very similarly in terms of phonotactics, participation in alternations, and general distribution, so I consider these two lexical classes as functionally similar for the purposes of this study. Finnic sound-symbolic word formation, with its particular inventory and phonotactics of consonants, might have been a relatively ancient mechanism that solicited the influx of certain foreign phonemes and contrasts alongside older loanwords, or, alternatively, the whole sound-symbolic mechanism could have been borrowed from Indo-European languages (see a brief overview in Laakso 2022, 252). This uncertainty is another reason why I do not draw a neat boundary between the two aforementioned lexical strata in the description below. In any case, like core phonemes, robust marginal consonants are contrastive and phonetically stable across speakers and sub-dialects.

The third dimension of marginality (c), considered here, concerns units that may or may not occur in non-native expressive lexicon but, due to specific FUNCTIONAL or PHONETIC properties, are contestable, with relatively easily acceptable alternative interpretations. These are, for example, transitory units occurring at intermediate stages of sound change (e.g. certain

palatalised consonants in Lower Luga Ingrian and Votic), or units in complementary distribution with other, structurally better-established units. For example, some singletons in Soikkola Ingrian, considered unclear marginal, occur only in positions of quantity contrast neutralisation, whereas corresponding geminates, considered marginal robust, occur in positions of quantity contrast. Some palatalised consonants, which only occur as a result of ongoing vowel reduction (cf. 1.2), are also considered unclear marginal. All cases of unclear marginal phonemes are addressed explicitly and are written in PLAIN IN PARENTHESES.

The fourth dimension (d), relevant for this work only to a limited extent, is the rarity of a phoneme across morphemes in general, and across native morphemes in particular. Limited occurrence of a phoneme, even if it is otherwise a phonetically stable and contrastive unit, renders it more marginal than a frequent unit. Rare consonants (appearing in only one or a few lexemes) are explicitly discussed below. However, if these units consistently appear in at least one native non-expressive lexeme, they are considered among core phonemes. The reason is that this work is focused on contacts, whereas such units in my data are rather manifestations of "latent" internal phonological tendencies in the system (Bybee 2001; Popov 2004, 76–91), although they may have subsequently reinforced their phonemic status through borrowed words. Another factor that is often considered relevant for phonemic analysis but irrelevant for the present study is the presence of a syllable boundary within geminate consonants. All types of geminates are treated as independent phonemes in this work.

Finally, the units which occur in the data only in a few very recent borrowings from Russian and were pronounced only by single speakers (while the general shape of these loanwords in a given language appears different), are not considered among phonemes. Such cases may often be seen as code-switching or code-mixing, given that almost all speakers of the languages of Ingria are fluent in Russian. Non-systemic elements occurring in certain expressive (especially vocative) words (e.g. the bilabial trill or the alveo-dental click, both used for stopping a horse and included in the consonant inventory of Lower Luga Votic by Viitso 1961, 151) are not included among phonemes either. Such non-phonemic elements are, however, briefly listed in dedicated sub-sections.

2. SOIKKOLA INGRIAN CONSONANTS

The consonant inventory of Soikkola Ingrian represents a more conservative evolutionary step than that of both Votic and Lower Luga Ingrian. Table 1 shows the consonant inventory of Soikkola Ingrian (revised from Kuznecova 2009, 169 and Kuznetsova, Brodskaya & Markus 2023, 4), one of the core Ingrian dialects. Generally, this is the most conservative consonant system among Ingrian dialects, as it still preserves a rare ternary quantity contrast, lost in other Ingrian dialects and even in some southernmost and easternmost Soikkola varieties (Kuznetsova 2015, 105).

2.1. Core phonemes

Core phonemes, marked in bold in Table 1, include all three quantity grades in most cases. Consonants of the second degree of quantity (C˙; so-called short, or secondary, geminates, or Q2) originate from singleton consonants, so in some descriptions they are not considered

Table 1. Soikkola Ingrian consonants

Manner	Place			
	Labial	Dental	Palatal(ised)	Velar
Stops	p pˑ pː	t tˑ tː	tʲ tʲˑ tʲː	k kˑ kː
Fricatives	(f) (fˑ) fː	s sˑ sː	sʲ sʲˑ (sʲː)	h hˑ hː
Affricates			(t͡ʃʲ) t͡ʃʲː	
Glides	v vˑ (vː)		j jˑ (jː)	
Nasals	m mˑ mː	n nˑ nː	nʲ nʲˑ	
Laterals		l lˑ lː	lʲ lʲˑ (lʲː)	
Trills		r rˑ rː	rʲ rʲˑ	

independent phonemes (Gordon 2009). However, they have phonologised in Soikkola Ingrian, and now contrast both with singletons (Q1) and with long geminates (Q3) in the same contexts (Kuznetsova, Brodskaya & Markus 2023; Kuznetsova 2023), so they are seen here as part of core phonemes.

Examples of Soikkola Ingrian core consonants in the intervocalic position after the stressed (first syllable) vowel, in the position of the maximal number of consonantal contrasts, include: *a**p**u* 'help', *na*ː*p*ˑ*uri* 'neighbour', *ta**p*ː*elo* 'fight', *sa**t**a* 'hundred', *sa**t*ˑ*a*ː 'hundred:ILL/PRT', *ha**t*ː*u*ː 'hat:ILL', *su**k**u* 'kin', *tai**k*ˑ*ina* 'dough', *ka**k*ː*u* 'pancake', *rü**s**ä* 'fishing net (type)', *rü**s*ˑ*ä*ː 'fishing net:ILL/PRT', *ka**s*ː*e* 'dew', *ja**m**ak*ː*a* 'sour milk', *ha**m*ˑ*a*ː*ra* 'poll, butt (part of the head of an axe)', *ha**m*ː*as* 'tooth', *ka**n**a* 'hen', *ka**n*ˑ*a*ː 'hen:ILL/PRT', *li**n*ː*a*ː 'city:ILL/PRT', *ka**l**a* 'fish', *ka**l*ˑ*a*ː 'fish:ILL/PRT', *po**l*ː*e* 'apron', *ke**r**a* 'with', *so**r*ˑ*i(ː)a* 'coarse', *pa**r*ː*as* 'edge, cliff', *sa**v**i* 'clay', *a**v*ˑ*ain* 'key', *o**j**a* 'creek', *o**j*ˑ*a*ː 'creek:ILL/PRT', *ra**h**a* 'money', *ra**h*ˑ*a*ː 'money:ILL/PRT'. The phoneme *t͡ʃʲ*ː, as in *lu**t͡ʃʲ*ː*ik*ː*ain* 'bedbug', is rare but included within core consonants, cf. 2.3.

2.2. Robust marginal phonemes occurring in expressive lexemes or in loanwords

In general, the consonantal periphery in Soikkola Ingrian is relatively compact and the status of most marginal consonants is quite robust. Some phonemes were attested only in loanwords from Russian, e.g. *f*-type phonemes (*f, fˑ, f*ː) and palatalised *l*ʲ-types before back vowels. Before front vowels, palatalised laterals are frequent. However, in this position, they freely vary with alveolar and, sometimes, with velar laterals (cf. Kuznecova 2009, 202–205 and also Rozhanskiy & Markus 2026, discussing laterals for Votic). In turn, *s*ʲː was attested only in a vocative word and in its verbal derivative (see 2.4).

Examples of loanwords and expressive lexemes containing marginal consonants include (for borrowings, their source word is indicated): *ko**n**ʲ**u**h*ː*a* ~ *ko**n**ʲ**u**hka* 'groom' (Rus. *кóнюх / konûx*), *fu**n**ʲ**o**i* ~ *hu**n**ʲ**o**i* 'silly person' (Rus. dial. *фýнька / fun'ka*), *fu**n**ʲ*ˑ*o(ː)ja* ~ *hu**n**ʲ*ˑ*o(ː)ja* 'silly:PRT' (cf. the previous example), *ka**f**ˑ**e*ː 'cafe' (Rus. *кафé / kafe*), *sar**f*ː*i*ː 'scarf:ILL' (Rus. *шарф / šarf*),

set^{j}olka ‘back pad (part of the horse harness)’[4] (Rus. *седёлка / sedëlka)*, *Fet$^{j\cdot}$a:* ~ *Fet$^{j\cdot}$a:ta* ‘Fed^{j}a:PRT’ (a personal name; from Rus. *Фёдя / Fedâ)*, *vont^{j}:u* ‘mentally slow, silly person’, *pos^{j}uk:ain* ‘little boy’, *Vas$^{j\cdot}$a:* ~ *Vas$^{j\cdot}$a:ta* ‘Vas^{j}a:PRT’ (a personal name; Rus. *Вáся / Vasâ*), *rul^{j}a* ‘steering wheel’ (Rus. *руль / rul’*), *Kol$^{j\cdot}$a:* ~ *Kol$^{j\cdot}$a:ta* ‘Kol^{j}a:PRT’ (a personal name; Rus. *Кóля / Kolâ*), *lar^{j}ok:a* ‘kiosk’ (Rus. *ларёк / larëk*), *Var$^{j\cdot}$a* ~ *Var$^{j\cdot}$ata* ‘Var^{j}a:PRT’ (a personal name; Rus. *Báря / Varâ*). More examples of palatalised consonants and *ƒ*-type phonemes in (a) expressive lexemes, (b) borrowed proper nouns, verbs, and adjectives, and (c) borrowed personal names are provided by Sovijärvi (1944, 33), Laanest (1986, 65) and Kuznecova (2009, 196, 286–287).

Notably, the phoneme *t^{j}* has entered the native lexicon of at least part of Soikkola Ingrian idiolects. It often appears in the innovative contractive stem of the verb *tul:a* ‘come:INF’. In the personal present tense active forms, this verb has a stem [t^{j}üe- ~ tüe- ~ t^{j}ö:- ~ tö:-] (<**tule-*), e.g. *t^{j}ö:n* [t^{j}ö:n] ‘come:1SG’. Variants with *ö:* are more innovative than those with *üe*. Palatalised *t^{j}* is attested both before *ö:* and before *üe*, most reliably in Viistinä (Rus. *Ви́стино / Vistino*) and in some villages north of it (Kuznecova 2009, 130–131, 196–197). The palatalised *t^{j}* before a new front long vowel or diphthong in this verb is in contrast with a position before original *ö* and *ü*, where *t* is never palatalised in Soikkola Ingrian, e.g. *tö:* ~ *tü:* [tö: ~ tǫ̈: ~ tü:] ‘work’, *tütök:äin* [ˈtüdöˑˌkːäi̯n] ‘little girl’ etc. A similar situation with *t^{j}* and *t$^{j\cdot}$* is observed before the diphthongs starting with *i*, which tend to turn into long vowels in Ingrian dialects in certain positions (e.g. *iu* > *u:*, *ia* > *a:*; cf. also 3.2.4). In Soikkola Ingrian, palatalised dental stops are systematically observed in such cases, e.g. *t^{j}iuk:u* [ˈt^{j}iukːu] ~ (in a few southernmost speakers) [ˈt^{j}uːkːu] ‘game’, *ka:t$^{j\cdot}$iat* [kaːt$^{j\cdot}$iat] ~ (in many speakers) [kaːt$^{j\cdot}$aːt] ‘trousers:PL’. This is in contrast to non-palatalised *t* in e.g. *te:* ~ *ti:* [te: ~ tę: ~ ti:] ‘road’, *tila* [ˈtiʟaː] ‘bed’, *kot·i(:)a* [ˈkot·iːa ~ ˈkot·ia] ‘home:PRT’ (the diphthong in the prosodic position as in the last example may be lengthened and is never transformed to a long monophthong).

2.3. Consonants with functionally or phonetically unclear phonological status

Four main groups of consonants, discussed in the following four sub-sections, present certain structural and/or phonetic challenges for their phonological interpretation: (a) palatalised geminates, (b) glides, (c) *ƒ*-types, and (d) affricates. The challenges mostly concern certain quantitative distinctions of consonants, given the complex ternary length contrast in the language, as well as distinctions between phonemes and consonant clusters.

2.3.1. Palatalised short geminates and long geminates. Palatalised phonemes are innovative elements in Finnic languages, occurring only in some expressive words and borrowings, and Soikkola Ingrian manifests this contrast only in dental consonants (unlike Votic and Lower Luga Ingrian, cf. Table 1 with Tables 2 and 3). Due to the productive morphological mechanism of so-called secondary gemination of original singletons (Kuznetsova 2023) in certain grammatical forms (e.g. *Fet^{j}a*, a personal name from Rus. *Фёдя / Fedâ* → in partitive *Fet$^{j\cdot}$a:* ~ *Fet$^{j\cdot}$a:ta*), Soikkola Ingrian has a full set of palatalised short geminates in addition to palatalised singletons.

The situation, however, is more complex with long geminates, which in native non-expressive lexemes are reflexes of original Proto-Finnic geminates or consonantal clusters. The most

[4]For palatalised *t^{j}*, Sovijärvi (1944, 33) also cites a native expressive word *mut^{j}uk:a* ‘an affective way of calling a child’.

robust cases of long palatalised geminates were attested only for *tʲː*, cf. *vontʲːu* in 2.2[5] and instances of **CjV* > *CʲːV* (see 3.2.4). The latter sound change has been established among modern Soikkola speakers in certain loanwords. For example, instead of *utʲjuːka* 'iron (for ironing clothes)' (Rus. *утюг / utûg*) and its derivative *utʲjuːkoitːa*: 'iron:INF' attested in Nirvi (1971, 626), only *utʲːuːkːa* ~ *utʲːukːa* and *utʲːukoitːa*: systematically occurred in my data.

Possible instances of *lʲː* included only two Russian loanwords: *kulʲːaitːa*: 'walk:INF, have:INF fun' (Rus. *гуля́ть / gulât'*) and *valʲːaitːa*: 'lie:INF idle, fool:INF around; make:INF felt boots, flatbreads' (Rus. *валя́ть / valât'*), attested across many speakers. However, sources also sometimes transcribe these words with a short geminate *lʲˑ* (Sovijärvi 1944, 33; Nirvi 1971, 216), along with a long geminate *lʲː* (Nirvi 1971, 636; Laanest 1986, 65), and no phonetic measurements of the durations of these palatalised geminates have yet been conducted. Short geminates before short diphthongs ending in *-i* (cf. Rozhanskiy & Markus 2026, on short and long diphthongs in Soikkola Ingrian) are generally untypical but attested in verbs of similar structure to the aforementioned borrowed verbs: *sekˑoitːa*: 'mix:INF', *sekˑoitetːi*: 'mix:IPS:PST' (my data and Nirvi 1971, 515, 517, but *sekoitetˑa*: 'mix:IPS' on p. 515), *tinˑoitːa*: (my data) ~ *tinoitːa*: (my data and Nirvi 1971, 586) 'cover:INF with tin'. Because it is unclear whether, in the case of the two loanwords in question, we observe **lj* > *lʲː* or **lʲ* > *lʲˑ*, the type *lʲː* is considered as a phoneme in this work only tentatively and is subject to further instrumental confirmation.

A similar consideration is valid for *sʲː*, which was tentatively attested only in the verb *vusʲːutːa*: 'entice:INF (a dog)', from a vocative interjection *vusʲː* ~ *vusʲ* 'a sound of enticing a dog' (Nirvi 1971, 682). There is a possibility that the word actually contains a short geminate *sʲˑ*, also subject to future measurements.[6]

Instead, *nʲː* is not included in the Soikkola Ingrian consonant inventory. It was attested only in our only speaker from Koskisenkülä (Rus. *Ко́сколово / Koskolovo)* as a result of the **CiV* > *CʲːV* change (cf. 3.2.4, 4.2.5) in Russian loanwords like *varenʲːa* 'jam' (Rus. *варе́нье / varen'je*) and native words like *minʲːa* (< *minja*) 'daughter-in-law'. However, the Koskisenkülä variety is transitory between Soikkola and Lower Luga Ingrian, with many particular traits (Kuznecova 2009, 20), and other Soikkola speakers pronounced these words as *varenia* [ˈvareˑnja ~ ˈvareˑnʲja ~ ˈvareˑn̥ja ~ ˈvareˑnʲaː] and *minja* [ˈminja ~ ˈminʲja], often without even palatalising *n*. Therefore, the realisations with *nʲː* are considered as absent from Soikkola Ingrian.[7]

[5]This word might be a Votic loanword, though, because only the Koskisenkülä speaker and one Soikkola speaker from Mäkkülä (Rus. *Ми́шино / Mišino*) knew it, while three speakers did not recognise it, whereas all five Lower Luga Votic speakers were familiar with this word. However, two of my south Lower Luga Ingrian consultants, from Kotko (Rus. *Орлы́ / Orly*) and from Vanakülä (*Ва́накюля / Vanakûlâ*), did not know the word, whereas Nirvi (1971, 678) mentions it for Laukaansuu (Rus. *Усть-Лу́га / Ust'-Luga*; north Lower Luga Ingrian).

[6]Note that in Votic (V) and Lower Luga Ingrian (LLI), all these words typically contain palatalised singletons rather than geminates: *utʲugəka* 'iron:COM' (LLI), *vusʲutən* ~ *vusʲitən* 'entice:1SG dog' (LLI) and *vusʲutːi* 'entice:PST:3SG dog' (V), *gulʲaitːama*: ~ *gulʲaitːəma* 'have:SUP fun' (LLI), as well as a variation *lʲ* ~ *lʲː* in the last word with forms with the short *lʲ* prevailing in Votic in Grünberg (2013, 240), including *gulʲaitːəma* in Tsvetkov (1995, 46).

[7]In Lower Luga Ingrian, *varenʲː* ~ *varenʲːa* and *minjə* ~ *minʲːa* were attested; in Votic, *vareni* and *minʲdʲə* ~ *minʲjə* ~ *minʲːə*.

2.3.2. Geminate glides. Glides present an issue similar to that outlined in 2.3.1 for palatalised geminates. Short geminate glides occur as a regular result of secondary gemination: *oja* 'creek' → *oj'a*: (ILL/PRT), *savi* 'mud' → *sav'e*: (ILL).

