THE PRONOMINAL SYSTEM OF SOIKKOLA INGRIAN: PERSONAL, DEMONSTRATIVE, REFLEXIVE AND RECIPROCAL PRONOUNS AND THEIR VARIATION

Anna Schwarz
Fedor Rozhanskiy[a]
[a]University of Tartu, EE
annakhor@gmail.com, handarey@yahoo.com

Abstract. This research analyses the pronominal system of Soikkola Ingrian. The main goal of the article is descriptive, because existing publications on Soikkola Ingrian give only a superficial analysis of pronouns and there are no detailed descriptions based on data from the 21st century. The article focuses on both interspeaker and intraspeaker variation in four groups of Ingrian pronouns: personal, demonstrative, reflexive and reciprocal. The field data collected by the authors in the 21st century are compared with the data presented in grammars written in the 19th and 20th centuries. The research concludes that although the pronominal system has not changed significantly, there are a number of innovations in Ingrian pronominal forms and their variation. The most frequent type of variation is the variation between individual speakers, but not between sub-dialectal groups. Free variation is also common, while other types of variation are less relevant.

1. Introduction

The subject of this paper is the pronominal system of Soikkola Ingrian. In the existing literature, there is no detailed description of Soikkola Ingrian pronouns. Porkka (1885: 78–87) lists the main classes of pronouns and gives several paradigms in section "Pronomina", but since his description compares several varieties and a significant part of the section is dedicated to the possessive suffixes, the information

about Soikkola pronouns is very brief. Junus (1936: 28–29, 98–101) and Laanest (1978: 250–253, 1986: 119–121[1]) give only the most basic information. In Saar (2017), only a few pronouns are mentioned. Generally, the Ingrian pronominal system is a blind spot rather than a well-studied topic.

Pronouns are usually the most irregular part of the nominal system. They often combine archaic features with innovations, and this leads to variation of forms in their paradigms. Irregularities are usually specific to a particular pronoun or small group of pronouns and often they are disregarded in general descriptions of a language. For these reasons, a description of the Ingrian pronominal system is warranted.

This article is aimed at describing and analysing pronominal paradigms of contemporary Soikkola Ingrian.[2] We classify the Ingrian pronouns into the following groups:[3]

(1) personal pronouns
(2) demonstrative pronouns
(3) reflexive-possessive, reflexive, and reciprocal pronouns
(4) relative/interrogative pronouns
(5) negative pronouns
(6) indefinite pronouns

In the current paper, we analyse only the first three groups; the other three groups will be studied in our further publications.

In our research, a special focus is put on the variation of forms and their competition in the speech of Ingrian native speakers. In previous publications on Ingrian, variants of pronominal forms were either mentioned without discussing the underlying reasons for the variation or were not listed at all. The significance of variation in language was increasingly acknowledged in the second half of the 20th century starting from Labov (1972, 1979 and others) and up to various recent publications (e.g. Dufter, Fleischer & Seiler 2009, Drager 2015,

---

1   Laanest's doctoral thesis (Laanest 1978) was written in Russian but later its Estonian version (Laanest 1986) was published. Below we refer to the Russian variant but the same information can be easily found in the corresponding section of the Estonian version.

2   By contemporary Soikkola Ingrian we mean the language used by the native speakers in the 21st century.

3   Various quantifiers ('all', 'every', 'some', etc.) that are often considered together with pronouns constitute a separate class of words which we do not analyse.

Tamminga, MacKenzie & Embick 2016, Van Olmen, Mortelmans & Brisard 2019). As is often the case, theoretical research is ahead of the practical application of theory. Although variability is already recognized as a universal phenomenon (Metslang, Habicht & Hennoste 2020) and an important feature concerning various aspects of language (Reppen, Fitzmaurice & Biber 2002: VII, Kiesling 2011: 14–17, Siemund 2011: 2–3), language-specific grammatical descriptions rarely focus on variability. Diversity of the existing variants used to be mostly ignored and even if the variants were listed, no information about their functioning in the language and the relations between them was given. A contemporary approach to grammatical description requires information on how forms function in speech, so the analysis of variation becomes an essential component of a modern grammar. However, a standard for describing variation has not yet been formulated. Although there were some attempts to create a general theory of variation (see, for example, Leib 1993), there is no generally accepted system of variation that could be applied by a grammarian. As a result of this, we have produced a classification suitable for our data which distinguishes the following types of variation:

a. Dialectal variation.