Long geminate glides *j*: and *v*: do not contrast in the language with the intervocalic sequences *ij* and *uv*. In cases when these glides originate from sequences **ij* and **uv*, as in *äj:ä* 'grandfather', *sav:u* 'smoke' (<Proto-Finnic **äijä*, **sauvu*[8]), they seem to be pronounced as long geminates. It is generally difficult to draw a phonetic distinction between [ij] and [j:] after a vowel. However, words like *sav:u* 'smoke' or *kav:an* 'for a long time' are consistently pronounced in modern Soikkola Ingrian with [v:] rather than with [uv], in contrast to varying types of transcriptions in sources, e.g. "pö**ŭ̆v̆**vü ~ pö**v̆v̆**ü ~ pö**β̆β̆**ü" 'sheepskin coat', "ra**ŭ̆β̆**$_{v}$**β̆**$_{v}$an" 'iron:GEN' (Sovijärvi 1944, 32–33; Nirvi 1971, 459), "ka**vv**an" and the like (Laanest 1986, 18). Laanest (ibid.) also mentions that *v*: is consistently pronounced in all such cases.

However, a phonetic distinction between long and short geminate glides (as in *kav:an* 'for a long time' vs *av'ain* 'key') has never been instrumentally tested, so long glides are distinguished here as independent phonemes only preliminarily.

Glides after long vowels, where they emerged as syllable boundary markers in place of lost stops, present an even more difficult case. In such words, *v*: occurs after long *o, ö, u, ü, a, ä*, while *j*: occurs after long *e, i*, e.g. *to:v:a* ~ *tu:v:a* 'bring:INF here', *sa:v:a* 'get:INF', *ve:j:ä* ~ *vi:j:ä* 'bring:INF there' (cf. Fin. *tuoda, saada, viedä*). A certain variability in the duration and quality of glides is attested in this position. For example, glides may sometimes be pronounced as singletons or short geminates, or [w:] may be pronounced in place of *v*:. However, the realisation with long glides is the most typical in these cases (see examples in Laanest 1986, 18–19; Kuznecova 2009, 254). On the other hand, there are other similar cases where clear singleton glides (or no glides at all) occur, e.g. *u:ven* 'new:GEN', *vo:ven* ~ *vu:ven* 'year:GEN', *oen:an* [ˈoeˑn:an ~ ˈoje̯ˑn:an] 'straighten:1SG'.

In sum, I tentatively postulate the presence of long geminate glides *j*: and *v*: in the consonant system, but this is a subject for future phonetic investigation.

2.3.3. Phonemes *f* – *f'* – *f*:. As mentioned in 2.2, *f*-type phonemes occur only in loanwords. Apparently, *f*-types used to be re-analysed in Soikkola Ingrian (as *v, h, hv*, or *p*:, depending on position), cf. e.g. *kohvi* 'coffee' (Rus. *кóфе/kofe*) as the only attested variant even among modern speakers. For Soikkola Ingrian, Nirvi (1971, 271, 275) cites, for example, "lo**f**ko" ~ "lo**uh**ka" 'apt' (Rus. *лóвкий / lovkij*), "Mar**pp**a" (Rus. personal name *Мáрфа/Marfa*), while my data include only *lofkoi* ~ *lofko* ~ *lofka* and *Marf:a.* Additionally, the personal name *Fetʲa* (or rather its full form Фёдор/Fëdor) has an older variant ***V**ö:t:öi* ~ ***V**ü:t:öi*; see also the variability ***f**unʲoi* ~ ***h**unʲoi* in 2.2.

At present, however, *f*-types have already entrenched in the consonant system of Soikkola Ingrian. They are the only phonemes which can be clearly opposed to other phonemes (*v*-types) in voicing, cf. *kaf'e*: 'cafe' (Rus. *кафé / kafé*) vs *sav'e*: 'mud:ILL', ***f**a:prik:a* 'factory' (Rus. *фáбрика / fabrika*) vs ***v**ari* 'hot' (but see also 2.4 on other emerging voicing contrasts).

Like in cases discussed in 2.3.1 and 2.3.2, here, too, there is a question about a clear distinction between long and short geminates. Tentatively, long geminates were established in all grammatical forms of the words *Marf:a* 'Marfa', *sarf:i* 'scarf', as well as in the two-foot tetrasyllabic words *saraf:ana* 'traditional women's dress' (Rus. *сарафáн / sarafan*) and *telef:ona*

[8]Cf. https://sanat.csc.fi/wiki/Etymologiadata:imsm:äijä, https://sanat.csc.fi/wiki/Etymologiadata:imsm:sauvu.

'telephone' (Rus. *телефо́н / telefon*), whereas short geminates are considered to occur, for example, in *ka**f**'e*: 'cafe' and *Mi**f**'o:dia* (personal name; Rus. *Мефо́дий / Mefodij*) (cf. also Nirvi 1971, 309). These distinctions are also to be instrumentally verified.

In contrast to palatalised consonants and glides, *f*-types present an issue with a singleton *f*. Singleton *f* can occur only in those positions where any geminates are phonetically impossible and quantity contrasts are neutralised, i.e. word-initially (***f**a:prik:a*) and in clusters with obstruents (*lo**f**koi*). A short phoneme *f* in these positions is proposed out of systemic and phonetic reasons, in parallel to most other consonants, which manifest a full ternary quantity contrast in quantity distinction positions (between vowels and after a sonorant before a vowel). In neutralisation positions, allophones phonetically corresponding to singletons occur for all consonants, so I postulate singletons in these positions also for unclear cases like *f* or *t͡ʃʲ* (cf. 2.3.4).

2.3.4. Affricates. Short and long affricates *t͡ʃʲ* and *t͡ʃʲ:* present the same kind of complementary distribution as *f*-types (see 2.3.3), and their phonemic status and relation to *ts* are challenging issues (Kuznecova 2009, 270-273). The difference from *f* is that short geminates and long geminates are not opposed in case of *t͡ʃʲ:*. This *t͡ʃʲ:* occurs word-internally between vowels or after a sonorant before a vowel, including in a few native non-expressive lexemes, e.g. *man**t͡ʃʲ**:ik:ain* 'strawberry', *lu**t͡ʃʲ**:ik:ain* 'bedbug', *klä:**t͡ʃʲ**:i ~ klä:**t͡ʃʲ**:ü* 'boreal vetch (*Vicia cracca*)', and is included in core consonants (2.1). Both long and short affricates occur in many native expressive lexemes and entrenched loanwords, e.g. *mä**t͡ʃʲ**:ik:ä ~ mä**t͡ʃʲ**:ük:ä ~ mä**t͡ʃʲ**:ik:ö* 'ball (for playing)' (Rus. *мя́чик/мя̂чик*), ***t͡ʃʲ**irk:ulain* 'sparrow' (an expressive lexeme referring to the sound of chirping), *po**t͡ʃʲ**ka* 'barrel' (from Rus. *бо́чка / bo**t͡ʃʲ**ka*).

The affricate occurring in quantity distinction positions phonetically behaves like a geminate consonant (or a consonant cluster) rather than as a singleton. In particular, in words like *lu**t͡ʃʲ**:ik:ain* [ˈlu**t͡ʃʲ**:iˌkːaˌin], the post-stress vowel after an affricate does not undergo phonetic lengthening, unlike it does in structures with the stressed light syllable (C)V followed by a singleton consonant, cf. with *harak:a*: [ˈhar**a**ːˌkːaː] 'magpie' (see Kuznetsova, Brodskaya & Markus 2023; Kuznetsova, Cornelio Tiburcio & Uchihara to appear). Short geminates in Soikkola Ingrian occur either before long vowels or in trisyllabic feet (Kuznetsova 2023), whereas *t͡ʃʲ:* occurs widely before short vowels (cf. examples above). Additionally, its duration is close to the duration of consonant clusters (Rožanskij & Markus 2020, 471), so the affricate *t͡ʃʲ:* is considered here as an equivalent of a long geminate.

The affricates *t͡ʃʲ* and *t͡ʃʲ:* can hardly be considered as clusters either. The long *t͡ʃʲ:* contrasts intervocalically to the cluster *ts*, as in *me**ts**ä* 'forest' (cf. Rožanskij & Markus 2020; Rozhanskiy & Markus 2026, discussing *ts* in detail). However, outside this context, *ʃ*-types and *s*-types in Soikkola Ingrian can be considered allophones within the single class of sibilants (unlike in Votic and in Lower Luga Ingrian; see 3.1 and 4.2.2). In particular, *ʃ*-types [ʃ/ʒ, ʃˑ, ʃː] together with *s*-types [s/z, sˑ, sː] make part of the allophonic pool of the phonemes *s, sˑ,* and *sː,* respectively. In the language of the 2000s we recorded, mostly *ʃ*-like sounds were observed in almost all positions, whereas in earlier sources more *s*-like sounds were attested (Ariste 1960, 50; Laanest 1966a, 54–58, 1986, 40–42, 1987, 291–293). The most typical exception is *ts*, where [s] instead of [ʃ] is still pronounced. This is also valid for sandhi, cf. an isolated pronunciation of *sa:tu* 'garden' (Rus. *сад / sad*) as [ˈʃaː**d**u], but a sandhi example with [ts] from my field data: *hän ei osta:nt sa:tˑua* [ˈoʃtaːn**t̪** ˈ**s**aːtˑua] 'he did not buy a garden'.

Furthermore, [s] instead of [ʃ] has been attested in my data also in other contexts, for example, before *t,* especially in cases when *t* or the whole cluster was palatalised, e.g. *astiata* [ˈaʃtiada ~ ˈaʃtʲada ~ ˈas**tʲ**ʲ̯ăda ~ ˈas**tʲ**ʲada] 'dishes:PRT'. Intra-speaker variation is also possible here, cf. *astioja* [ˈaʃtioja ~ ˈaʃtʲioja] (PL:PRT) but *astiois* [ˈas**ʲ**t**ʲ**ojʃ] (PL:IN) produced by a speaker from Viistinä. The presence of [s] instead of [ʃ] is likely related to a certain degree of palatalisation of *s* in such contexts. Note that palatalised phonemes *sʲ* and *sʲ*: (see examples in 2.2, 2.3.1) are rarely realised as *ʃ*-types. Other contexts where *s*-types have also been frequently attested in free variation with *ʃ*-types include clusters with sonorants and glides like *sr, sv*, *sl,* as in *osra* 'barley', as well as those Russian loanwords which contain *s* in Russian, e.g. *se**ʲ**olk(ə)* 'back pad' from Rus. *седёлка / sedôlka* (see more examples in Kuznecova 2009, 235–239). Additionally, *ʃ*-types do not occur in the southernmost transitory Soikkola-Lower Luga variety of Koskisenkülä (see 2.3.1) and are less typical of the adjacent Säätinä (Rus. *Слобóдка / Slobodka*) (cf. also Laanest 1966a, 56), both of which are geographically the closest to the Lower Luga Votic area.

Importantly, in all such cases, *s*-types and *ʃ*-types never contrast.[9] Therefore, treating *t͡ʃʲ*: as a cluster would create a problem for the phonological interpretation of the second component of this cluster, as contrasted to the second component of *ts.*

As for the short *t͡ʃʲ*, there are no reliable examples of its contrast to *ts* in any position (cf. also Laanest 1966b, 104), although rare occasional occurrences of word-initial *ts* instead of expected *t͡ʃʲ* have been attested, mostly where *ts* occurs in Russian words, cf. [ˈt͡sepaː] instead of regular *t͡ʃʲepa* 'chain', from Rus. *цепь / cep'*. It should also be noted that at least *t͡ʃʲ*: has apparently developed from *ts* in Soikkola Ingrian, and before a massive recent influx of Russian loanwords, the distance between *t͡ʃʲ*: and *ts* was likely smaller. In some words, especially in earlier borrowings from local Russian dialects and in expressive lexemes, variation *t͡ʃʲ*: ~ *ts* word-internally was attested, including cases when Russian *t͡ʃʲ* was rendered as *t͡s* in Ingrian, e.g. *la**t͡ʃʲ**:urikain* ~ *la**t͡s**urkain* 'flattened, fine' (Nirvi 1971, 252; while in my data and in Laanest 1986, 40 only the first variant was recorded) or "ku**t͡s**evekka" ~ "ku**t͡s**evēk**k**a" (Nirvi 1971, 223) ~ *ku**t͡ʃʲ**:evik:a* (my data) 'women's short coat' (Rus. *кацавéйка / kacavejka,* here probably *куцавéйка / kucavejka*) (see other examples of variation *t͡ʃʲ* – *t͡s* word-initially and *t͡ʃʲ*: – *ts* word-internally in Kuznecova 2009, 270; Rozhanskiy & Markus 2026). It is also possible that the north-western Russian dialects from which Ingrian borrowed some of its lexicon belonged to the area of *цóканье/ cokan'je* (Borkovskij & Kuznecov 1963, 84), where *t͡ʃʲ*-types and *t͡s*-types were not distinguished and were pronounced close to [t͡s / t͡sʲ]. By now, however, *t͡ʃʲ*-types have gained significant independence from *ts*, probably because *ʃ*-types have generally become more prominent in Soikkola Ingrian likely due to the influx of Standard Russian words, in which *t͡ʃ*-types and *t͡s*-types are contrasted.

2.4. Elements not considered as phonemes

In this section, I briefly mention elements which present the degree of phonemicity too weak to be recognised as phonemes according to the criteria adopted here. First of all, these include the already mentioned [nʲ:] (the Koskisenkülä variant of *nj* ~ *ni*, see 2.3.1), [ts] (which is considered

[9]In Koskisenkülä, *ʃ*-types actually do occur in loanwords, e.g. in the verb 'lie' from Rus. *лежáть / ležat'*: [leʒiːn] (1SG), [leʃˑiː] (3SG), and, therefore, are in contrast with *s*-types, like in Votic and Lower Luga Ingrian. However, as said in 2.3.1, idiosyncratic Koskisenkülä facts are not taken into account while modelling the general Soikkola phonological system.

as a cluster rather than a phoneme word-internally and as a marginal irregular variant of *t͡ʃʲ* word-initially, see 2.3.4), and a possible phonemic contrast of *s*-types and *ʃ*-types (these sounds do not contrast outside the *t͡ʃʲ*: – *ts* opposition, although for some words or contexts *s*-types might be more typical than *ʃ*-types, see 2.3.4).

As said in 2.3.3, the contrast of *f*-types and *v*-types is considered here as the only true voicing contrast in Soikkola Ingrian. Otherwise, Soikkola Ingrian does not distinguish between voiced and voiceless obstruents, unlike Votic and Lower Luga Ingrian (cf. 3.1, 4.2.1). Voicing in Soikkola Ingrian is an additional phonetic feature, optionally appearing in singleton stops *p, t, k* and fricatives *s, h* in the positions of quantity contrasts (intervocalic and after sonorants). In other positions, where all quantity contrasts are neutralised (word-initially, word-finally, and in other types of clusters), singletons appear as short and either voiceless or optionally voiced depending on the surrounding phonetic context (see Laanest 1966a, 50–54, 1987; Kuznecova 2009, 171–173). For example, word-initially and word-finally, short voiceless obstruents may become voiced in sandhi after a word that ends or begins in a vowel or a sonorant.

However, in some cases, voiced obstruents have been attested word-initially also outside sandhi. First, this is the variant of the lexeme ‘raspberry’ [ˈ**b**aːbuˌkːai̯n] attested as the only variant in Koskisenkülä and along with [ˈvaːbuˌkːain] from a Viistinä speaker (*vaːpukːain* ~ *vaːpukːa* are the regular Soikkola variants). Especially in the case of Koskisenkülä, this might potentially be a Votic influence. Second, voiced *b* has been attested in vocative words for calling a ram [ˈ**b**orʲaː ˈ**b**orʲaː] (recorded from eight speakers; including once in a form [ˈ**b**orʲaː ˈporʲaː]; cf. with *poroi* ~ *porʹoːna* ‘ram’, Rus. dial. *бóран / boran*), and for calling a sheep [ˈ**b**asʲuː ˈ**b**asʲuː] (from one speaker, while Rozhanskiy (p.c.) also attested *vasʲu vasʲu*; cf. with *pässi* ‘ram’). The factor here could be a multiple repetition of the vocative. Starting from the second iteration, the word-initial consonant in vocatives might have been frequently voiced in sandhi, cf. “pässi bässi bässi” for calling a ram in Nirvi (1971, 452). The voiced realisation might have spread also to the first iteration. Third, modern speakers sometimes pronounce initial voiced obstruents, especially *b*, in those Russian loanwords where there is a voiced obstruent in Russian, e.g. in *pot͡ʃʲka* ‘barrel’ [ˈpot͡ʃʲka ~ ˈ**b**ot͡ʃʲka] (from Rus. *бóчка / bočka*), *klaːtitːaː* [ˈklaːd(ʲ)iˌtːaː ~ ˈ**g**laːd(ʲ)iˌtːaː] ‘iron:INF’ (Rus. *гла́дить / gladit’*). It is not possible to account for the word-initial voiced pronunciations of obstruents in Soikkola Ingrian within the framework adopted in the current description, and they are left outside the Soikkola phonemic system.

3. LOWER LUGA VOTIC CONSONANTS

The consonant system of Lower Luga Votic differs significantly from that of Soikkola Ingrian. The primary differences in Votic relate to the following aspects:

(1) more common binary quantity instead of ternary quantity (secondary geminates originating from singletons do exist in Lower Luga Votic but have the same duration as original geminates, cf. Viitso 1964; Kuznecova 2015);
(2) two separate contrasts, in length and in voicing, instead of a single lenis/fortis contrast simultaneously based on duration and voicing (cf. 2.4), for *p-, t-, k-, s*-types (i.e. for both singleton and geminate obstruents; see examples in 3.1);
(3) two separate phonemic types, *s*-types and *ʃ*-types (both singletons and geminates; see examples in 3.1) instead of a single type of sibilants (cf. 2.3.4);

(4) a richer system of affricates: two *t͡ʃʲ*-types (a singleton and a geminate, cf. 3.1) and two *t͡s*-types (also a singleton and a geminate, see 3.1), more clearly monophonemic than in Soikkola Ingrian (cf. 2.3.4), and an affricate *d͡zʲ* (see 3.1.2);
(5) a richer inventory of palatalised phonemes due to the assimilation, syncope, and apocope of the following short front vowels and *j*, some of these consonants occurring also in native non-expressive lexemes (cf. 3.1, 3.2.3, 3.2.4).

The Lower Luga Votic consonant inventory, given in Table 2, is built based on Markus & Rožanskij (2011, 16), supplemented with my field data and following the same classification principles (see 1.3) as those applied to Soikkola Ingrian above. Data from Tsvetkov (1995) on Jõgõperä and from Viitso (1961) on Luutsa-Liivčülä were also taken into account.

Differences between Table 2 and the Lower Luga Votic consonant inventory provided by Markus & Rožanskij (2011, 16) mostly arise from: (a) additional data on consonants obtained from my fieldwork on Votic phonology and from the existing literature, and (b) partially different principles of phonemic classification. All these deviations are discussed below.

Table 2. Lower Luga Votic consonants

		Place						
Manner		Labial		Dental		Post-alveolar	Velar	
		Plain	Palatalised	Plain	Palatal(ised)		Plain	Palatalised
Stops	voiceless	p pː	(pʲ pʲː)	t tː	tʲ tʲː		k kː	(kʲ kʲː)
	voiced	b bː	(bʲ)	d (dː)	(dʲ) dʲː		g	
Fricatives	voiceless	f fː	(fʲ)	s sː	sʲ sʲː	ʃ ʃː	h hː	(hʲ)
	voiced			z zː	zʲ (zʲː)	ʒ (ʒː)		
Affricates	voiceless			t͡s t͡sː		t͡ʃʲ t͡ʃʲː		
	voiced					(d͡ʒʲ)		
Glides		v vː	(vʲ vʲː)		j jː			
Nasals		m mː	(mʲ mʲː)	n nː	nʲ nʲː			
Laterals				l lː	lʲ lʲː			
Trills				r rː	rʲ rʲː			

3.1. Core consonants

Lower Luga Votic (Luutsa-Liivcülä) phonemes are described in detail by Viitso (1961); see also examples of core consonants in Markus & Rožanskij (2011, 51–53, 56). Core consonants are generally similar to those in Lower Luga Ingrian (see 4.1 below). Hence, I do not provide examples of Votic core consonants that are identical to those of Lower Luga Ingrian, focusing only on the distinctive properties of Votic. Structural and functional differences from Lower

Luga Ingrian concern the sub-systems of (a) voiced stops and fricatives, (b) *s*-types vs *ʃ*-types, (c) affricates, and (d) certain palatalised consonants.

Voiced phonemes are a common feature of the native non-expressive lexicon in Votic, due to a specific system called stem gradation. In this system, voiceless consonants in the "strong" grade regularly alternate with voiced consonants in the "weak" grade. For example, the following alternations with voiced phonemes are possible in Votic but are impossible in Lower Luga Ingrian and, in fact, in most other Finnic languages (examples from Markus & Rožanskij 2011, 52–53): *i**t**k**ə**ma* 'cry:SUP' — *i**d**g**ə**n* 'cry:1SG', *i**s**ä* 'father' — *i**z**ä* 'father:GEN', *la**s**k**ə**ma* 'release:SUP' — *la**z**:ə**n* 'release:1SG', *ʃi**ʃ**k* 'rag' — *ʃi**ʒ**g**ə**d* 'rag:PL', *au**t͡ʃʲ**i* 'pike' — *au**d͡ʒʲ**id* 'pike:PL'. Examples of other voiced phonemes in native non-expressive lexicon from my field data include *ku**z**a* 'where' — *ku**z**:a* 'nowhere', ***b**a:**b**uk**ə**z* 'raspberry', *õ**ʒ**:ă* ~ *e**ʒ**:* 'horsetail (Equisetum)', *mi**n**ʲ**d**ʲ**ə* ~ *mi**n**ʲ**j**ə* ~ mi**n**ʲ:ə 'daughter-in-law', *la**d**ʲ:a* 'wide:PRT'.

The status of *ʃ*-types and affricates in Votic is also solid, because both groups can occur in the native non-expressive lexicon (unlike in Lower Luga Ingrian, cf. 4.2.2). Additionally, both *ʃ*-types and affricates alternate between singletons and geminates in the paradigms of native words, e.g. *ku**t͡s**:u**m**a* 'call:SUP' — *ku**t͡s**u**n* 'call:1SG', *pi**t͡ʃʲ**:i* 'long' — *pi**t͡ʃʲ**ä*[10] 'long:GEN', *karju**ʃ**:i* 'shepherd' — *karju**ʃ**il* 'shepherd:AD' (Markus & Rožanskij 2011, 51).

Both voicing and *s*/*ʃ* distinctions are apparently quite ancient contrasts in Votic, especially the former. Some of the relatively early Germanic loanwords appear in Votic, unlike in other Finnic languages, with their original voiced consonants and *ʃ*-types. For example, initial and middle voiced obstruents are observed in Lower Luga Votic ***g**lazi* 'glass', appearing as *klasi* in Soikkola Ingrian (Nirvi 1971, 176) and as *lasi* in Finnish, which is traced to Middle Low German (1200–1500 AD) *glas* (Bentlin 2008).[11] An even more interesting example is Lower Luga Votic ***g**laizə̆*[12] 'club-rush (Scirpus)', which appears in all other Finnic languages with initial *k*- (e.g. Finnish *kaisla)*, but is derived from Proto-Germanic **gaisilaz* 'staff, whip', or even **gaizilaz*, from **gaizaz* 'spear, tip' and **-ilaz* 'a diminutive suffix' (cf. *gaiza-* in Kroonen 2013, 164).[13] Viacheslav Kuleshov (p.c.) attributes the borrowing of this word to a period between late Proto-Norse and early Old Norse, before a sound change **ai* > *ei* took place in Old Norse, i.e. to the Vendel period or the early Viking Age (~500–900 AD). This example demonstrates that Votic was already a separate language at that time and that it already had its own system of voicing contrasts, in which Germanic lenis obstruents were borrowed as voiced consonants in certain positions.

The *s*/*ʃ* contrast is not evidenced in such early loanwords, but it can be traced to at least Middle Low German, due to borrowings with *ʃ* like ***ʃ**peili* 'mirror' (Grünberg 2013, 1238),

[10]Votic is the only Finno-Ugric language featuring regular palatalisation **k* > *t͡ʃʲ* / *t͡ʃʲ*: before front vowels, but in modern language *k* can contrast to *t͡ʃʲ* in this context (mostly due to Ingrian loanwords without this palatalisation), e.g. *pi**t͡ʃʲ**ä* 'long:PRT' — *mi**k**ä* 'what' (Markus & Rožanskij 2011, 43).

[11]Ultimately traced to Proto-Germanic *glasą ~ *glazą (Kroonen 2013, 180); https://en.wiktionary.org/wiki/lasi#Etymology_2, https://en.wiktionary.org/wiki/Reconstruction:Proto-Germanic/glasą.

[12]For these two and other similar examples, Grünberg (2013) often provides examples of variation between word-initial voiced and voiceless obstruents. I also did attest this variation e.g. for *glaiza*, but this is apparently a phonetic effect, as my Luutsa consultant clarified the word-initial voiced sound in clear speech.

[13]https://en.wiktionary.org/wiki/Reconstruction:Proto-Finnic/kaisila, https://en.wiktionary.org/wiki/Reconstruction:Proto-Germanic/gaisilaz, https://en.wiktionary.org/wiki/Reconstruction:Proto-Germanic/gaizaz.

cf. Soikkola Ingrian *speili,* Finnish and Karelian *peili*, Estonian *peegel*. This word reflects a specific Middle Low German sound change **ēg* > *ej* (Kuleshov, p.c.; cf. Middle Low German variants *spēgel, speyl, spīgel, speigel, speiel*[14]). In native words, *ʃ* is mostly restricted to expressive lexemes (cf. 4.2.4). Words like ***ʃiʃk*** 'rag', cited above, ***ʃam*:*ama*** 'smoke:SUP (pipe etc.)', or ***ʃimar*** 'wrinkle' (Tsvetkov 1995, 319) do not appear to have expressive meaning synchronically but might have contained an expressive semantic component originally (e.g. one Luutsa speaker noted on ***ʃam*:*ama*** that this is a word to use "for fun", while the neutral word for smoking is *põlõt*:*əm*:*a*, lit. 'burn:SUP'). Therefore, the phonemic *ʃ*-types are included within core phonemes tentatively, with more etymological work needed here.

Before considering marginal phonemes, specific comments should be made on particular groups of core phonemes: (a) ***ʒ***: and *z*^{j}:; (b) *d͡ʒ*^{j}; (c) *d*^{j}, *d*^{j}: and *n*^{j}:; (d) *l, l*:, *l*^{j}, *l*^{j}:, and *t*^{j}. Some of the palatalised phonemes (*n*^{j}:, *l*^{j}) are considered marginal; palatalised consonants are further discussed in 3.2.3, 3.2.4.

3.1.1. Phonemes *ʒ*: and *z*^{j}:. Phoneme *ʒ*: is missing from Markus & Rožanskij (2011, 16), and relevant examples are also absent from Viitso (1961).[15] Tsvetkov (1995, 38), however, contains at least one entry with a native-like word with *ʒ*: in Jõgõperä, namely *õʒ*: ("ę̄ž̄") 'horsetail (Equisetum)'. Viitso (1961, 165) cites this word for Luutsa as *õz*^{j}*:ă* ("ę́ź́ă"), and this is also the most frequent variant ("õźźa") cited by Grünberg (2013, 1590), along with "õźja", "õzja", "õžža", "õžž", and "õzzi". The variants attested in my data included *eʒ*: for Jõgõperä and *õʒ*:*ă* for Luutsa (each from one speaker).

This word (Proto-Finnic **hosja*) is attested widely across Finnic languages, in some of them with *s* (Fin. *hosia*) but in others also with *ʒ* or *d͡ʒ*, or *z*^{j}, and is also reconstructed to Proto-Saami with **ʃ*:,[16] apparently as a Finnic loanword. The Finnic word appears related to Slavic **xvoščъ*[17] (cf. Rus. *хвощ / xvoŝ)*. Kuleshov (p.c.), however, considers it to belong to an earlier period. Two separate proto-forms can be reconstructed for North and South Finnic, indicating that this word was borrowed before the Proto-Slavic period. Kuleshov notes a parallel with Latvian *aste* 'tail',[18] but attributes the word to an even earlier, Proto-Iranian, period, because its shape across Finnic languages better corresponds to a Proto-Iranian form containing an initial laryngeal consonant than to a Proto-Baltic form without a laryngeal.

For this work, focused on contacts between Finnic daughter languages, I consider loanwords from the Proto-Finnic stage and earlier as part of the native lexicon (cf. 1.3). For this reason, both *ʒ*: and *z*^{j}: are included among the core phonemes. They are given in parentheses in Table 2 because they seem to be in complementary distribution across Jõgõperä and Luutsa. For Jõgõperä, Kettunen (1915, 54) specifically mentions one more word containing *ʒ*: <**zj* <**sj*, a form

[14]https://en.wiktionary.org/wiki/peili, https://en.wiktionary.org/wiki/Reconstruction:Proto-West_Germanic/spēgl.

[15]Viitso does not list geminates as separate phonemes but still provides examples of geminates for all consonant types.

[16]Cf. https://en.wiktionary.org/wiki/Reconstruction:Proto-Finnic/hosja, https://sanat.csc.fi/wiki/Etymologiadata:imsm:(h)osja, https://kaino.kotus.fi/algu/index.php?t=sanue&lekseemi_id=59862&hakusana=hosia&sanue_id=40833.

[17]Cf. https://en.wiktionary.org/wiki/Reconstruction:Proto-Slavic/xvoščь.

[18]Cf. https://en.wiktionary.org/wiki/aste#Latvian.

leʒ:es:i ‘widow:TRL’ (from *lehtʃʲi*, cf. Fin. *leski)*, while for other Votic varieties *lez:es:i* or *lezdʲes:i* is cited (ibid.).

In turn, another word containing *zʲ:* as a reflex of **sj* in Luutsa Votic is *azʲ:a* ‘affair, business’. In my data, it was recorded from two Luutsa speakers (one of them also mentioning *asʲ:a*, which is rather an Ingrian variant). Contrarily, none of my two Jõgõperä consultants recognised the word in the form *azʲ:a* (one of them mentioning *asʲ:a* specifically as an Ingrian form). Viitso (1961, 165) cites *azʲ:ad* for Luutsa Votic, but Tsvetkov (1995, 11) mentions *asi* for Jõgõperä Votic (unfortunately, this second variant was not checked in my questionnaire).

3.1.2. Phoneme *d͡ʒʲ*.

Both Viitso (1961, 151) and Rozhanskiy & Markus (2026) observe that *d͡ʒʲ* appears as a weak counterpart of *t͡ʃ* (cf. *au**t͡ʃʲ**i — au**d͡ʒʲ**id* in 3.1) in modern Lower Luga Votic words very rarely, mentioning a tendency to level stem gradation in such words. As a weak grade of e.g. *bo**t͡ʃʲ**ka* ‘barrel’ (from Rus. *бóчка / bočka*), speakers tend to pronounce *bo**t͡ʃʲ**kad* (PL) rather than *bo**d͡ʒʲ**kad*, while also acknowledging the latter variant as correct. In my data, *au**d͡ʒʲ**id* was recorded from one Jõgõperä speaker, while the other Jõgõperä speaker pronounced *au**t͡ʃʲ**id* (and one Luutsa speaker used the Ingrian-like forms *auki – auvəd*). However, the same Luutsa speaker also pronounced *mänʲ**d͡ʒʲ**üd* ‘game:PL’ (but also *mänʲ**t͡ʃʲ**üd*; from *mänʲ**t͡ʃʲ**ü*), and another Luutsa speaker produced *sän**d͡ʒʲ**ü* ‘bed:GEN’ (from *sän**t͡ʃʲ**ü*).

Apart from levelling stem gradation, another reason for the rarity of *d͡ʒʲ* might be that *d͡ʒʲ* in the weak grade in west Votic was not very common overall (Ariste 1968, 10). Kettunen (1915, 49–55), for example, mentions forms with *d͡ʒʲ* for east Votic, but different types of forms for west Votic, including *au**dʲ**id* for ‘pike:PL’; cf. also *mä**nʲ**:üˑd* ‘game:PL’, *sä**nʲ**:üˑd* ‘bed:PL’, cited by Tsvetkov (1995, 184, 316) for Jõgõperä and *mänd**ʲ**üd* mentioned by Viitso (1961, 164) for Luutsa. Therefore, *d͡ʒʲ* in the weak grade might have been a late analogical innovation in Lower Luga Votic, in parallel with voicing-based stem gradation in other consonants (cf. 3.1), and now tends to disappear again due to stem levelling. Following the methodology of this study, I consider *d͡ʒʲ* a core phoneme, as it appears in some native non-expressive lexemes. Still, it is placed in parentheses in Table 2, because its functional status in Lower Luga Votic is unclear.

3.1.3. Phonemes *dʲ – dʲ:* and *nʲ:*.

The phonemes *dʲ* and *dʲ:* are relatively uncommon in Votic, especially in modern speech. They historically appeared mainly in west Votic, including some weak grade contexts where east Votic features *d͡ʒʲ* (see examples in 3.1.2). They originate from Proto-Finnic **k* > [j] before front vowels and after a vowel or a consonant (usually **n,* **r,* **l*, but also **h,* **s*), as well as in some other clusters containing phonetic [j:], like **ki* > [gj] > [j:], **ni* > [nj], **it* > [jd] > [j:] (Kettunen 1915, 49–54, 56–57). Short [j] changed into *dʲ*, and long [j:] into *dʲ:*, which was apparently a case of glide fortition (cf. Culhane & Edwards 2025, 189 on a similar process in Austronesian languages). In modern data, these phonemes are less frequent, first, because they are being replaced by *j* and *j:*, likely under Ingrian influence, and second, due to the levelling of a grade alternation *t͡ʃʲ* : *dʲ* (cf. also 3.1.2).

My own data and some published findings suggest that *dʲ* and *dʲ:* might have been more typical of Luutsa-Liivčülä Votic than of Jõgõperä Votic. In particular, I observed the following examples of correspondences between Luutsa and Jõgõperä, respectively: *mänd**ʲ**üd — mänʲ:üˑd* (see 3.1.2), *minʲd**ʲ**ə — minʲjə / minʲ:ə* ‘daughter-in-law’, *enʲd**ʲ**e — en(ʲ):e* ‘soul:GEN’, *võd**ʲ**:ə võd**ʲ**:əl —*

võjːə võjːəl ‘spread:IMP butter:AD’ (‘spread butter!’), *laⁿdʲːə* — *lajːə* ‘wide’; cf. also the alternation *nt͡ʃʲ* : *ndʲ*, attested in Luutsa-Liivčülä but not in Jõgõperä Votic in Markus & Rožhanskij (2011, 54).

Variants without *dʲ*, *dʲː* were, however, attested also in Luutsa, e.g. *lajːə*, *t͡ʃʲenːɛd* ‘shoe:PL’, *äjːɛ* ‘grandfather’ (vs *äjːɛ* ~ *äⁿdʲːɛ* in Jõgõperä), and *märjɛd* ‘wet:PL’. Moreover, one Luutsa speaker, who himself produced *lajːə* and *mejːe* ‘we:GEN’, mocked the use of *dʲː* in the speech of his former neighbour from Jõgõperä. The neighbour offered him a cake, accompanied with a phrase full of *dʲː*, ‘*na-ka meⁿdʲːe vaⁿdʲːa kakːua, sʲeːmvõika võⁿdʲːětːua*’ ‘take our Votic cake, coated with [sunflower] seed oil’. Tsvetkov (1995) also cites several words with *dʲː* for Jõgõperä, (e.g. “лаd’ ~ лad’d’Ę” ‘wide’, p. 137, “äd’d’E” ‘grandfather’, p. 420, “med’d’ė” ‘we:GEN’, p. 183, “vęd’d’ę ~ vęd’d’Ę” ‘butter; cream’, p. 397, “mad’d’ęs” ‘dainty’, p. 164, “od’d’ęī” ‘protect:INF’, p. 204) but I was unable to find any native non-expressive words with short *dʲ* in that dictionary.

In sum, it seems that at least a short *dʲ* might be missing from Jõgõperä. I included both *dʲ* and *dʲː* in core phonemes, but *dʲ* is placed in parentheses in Table 2 due to its unclear dialectal distribution.