Dialectal variation differentiates between variants, each typical of a particular dialect, and is a subtype of geographical variation. In its wider application, this notion also covers sub-dialectal variation that opposes smaller varieties (e.g. varieties of particular villages). In the case of contemporary Soikkola Ingrian, a detailed analysis of sub-dialects is problematic because often we have only one or two speakers that represent a particular village. Moreover, many speakers have moved away from their home villages, and it is not obvious which variety they represent. As a result, the border between sub-dialectal and idiolectal variation (the latter is described below) becomes vague. However, Soikkola Ingrian is a relatively homogeneous dialect compared to Lower Luga Ingrian, and we would not expect significant diversity between the varieties from neighboring villages. Instead, we assume that there might be differences between the main dialectal zones of Soikkola Ingrian. Kuznetsova (2009: 19) distinguishes the northern and southern zones – to the north and to the south of the village of Vistina, respectively – and

a transitional Vistina variety. Accordingly, we postulate dialectal varia-
tion when one variant is used by the representatives of the northern zone
and another variant is used by the representatives of the southern zone
(possibly with some fluctuation in the village of Vistina).

b. Idiolectal variation.

In the case of dialectal variation, every group of speakers using
a particular variant consists of representatives of the same territory
(a village, a dialectal zone, etc.). If there is no correlation between
groups and territories, we consider such variation as idiolectal.
Cross-linguistically, interspeaker variation without any correlation with
territories is often due to sociolinguistic reasons. In the case of Soikkola
Ingrian, however, we do not consider sociolinguistic variation. First, the
number of speakers remaining is too small to allow them to be divided
into representative groups with different sociolinguistic characteristics.
Besides, most of our speakers belong to the same society and generation,
and there is no significant difference in their social status (although they
vary in education level and linguistic biographies). In this situation, it is
more appropriate to consider the existing differences as individual and
therefore analyse them as idiolectal variation.

Dialectal and idiolectal variation are two types of interspeaker
variation. The other types come under intraspeaker variation.

c. Contextual variation. If the choice of a variant is defined by its
linguistic context at any level – phonetic, morphological, syntactic or
semantic – we identify this variation as contextual.

d. Register variation. If different variants appear in the speech of a
single speaker but their distribution depends on the speech register, it is
register variation. In our data, it is usually the difference between fluent
speech and careful speech. The latter appears when the native speaker
pronounces a sentence very distinctively. It often happens during
elicitation sessions, when at first some utterance is produced fluently
and then it is repeated more slowly and more distinctively.

e. Free variation. If different variants appear in the speech of a single
speaker without any motivation (i.e. there are no evident features that

condition the choice of a particular variant), this can be considered as free variation. Certain cases may belong to this type due to the lack of data that could reveal factors defining the distribution.

There are some other types of variation that we do not discuss in this article. Besides the abovementioned sociolinguistic variation, we do not consider cross-linguistic variation (we analyse only one language), lexical variation (we work with a very limited number of lexemes), or diachronic variation (although we refer to earlier sources on Ingrian, the analysis of diachronic variation requires a thorough study of all text collections, which is not possible in this article).

The listed types of variation are not mutually exclusive. Moreover, pure types of variation occur less often than mixed types. For example, one of the most frequent combinations is when there is free variation within idiolectal variation. In other words, variants are distributed among speakers, but for certain speakers variants are freely interchangeable. It is neither simple free variation, because most speakers prefer only a particular variant, nor simple idiolectal variation, because some speakers use more than one variant. A more sophisticated situation is a combination of three types of variation. For example, variants are mostly distributed among speakers (idiolectal variation) but there are two different groups of speakers: one group displays contextual or register variation while another group exhibits free variation.

It is often the case that one type of variation is dominant, and only traces of the other types are attested. For example, typically the choice of a variant is speaker-dependent (idiolectal variation) but it may be the case that one or two speakers use more than one variant without motivation (free variation). In this case, we would speak about idiolectal variation with elements of free variation.

This article has the following structure: section 2 describes the data and methods of analysis; sections 3, 4, and 5 are dedicated to different types of pronouns (personal, demonstrative and reflexive plus reciprocal, respectively); and section 6 contains the conclusions. Two appendices describe the experiments which we used for analysing two particular cases of variation.

2.   Data and methods

The main source of data used for the current research are the materials collected by the authors and their colleagues on the Soikkola peninsula beginning from 2006.

Our data include a corpus of spontaneous speech samples and a corpus of elicitations. The collection of spontaneous speech samples consists of narratives and, to a lesser extent, dialogues recorded by 24 Soikkola Ingrian speakers. The total amount of recordings in this collection is about three and a half hours (more than 20,000 words). We refer to this collection as the Spontaneous Speech Corpus.

The corpus of elicitations includes various questionnaires (more than 700 hours of recordings). In this study we frequently refer to a questionnaire that was originally developed for the initial stage of our research on Ingrian. It consists of the main part (150 sentences) and an additional part (25 sentences) aimed at exploring the main properties of Ingrian morphology, syntax and phonetics. The main part of the questionnaire was recorded by 33 Soikkola Ingrian speakers (plus partial recordings were made by 3 more speakers). The additional part was recorded by 16 speakers. This questionnaire is representative of the analysis of dialectal and idiolectal variation. It is referred to as the Basic Grammar Questionnaire. Of the 38 speakers who were recorded for one or both parts of this questionnaire, 8 lived in the northern dialectal zone, 18 in the southern dialectal zone and 12 in the transitional zone (village of Vistina). In the first of these groups we do not have speakers who relocated from another zone or whose parents were born in different dialectal zones. In the second group, only 2 such cases are attested.