Additionally, *nʲdʲ* (typical of Luutsa) varies with a consonant marked as *nʲː* (“ńń”) in Tsvetkov’s dictionary for Jõgõperä, cf. examples in 3.1.2 and above, as well as *eⁿʲːeˑ* ‘soul:GEN’, *eⁿʲːähtab* ‘breath:3SG’, as opposed to *enːě*, *enːepʲ* ‘before’ (Tsvetkov 1995, 30). Kettunen (1915, 51) mentions “e**n̯i**ähtäb ~ e**nń**ähtäb” for Luutsa and Jõgõperä, with a partially palatalised nasal, but “tše**nn**äd” without palatalisation, whereas Tsvetkov (1995, 349) cites “tšeńńed”. My data show that there is no clear distinction between palatalised and non-palatalised nasals in such words, at least in modern speech. Neither Luutsa nor Jõgõperä speakers pronounced any clear palatalisation in *t͡ʃʲenːɛd*, and one Jõgõperä speaker pronounced a half-palatalised nasal before a front vowel in [ˈenʲnːe] ‘soul:GEN’ (vs [enː] ‘before’). Therefore, *nʲː* was not included in core phonemes (cf. 3.2.4 on processes related to palatalisation).

3.1.4. Phonemes *l* – *lː*, *lʲ* – *lʲː*, and *tʲ*.

Votic laterals are discussed by Markus & Rozhanskiy (2026; cf. also Markus & Rožanskij 2011, 19–20). The key issue is that phonemes *l* – *lː* used to have two harmonic allophones: the velar one before back vowels and the alveolar/palatalised one before front vowels (similar to Soikkola Ingrian). This pattern still applies to Luutsa Votic, where there is no complete apocope of reduced vowels (see Rozhanskiy & Markus 2026). However, it becomes less clear in Jõgõperä Votic, where back and front schwa (marked as *ə* and *ɛ* here) often fully elide, allowing for quasi minimal pairs like [iʟm] ‘weather’ vs [si**l**(ʲ)m] ‘eye’ (in Luutsa, [iʟmə ~ iʟmă] vs [si**l**(ʲ)mɛ ~ si**l**(ʲ)mă̆] respectively).

The original alveolar/palatalised allophones do not phonologically contrast with the palatalised *lʲ* – *lʲː*: occurring before back vowels, including as a result of former or ongoing processes involving *j* and *i* (see 3.2.4), e.g. *paⁱlʲːo* ‘a lot’, *lʲuːkuma* (~ *lʲiukuma*) ‘slide:SUP downhill’. It is unclear whether there is any significant phonetic difference between these two categories, either. Therefore, when former alveolar/palatalised allophones appear in non-conditioned contexts due to vowel reduction and loss, they are written as palatalised *lʲ* – *lʲː*, i.e. *ilm* vs *silʲm*. If non-conditioned alveolar/palatalised allophones are considered as palatalised phonemes, the question arises: what phonemes should be postulated before non-reducing front vowels, as in [si**l**(ʲ)mä ~ si**l**(ʲ)mäˑ] ‘eye:PRT’? In my view, alveolar and palatalised laterals in vowel-conditioned contexts might also be regarded as palatalised phonemes. Impressions suggest no difference in the quality of laterals between, for

example, *väl'*:*ä* 'out' (<**lj*) and *t͡ʃ'ül'*:*ä* 'village:PRT' (original alveolar lateral), where in the first case, Rozhanskiy and Markus assume a palatalised /l'ː/ (cf. Markus & Rožanskij 2011, 20) and in the second case, apparently, a palatal /lː/ (cf. *t͡ʃ'ül'ä* 'village' with a short /l/ in Rozhanskiy & Markus 2026). However, more phonetic research is required, and palatal and velar laterals may sometimes freely vary in Votic in sandhi with the initial vowel of the following word (see Rozhanskiy & Markus 2026). In the transcription adopted here, palatalisation of laterals is marked only if not conditioned by the following front vowel.

At least *l'*: may be included in core phonemes, due to a series of Lower Luga Votic words where Proto-Finnic **lj* before back vowels consistently surfaces as *l'*:, without any variation, e.g. *pal'*:*o* cited above, *kal'*:*uma* 'cry:SUP', *il'*:*əkoitta* 'quietly, slowly', *(t͡ʃ'ivi)kal'*:*o* 'rock' (also the name of a local spring in Jõgõperä). In some other words, however, synchronic variation between *l'j* and *l'*: is still observed, as in *nal'ja* ~ *nal'*:*a* 'joke:PRT', *nel'*:*ə* ~ *nel'jə* 'four'. Contrarily, short *l'* is considered in this study as a marginal phoneme, because I did not find similar clear cases where *l'* <**li* would consistently surface before back vowels in native non-expressive words (cf. variable *l'u*:*kuma* ~ *l'iukuma* cited above and 3.2.4).

Instead, at least one instance of consistent **ti* > *t'* was observed in the word *lait'o* 'sleigh', so *t'* was included in the core phonemes. Note also that *t'* in Russian loanwords can contrast with *t* in native words also before *e*, e.g. *mat'eri* 'tissue' (Rus. *мате́рия* / *materiâ*) vs *lähted* 'spring:PL'.

3.2. Marginal consonants

Robust marginal consonants in Lower Luga Votic include those found in expressive lexicon and loanwords (examples from my field data): *pos'o* ~ *pot'u* 'little boy' — *poz'o* ~ *pos'o* (GEN) — *pos'*:*oa* ~ *pot'*:*ua* (PRT), *vont'*:*u* 'silly', *Kuz'*:*a* 'Kuz'a:PRT' (personal name; Rus. *Ку́зя* / *Kuzâ*), *kon'o* 'lol' — *kon'*:*oa* (PRT), *i*:*l'o* ~ *i*:*l'u* 'the day of St. Elijah' (Rus. *Илья́* / *Il'jâ*), *Mar'o* 'Mary' — *Mar'*:*oa* (PRT) 'Mary' (personal name; Rus. *Ма́рья* / *Mar'jâ*), *furskima* ~ *furskama* ~ *furk*:*ima* ~ *furt͡sima* 'snort, sniff:SUP' (onomatopoeic, but cf. also Rus. *фы́ркать* / *fyrkat'*, dial. *фурча́ть* / *furčat'*), *kof*:*ia*[19] 'coffee:PRT' (Rus. *ко́фе* / *kofe*), *grib*:*a* 'mushroom:ILL/PRT' (Rus. *гриб* / *grib*).

The sections below discuss marginal consonants with less clear status.

3.2.1. Phoneme *d*:. Phoneme *d*: is absent from the consonant inventory in Markus & Rožanskij (2011, 16). Unlike almost all other consonants, *d* usually does not undergo secondary gemination, e.g. *ahjo vo**d**a* 'furnace vault' (Luutsa; Tsvetkov 1995, 327 mentions *zvoda* for Jõgõperä) (Rus. *свод* / *svod*) remains *voda* in partitive and in illative, where singletons normally geminate (see 2.3.1).

Viitso (1961, 164), however, mentions *ho**d**:ua* 'move:PRT' (from *hodu;* Rus. *ход* / *xod*) with gemination for Luutsa Votic. I attested also *be**d**:a* 'trouble:PRT' (from *beda*; Rus. *беда́* / *beda*) from a Luutsa speaker and *paraho**d**:a* 'steamship:PRT' (from *parahoda;* Rus. *пароха́д* / *paroxod*) from a Jõgõperä speaker (along with a more regular *paraho**d**a* from two Luutsa and one Jõgõperä speaker). Geminated *d*: might either be an Ingrian influence (Lower Luga Ingrian

[19]For Luutsa Votic, Viitso (1961, 164) mentions also *ga*:*f*:*ali* 'gaff (nautical)' (possibly from Middle Low German *gaffel* or later, cf. https://en.wiktionary.org/wiki/Reconstruction:Proto-West_Germanic/gabulu), where long *f*: occurs before a back vowel.

regularly geminates *d*ː, cf. an example in 4.2.1) or an internal analogical development (cf. with *d͡ʒʲ* in 3.1.2). This unit is included in marginal phonemes with unclear status.

3.2.2. Phonemic *f*-types. Unlike Soikkola Ingrian, Votic contains not only loanwords but also some expressive (onomatopoeic) words with *f*-types, cf. ***f****urskima* 'snort:SUP' (and many other variants in 3.2 and in Grünberg 2013, 225–227), also ***f****utʲku* '1) spotted crake (*Porzana porzana*); 2) pejorative nickname of Liivčülä inhabitants by Jõgõperä inhabitants' (Grünberg 2013, 227). In some cases, there is variation between initial *f* and *v* in such words, cf. ***f****uris*ː ~ ***f****urisa* 'make:SUP noise during rotation, flying' in earlier sources (Tsvetkov 1995, 41; Grünberg 2013, 227), but ***v****urizəb* 'make:3SG noise', produced by all five speakers in my data. Also word-internally, especially in older loanwords, sometimes there is a variation between *f*-types and *hv*, e.g. *kof****f****ːi* ~ *kohvi* 'coffee' (Rus. *кóфе / kofe*) or *u****f****atk* (Tsvetkov 1995, 373) ~ *u****f****ːatka* (the main variant in Grünberg 2013, 1430) ~ *u****hv****ətkă* (my data from two Luutsa speakers) 'fork for taking pots out of furnace' (Rus. dial. *ухва́тка / uxvatka*).

Unlike Ingrian (both Soikkola and Lower Luga), *f* and *f*ː (in loanwords) are also a regular part of grade alternation in Votic, e.g. *sli****f****koj* 'cream:PL:PRT' — *sli****v****gəd* 'cream:PL' (Rus. *сли́вки / slivki*) (Markus & Rožanskij 2011, 52), *ʃar****f****ːi* 'scarf' — *ʃar****f****ːi˙* 'scarf:GEN' (Rus. *шарф / šarf*; one Jõgõperä and two Luutsa consultants in my data).

Votic also contains Germanic loanwords preserving original their *f*-types, e.g. ***f****rovːă* ~ ***f****rovːə* ~ ***f****rov*ː 'madam, lady' (cf. *provːa* ~ *rovːa* in Soikkola Ingrian and *rouva* in Finnish). This word appears to have been borrowed after the Middle Low German period (*vrûwe, vrouwe*[20]), because Middle Low German loanwords would rather contain *pl* < **fl* in Votic, as in ***p****läkːi* 'spot, stain'.[21] Kuleshov (p.c.) attributes this word (along with the parallel *herra* 'lord, master') to the late 16th–18th cc., possibly from Old High German *frouwa, frowa* (cf. footnote 20), *hër*.[22] Phonemic *f*-types also appear in the nautical terminology of Germanic origin in Votic, e.g. *gaː****f****ːali* 'gaff (nautical)' (see footnote 19). However, it remains unclear whether and which of this terminology came to Votic directly from Germanic languages, as Viitso (1961, 155) suggests, or later through Russian.

As further discussed in 3.2.3, palatalised *fʲ* can be considered as an (unstable) marginal phoneme, but there are no examples where long *fʲ*ː would appear before back vowels.

3.2.3. Unstable palatalised consonants in loanwords. Several palatalised consonants before back vowels are documented in my Russian loanword data, but they still remain classified as marginal consonants with less definite status. In some cases, it was unclear during the interview whether a loanword was genuinely an established borrowing or a result of code mixing

[20]Cf. https://en.wiktionary.org/wiki/rouva, https://en.wiktionary.org/wiki/Reconstruction:Proto-Germanic/frawjǭ.

[21]In Soikkola Ingrian *pläkːi* ~ *plekːi*; in Lower Luga Ingrian *pläkːĭ* (see 4.2.1), in Estonian *plekk,* in Finnish *pläkki* ~ *läkki;* cf. Middle Low German *flek, vlek, vlak, vlack* in https://en.wiktionary.org/wiki/Reconstruction:Proto-Germanic/flekka-; also https://en.wiktionary.org/wiki/pläkki.

[22]Middle Low German form would be *hê*, *hî*, cf., https://en.wiktionary.org/wiki/Reconstruction:Proto-Germanic/hiz, also https://en.wiktionary.org/wiki/herra.

or switching. Additionally, it was sometimes uncertain whether the following vowel was front or back (the same variation is also observed in Votic dictionaries).

These phonemes include labial and velar palatalised consonants: ***m*ʲ*aukib*** ~ ***m*ʲ*aukub*** ~ *n*ʲ*aukub* 'meow:3SG' (Rus. *мяу́кать / **m**âukat'*),[23] ***s*ʋʲ*at:oi*** ~ ***s*ʋʲ*atoi*** 'saint' (Rus. *свято́й / svâtoj*), ***t*ʲ*u*ʃʲ*ak:ə*** ~ ***t*ʲ*u*ʃʲ*ak:ă*** ~ ***t*ʲ*ufäk:a*** (in the expression *siä õ:d niku **t**ʲ**u**ʃʲ**ak:ə*** 'you are so clumsy', from Rus. *мюфя́к / tûfâk* 'traditional mattress'), ***le*ʃʲ*oʃk*** 'flatbread' (Rus. *лепёшка / lepëška*),[24] ***li*ʲ*ora*** ~ *liköra* ~ *likörä* 'liquor:PRT' (Rus. *ликёр / likër*). The most consistent combination of a palatalised consonant and a back vowel occurred in cases of *ʋʲ* and *ʃʲ*.

3.2.4. Unstable palatalised consonants resulting from ongoing sound changes. The highest number of marginal consonants with unclear status in modern Lower Luga Votic is linked to the processes of reduction and loss of front vowels and *j* triggering the emergence of palatalised consonants no longer conditioned by the context.

One of such processes, a transformation of Proto-Finnic **Cj/Ci* [Cj] into *C*ʲ*:*, *C:*, or *C*(ʲ)*C*ʲ, is apparently very old in Votic (unlike in Lower Luga Ingrian, cf. 4.2.6 below). Some of the consonants which emerged along this path in Votic dialects were included in core phonemes, cf. **sj* > *ʒ:* ~ *z*ʲ*:* in Lower Luga Votic (also *z:*, *zd*ʲ in other Votic varieties, cf. 3.1.1), **nk* > *nj* > *n*ʲ*d*ʲ ~ *n:*(ʲ) (and other clusters with **k*, cf. 3.1.3), **ij* ~ *j:* > *d*ʲ*:* ~ *j:* (3.1.3), **lj* > *l*ʲ*:* (3.1.4). Some older Russian loanwords also contain relatively robust palatalised consonants that evolved this way, e.g. **platja* > *pla**t**ʲ**:ə*** 'dress' (Rus. *пла́тье / plat'je*), recorded from one Luutsa and one Jõgõperä speaker, the latter also producing a plural form *pla**t**ʲ**ad*** with gradation.

A transformation of **CiV* [CiV] into *C*ʲ*V:* appears as a more recent process because it manifests some synchronic variation, cf. the uniformly produced *lai**t**ʲ**o*** (<**laitio*) 'sleigh', but [**l**ʲu:kuma] 'slide:SUP downhill', which was spelled in clear speech as [ˈ**l**ʲu:kuma] by one Luutsa and one Jõgõperä speaker, but as [ˈ**l**ʲiukuma] by another Luutsa speaker (see also 3.1.4). In *lai**t**ʲ**o***, we can consider the sound change as complete and therefore can include *t*ʲ in the core consonants. However, in ***l**ʲ**u:kuma*** ~ ***l**ʲ**iukuma***, the sound change is apparently still ongoing, so it is unclear whether the palatalised consonant can be considered as already unconditioned by the front vowel or not yet.

A further issue is that in some consonant types, palatalisation can be further lost, as e.g. discussed in 3.1.3 for *t͡ʃʲen:ɛd* (< [*kenjät] < **kenkät*). Tsvetkov (1995, 349) mentions "tšeńńed", with a palatalised consonant, for this word (and similar words, cf. 3.1.2, 3.1.3), but in my data the nasal was either non-palatalised (in two Luutsa speakers) or palatalised only very weakly (in one Jõgõperä speaker). Variability between a palatalised and a non-palatalised variant can appear both among speakers and in the pronunciation of the same speaker. These factors lead to the conclusion that the inventory of palatalised consonants cannot be established for Lower Luga Votic (as well as for Lower Luga Ingrian, cf. 4.2.6) with certainty, and Table 2 presents the maximal possible inventory.

[23]Note that Tsvetkov (1995, 169, 191) provides parallel forms *m*ʲ*auk:uma* and *n*ʲ*auk:uma* 'meow:SUP', while in Lower Luga Ingrian only forms with *n*ʲ were attested (see 4.2.5). It is, therefore, unclear whether and to which extent Votic forms with *m*ʲ are a result of recent Russian influence.

[24]Note also ***p**ʲ**otr*** 'Peter' (Rus. *Пётр / Pëtr)* in Tsvetkov (1995, 241).