The corpus of elicitations beyond the Basic Grammar Questionnaire is referred to as the Corpus of Elicitations. It includes, inter alia, data from a number of questionnaires specifically concerning pronouns that were collected over the course of our fieldtrips. The structure of our data used in this article is plotted in Figure 1.
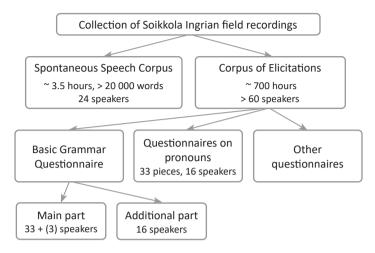
Figure 1. The structure of data used in the article.

Our research focuses on the Soikkola Ingrian of the 21st century, so the data that were collected by previous researchers are only referred to in order to highlight the observed innovations. In fact, the amount of available data on Soikkola Ingrian is not large. There are only a few published samples of Soikkola Ingrian spontaneous speech: two tales in Porkka (1885), a number of samples recorded mostly by one speaker (born in Saarove, lived in Tarinaisi) in Ariste (1960), and a collection of texts (Laanest 1966a) featuring Soikkola material recorded by one speaker from Voloitsa. We do not consider the numerous schoolbook texts published in the 1930s, because those were created with the aim of developing a literary language and were most likely edited according to an artificially designed standard. A vast number of examples can be found in the dictionary of Ingrian (Nirvi 1971); however, they only specify the place where a particular example was recorded, providing no information on either the exact time of recording or the linguistic background of the speaker, and thus they lack the information relevant for our study.

Some differences in the phonetic shape of pronouns required instrumental techniques to make a reliable decision about the quantity and quality of segments. We used the Praat software package (Boersma & Weenink 2020) for phonetic analysis.

3.   Personal and possessive pronouns

The system of Soikkola Ingrian personal pronouns distinguishes 2 numbers and 3 persons, resulting in 6 basic forms. According to the typology of pronouns provided in Cysouw (2009: 106–107), such systems belong to the "Latin-type" and are quite typical for the Uralic family. Table 1 presents the forms of Soikkola Ingrian personal pronouns attested in our material. Variants are separated with a tilde; the first is the main (most common) variant.

Table 1. Paradigms of personal pronouns.

		Number and person					
		Singular			Plural		
		1	**2**	**3**	**1**	**2**	**3**
Case	NOM	miä	šiä	hää ~ hään	mȫ̈ ~ möö ~ müü	tȫ̈ ~ töö ~ tüü	hȫ̈ ~ höö ~ hüü ~ hȫ̈d ~ hôôd ~ hüüd
	GEN	miun	šiun	hänen	meijen	teijen	heijen
	ACC				meijed	teijed	heijed
	PART	miňnua	šiňnua	händ ~ händä	meidä ~ meid	teidä ~ teid	heidä ~ heid
	ILL	miuha ~ miuhe ~ miuhu	šiuha ~ šiuhe ~ šiuhu	häňnee	meihe	teihe	heihe
	INE	miuž	šiuž	hänež	meiž	teiž	heiž
	ELA	miušt	šiušt	hänešt	meišt	teišt	heišt
	ALL	miule ~ miulle	šiule ~ šiulle	hänele ~ hänelle	meile ~ meille	teile ~ teille	heile ~ heille
	ADE	miul	šiul	hänel	meil	teil	heil
	ABL	miuld	šiuld	häneld	meild	teild	heild
	TRANS	miukš	šiukš	hänekš	meikš	teikš	heikš

3.1. General description

The 1st and 2nd person pronouns have the same origin in all Finnic languages while the 3rd person pronouns differentiate the northern Finnic languages (Finnic, Ingrian, Karelian, Ludic, and Veps) from the southern ones (Estonian, South Estonian, Votic, and Livonian).[4]

The phonetic structure of the 1st and 2nd singular personal pronouns is very similar, in contrast with the 3rd singular pronoun. An exceptional feature of the 1st and 2nd singular personal pronouns is the change of the harmonic type within the paradigm: both nominative forms (*miä*, *šiä*) are front-vocalic while the remaining forms in the paradigm are back-vocalic.[5] In the stem of these pronouns, the historical second consonant *n* is lost in all cases except for the partitive (Laanest 1978: 250), cf. Ingrian *miä* '1sg' with Estonian *mina* and Finnish *minä*.

The plural pronouns are uniform in all 3 persons. They are the only Ingrian words that have a special accusative form,[6] which appears in the position of a total object.[7] However, even in contexts typical for a total object, many Ingrian speakers prefer the partitive rather than accusative pronominal forms. For example, the sentence *We have driven them away* was translated with *heidä* '3pl.part' by 19 speakers, and only 14 speakers used *heijed* '3pl.acc'. In the sentence *They have driven us away*, the accusative form was even rarer: 24 partitive vs 7 accusative forms.

---

4   However, Vaipooli Votic borrowed the 3rd plural pronoun from Ingrian, so it has idiolectal variation between the original Votic *nämä(d)* and borrowed *hüü* 'they'. In published Votic texts, one also finds the Ingrian 3rd singular pronoun *hää* attested in the speech of Votic native speakers (see, for example, Heinsoo & Kuusk 2002: 114). However, in our field data collected from 8 Votic speakers in the 21st century, the Ingrian 3rd singular pronoun *hää* was not attested, while most of these speakers used the 3rd plural pronoun *hüü* borrowed from Ingrian. We fully agree with the evaluation by Blokland (2012: 22) that "*hän* is most likely an occasional Ingrianism".

5   The only other such word in Ingrian that we know of is the verb *tulla* 'to come'. It has front-vocalic forms in the 1st and 2nd person present forms (*töön* 'come.prs.1sg', *tööd* 'come.prs.2sg', etc.) and back-vocalic forms in the rest of the paradigm.

6   Note that there is a structural and functional similarity between the pronominal accusative and substantive nominative plural forms: both are derived from the genitive singular and encode a total object. Laanest (1982: 190) assumes that *d* in the accusative can be related to the nominative plural marker *d*.

7   On the notion of the total object see, for example, Erelt (2003: 96–97).

The list of pronominal case forms provided in Laanest (1966b: 108) also includes the essive (*minnūn* '1SG.ESS', *hännēn* '3SG.ESS', *mein* '1PL.ESS') and abessive (*miuᴅa* '1SG.ABE', *häneᴅä* '3SG.ABE', *meiᴅä* '1PL.ABE') forms. These, however, were not attested in our material.[8]

The excessive case forms of personal pronouns are not attested in either the existing literature or our material.

In Ingrian, there is no special class of possessive pronouns. Instead, the genitive forms of personal pronouns are used in the possessive function (which is typical for the Finnic languages), e.g. *miun* 'my', *šiun* 'your', etc. The reflexive-possessive pronoun *oma* 'own' is discussed in section 5.

3.2. Variation

3.2.1. 1st and 2nd person singular illative forms

The illative of monosyllabic words has a specific marker that consists of the consonant *h* and a vowel, whose quality is defined by the first syllable vowel. The first vowel of monosyllabic nouns is either a long vowel (*maa* 'land', *puu* 'tree', *tęę* 'road', *pii* 'prong', etc.) or a diphthong. For long vowels, the rule for defining the vowel of the illative marker is simple: for all stem vowels except *i* the stem vowel and the illative affix vowel have the same quality:[9] *maaha* 'land.ILL', *puuhu* 'tree.ILL', *tęęhe* 'road.ILL'. The stem vowel *i* requires *e* in the illative marker: *piihe* 'prong.ILL'.[10] A general rule for diphthongs is hard to formulate, as there are too few matching words. Across our data,

---

8   In our data, we have only observed the abessive forms of the supine (e.g. *kündä-mä-dä* 'plow-SUP-ABE') but never of nouns, adjectives, or pronouns. However, Nirvi (1971: 312) gives an example *ilmam miu-ᴅa* 'without 1SG-ABE', which is further evidence that the abessive forms of personal pronouns were in use in Ingrian. The same example is quoted in Nirvi (1971: 325), and possibly in Laanest (1978: 250) and Nirvi (1971: 91), because the same village is indicated for this example. Multiple instances of the same example lead us to assume that such forms were not widespread.

9   Some details on merging the long mid and long high vowels in certain idiolects are beyond the scope of this discussion.

10 Cf. also the rule for the possessive 3Sg marker *hV* described in Porkka (1885: 79): the vowel in the suffix has the same quality as the preceding vowel except the vowel *i* that triggers *e* in the suffix.

there are only 2 monosyllabic nouns containing a diphthong (*voi* 'butter' and *täi* 'louse').[11] Both have *i* as the second part of diphthong and *e* in the illative marker (*voihe* 'butter.ILL' and *täihe* 'louse.ILL'). Judging from these examples, we can assume that the second component of a diphthong defines the vowel in the illative marker (cf. Standard Finnish where the vowel in the illative singular marker always has the same quality as a long vowel or the second part of a diphthong of a mono-syllabic stem).

The 1st and 2nd singular personal pronouns contain the diphthong *iu*, which never appears in monosyllabic nouns. Most existing sources (Porkka 1885: 78, Laanest 1966b: 108, 1978: 250, Nirvi 1971: 312, 530, Saar 2017: 127) cite the illative forms with a final *e*: *miuhe* '1SG.ILL', *siuhe* '2SG.ILL'. Based upon these pronominal forms, one can infer that diphthongs (at least those containing *i*) require *e* in the illative marker. However, Junus (1936: 41)[12] gives the form with a final *u, miuhu* '1SG.ILL', which appears to be built according to a rule that the last stem vowel (here, the second part of a diphthong) defines the quality of the vowel in the illative marker.