The palatalised consonants with uncertain status, which do not appear outside the context of (potentially transient) results of ongoing processes, include the following labial and velar units in Votic: *pʲ, pʲː, bʲ, vʲː, mʲː, hʲ, kʲ, kʲː*. Many other consonants, which appear in expressive lexemes and loanwords as shown in earlier sections, participate in these processes as well, exhibiting the same kind of unstable palatalisation in this particular context. The only exceptions are *lʲ, lʲː*, which never velarise after vowel loss (cf. 3.1.4). Examples of variability are found in the following cases (individual attested pronunciations of tokens are provided; palatalised consonants are only marked in bold before back vowels, *e*, or word-finally):

(a) ongoing transformation of the **C(ː)iV* and **C(ː)jV* sequences into *Cʲ(ː)V* (variation between speakers is illustrated): [ˈlapːiata] ~ [lappʲiat(ə)] 'shovel:PRT' (cf. Fin. *lapio*) — [la**pʲ**ːajt(ə)] (PL:PRT), [ˈtikːiˌtakː] 'greater burdock (*Arctium lappa*)' (cf. Fin. *isotakiainen*) — [ˈtikːiˌtak**kʲ**$_{(i)}$a] (PRT), [ˈaŋgʲĕrʲi ~ ˈaŋgʲe**rʲ**es ~ ˈaŋge**rʲ**a] 'eel' (cf. Fin. *ankerias*) — [ˈaŋgə**rʲ**a ~ ˈaŋgʲe**rʲ**at ~ ˈaŋgʲe**rʲ**ata] (PRT) — [ˈaŋgʲerʲit ~ ˈaŋgʲe**rʲ**ajt ~ ˈaŋgʲe**rʲ**ajt ~ ˈaŋgʲ(e)**rʲ**ejt] (PL:PRT);

(b) consonant palatalisation retained after the final vowel reduction into *ə* and its ultimate loss, i.e. *CV̆* (in Votic, usually only *ä*, rarely *e* and, in certain grammaticalised suffixes, *i*)[25] > *Cʲə* > *Cʲ* > *C* (both inter-speaker and intra-speaker variation is illustrated): [ˈlʲeɨpʲă ~ lʲeɨpʲɛ ~ lʲeɨ**pʲ**] 'bread', [pʲiä**bʲ** ~ pʲiäb] 'must:3SG' (from the verbal suffix **-pi*), [ˈteräv**ʲ** ~ teräv] 'sharp' (from **terävä*), [ˈt͡ʃʲävʲːă̆ ~ ˈt͡ʃʲäüvːă̆ ~ ˈt͡ʃʲävʲːɛ ~ ˈt͡ʃʲäü**vʲ** ~ t͡ʃʲävː ~ t͡ʃʲäüvː] 'goːINF' (cf. Fin *käydä*), [ˈmüːhʲɛ ~ müː**hʲ** ~ müːh] 'late', [ˈit͡ʃʲem**ʲ** ~ ˈit͡ʃʲem] 'gum', [ˈlʲehʲmă̆ ~ lʲehm] 'cow', [ˈämːă̆ ~ ˈämʲːɛ ~ ämː] 'grandmother',[26] [näimʲːɛ ~ näimː] 'see:1PL'.

In case (b), labial consonants are particularly susceptible to losing palatalisation after vowel loss. In Jõgõperä, where vowel loss is more advanced, there were many tokens in my data with word-final non-palatalised *v, vː, m, mː, b* remaining after the loss of the following front vowel, even in words with harmonic front vowels in the stem. In my data, one of the two Jõgõperä speakers almost never produced these consonants as palatalised, typically realising variants like [t͡ʃʲäüvː].

3.3. Elements not considered as phonemes

Some palatalised sounds were not included in Table 2, because they were observed before the front vowels that do not undergo loss. These units include [fːʲ], [hːʲ], and [ʃːʲ]. The sound [fːʲ] was attested, for example, in [ˈkofʲːia] 'coffee:PRT', [ˈʃarfʲːi] 'scarf' cited below (long *f*ː before non-front vowels was not recorded in my data at all, but see footnote 19), while [hːʲ] appeared in [ˈlʲühʲːe] 'short:GEN'.

Regarding palatalised *ʃ*-types, Markus & Rožanskij (2011, 16–18) included *ʃʲ, ʃːʲ* as separate phonemes in Votic. Their decision was based on the fact that, in Russian loanwords, variability between a palatalised and a non-palatalised variant may occur, as seen in [ˈpereˌvo**ʃʲ**ːikː ~ ˈpereˌvo**ʃ**ːikː] 'ferryman' (Rus. *перево́зчик / perevozčik*), whereas in native words like *karju**ʃ**ːi*

[25] Cf. also e.g. *üväs**ʲ**ːi ~ üvä**sʲ**ː ~ üväsː* 'well' (Markus & Rožanskij 2011, 46).

[26] Tsvetkov (1995, 420) mentions variation *ämːʲə ~ ämːɛ* in Jõgõperä. The words 'cow' and 'grandmother' may be opposed e.g. to *rihmə ~ rihm* 'rope' and *tamːă ~ tamːə* 'mare' with plain consonants.

'shepherd' only non-palatalised *ʃ*-types are present. My data do not quite confirm this, however, because I attested some variability between [ʃː] and [ʃʲː] also in native lexemes like *korkuʃːi* 'three-spined stickleback (*Gasterosteus aculeatus*)'. In fact, *ʃ*-types in words like *korkuʃːi* and *perevoʃːik* were pronounced very similarly in my data.

Additionally, I did not attest loanwords with short [ʃʲ]. Markus & Rožanskij (2011, 18) mention *jaʃʲik* 'wooden box' (Rus. *я́щик / âšik*), but I recorded variants [ˈjaʃːikːə] (one Luutsa speaker) ~ [ˈjaʃʲːikː] (one Jõgõperä speaker) with a long consonant. Note that Rozhanskiy (2010, 64) also wrote this word with a long non-palatalised *ʃ*-sound and that Russian *ʃʲ* is actually phonetically pronounced as a long sound [ʃːʲ] (cf. Popov 2004, 45). The Jõgõperä variant [ˈjaʃʲːikː] is the only token in my data where a palatalised *ʃ*-sound occurred in a Russian loanword, and I consider it rather as a case of code-mixing. In sum, [ʃʲ, ʃʲː] are not seen as independent phonemes in the present description.

4. LOWER LUGA INGRIAN CONSONANTS

Table 3 shows the consonant inventory of Lower Luga Ingrian (revised from Kuznecova 2009, 170). As discussed in 1.1, Lower Luga Ingrian is a variety born in contact between Ingrians and Votes, the main mechanism being Votes shifting into Ingrian. As a result, the Lower Luga Ingrian consonant inventory is structurally closer to that of Lower Luga Votic than to that of Soikkola and other core Ingrian dialects.

Table 3. Lower Luga Ingrian consonants

		Place						
Manner		Labial		Dental		Post-alveolar	Velar	
		Plain	Palatalised	Plain	Palatal(ised)		Plain	Palatalised
Stops	voiceless	**p p:**	(pʲ pʲː)	**t t:**	tʲ tʲː		**k k:**	(kʲ kʲː)
	voiced	**b b:**	(bʲ)	**d d:**	dʲ dʲː		**g**	(gʲ)
Fricatives	voiceless	**f f:**	(fʲː)	**s s:**	sʲ sʲː	**ʃ ʃː**	**h h:**	(hʲ hʲː)
	voiced			**z z:**	(zʲ zʲː)	**ʒ ʒː**		
Affricates					**t͡sʲ**	**t͡ʃ t͡ʃː**		
Glides		**v vː**	(vʲ vʲː)		**j jː**			
Nasals		**m mː**	(mʲ mʲː)	**n nː**	nʲ nʲː			
Laterals				**l lː**	lʲ lʲː			
Trills				**r rː**	rʲ (rʲː)			

In particular, nearly all the dimensions that differentiate Lower Luga Votic from Soikkola Ingrian (listed at the beginning of Section 3) are characteristic also of Lower Luga Ingrian. These dimensions are detailed in the subsections of 4.2. The main structural differences between Lower Luga Ingrian and Votic concern the sub-system of affricates. Lower Luga Ingrian lacks the voiced affricate *d͡ʒʲ*, does not have a contrast in quantity for the *t͡s*-type affricate, and realises the *t͡ʃ*-type affricates as non-palatalised (in contrast to both Votic and Soikkola Ingrian). Additionally, the following phonemes are considered here as CORE consonants for Votic, because they can occur in non-expressive native lexicon, but as MARGINAL consonants for Lower Luga Ingrian: *ʃ*, *ʃ*ː, *ʒ*, *ʒ*ː, *z*ː, *t͡s*ː (Votic *t͡s* may be roughly equated to Lower Luga Ingrian *t͡sʲ*), *t͡ʃʲ*, *t͡ʃʲ*ː (the latter two being non-palatalised in Lower Luga Ingrian), *tʲ*, *dʲ*, *dʲ*ː, *lʲ*ː, *zʲ*ː. Finally, Lower Luga Ingrian has more marginal consonants with unclear phonological status (marked as plain in parentheses) than Lower Luga Votic or Soikkola Ingrian, and all of these are palatalised consonants.

4.1. Core consonants

Examples of core consonants in Lower Luga Ingrian (in the positions of quantity and voicing contrasts; see further comments in 4.2.1) include: *na*ː***p***ərĭ ~ *na*ː***p***ərĭ ~ *na*ː***p***uri ‘neighbor’, *na****p***ːŭ ‘button’, *la****b****j(ə)* ‘shovel’, *sa****t****a* ‘hundred’, *ha****t***ːŭ ‘hat’, *a****d****r(ə)* ‘plough’, *ma****k****u* ‘taste’, *i****k***ː*un(ə)* ‘window’, *sʲe*ː***g****əl(ə)* ‘sieve’, *rü****s****ä* ‘fishing net (type)’, *ka****s***ː*e* ‘dew’, *ra****z****v(ə)* ‘grease, fat’, *va****h****a* ‘wax’, *ra****h***ː*a* ~ *ra****h***ː*a* ‘money:ILL/PRT’, *ja****m****ak*ː*(ə)* ‘soured milk product’, *la****m***ː*əz* ‘lamb’, *ka****n****a* ‘hen’, *li****n***ː*(ə)* ‘city’, *ka****l****a* ‘fish’, *ve****l***ː*o* ~ *ve****l***ːŏ ‘brother’, *ha****r****av(ə)* ‘rake’, *ke****r***ː*əks* ‘once’, *ka****v****al* ‘cunning’, *a****v***ː*ajn* ‘key’, *opet*ː***aj****a* ‘teacher’, *opet*ː***aj***ːa ‘teacher:ILL/PRT’, *man****t͡sʲ****ikəz* ‘strawberry’.

4.2. Marginal consonants

Marginal consonants in Lower Luga Ingrian result from the same borrowing mechanisms and vowel reduction processes as outlined above for Lower Luga Votic.

Most of the marginal consonants in Lower Luga Votic and Lower Luga Ingrian are similar. It is also possible that some differences between palatalised consonants in Tables 2 and 3 may be due to gaps in description. Several important structural differences that still exist between the systems of core and marginal consonants in the three varieties discussed in this paper are addressed below.

4.2.1. Length and voicing in Lower Luga Ingrian vs lenis/fortis contrast in Soikkola. As discussed in 2.4, Soikkola Ingrian features a single contrast between lenis and fortis consonants, which are phonetically distinguished by duration, optionally voicing, or both, depending on their position (the only true voicing contrast is between *f*-types vs *v*-types, cf. 2.3.3). In Lower Luga Ingrian, however, binary quantity (since Q2 and Q3 are not distinguished, unlike in Soikkola) and voicing are two separate contrasts, similar to Votic. To illustrate the structural differences between Soikkola Ingrian, on the one hand, and Votic and Lower Luga Ingrian, on the other, consider the following two examples in phonological and phonetic transcription. The first one reflects Proto-Finnic **sata* ‘hundred’, while the second one includes adaptations of Russian *cað* / *sad* ‘garden’.

Variety	'hundred' (NOM – ILL)	'garden' (NOM – ILL)
Soikkola Ingrian	*sata* [ˈʃada:] – *satʼa:* [ˈʃatʼa:]	*sa:tu* [ˈʃa:du] – *sa:tʼu:* [ˈʃa:tʼu:]
Lower Luga Votic	*sata* [ˈsata:] – *sat:a(:)* [ˈsat:a(ˑ)]	*sa:du* [ˈsa:du] – *sa:du(:)* [ˈsa:du(ˑ)]
Lower Luga Ingrian	*sata* [ˈsata:] – *sat:a(:)* [ˈsat:a(:)]	*sada* [ˈsada:] – *sad:a(:)* [ˈsad:a(:)]

Voicing is generally more characteristic of Votic than of Lower Luga Ingrian due to a partially different system of stem gradation in the two varieties. In Lower Luga Votic, alternations between voiced and voiceless obstruents and affricates in native lexemes are very common (cf. 3.1). Conversely, for Lower Luga Ingrian, they have been attested only in the cluster **nk*, e.g. *reng**k**əz* 'circle' — *ren**g**ad* 'circle:PL', *pan**g**is:(ə)* 'bucket:PL:IN' — *pan**k**is:(ə)* ~ *pan**k**es:(ə)* 'bucket:IN'.

Moreover, short voiced obstruents are more common in Lower Luga Ingrian than the long voiced ones. Short consonants (especially *g*) may occur in native non-expressive words (see examples in 4.1). Certain voiced obstruents (*b, z, d*) occur only in clusters with inherently voiced consonants (sonorants and glides) in native words. However, as shown above, the Lower Luga Ingrian system generally manifests a contrast in voicing. Furthermore, if loanwords and expressive lexemes are taken into account, voiced obstruents can contrast with voiceless obstruents in certain positions, including in clusters with sonorants and glides: *a**d**r(ə)* 'plow' — *te**t**retʲ:ĭ* 'notebook' (Rus. *мемра́дь / tetradʲ*),[27] ***s**ve:ž(ə)* 'fresh' (Rus. *све́жий / svežij*) — ***z**ve:rĭ* 'animal' (Rus. *зверь / zverʹ*), ***p**läk:ĭ*[28] 'spot, stain' — ***b**lina* 'pancake' (Rus. *блин / blin*). Therefore, all voiced obstruents are regarded as independent phonemes (similar to Votic and unlike in Soikkola Ingrian), and those appearing in native non-expressive lexemes are included in the core phonemes listed in Table 3.

Long voiced obstruents, on the contrary, are rare (long *g*: is absent from my data) and appear mostly in forms derived through secondary gemination (cf. 2.3.1) from loanwords. These are typically Russian loanwords, like *sa**d**a* → *sa**d**:a* above, or *gri**b**a* → *gri**b**:a* 'mushroom:ILL/PRT' (Rus. *гриб / grib*), *paravo**z**a* → *paravo**z**:a* 'locomotive:ILL/PRT' (Rus. *паровóз / parovoz*), *le**ʒ**i-n* 'lie-1SG' → *le**ʒ**:i* 'lie:3SG'. However, there are also older Germanic loans, such as *glazi* 'glass' → *glaz:ia* ~ *glaz:e* (PRT). Such loanwords appear in Lower Luga Ingrian in the form in which they were borrowed from Germanic languages to Votic, reflecting the same voiced or voiceless obstruents. For example, Votic has ***g**lazi* and ***g**laizə̆* (see 3.1), and Lower Luga Ingrian also manifests ***g**lazi* and ***g**laizə̆* ~ ***g**laizlə̆* (Nirvi 1971, 176), whereas Soikkola Ingrian features *klasi* [ˈkʟaʒi:] → *klasʼi(:)a* [ˈkʟaʃʼi(:)a] (PRT), parallel to the 'garden' above, and *klaiza* [ˈklaiʒa] (Nirvi 1971, 176). Loans from other Germanic periods, ***t**reng**ĭ*** 'worker at a farm or a ship' (with a voiceless

[27]If Lower Luga Ingrian variants with the total loss of *ə* in certain positions are considered, *tr* can also occur in native non-expressive lexemes, cf. *pe:nt(ə)rəs*:[ˈpʲe:n**t**rəs:] 'garden bed:IN'.

[28]Same as in Votic (see 3.2.2); apparently a Middle Low German loanword, cf. footnote 21.

initial and voiced middle obstruent; Tsvetkov 1995, 344; Nirvi 1971, 597)[29] and ***frov:ă*** (cf. 3.2.2), also look the same in Votic and in Lower Luga Ingrian (in Soikkola Ingrian, the second word is *prov:a* ~ *rov:a*, see 3.2.2). Other long voiced geminates are attested, for example, in a Votic substrate lexeme *va**d**ʲ:(ə)* ~ *va**d**ʲ:alajn* 'Vote', and in words such as *koz**ʲ**:oma* 'woo:SUP', a result of late palatalisation development (see 4.2.5).

Phonemes *f* and *f:*, opposed to *v* and *v:*, are a regular part of the length and voicing correlations in Lower Luga Ingrian. In the positions of quantity contrast, a long *f:* usually appears (like in Votic), e.g. *sar**f:**ĭ* 'scarf', *ko**f:**e* 'coffee:ILL' (Rus. *кóфе / kofe*). In tetrasyllabic words, inter-speaker variation between short and long consonants is documented. For example, two speakers pronounced the following words with *f*, while two others used *f:*, i.e. *sara**f**ana* ~ *sara**f:**ana* 'traditional women's dress' (Rus. *сарафáн / sarafan*), *tʲele**f**ona* ~ *tʲele**f:**ona* 'telephone' (Rus. *телефóн / telefon*). The same kind of variation between *f, v,* and *hv* in older loanwords as seen in Soikkola Ingrian and in Votic was also attested in Lower Luga Ingrian in a few cases, e.g. "uffatka" ~ "ūvotka" ~ "uhvotka" 'a two-pronged iron fork attached to a wooden handle used to lift pots from the oven' (Rus. dial. *yxвáтка / uxvatka*) (Nirvi 1971, 620); in my data, *u**f:**ətkə* ~ *u**hv**atka* ~ *u**hv**ət:*.

4.2.2. Contrasts of *s*-types and *ʃ*-types. Soikkola Ingrian does not distinguish between *s*-types and *ʃ*-types (see 2.3.4), whereas in Lower Luga Ingrian, *s*-types and *ʃ*-types are independent phonemes, similar to Votic (cf. Ariste 1960, 50–51; Laanest 1966a, 54–58; Laanest 1987, 291–293; Kuznecova 2009, 235–240). They can be contrasted in many positions, but *ʃ*-types appear only in expressive and borrowed lexemes, e.g. ***s**ika* 'pig' vs ***ʃ**ikertel:öd* 'chirp:3PL' (in Soikkola Ingrian, *t͡ʃʲ* often appears in such onomatopoeic words, cf. *t͡ʃʲirk:ulain* 'sparrow' in 2.3). As further discussed in 4.2.3, word-initial *ʃ* in south Lower Luga Ingrian is apparently part of Votic phonological influence. Some words with *ʃ*-types, e.g. ***ʃ**iʃk(ə)* ~ ***ʃ**iʃkă* 'rag' widely occurring across the dialect, are part of the substrate Votic lexicon in Lower Luga Ingrian, cf. *ʃiʃk(a)* in Lower Luga Votic (see 3.1, also Grünberg 2013, 1232), but *hohla* in Soikkola Ingrian (Nirvi 1971, 65; *siska* was attested only in Säätina, one of the southernmost villages bordering on Votic, p. 528).