In our data, we observe variation between three vowels in the illative marker of the 1st and 2nd singular pronouns recorded by 5 speakers. One of these speakers used forms with a final *e* (*miuhe*, *siuhe*), another speaker used forms with a final *u* (*miuhu*, *siuhu*)[13] and 3 other speakers used forms with a final *a* (*miuha*, *siuha*). We have not observed any unambiguous correlation of these variants with the speakers' place of birth although forms with the final *-a* were all recorded in Viistina (un-like the forms with *-u* and *-e*). Also unlike the forms ending in *-e* and *-u*, the forms with a final *-a* were never mentioned in previous research.

---

11 Some monosyllabic nouns (e.g. *koi* 'moth') were not recognized by the native speakers that we worked with. The words *rae* 'hail', *tae* 'dung' and such, where the combination of vowels in the nominative resulted from the null grade of alternation (cf. *raǩkeehen* 'hail.GEN', *taťteehen* 'hail.GEN') are not considered monosyllabic (cf. the genitive forms of "true" monosyllabic nouns: *voin* 'butter.GEN', *täin* 'louse.GEN').

12 Though here and below we refer to Junus (1936), this grammar should not be considered a reliable source of data on Soikkola Ingrian. First, it is a prescriptive rather than descrip-tive grammar, and second, it combines features of Soikkola and Lower Luga Ingrian.

13 This native speaker was born in 1935 so the form with a final *u* cannot be the direct influence of school education – Ingrian materials prepared by Junus and his colleagues were only in use until 1938.

We do not have a reliable explanation for how the -*a* forms emerged.[14] There is insufficient data to determine whether the given variation is dialect-specific. Thus, we classify it as purely idiolectal.

This kind of variation is not attested in the illative plural forms of personal pronouns.

3.2.2. 3rd person singular pronoun: nominative

The form *hää* is provided in all existing sources as a 3Sg pronoun. Laanest (1966b: 108) spells this form as *hā̈(n)* for Soikkola Ingrian. This final *n* appears in some other sources, too, but mostly referring to other Ingrian dialects: there are examples with *hän ~ hā̈n* from Lower Luga and Hevaha Ingrian in Nirvi (1971: 84). Porkka (1885: 78–79, 145) gives the form *hää(n)* (not specifying a dialect) but, in his text collection, the form *hään* is only observed in the Lower Luga dialect. Junus (1936: 29) and Saar (2017: 127) spell this pronoun without the final *n*. In the Basic Grammar Questionnaire, the form with the final *n* was given by 5 native speakers; 2 of them exclusively use this form, while 3 others demonstrate *hää ~ hään* variation to varying degrees.[15] Our data do not give any reliable evidence to suggest that the form with the final *n* is a feature of a particular sub-dialect of Soikkola Ingrian. Those speakers who use this form without variation belong to the southernmost and northernmost varieties, while the other 3 speakers also belong to different dialectal groups. In the Spontaneous Speech Corpus, *hään* is mostly observed in recordings from the same speakers. However, episodically *hään* occurs in the data recorded by other speakers as well. This variation seems rather a feature of an idiolect than that of a variety, allowing free variation within an idiolect in some cases.

---

14 A possible explanation is an incorrect restoration of the reduced vowel. In Soikkola Ingrian, a short *a* is usually reduced, and its pronunciation varies between *ă* and *e*. In careful speech (i.e. when a native speaker repeats sentence or form slowly and distinctively), the original *a* is usually restored. However, we noticed that some speakers also pronounce the final *e* as *ă* in markers, e.g. the illative marker *šše* may sound as *ššă*. We explain it as a kind of hypercorrection when restoring a full vowel happens in the wrong place. On the other hand, this pattern is not attested in the speech of the native speakers who use the illative forms *miuha* and *šiuha,* so this hypothesis is questionable.

15 One of the speakers prefers *hään*; the form without the final consonant appears only episodically in her speech. The second speaker uses *hään* approximately half as often as *hää*, and the third speaker prefers *hää* but still uses *hään* from time to time.

The duration of the vowel in monosyllabic words tends to shorten in fast speech, especially in the case of auxiliary words. In the pronominal forms *hää* and *hään*, the duration of the vowel also varies between short and long. Although some speakers shorten this vowel more often than others, typically both short and long vowels can occur in the same speaker's speech. Given that such shortening is typical for monosyllabic words and applies to other pronouns of the same structure (e.g. the nominative plural forms of the personal pronouns), we do not mark it in our transcription, and always spell such pronouns with a long vowel.

3.2.3. 3rd person singular pronoun: partitive

There is variation between *händ ~ händä* in the partitive form of the 3Sg pronoun. The main variant, *händ*, does not have a final vowel. The variant *händä*, ending with a vowel, is used by only a few speakers. In the Basic Grammar Questionnaire, the partitive forms with the final *ä* were given by 5 out of 34 speakers, and these 5 speakers belong to different dialectal zones of the Soikkola peninsula. In all such examples, *händä* occurs in sentence-final position. The Spontaneous Speech Corpus reveals that only one of these 5 speakers uses *händä* as the main variant (irrespective of its position), whereas the other 4 speakers switch between *händä* and *händ.* Thus, this variation is neither dialectal nor purely idiolectal. Some idiolects do have strong preferences, but in others there is free variation with positional limitations: the variant with the final vowel appears only in sentence-final position.

This variation is mentioned neither in Porkka (1885: 78) nor in Junus (1936: 98); they both give only the forms *händä* and *häntä*, respectively. However, Nirvi (1971: 49, 77, 84, 89, 92, etc.) provides several examples containing *hänD* recorded in different villages; Saar (2017: 127) spells this form as *hänD(ä).*

3.2.4. Nominative plural: quality of the vowel

The nominative plural forms of personal pronouns consist of an initial consonant and a long vowel: *möö* '1PL', *töö* '2PL', *höö* '3PL'. The historical long mid vowels of the first syllable (*ee, oo, öö*) demonstrate variation in Soikkola Ingrian: they can be pronounced as mid, high-mid,

or high vowels. This fact was already mentioned by Sovijärvi (1944: 105–106) and Laanest (1966c: 77–80) but a detailed analysis was first provided by Kuznetsova (2009: 125–156). Following her predecessors, Kuznetsova also concludes that the reflexes of these vowels are variety-dependent: the higher vowels (*i̮i̮ ~ ii, u̮u̮ ~ uu, ü̮ü̮ ~ üü*) are more typical of the southern varieties of Soikkola Ingrian, while variants which are closer to the original mid vowels are better preserved in the northern varieties. However, particular idiolects and/or particular words can deviate from this system. Kuznetsova (2009: 130–131) lists phonetic variants of the pronouns *mȫ* '1PL', *tȫ* '2PL', and *hȫ* '3PL' collected from various speakers. Kuznetsova's evaluation is based exclusively on her own perception; she did not make acoustic measurements. In order to verify her results experimentally, we analysed the quality of the vowel in *mȫ* '1PL' in 12 idiolects using instrumental techniques (see Appendix 1 for a detailed description). Out of these 12, the vowel in 1Pl is of the same quality as the singleton *ö* in 4 idiolects, of the same quality as *üü* (*ü*) in another 4 idiolects, 2 speakers distinguish all three vowels *ȫ, ö* and *üü* (*ü*), and another 2 speakers represent borderline cases (the difference is on the border of statistical significance so the type cannot be defined unambiguously). Each group contains idiolects belonging to different zones of the Soikkola peninsula, so we consider this variation in vowel quality in plural pronouns to be idiolectal rather than dialectal.

3.2.5. Nominative of the 3rd plural pronoun: the final consonant and the vowel length

Laanest (1966c: 114) lists several variants for the 3rd plural perso-nal pronoun. In addition to the variation in vowel quality (see 3.2.4), the attested forms differ with regard to the presence/absence of a final consonant: there are forms (a) without a final consonant, *hȫ*, (b) with a final -*d*, *hȫd*, and (c) with a final -*n*, *hȫn*. In Laanest's list, there are no villages whose speakers would use (b) or (c) and not use (a). Type (c) is only attested in the village of Viistina,[16] and type (b) occurs in several villages located in different zones of the Soikkola peninsula.

---

16   Laanest also gives a variant with a final -*n* in Hevaha dialect (the village of Kangaspää). Porkka (1885: 79) mentions this variant for the Äyrämöis dialect of Finnish and says that he also heard it in Lavansaari.

In the Basic Grammar Questionnaire, the type without the final consonant definitely dominates: 35 out of 41 native speakers prefer it, and only 6 speakers use the variant with the final -*d*. The distribution is rather clear: only one of these 6 speakers switches between forms with and without the final consonant. He tends to use the form without the final consonant in fluent speech while in careful speech the final consonant is always retained. The speakers who use type (b) come from different zones of the Soikkola peninsula, so we cannot claim a dialectal distribution. Thus, this is most likely a pure example of idiolectal variation involving some elements of register variation within certain idiolects. None of the native speakers prefer the form with the final -*n*. Although occasional examples with some indistinct nasal sounding, which can be interpreted as the final -*n*, have been attested throughout our data, we do not have any unambiguous examples of this type. For that reason, type (c) with the final -*n* is not included in Table 1.

It looks like there is a subtle correlation between the presence of a final consonant in 3rd singular and 3rd plural forms. Two of the speakers who use the variant *hȫd* prefer the variant *hään* as well, and the speaker who exhibits *hȫ* ~ *hȫd* variation also switches between *hää* and *hään*.

The vowel duration in forms with the final -*d* varies in a way similar to that of the forms without it. We spell such forms with a long vowel, although the duration of a vowel in fluent speech is sometimes closer to that of a singleton (see section 3.2.2).