Another Lower Luga Ingrian example, *mer**ʃ:**ŭ* 'foodsack put on the horse's muzzle to feed it while walking', is potentially a Germanic loanword (a similar word *merssi* exists also in Soikkola Ingrian in a sense 'sack made of net', cf. Nirvi 1971, 306). Nirvi (ibid.) compares it to Estonian *märss* 'satchel' (cf. also Finnish *märssy* 'top (sail)'[30]) and marks it as a Swedish loanword. However, Swedish *märs* 'top (sail), crow's nest', as well as German *Mars*, are themselves considered Middle Low German loans, originating from *marse*, *merse*.[31] The word could have carried a

[29]Cf. with Proto-Germanic *drangijaz* (https://en.wiktionary.org/wiki/Reconstruction:Proto-Germanic/drangijaz), appearing only in Proto-West Germanic and Old Norse (and their descendants). It is unclear, when and how this word was borrowed to Finnic languages, including Votic *trengi* (Grünberg 2013, 1303–1304), Soikkola Ingrian *trenki* and Lower Luga Ingrian *trengi* (Nirvi 1971, 597), Finnish and Karelian *renki*; Finnish also *trenki* in the 16th c. texts (https://en.wiktionary.org/wiki/renki). Kuleshov (p.c.) considers it a loanword from the late Viking Age, but, given the initial cluster *tr* (rather than *dr*) in Votic, it could have entered at least Votic later and through North Finnic languages.

[30]Cf. https://en.wiktionary.org/wiki/märssy.

[31]Cf. http://germanic-studies.org/Middle-Low-German-loanwords-in-the-Scandinavian-languages.htm, https://en.wiktionary.org/wiki/Mars#Etymology_2_3.

generic meaning of 'top' in Middle Low German and been borrowed into different languages, where it acquired specific meanings. Lower Luga Votic dictionaries do not attest this word, but the presence of *ʃ*: in the Lower Luga Ingrian lexeme alludes to its possible Votic origin in this Ingrian variety, much like in the case of *glazi* and other similar words discussed in 4.2.1. However, this case would require further studies.

In summary, the contrast between *s*-types and *ʃ*-types in Lower Luga Ingrian appears to be a manifestation of Votic substrate phonological influence on Ingrian, originally unrelated to Russian influence, but later reinforced by the influx of Russian loanwords.

4.2.3. Contrasts of *t͡s*-types. The subsystem of affricates is what differs the most between the three varieties in general and between Lower Luga Ingrian and Votic in particular (cf. the beginning of Section 4). This subsystem in Lower Luga Ingrian presents major challenges for phonological interpretation.

First, Lower Luga Ingrian has short phonetically palatalised [t͡sʲ] in at least one native lexeme: *man**t͡sʲ**ikːa* ~ *man**t͡sʲ**ikː(ə)* ~ *man**t͡sʲ**ikəz* ~ *man**t͡sʲ**ikajz* 'strawberry'. The brevity of this affricate in Lower Luga Ingria is marked also in some variants attested in Nirvi (1971, 296) as a syllable boundary marker before it, e.g. "man‿**ts**iǩka". The same short palatalised [t͡sʲ] occurs word-initially in at least one Russian loanword ***t͡sʲ**epa* (cf. with *t͡ʃʲepa*, *manʧʲːikːain* in Soikkola Ingrian in 2.3.4). In Jõgõperä Votic and in more eastern Votic varieties, 'strawberry' is *maːzikəz* ~ *maːzikːainə*; while the variants *mansikaz* and *man͡sikaz* were attested only in Luutsa and Liivčülä Votic (Tsvetkov 1995, 171; Grünberg 2013, 687); I also recorded *man**t͡sʲ**ikəz* from one Luutsa Votic speaker. The variant with [t͡sʲ] may represent an independent Lower Luga Ingrian development and its secondary contact influence on Luutsa and Liivčülä Votic.

The word-internal short palatalised [t͡sʲ] in 'strawberry' contrasts with a long [t͡sː] ~ a cluster [ts] appearing in all other words, including e.g. [ˈme**t͡s**ː(ə)] 'forest', expressive lexemes like [ˈLa**t͡s**ːərkə ~ ˈLa**t͡s**ːorˌkajn] 'flat', and numerous loanwords like [ˈtan**t͡s**ːu̯] 'dance' (Rus. *máнец* / *tanec*). Kuznecova (2009, 273–281), who analysed Lower Luga affricates in detail, found that the average duration of [t͡sʲ] was about 40 ms shorter than the duration of [t͡sː ~ ts] when measured in comparable structures. In particular, the duration was 131 ms (SD 25.7 ms, 16 tokens) in trisyllabic structures like [ˈman**t͡s**ʲiˌkːasː] 'strawberry:IN' vs 173 ms (SD 45.2 ms, 12 tokens) in structures like [ˈtan**t͡s**ːuˈsːa] 'dance:ILL:until'. Furthermore, the duration was 152 ms (SD 18.2 ms, 7 tokens) in disyllabic structures like [ˈman**t͡s**ʲikː] 'strawberry' vs 193 ms (SD 43.2 ms, 24 tokens) in structures like [ˈtan**t͡s**ːuLː] 'dance:AD'.[32] Based on these (still preliminary) data, it has been established that at least the short palatalised [t͡sʲ] can be considered an independent phoneme *t͡sʲ* in Lower Luga Ingrian. Since it occurs in at least one native non-expressive lexeme, it was included in the core consonants.

The long [t͡sː] varies phonetically with the cluster [ts], sometimes within the same words or roots. No specific phonetic investigations have been carried out here yet. However, impressionistically, more cluster-like phonetic realisations tend to occur more frequently before a front vowel, when the second element is palatalised while the first one is not. This is particularly evident in shorter structures, where the duration of [t͡sː] ~ [ts] is longer, e.g. [ˈvʲe‿i‿**ts**ʲĭ̯] 'knife',

[32]The duration of Soikkola Ingrian cluster *ts* was about 170–200 ms before a short vowel in disyllabic forms in the study by Rožanskij & Markus (2020, 475), which is comparable to the duration of [t͡sː ~ ts] in Lower Luga Ingrian.

[ˈkʲitsʲi̯] ‘she-goat’ show only cluster-like realisations, but there is variation in longer structures, where sound durations are more compressed, such as [ˈtan**tsʲ**i(ˌ)ma] ‘dance:SUP’ ~ [ˈtan**t͡sʲ**:i(ˌ)tʲ:i] ‘dance:IPS:PST’.

Lower Luga Ingrian does not display a phonological contrast between [t͡s:] and [ts], nor a grade alternation mechanism that would indicate a functional connection between the long and the short affricate (unlike in Votic *ku**t͡s**:uma* — *ku**ts**un* in 3.1). Since cluster-like phonetic realisations are common and following Occam’s razor principle of economy, [t͡s:] ~ [ts] are regarded here as representing the cluster *ts* (or *ts*^{j} before an elided front vowel).

4.2.4. Contrasts of *t͡ʃ*-types. The interpretation of *t͡ʃ*-types is also far from straightforward. In Soikkola Ingrian, *t͡ʃ*^{j} and *t͡ʃ*^{j}*:* are phonetically palatalised, whereas in Lower Luga Ingrian, they used to be non-palatalised. Nirvi’s (1971) dictionary renders Soikkola *t͡ʃ*-types as palatalised “tš́”, while the Lower Luga ones as non-palatalised “tš”, and Laanest (1986, 40) also highlights this difference.[33] Word-initially, *t͡ʃ*-types are much less typical of Lower Luga Ingrian than of Soikkola Ingrian, especially in the south. In southern varieties, word-initial *ʃ* usually occurs in native expressive lexemes rather than *t͡ʃ* (the same was observed by Mägiste 1925, 39–40). Consequently, a variation like ***t͡ʃ**ik:ara* (north) ~ ***ʃ**ik:ěrĭ* (south) ‘curl’, ***t͡ʃ**ikertä:* (north) ~ ***ʃ**ikertel:ö* (south) ‘chirp:3SG’ is typical across Lower Luga varieties. However, in Russian loans, word-initial *t͡ʃ* is typical in all varieties, e.g. ***t͡ʃ**as:ona* ~ ***t͡ʃ**asona* ~ ***t͡ʃ**as:on* ‘chapel’ (Rus. *часо́вня / časovnâ*).

The presence of word-initial *ʃ* in southern varieties might apparently be attributed to the Votic substrate phonological influence. Due to the sound change of **k* into *t͡ʃ*^{j} and *t͡ʃ*^{j}*:* (cf. footnote 10), *t͡ʃ*^{j}-types are not exclusively connected to sound symbolism (and to loanwords) in Votic, unlike in Ingrian. This could explain why *ʃ*-type sounds — rather than *t͡ʃ*^{j}-types — are so common in Votic expressive lexemes, both at the beginning and inside words, e.g. ***ʃ**igrimä* ‘rummage:SUP (in the ground), tossle:SUP (hair)’ (cognate of the Ingrian word for ‘curl’ above, cf. Tsvetkov 1995, 318–319), ***ʃ**ulʲp:uma* ‘splash:SUP’, *mur**ʃ**:ima* ‘hustle:SUP, make:SUP a mess’ from my data, *mu**ʃ**:u* – *mu**ʃ**u* (NOM – GEN), ‘stupid’ from Tsvetkov (1995, 181) and various examples with word-initial *ʃ* in Tsvetkov (1995, 317–322) and Viitso (1961, 165).

Word-internally, a contrast between short *t͡ʃ* and long *t͡ʃ:* seems to be marginally possible in Lower Luga Ingrian (cf. Kuznecova 2009, 280–286). In general, their distribution is nearly complementary, similar to Soikkola Ingrian *t͡ʃ*^{j} and *t͡ʃ*^{j}*:* (see 2.3.4). Long *t͡ʃ:* occurs in the positions of quantity contrasts in loanwords, e.g. *mä**t͡ʃ**:ik:ä* ~ *mä**t͡ʃ**:ik:(ə)* (Rus. *мя́чик/mâčik*), *klʲa:**t͡ʃ**:(ə)* ‘old horse’ (Rus. *кля́ча / klâča*), *u**t͡ʃ**:ĭtʲelʲa* ‘teacher’ (Rus. *учи́тель/učitel’*), while no native lexemes with *t͡ʃ:* in these positions have been attested. Instead, in the positions of quantity neutralisation, short *t͡ʃ* occurs. Usually, these cases involve expressive lexemes like northern ***t͡ʃ**ik:ara* ‘curl’ mentioned above or loanwords like *bo**t͡ʃ**ka* ~ *bo**t͡ʃ**k(ə)* ‘barrel’ (Rus. *бо́чка / bočka*). However, *t͡ʃ* has been attested in at least one native-like word in the archive recordings from the 1960s in a speech of a north Lower Luga Ingrian speaker, Irina Filimonova (born in 1904 in Pärspää, Rus. *Ли́пово / Lipovo*): [ˈha:ro**t͡ʃ**ˈpu:] ‘a pole with one or more forks at the top, on which nets are stretched during production’. However, Nirvi (1971, 49) provides the variant *har**t͡ʃ**:upu:*, apparently from the same speaker or her son (see data on speakers in Nirvi 1971, X). In Soikkola

[33] Under the influence of Russian, which only has palatalised *t͡ʃ*^{j}, palatalised *t͡ʃ*-types may also sometimes occur in modern Lower Luga Ingrian, freely varying with non-palatalised *t͡ʃ*-types.

Ingrian, the closest counterpart is *hark:i* (in Nirvi 1971, 49 and in my field data). Therefore, the precise phonological shape of this word in Lower Luga Ingrian and the origins of the *t͡ʃ*-affricate in it remain uncertain.

Instead, a couple of likely older loanwords from local Russian dialects seem to contain a short *t͡ʃ* intervocalically, i.e. in the position of the quantity contrast. The first one, attested in the literature, is *brägä**t͡ʃ*** 'buckthorn (*Frangula*)' (Laanest 1986, 40), cf. also "bräga‹**tš**imarjanpū" (Nirvi 1971, 26), where the syllable boundary marker before the affricate indicates its brevity (Nirvi compares this word to Rus. dial. *бредня́га / bredn̂aga*). The second one was attested in audio data: as *brä**t͡ʃ***ägä in the same northern idiolects of Filimonova and her son from the archive recordings (both of them discussing how *gouru**ʃ**:id,* small dark fishes, could be caught there) and as *brä**t͡ʃ***äg^{j}*(ə)* 'a small pond or pit of muddy, silty water; swale' in my field data (one speaker from Kotko). All five of my Lower Luga Votic consultants also knew this word (as *brä**t͡ʃ***äg in Jõgõperä and *br^{j}a**t͡ʃ^{j}***ag in Luutsa), because there was a particular toponymic object in Jõgõperä called in this way (the Kotko Lower Luga speaker, who lives in Kukkuzi, was also familiar with it). Apparently, this word derives from something like Russian dialectal *брячаг(а) / brâčag(a)* (modern Standard Russian has only *бочáг / bočag* in a similar meaning).

Measurements in Kuznecova (2009, 282–284) showed that the *t͡ʃ* in this Lower Luga Ingrian word is much shorter (105 ms, SD 14.2 ms, 8 tokens) than the long *t͡ʃ:* in structurally comparable recent Russian loanwords like *mä**t͡ʃ**:ik:ă̆ ~ mä**t͡ʃ**:ik:(ə)* and *u**t͡ʃ**:ĭt^{j}el^{j}a* (167 ms, SD 20.3 ms, 19 tokens). Additionally, a short post-stress vowel is not reduced, but instead undergoes lengthening in *brätʃägä ~ brätʃäg^{j}(ə)* (97 ms, SD 21.5 ms, 8 tokens), which is typical after a light stressed CV syllable (cf. 2.3.4). Conversely, in words like *mät͡ʃ:ik:ă̆ ~ mät͡ʃ:ik:(ə)* and *ut͡ʃ:ĭt^{j}el^{j}a*, where the first syllable is heavy CVC, the post-stress vowel undergoes expected reduction (39 ms, SD 9.9 ms, 15 tokens; in 4 out of 5 tokens of *ut͡ʃ:ĭt^{j}el^{j}a* the post-stress vowel elided completely).

Therefore, a contrast between *t͡ʃ* and *t͡ʃ:* appears to exist in the positions of quantity distinctions in Lower Luga Ingrian, unlike in Soikkola Ingrian, where only long *t͡ʃ^{j}:* occurs in these contexts (see 2.3.4). The difference from Lower Luga Votic, in turn, is that *t͡ʃ:* and *t͡ʃ* (like *ts* and *t͡s*) in Lower Luga Ingrian do not participate in grade alternation (cf. with Votic *pi**t͡ʃ^{j}**:i — pi**t͡ʃ^{j}**ä* in 3.1). I tentatively consider not only *t͡ʃ*, but also *t͡ʃ:* as monophonemic, because, impressionistically, no cluster-like realisations were attested for *t͡ʃ:* (unlike for *ts*), which potentially correlates with the fact that *t͡ʃ*-types are not palatalised in Lower Luga Ingrian, unlike *t͡s*-types. Additionally, *t͡ʃ:* was much shorter than *ts*, when measured in maximally comparable structures like [ˈmä**t͡ʃ**:ik:] 'ball' vs [ˈme**ts**əs:] 'forest:IN' from the same speaker (a male from Vanakülä): 171 ms (SD 16.5 ms, 11 tokens) vs 223 ms (SD 21.1 ms, 15 tokens) (Kuznecova 2009, 279, 283), although more phonetic studies would be needed here.

4.2.5. Robust palatalised consonants. Robust palatalised consonants, present in all Lower Luga varieties, appear before non-front vowels in expressive lexemes and loanwords. These are the same instances where palatalised consonants occur also in Soikkola Ingrian (cf. 2.2), e.g.: *ʃu**l^{j}**paht^{j}ĭ* 'flop:PST:3SG down', *ru**l^{j}**:a* 'steering wheel:ILL/PRT' (Rus. *рупь / rul'*), *po**t^{j}**ok:(ə) ~ po**t^{j}**ok:ajn* 'little boy', *Pe**t^{j}**:a* 'Pet^{j}a:ILL/PRT' (personal name; Rus. *Пётя / Petâ*), *se**d^{j}**olk(ə)* 'back pad (part of horse harness)' (Rus. *седёлка / sedëlka*), *Fe**d^{j}**:a* 'Fed^{j}a:ILL/PRT' (personal name; Rus. *Фёдя / Fedâ*), ***n^{j}**auk:a ~ **n^{j}**auk:u* 'meow:3SG', *Va**n^{j}**:a* 'Van^{j}a:ILL/PRT' (personal name; Rus. *Ва́ня /*

Vanâ), *ves^{j}ol(ə)* ‘merry, cheerful’ (Rus. *весёлый / vesëlyj*), *Vas^{j}:a* ‘Vas^{j}a:ILL/PRT’ (a personal name; Rus. *Вáся / Vasâ*), *lar^{j}ok:ă ~ lar^{j}k(ə)* ‘kiosk’ (Rus *ларёк / larëk*).

4.2.6. Unstable palatalised consonants. Palatalised phonemes with unclear status result from the same processes of reduction and loss of front vowels and *j*, as outlined for Lower Luga Votic in 3.1.4. A greater number of unstable palatalised consonants in Lower Luga Ingrian compared to Votic is related to two factors. First, in Votic, many such consonants exist not only due to these processes but also for other reasons (see e.g. 3.1.1 on *z^{j}:*, 3.2 on *z^{j}*). Second, individual Lower Luga Ingrian varieties may have very different inventories of palatalised consonants. In northern varieties, where vowel reduction is weak, the number of palatalised consonants is comparable to that in Soikkola Ingrian. In southern varieties, where vowel reduction is the strongest, almost all consonants (both long and short) have a palatalised correlate. Table 3 presents the maximal inventory, reflecting the situation in southern varieties.

Unstable palatalised consonants (placed in parentheses in Table 3) include variability, for example, in the following cases (individual attested pronunciations of tokens are provided; palatalised consonants are only marked in bold before back vowels or word-finally):

(a) ongoing transformation of **C(:)iV* and **C(:)jV* into *C^{j}(:)V*, which in some cases also involves the loss of palatalisation, i.e. a further change into *C(:)V* (variation between speakers is illustrated): [ˈko**z^{j}**:oma ~ ˈko**z^{j}**:uma ~ ˈkoi̯zuma] (sic!) ‘woo:SUP’ (cf. Fin. *kosiomaan*), [ˈso**r^{j}**:a ~ ˈsor:e][34] ‘coarse’ (cf. Fin. *sorea*), [ˈma**k^{j}**:a ~ ˈmak^{j}:e ~ ˈmak:e] ‘sweet’ (cf. Fin. *makea*), [ˈvaʟk^{j}iad ~ ˈvaʟk^{j}ed] ‘white:PL’ — [ˈvaʟ**k^{j}**aj(t) ~ ˈvaʟk^{j}ej(t)] ‘white:PL:PRT’ (cf. Fin. **valkea*), [ˈsup^{j}ĭ:i] ‘soup’ — [ˈsup^{j}:ia ~ ˈsu**p^{j}**:a] ‘soup:PRT’, [ˈpaʟ^{j}jo ~ ˈpa**ʟ^{j}**:o ~ ˈpa**ʟ^{j}**:ŏ ~ ˈpa**ʟ^{j}**:u ~ ˈpa**ʟ^{j}**:ŭ ~ pa**ʟ^{j}**:] ‘a lot’ (cf. Fin. *paljon*);

(b) consonant palatalisation kept after the reduction of a vowel into *ə* and its ultimate loss, i.e. *CV̆* (*ä, i, ü*, rarely *e, ö*) > *C^{j}ə* > *C^{j}* > *C*, e.g. (inter- and intra-speaker variation is illustrated): [ˈsö:p^{j}i ~ süö**p^{j}** ~ sö:**b^{j}**] ‘eat:3SG’, [ˈbrät͡ʃä˙gä] ~ [ˈbrät͡ʃä˙**g^{j}**] ‘pit of muddy water’, [ˈma:**z^{j}**ə] ‘cream’ (Rus. *мазь / maz^{j}*), [ˈu:s̯̆̊ ~ u:**z^{j}**] ‘new’, [ˈ^{j}ehmä ~ ^{j}e**h^{j}m^{j}** ~ ^{j}ehm] ‘cow’, [ˈsüötə**m^{j}**: ~ ˈsüötəm:] ‘feed:1PL’ (from earlier Ingrian **sö:täm:ä*), [ˈv^{j}ikah**^{j}**:ĕ ~ ˈv^{j}ika**h^{j}**:] ‘scythe’ (cf. Fin. *viikate*), [ter**v^{j}**^{ə}] ‘healthy’ (from **terve*), [ˈsü**v^{j}**:ə ~ sö**v^{j}**: ~ s^{j}öv: ~ süv:] ‘eat:INF’ (from **süv:ä ~ söv:ä*), [ˈkof:ĩ̯ ~ ˈkof^{j}:ĕ ~ ko**f^{j}**:] ‘coffee’.

Similar to Lower Luga Votic, consonants in case (b), i.e. word-finally, are more prone to lose palatalisation than those occurring word-internally in case (a). Consonants *l^{j}, l^{j}:* never depalatalise, as in Votic. Like in Votic (3.1.4) and unlike in Soikkola Ingrian (2.2), these alveolar/palatalised *l^{j}, l^{j}:*, which have resulted from vowel loss, are here phonemically identified with palatalised *l^{j}, l^{j}:* found in other contexts (such as in expressive and borrowed words).

For *t^{j}, t^{j}:*, the situation varies across Lower Luga Ingrian sub-dialects, forming one of the intradialectal isoglosses (see Kuznecova 2009, 210–213; Kuznetsova & Verkhodanova 2019, 227). In many central Lower Luga varieties, *t* and *t:* before *i, ü* undergo only weak secondary palatalisation (especially in their eastern part, where Lower Luga Ingrian is in direct contact with

[34]Proto-Finnic **Ceta /*Cetä* and later Proto-Ingrian **Cea/*Ceä* or **Cia/*Ciä* sequences (cf. Laanest 1986, 91–92) have remained as such or are being transformed into *C^{j}:a/C^{j}:ä* in the varieties on the right bank of the Luga river, while the sound change into *C^{j}:e ~ C:e* is characteristic of its left bank.

Votic), whereas both in northern and in southern varieties, primary palatalisation of *t* and *t:* in this position is observed, e.g. *lustʲi* (north), *lustʲˇi* (south) — *lustˇi* (central east) ~ *lustʲˇi* (central west). However, this palatalisation does not happen before the *i, ü* that are the result of a transformation of the original long **e:, *ö:* into *i:, ü:* (processes typical of all Ingrian varieties) or diphthongisation into *üö* and *ie* (found mostly in Lower Luga Ingrian on the left bank of the Luga river, under Finnish influence). Therefore, minimal pairs such as ***tʲ**üö* 'come:IMP' (<**tule*, cf. 2.2 on Soikkola Ingrian) vs ***t**üö* (work, from **tö:*) are possible in certain Lower Luga varieties (this pair was noticed by a speaker of south Lower Luga Ingrian from Vanakülä on the left bank of the Luga river, cf. Kuznecova 2020, 281).

5. DISCUSSION

5.1. Comparison of the Soikkola Ingrian, Lower Luga Ingrian and Votic consonant systems

Contacts between Votic and Ingrian, which led to the development of the Lower Luga dialect of Ingrian, had a significant influence on its consonant system. Lower Luga Ingrian can be metaphorically described as Ingrian spoken with a strong Votic accent.

Generally, the consonant inventory of Lower Luga Ingrian resembles that of Lower Luga Votic more than it does the consonant inventories of core Ingrian dialects (exemplified with Soikkola Ingrian in this paper). The similarities between the consonant systems of Lower Luga Ingrian and Lower Luga Votic are seen in several features common to both, setting them apart from Soikkola Ingrian:

(a) a binary quantity contrast instead of a ternary one;
(b) independent contrasts in voicing and length instead of a single lenis/fortis contrast;
(c) independent phonemic *s*-types and *ʃ*-types instead of a single sibilant type;
(d) a much richer system of affricates (although the Lower Luga Ingrian system and the Votic system are partially different in this respect);
(e) reduction and loss of unstressed vowels and *j* in certain prosodic positions, which leads to a significant expansion of the inventory of palatalised consonants (many of these consonants having an unclear status due to ongoing vowel loss processes).

The three systems examined here also exhibit some significant structural differences in both core and marginal consonants. The consonants identified as core in this work (i.e., occurring in the non-expressive native lexicon) include 31 units in Soikkola Ingrian, 41 units in Lower Luga Votic, and 27 units in Lower Luga Ingrian.

On the one hand, Soikkola Ingrian has short (secondary/Q2) geminate consonants as independent phonemes. The other two varieties only have a binary contrast of consonants because secondary geminates have merged with original geminates (Q3 class in Soikkola Ingrian) into a single class of geminates, contrasted with singletons (Q1). On the other hand, Lower Luga Ingrian and Votic have developed phonemic voiced obstruents in opposition to voiceless obstruents, as well as phonemic *s*-types as opposed to *ʃ*-types, unlike Soikkola Ingrian, and have enriched their subsystems of affricates.

However, Lower Luga Votic has notably more voiced obstruents, ʃ-types, and affricates within its core phonemes than Lower Luga Ingrian. Reasons for this include a different system of consonant gradation (which involves voiced obstruents, including affricates, in Votic), the presence of ʃ-types in the non-expressive native lexicon synchronically, and certain sound changes absent in Ingrian: the development of *dʲ*, *dʲː*, *d͡ʒʲ*, *ʒː*, *t͡ʃʲ*, *t͡ʃʲː*, *zʲː*, *tʲ*, *lʲ*: from specific palatalised or palatal sounds. Concerning the status and number of voiced obstruents, ʃ-types, and affricates, Lower Luga Ingrian thus occupies an intermediate position between Soikkola Ingrian and Lower Luga Votic.

The list below provides a numerical estimate of the similarities and differences between the core consonant inventories of the three varieties in question. It presents pairwise comparisons of shared core consonants between the varieties (with Lower Luga Ingrian and Votic geminates equated to Soikkola Ingrian long Q3 geminates) and the number of consonants that differ in each pair (present in one system but absent from the other).

(a) Soikkola Ingrian and Lower Luga Votic: 20 shared, 11 only in Soikkola, 21 only in Votic;
(b) Soikkola Ingrian and Lower Luga Ingrian: 19 shared; 12 only in Soikkola, 8 only in Lower Luga Ingrian;
(c) Lower Luga Ingrian and Lower Luga Votic: 27 shared; 0 only in Lower Luga Ingrian; 14 only in Votic.

Soikkola Ingrian has almost the same number of shared and different core consonants with both Votic and Lower Luga Ingrian, and the types of these consonants are nearly identical. The differences between 20 vs 19 "shared" and 11 vs 12 "only in Soikkola" in pairs (a) and (b) arise only from the presence of *t͡ʃʲː* in a few native Soikkola Ingrian lexemes (*mant͡ʃʲːikːain* 'strawberry' and some others) but of *t͡sʲ* in *man**t͡sʲ**ikəs* 'strawberry' in Lower Luga Ingrian (see 4.2.3). These differences are, therefore, minor.

Lower Luga Ingrian shares all its core consonants with Lower Luga Votic. However, the reverse is not true, as Votic has 14 additional core consonants that Lower Luga Ingrian lacks. On the other hand, Votic has 21 core consonants not shared with Soikkola Ingrian.

These estimates provide a rough measure of the direction of contacts in the varieties in question. Soikkola Ingrian and Votic are maximally distant in terms of core consonantal inventories. That the system of Lower Luga Ingrian core consonants results from Votic influence is reflected in the fact that Lower Luga Ingrian shares all its core consonants with Votic. However, a series of contrasts related to voicing, ʃ-types, and affricates was not transferred into the native non-expressive lexicon of Lower Luga Ingrian from Votic, partly due to deep systemic differences in consonant gradation and in the organisation of *s/ʃ* contrasts between (core) Ingrian and Votic. Note also that the core consonants present only in Votic are not the most central to the Votic native non-expressive lexicon either (they typically appear either in inflected forms, or in very few lexemes, sometimes even not in all Lower Luga Votic varieties). Consequently, they had fewer chances to be transferred to Lower Luga Ingrian.

On the other hand, the original complex system of ternary quantity, which is present in core Ingrian varieties, was also simplified to binary in Lower Luga Ingrian, a common development for ternary quantity contrasts in general (Blevins 2004, 202; Remijsen & Gilley 2008, 340–341; Kuznetsova 2015, 105–107) and apparently, at least partially, a language convergence development in the Lower Luga area (Laanest 1966a, 28; 1986, 147-149; Kuznecova 2015, 229-232). Overall, therefore, Lower Luga Ingrian has the smallest inventory of core consonants (27 units)

among the three varieties. This aligns with the fact that structural simplification is particularly typical of contact varieties (Trudgill 2001; Gil, Trudgill & Sampson 2009).

Regarding marginal phonemes, Lower Luga Ingrian also shows much greater similarity to Lower Luga Votic than to Soikkola Ingrian. The consonantal periphery of Soikkola Ingrian (20 units) includes dental palatalised consonants, such as the short $\widehat{tʃ^j}$, which is in complementary distribution with the long $\widehat{tʃ^j}$ː, borrowed *f*-types, and tentative long glides *j*ː and *v*ː, which have emerged as syllable boundary signals following the loss of intervocalic stops. Notably, only few marginal phonemes — the two long glides, the two long palatalised phonemes (l^jː, s^jː), and short $\widehat{tʃ^j}$ — have unclear phonological status across the entire lexicon, including loanwords and expressive lexemes.

The sub-system of marginal consonants in Lower Luga Votic, with 24 comparable units, is not much larger than that in Soikkola. By comparison, this study identified a total of 38 similar marginal units in Lower Luga Ingrian. The overall typology of marginal phoneme inventories is quite similar in the two Lower Luga varieties: it includes certain palatalised consonants, *f*-types, and, in Lower Luga Ingrian, also *ʃ*-types and affricates.

Lower Luga Ingrian still stands out against both Soikkola Ingrian and Lower Luga Votic in the ratio of core to marginal consonants within its system. It has the greatest number of marginal consonants among the three varieties, while the number of core consonants is the lowest (38 vs 27). Notably, the number of marginal phonemes in Lower Luga Ingrian exceeds the number of core phonemes, unlike in Soikkola Ingrian (20 vs 31) and Votic (24 vs 41).

There seem to be two main reasons for this. First, many of those phonemes that can be considered core in Votic but do not exist in Soikkola Ingrian (most affricates, *ʃ*-types, some voiced obstruents) also appear in Lower Luga Ingrian, but as marginal rather than core phonemes. Most of these have likely been borrowed into Lower Luga Ingrian from Votic rather than from Russian, although subsequent Russian influences have further reinforced the phonological status of these units. At least one phoneme ($\widehat{ts^j}$) might also have developed independently in Lower Luga Ingrian.

Regarding the palatalised consonants, apparently none of them were directly adopted by Lower Luga Ingrian from Votic as core phonemes, i.e. along with the native lexicon. Their rise is related to the expressive lexicon and loanwords, as well as to the internal processes of vowel shortening, reduction, and loss, which are prominent in Votic but much weaker in Soikkola Ingrian. Lower Luga Ingrian exhibits high variability in this regard, with reduction becoming stronger from north to south within the dialect (Kuznecova 2009, 93–104). The northernmost system reflects the extent of reduction as in Soikkola Ingrian, the central system resembles that of Lower Luga Votic (particularly of Luutsa-Liivčülä), and the southernmost system exhibits almost complete loss of unstable short vowels and entire shortening of formerly long vowels (this is akin to Jõgõperä Votic, but even more radical). Due to greater variability in Lower Luga Ingrian than in Votic regarding these processes, the status of the emergent palatalised consonants appears generally less definite in Lower Luga Ingrian than in Votic.

5.2. Stability of the Lower Luga Ingrian system and implications for theory and typology

As mentioned, many reduction processes leading to the formation of palatalised consonants are still ongoing in Lower Luga Ingrian. Additionally, the formation of affricates might still be

incomplete. Conversely, as stated, many word-final consonants tend to lose palatalisation after the following vowel is lost. The most stable palatalised consonants are only dental consonants, the same types as in Soikkola Ingrian and Votic.

The current consonant inventory of Lower Luga Ingrian appears, therefore, relatively unstable and moving towards further simplification. Generally, such a contact-induced inventory with a "hyper-inflated" periphery of ambiguous marginal phonemes might simply be a short-lived phase in the history of a vital language. However, the Ingrian language is nearly extinct, so there is a high likelihood that this stage will be the one documented in descriptions of Lower Luga Ingrian and used for typological comparisons as a linguistic rarity (cf. Kuznetsova & Markus 2022). Typology, therefore, should consider sociolinguistic factors related to language contact and loss more carefully. Otherwise, these factors might cause biases in the data, leading to controversial statements such as "small indigenous languages concentrate more linguistic rarities" (cf. Anderson, Kuznetsova & Easterday 2025).

6. CONCLUSION

This study explored the system of core and marginal consonants in Lower Luga Ingrian, a contact variety between two closely related languages, Ingrian and Votic. This contact system was compared with one of the core Ingrian systems (Soikkola Ingrian) and with the Lower Luga Votic system along the same criteria and on comparable material. The consonant inventory of Lower Luga Ingrian generally resembles that of Lower Luga Votic more than that of Soikkola Ingrian, due to a series of systemic innovations transferred from Votic to Lower Luga Ingrian.

The inventory of Lower Luga Ingrian still exhibits peculiar traits likely related to its specific status as a contact variety. Notably, the sub-system of core consonants is the simplest among the three varieties considered, while the sub-system of marginal consonants is the largest in Lower Luga Ingrian. The smaller core consonant system results from the simplification of some original Ingrian contrasts, such as the rare ternary quantity, and the failure to borrow certain contrasts from Votic into the Lower Luga native lexicon. Most phonemes present in Votic but absent from Ingrian have still entered the Lower Luga Ingrian inventory, but chiefly as marginal rather than core consonants — that is, only within expressive lexicon and loanwords. Furthermore, Lower Luga Ingrian has experienced similar vowel reduction and loss processes as Votic, which have led to the development of new palatalised consonants. Consequently, the number of marginal consonants in Lower Luga Ingrian surpasses the number of core consonants — which distinguishes it from both Soikkola Ingrian and Lower Luga Votic.

This inventory seems unstable and prone to further simplifications, which are unlikely to be completed due to language obsolescence. Because of high internal variability within Lower Luga Ingrian and numerous marginal phonemes with unclear status, creating a standard orthography for this dialect appears particularly challenging.

ABBREVIATIONS

1 – 1st person
3 – 3rd person

AD – adessive
COM – comitative
GEN – genitive
ILL – illative
IMP – imperative
IN – inessive
INF – infinitive
IPS – impersonal
PL – plural
PRT – partitive
PST – past
SG – singular
SUP – supine
TRL – translative
Fin. – Finnish
Rus. – Russian

REFERENCES

Agdzhoyan, Anastasia, Ponomarev Georgy, Pylev Vladimir, Autleva (Kagazezheva) Zhaneta, Gorin Igor, Evsyukov Igor, Pocheshkhova Elvira, Koshel Sergey, Kuleshov Viacheslav, Adamov Dmitry and Kuznetsova Natalia. 2024. The Finnic peoples of Russia: genetic structure inferred from genome-wide and Y-chromosome data. Genes 15(12). 1610. https://doi.org/10.3390/genes15121610

Anderson, Cormac, Kuznetsova Natalia and Easterday Shelece. 2025. Rarities in phonetics and phonology. In N. Kuznetsova, C. Anderson and S. Easterday (eds.) Rarities in phonetics and phonology: structural, typological, evolutionary, and social dimensions (Topics in Phonological Diversity 5). Berlin: Language Science Press. 3–28. https://doi.org/10.5281/zenodo.15148162

Ariste, Paul. 1947. Proishoždenije vodskogo jazyka [Origin of the Votic language]. In Filoloogilisi ettekandeid: peetud soome-ugri teaduste konverentsil Leningradis 1947. a. [Philological presentations: held at the conference of Finno-Ugric researchers in Leningrad in 1947] (Eesti NSV Tartu Riikliku Ülikooli Toimetised 4: Filoloogilised teadused, 3: Nõukogude Soome-Ugri Teadused). Tartu: RK "Teadlik kirjandus". 35–47.

Ariste, Paul. 1960. Isuri keelenäiteid [Ingrian language examples]. Keele ja kirjanduse instituudi uurimused V. 7–68.