3.2.6. Partitive of the plural pronouns: the final vowel

The most common variants of the partitive forms of plural personal pronouns are *meidä* '1PL.PART', *teidä* '2PL.PART', *heidä* '3PL.PART'. There are, however, alternative forms of the partitive plural, that do not contain a final vowel. In some sources, such forms are not mentioned: one can find only *meidä* in Porkka (1885: 78) and *meitä* and *heitä* in Junus (1936: 98). However, the apocopated forms *meiD* and *heiD* can be found in Nirvi (1971: 517, 208). Saar (2017: 127) gives only *meiDä* in the table of pronominal forms, but the apocopated form *meiD* is observed in an example (Saar 2017: 152).

In the Basic Grammar Questionnaire, the ratio of forms is as follows: 19 occurrences of *meidä* to 4 of *meid*, and 18 occurrences of *heidä* to 1 of *heid.* The Spontaneous Speech Corpus shows that forms with the

apocopated final vowel appear occasionally in the speech of different speakers. Two native speakers, who used to live in Estonia, use apocopated forms more often than the others, but neither of them prefers these forms to the non-apocopated variant. The partitive of the 2nd plural pronoun is not attested in the Basic Grammar Questionnaire nor in the Spontaneous Speech Corpus, but it occurs in the Corpus of Elicitations with the final vowel in nearly all cases.

The variation of the partitive plural forms belongs to a mixed type as it demonstrates features of idiolectal and free variation.

3.2.7. *Allative forms*

According to Porkka (1885: 78), the allative suffix of Soikkola Ingrian personal pronouns contains a geminated *l*: *miulle* '1SG.ALL', *meille* '1PL.ALL'. *Miulle* and *meille* are spelled in the same way by Junus (1936: 98), and Saar (2017: 90, 127) also has the spelling *miulle*. Laanest (1978: 250) lists two forms with a singleton (*miule* '1SG.ALL', *šiule* '2SG.ALL') and one with a geminate (*miulle* '1SG.ALL') but the latter is cited from Nirvi (1971), where numerous examples of both types (i.e. with or without a geminate) can be found, with the geminate forms being more prevalent.

Our analysis (see Appendix 2) proves that the duration of *l* in the allative pronominal forms as produced by contemporary speakers corresponds more closely to a singleton than to a geminate. However, sometimes this *l* is pronounced more intensively, as in a geminate, but such examples were recorded by only a few speakers. Geminate forms are more typical of careful rather than fluent speech though sometimes there is no obvious motivation for the distribution of these forms. We consider the allative forms with a singleton to be the major variant, and forms with a geminate the minor one. The different allative forms represent both idiolectal and register variation with elements of free variation.

4.   Demonstrative pronouns

Demonstrative pronoun systems can be roughly divided into distance-oriented and speaker-oriented systems, with the number of values within these varying across the world's languages (Diessel 1999: 39–40). Soikkola Ingrian apparently used to belong to the distance-oriented type

with a three-way contrast. Porkka (1885: 83–85) describes a tripartite system of Ingrian demonstrative pronouns consisting of *tämä ~ tää*, *se* and *tō*, and translates their meanings into German as 'dieser', 'der', and 'jener' respectively. In contemporary Soikkola Ingrian, this system has not been preserved: the pronoun *too* has been completely lost. It is difficult to say when this change happened. In the tale "Der goldene Vogel" recorded in the village of Tarinaisi and published in Porkka's grammar, we observe 8 forms of *tämä*, 29 forms of *še* and none of *too*.

The pronoun *tämä* has the meaning 'this' while *še* can be translated both as 'this' and 'that' (below we will use 'that' as a gloss for *še*). The opposition of these pronouns is therefore different from that in English or Russian, for example. In these languages, 'this' is the unmarked pronoun – it is more frequent[17] and there are contexts where 'that' appears just as a marked counterpart to 'this'. In Ingrian, the unmarked pronoun is *še*. In the abovementioned tale, *še* is almost 4 times more frequent than *tämä*. In our Spontaneous Speech Corpus, the ratio *še*/*tämä* is even higher than in "Der goldene Vogel".

Table 2 contains the paradigms of the demonstrative pronouns.

Table 2. Paradigms of demonstrative pronouns.