Ariste, Paul. 1967. Über wotische Orstnamen. Sovetskoje finno-ugrovedenije 2. 77–84.

Ariste, Paul. 1968. A grammar of the Votic language. Bloomington, The Hague: Indiana University.

Bazell, Charles Ernest. 1952. Phonemic and morphemic analysis. Word 8(1). 33–38. https://doi.org/10.1080/00437956.1952.11659417

Bazell, Charles Ernest. 1954. The choice of criteria in structural linguistics. Word 10(2–3). 126–135. https://doi.org/10.1080/00437956.1954.11659518

Bentlin, Mikko. 2008. Niederdeutsch-finnische Sprachkontakte: der lexikalische Einfluß des Niederdeutschen auf die finnische Sprache während des Mittelalters und der frühen Neuzeit (Suomalais-Ugrilaisen

Seuran toimituksia = Mémoires de la Société Finno-Ougrienne 256). Helsinki: Suomalais-Ugrilainen Seura.

Blevins, Juliette. 2004. Evolutionary phonology: The emergence of sound patterns. Cambridge: Cambridge University Press.

Borkovskij, Viktor Ivanovič and Kuznecov Pëtr Savvič. 1963. Istoričeskaja grammatika russkogo jazyka [Historical grammar of the Russian language]. Moskva: Izdatel'stvo Akademii nauk SSSR.

Bybee, Joan. 2001. Phonology and language use. Cambridge: Cambridge University Press.

Culhane, Kirsten and Edwards Owen. 2025. Consonant epenthesis in Meto: Typologically rare but diachronically explicable. In N. Kuznetsova, C. Anderson and S. Easterday (eds.) Rarities in phonetics and phonology: structural, typological, evolutionary, and social dimensions (Topics in Phonological Diversity 5). Language Science Press. 165–217. https://doi.org/10.5281/ZENODO.15148170

Gil, David, Trudgill Peter and Sampson Geoffrey. 2009. Language complexity as an evolving variable. Oxford: Oxford University Press.

Gordon, Matthew K. 2009. Prominence and gemination in Ingrian. Linguistica Uralica 45(2). 81–100. https://doi.org/10.3176/lu.2009.2.01

Grünberg, Silja (ed.). 2013. Vadja keele sõnaraamat = Vad′d′aa tšeelee sõna-tširja = Slovar′ vodskogo jazyka [Dictionary of the Votic language]. 2., täiendatud ja parandatud trükk. Tallinn: Eesti keele sihtasutus; Eesti keele instituut.

Holopainen, Sampsa. 2025. Studies in early Indo-European loans in Uralic part II. Suomalais-Ugrilaisen Seuran Aikakauskirja 101. 83–111. https://doi.org/10.33340/susa.153435

Hvoš'inskaja Natalija Vadimovna. 2023. Ètnogenez i istoričeskaja sud′ba vodi [Ethnogenesis and the historical fate of Votes]. Arheologičeskije vesti (39). 204–219. https://doi.org/10.31600/1817-6976-2023-39-204-219

Juškova, Marija and Kulešov Viačeslav. 2011. Kyorstovo 1: a new burial ground of the period of Roman influences in North-Western Russia. Archaeologia Lituana 12. 99–121. https://doi.org/10.15388/ArchLit.2011.12.5136

Kallio, Petri, Metsäranta Niklas and Honkola Terhi. 2025. The language-contact landscape in the Baltics during the first millennium BC. Finnisch-Ugrische Forschungen (70). 27–80. https://doi.org/10.33339/fuf.147292

Kettunen, Lauri. 1915. Vatjan kielen äännehistoria [Historical phonetics of the Votic language]. Helsinki: Suomalaisen kirjallisuuden seura.

Kochetov, Alexei. 2011. Palatalisation. In M. van Ostendoorp, C. J. Ewen, K. Rice and E. Hume (eds.). The Blackwell companion to phonology 3. 1666–90.

Köppen, Peter von (ed.). 1849. Ethnographische Karte des St. Petersburgischen Gouvernements. St. Petersburg: K. Akademie der Wissenschaften.

Krämer, Martin and Urek Olga. 2016. Perspectives on palatalization. Glossa: A Journal of General Linguistics 1(1). 31.1–17. https://doi.org/10.5334/gjgl.226

Kroonen, Guus. 2013. Etymological dictionary of Proto-Germanic (Leiden Indo-European Etymological Dictionary Series II). Leiden: Brill.

Kuznecova, Natal'ja Viktorovna. 2009. Fonologičeskije sistemy ižorskih dialektov [Phonological systems of Ingrian dialects]. Diss. …kand. filol. nauk [Candidate of sciences thesis]. Institut lingvističeskih issledovanij RAN [Institute for Linguistic Studies, Russian Academy of Sciences], Sankt-Peterburg.

Kuznecova, Natal'ja Viktorovna. 2012. Prosodika slovoformy v nižnelužskom dialekte ižorskogo jazyka [Word prosody in the Lower Luga dialect of the Ingrian language]. Acta Linguistica Petropolitana XVIII(1). 43–103; 506–521.

Kuznecova, Natal'ja Viktorovna. 2014. Ob istorii, suš'nosti i izmerenijah fonemy [On the history, nature, and dimensions of the phoneme]. In V. F. Vydrin and N. V. Kuznecova (eds.) Ot Bikina do Bambal'umy, iz var'ag v greki: Ekspedicionnyje et'udy v čest' Eleny Vsevolodovny Perehval'skoj [From Bikin to Bambaluma, from Varangians to Greeks: Field sketches in honour of Elena Vsevolodovna Perehval'skaja]. Sankt-Peterburg: Nestor-Istorija; Institut lingvističeskih issledovanij RAN; Grečeskij institut Filologičeskogo fakul'teta SPbGU. 406–442.

Kuznecova, Natal'ja Viktorovna. 2015. Dve fonologičeskije redkosti v dialektah ižorskogo jazyka [Two phonological rarities in the dialects of the Ingrian language]. Acta Linguistica Petropolitana XI(2). 201–257.

Kuznecova, Natal'ja Viktorovna. 2020. Rukopisnyje teksty N. D. Põdera na nižnelužskom dialekte ižorskogo jazyka (r. Rosson') s kommentarijami [Manuscripts of N.D. Põder in the Lower Luga dialect of the Ingrian language (Rosona river) with commentary]. Acta Linguistica Petropolitana XVI(3). 184–288. https://doi.org/10.30842/alp2306573716309

Kuznecova, Natal'ja Viktorovna and Sidorkevič Dar'ja Viktorovna. 2012. Priloženije 6: Karty k stat'jam [Appendix 6: Maps to the articles]. Acta Linguistica Petropolitana VIII(1). 559–67.

Kuznetsova, Natalia. 2015. Two phonological rarities in Ingrian dialects. In M. Hilpert, J. O. Östman, C. Mertzlufft, M. Riessler and J. Duke (eds.) New trends in Nordic and general linguistics (Linguae & Litterae 42). Berlin, Boston: De Gruyter. 91–117. https://doi.org/10.1515/9783110346978.91

Kuznetsova, Natalia. 2016. Evolution of the non-initial vocalic length contrast across the Finnic varieties of Ingria and adjacent areas. Linguistica Uralica 52(1). 1–25. https://doi.org/10.3176/lu.2016.1.01

Kuznetsova, Natalia. 2023. Typology, chronology, and phonetic mechanisms of Finnic secondary gemination in light of Soikkola Ingrian acoustic data. In O. Niebuhr and M. S. Lundmark (eds.) Proceedings of the 13th Nordic Prosody Conference: Applied and Multimodal Prosody Research, Sønderborg, Denmark. Warsaw: Sciendo/de Gruyter. 71–84. https://doi.org/10.2478/9788366675728-005

Kuznetsova, Natalia, Balanovsky Oleg, Muslimov Mehmed and Kuleshov Viacheslav. 2022. Contact of Ingrians and Votes in the light of linguistic, archaeological and genetic data. Orally presented at the Congressus XIII Internationalis Fenno-Ugristarum, Aug 21–26, 2022. https://cifu13.univie.ac.at/programme/symposia/e2-the-interdisciplinary/.

Kuznetsova, Natalia, Brodskaya Irina and Markus Elena. 2023. Compensatory effects of foot structure in segmental durations of Soikkola Ingrian disyllables and trisyllables. Journal of Phonetics 100. 101246. https://doi.org/10.1016/j.wocn.2023.101246

Kuznetsova, Natalia, Cornelio Tiburcio Oscar and Uchihara Hiroto. To appear. Stress-induced lengthening outside stressed syllable nucleus in Finnic and Ayutla Tlapanec: a challenge for theory and typology. In H.-S. Jeon, J. Choi, J. Setter and L. White (eds.) Language and Speech (special issue): Interdisciplinary approaches to Speech Prosody.

Kuznetsova, Natalia and Markus Elena. 2022. Ongoing vowel shortening in vanishing Soikkola Ingrian: challenges for description, codification, and typology. In S. Frota and M. Vigário (eds.) Proceedings of the 11th International Conference on Speech Prosody, 23–26 May, 2022. Lisboa: Universidade de Lisboa. 327–331. https://doi.org/10.21437/SpeechProsody.2022-67

Kuznetsova, Natalia, Markus Elena and Muslimov Mehmed. 2015. Finnic minorities of Ingria: the current sociolinguistic situation and its background. In H. F. Marten, M. Rießler, J. Saarikivi and R. Toivanen (eds.) Cultural and linguistic minorities in the Russian Federation and the European Union: Comparative studies on equality and diversity (Multilingual Education 13). Berlin: Springer. 127–167. https://doi.org/10.1007/978-3-319-10455-3_6

Kuznetsova, Natalia and Verkhodanova Vasilisa. 2019. Phonetic realisation and phonemic categorisation of the final reduced corner vowels in the Finnic languages of Ingria. Phonetica 76(2–3). 201–233. https://doi.org/10.1159/000494927

Laakso, Johanna. 2022. Finnic: General introduction. In M. Bakró-Nagy, J. Laakso and E. Skribnik (eds.) The Oxford guide to the Uralic languages. Oxford: Oxford University Press. 240–253. https://doi.org/10.1093/oso/9780198767664.003.0014

Laanest, Arvo. 1966a. Ižorskije dialekty: Lingvogeografičeskoje issledovanije [Ingian dialects: Lingvogeographic study]. Tallin: Akademija nauk Estonskoj SSR [Academy of Sciences of the Estonian Republic].

Laanest, Arvo. 1966b. Ižorskij jazyk [Ingrian language]. In V. I. Lytkin and K. E. Majtinskaja (eds.) Jazyki narodov SSSR. III: Finno-ugorskije i samodijskije jazyki [Languages of the peoples of the USSR. III: Finno-Ugric and Samoyedic languages]. Moskva: Nauka. 102–117.

Laanest, Arvo. 1986. Isuri keele ajalooline foneetika ja morfoloogia [Historical phonetics and morphology of the Ingrian language]. Tallinn: Valgus; ENSV Teaduste Akadeemia, Keele ja Kirjanduse Instituut.

Laanest, Arvo. 1987. Zur Phonologie der Ingrischen [On the phonology of Ingrian]. In Studien zur Phonologie und Morphonologie der uralischen Sprachen: Eisenstadt, 28. Juni — 1. Juli 1984: Akten der dritten Tagung für uralische Phonologie. Wien: Verband der wissenschaftlichen Gesellschaften Österreichs. 285–293.

Mägiste, Julius. 1925. Rosona (Eesti Ingeri) murde pääjooned [Main features of the Rosona (Estonian Ingrian) dialect] (Acta et Commentationes Universitatis Tartuensis (Dorpatensis) B: Humaniora, VII), Vol. 3. Tartu: K.-Ü. "Postimehe" trükk.

Markus, Elena Borisovna and Rožanskij Fëdor Ivanovič. 2011. Sovremennyj vodskij jazyk: Teksty i grammatičeskij očerk. Tom II: Grammatičeskij očerk i bibliografija [Contemporary Votic language: Texts and grammatical sketch. V. II: Grammatical sketch and bibliography], Vol. 2. Sankt-Peterburg: Nestor-Istorija.

Markus, Elena and Rozhanskiy Fedor. 2012. Votic or Ingrian: new evidence on the Kukkuzi variety. Finnisch-Ugrische Mitteilungen 35. 77–95.

Muslimov, Mehmed Zakirovič. 2005. Jazykovyje kontakty v Zapadnoj Ingermanlandii (nižneje tečenije reki Lugi) [Language contacts in West Ingria (the lower reaches of the Luga river]. Diss. …kand. filol. nauk [Candidate of sciences thesis]. Institut lingvističeskih issledovanij RAN [Institute for Linguistic Studies, Russian Academy of Sciences], Sankt-Peterburg.

Muslimov, Mehmed Zakirovič. 2020. Ob idiome der. Kurovicy (Kukkõsi / Kukkuzi) [On the variety of the Kukkuzi village]. Acta Linguistica Petropolitana XVI(3). 330–376. https://doi.org/10.30842/alp2306573716312

Nirvi, Ruben Erik. 1971. Inkeroismurteiden sanakirja [The dictionary of Ingrian dialects] (Lexica Societatis Fenno-Ugricae 18). Helsinki: Suomalais-Ugrilainen Seura.

Popov, Mihail Borisovič. 2004. Problemy sinhroničeskoj i diahroničeskoj fonologii russkogo jazyka [Problems of synchronic and diachronic phonology of the Russian language]. Sankt-Peterburg: Filologičeskij fakul'tet SPbGU.

Posti, Lauri. 1980. Kukkosin murteesta ja sen asemasta [On the Kukkusi dialect and its status]. In L. Posti and S. Suhonen (eds.) Vatjan kielen Kukkosin murteen sanakirja [Dictionary of the Kukkusi dialect of the Votic language]. Helsinki: Suomalais-ugrilainen seura. XV–XXI.

R'abinin, Evgenij Aleksandrovič. 1997. Finno-ugorskije plemena v sostave Drevnej Rusi: K istorii slav'ano-finskih kul'turnyh sv'azej: Istoriko-arheologičeskie očerki [Finno-Ugric populations as part of Ancient

Rus': To the history of Slavic-Finnic cultural ties: Historical-archaeological sketches]. Sankt-Peterburg: Sankt-Peterburgskij gosudarstvennyj universitet; Institut istorii material'noj kul'tury RAN.

Remijsen, Bert and Gilley Leoma. 2008. Why are three-level vowel length systems rare? Insights from Dinka (Luanyjang dialect). Journal of Phonetics 36(2). 318–344. https://doi.org/10.1016/j.wocn.2007.09.002

Rožanskij, Fëdor Ivanovič. 2010. Ižorskij jazyk: problema opredelenija granic v uslovijah jazykovogo kontinuuma [The Ingrian language: a problem of defining boundaries in the context of a language continuum]. Voprosy jazykoznanija 1. 74–93.

Rožanskij, Fëdor Ivanovič and Markus Elena Borisovna. 2020. O fonologičeskom statuse affrikaty ts v sojkinskom dialekte ižorskogo jazyka [On the phonological status of the affricate ts in Soikkola Ingrian]. Acta Linguistica Petropolitana XVI(3). 440–479. https://doi.org/10.30842/alp2306573716315

Rozhanskiy, Fedor and Markus Elena. 2026. How do marginal phonemes emerge? Case studies from minor Finnic varieties. Acta Linguistica Academica 73(2). 267–285. https://doi.org/10.1556/2062.2025.01004

Rozhanskiy, Fedor and Zhivlov Mikhail. 2019. Votic and Ingrian core lexicon in the Finnic context: Swadesh lists of five related varieties. Linguistica Uralica 55(2). 81. https://doi.org/10.3176/lu.2019.2.01

Sovijärvi, Antti. 1944. Foneettis-äännehistoriallinen tutkimus Soikkolan inkeroismurteesta [Phonetic and historical-phonological study on the Soikkola Ingrian dialect] (Suomi-kirjan eripainoksia 5). Helsinki: Suomalaisen kirjallisuuden seura.

Stas'uk, Ivan Vadimovič. 2023. Proishoždenie i rann'aja istorija vodi: starye diskussii i novyje nahodki [Origin and early history of Votes: old discussions and new findings]. Etnografia 3(21). 68–89. https://doi.org/10.31250/2618-8600-2023-3(21)-68-89

Suhonen, Seppo. 1985. Wotisch oder Ingrisch? In W. Veenker (ed.) Dialectologia Uralica: Materialien des ersten internationalen Symposions zur Dialektologie der uralischen Sprachen 4.-7. September 1984 in Hamburg (Veröffentlichungen Der Societas Uralo-Altaica 20). Wiesbaden: Harrassowitz Verlag. 139–148.

Suni, Lev Val'terovič. 2003. Finny-ingermanlandcy: istoričeskij očerk [Ingrian Finns: historical sketch]. In E. I. Klement'jev and N. V. Šlygina (eds.) Pribaltijsko-finskije narody Rossii [Finnic peoples of Russia] (Narody i kul'tury). Moskva: Nauka. 469–484.

Trudgill, Peter. 2001. Contact and simplification: Historical baggage and directionality in linguistic change. Linguistic Typology 5(2–3). 371–374. https://doi.org/10.1515/lity.2001.002

Tsvetkov, Dmitri. 1995. In J. Laakso (ed.) Vatjan kielen Joenperän murteen sanasto [Vocabulary of the Jõgõperä dialect of the Votic language] (Lexica Societatis Fenno-Ugricae 25; Kotimaisten Kielten Tutkimuskeskuksen Julkaisuja 79). Helsinki: Suomalais-ugrilainen seura; Kotimaisten kielten tutkimuskeskus.

Viitso, Tiit-Rein. 1961. Vadja Luutsa-Liivčülä murraku fonoloogia [Phonology of the Luutsa-Liivčülä sub-dialect of Votic]. Emakeele seltsi aastaraamat 1961 VII. 142–174.

Viitso, Tiit-Rein. 1964. Hilisgeminatsioonist vadja keeles [On late gemination in the Votic language]. In E. Pajusalu-Adler (ed.) Töid läänemeresoome ja volga keelte alalt [Works on Finnic and Volga languages]. Tallinn: Keele ja kirjanduse instituut, Eesti NSV teaduste akadeemia. 24–35.