Number / Case	'this'		'that'	
	Singular	**Plural**	**Singular**	**Plural**
NOM	tämä	nämäd ~ nämä	še ~ šee	need ~ nee ~ ned ~ ne
GEN	tämän	näijen	šenen ~ šen	niijen
PART	tädä	näidä	šidä	niidä
ILL	tähä	näihe	šihe ~ šiihe	niihe
INE	täž	näiž	šiin	niiž
ELA	täšt	näišt	šiind	niišt
ALL	tälle	näille	šille	niille
ADE	täl	näil	šil	niil
ABL	täld	näild	šild	niild
TRANS	täkš	näikš	šikš	niikš

---

17 See, for example, Bender (1998: 58) about higher frequency as a feature of unmarked categories.

4.1. General description

Both demonstrative pronouns have suppletive paradigms: different stems distinguish singular and plural forms.[18]

Table 2 does not contain the abessive, essive and excessive forms. The historical essive and excessive forms of *še* 'that' serve as the inessive and elative forms.[19] The form *tänä* (the essive of *tämä* 'this') is lexicalized, and functions as a modifier in temporal constructions (e.g. *tänä kežäl* 'this summer') but not as a typical pronominal form. No other abessive, essive or excessive forms occur in our material; the forms *tämännä* 'this.ESS', *senennä* 'that.ESS', *näinnä* 'this.PL.ESS', *niinnä* 'that.PL.ESS' given by Junus (1936: 99) are also unattested.

The paradigms of the demonstrative pronouns contain some irregularities typical of Finnic languages. The first of them concerns the alternation of vowels in the stem, cf. *šenen* 'that.GEN' vs *šidä* 'that.PART', *šil* 'that.ADE', or *need* 'that.PL.NOM' vs *niidä* 'that.PL.PART', *niille* 'that.PL.ALL'.

The second irregularity is the combination of "full" and "short" stems in the paradigm, cf. *tämä* 'this.NOM', *tämä-n* 'this-GEN' vs *tä-dä* 'this-PART', *tä-hä* 'this-ILL', or *nämä* 'this.PL.NOM' vs *nä-i-dä* 'this-PL-PART'. The full and short stems are distributed depending on the morphological form, and for most forms there is no variation. Although Porkka (1885: 83), Nirvi (1971: 616), and Saar (2017: 127) mention the variation of the nominative forms *tämä ~ tā̄* 'this', we did not encounter the variant *tā̄* in our materials. Porkka (1885: 83) also indicates a similar variation for the genitive forms: *tämän ~ tā̄n* 'this.GEN' but it is neither mentioned by other researchers nor attested in our data. However, in our Corpus of Elicitations we observe occasional deviations, e.g. the form *tämäl* 'this.ADE' (which occurred 4 times in the data from 2 native speakers) and *tämäz* 'this.INE' (which occurred only once). We consider that such forms were likely constructed by analogy with the genitive form *tämän*. In the Spontaneous Speech Corpus, such forms are not attested.

---

18 An alternative interpretation is to consider the singular and plural demonstrative pronouns as separate lexemes. For the current description, it is not significant which interpretation is chosen.

19 Laanest (1978: 252) notes that *sīnD* is the excessive functioning as the elative, and translates it as both a pronominal form ('в этот') and an adverb ('оттуда'). However, in our material, only *šęęld* is used in adverbial function.

4.2. Variation

4.2.1. Nominative plural of 'this'

The existing sources give different variants for the nominative plural form of the pronoun 'this', in particular *nämä*(*t*) ~ *nää* in Porkka (1885: 83), *nämäᴅ* in Nirvi (1971: 355), *nämät* in Junus (1936: 99), and *nämä* in Laanest (1978: 252). Saar (2017) does not mention this form. In our material, the variant *nämäd* definitely dominates. In the Basic Grammar Questionnaire, 20 native speakers use the variant *nämäd*, and only 4 speakers (from different dialectal zones of the Soikkola peninsula) prefer the variant *nämä* without the final consonant. Moreover, in the speech of 2 of these 4 speakers the variant *nämäd* was also attested. In the Spontaneous Speech Corpus, *nämä* is attested only once in the speech of the native speaker who uses *nämäd* in the Basic Grammar Questionnaire. We consider the variation of *nämäd* ~ *nämä* as idiolectal with elements of free variation. The short form *nää* mentioned by Porkka was not attested in our data.

4.2.2. Nominative singular of 'that'

Porkka (1885: 84), Junus (1936: 99), Nirvi (1971: 514), Laanest (1978: 252), and Saar (2017: 127) spell *še* 'that' with the short vowel. However, our data shows that some speakers pronounce a long vowel in this pronoun. The variant with a long vowel should be considered as an innovation; otherwise, it is problematic to explain why there is no raising of the long mid vowel *ee* > *ẹẹ* ~ *ii* in the first syllable (see section 3.2.4).

The distribution of variants among native speakers is rather clear but still some speakers alternate between forms with a short and a long vowel. Thus, this is an idiolectal variation with elements of free variation.

4.2.3. Genitive singular of 'that'

Most researchers mention the variation in the genitive singular forms: *sen* ~ *senen* 'that.GEN' (Nirvi 1971: 514, Laanest 1978: 252, Saar 2017 :127). Porkka (1885: 84) identifies *senen* as a specific Soikkola variant. In our material, *šenen* is the dominant variant and *šen* occurs only sporadically: once in the Spontaneous Speech Corpus and several times in the Corpus of Elicitations in the speech of 3 native speakers.

All these speakers use *šenen* as the main variant. The small number of examples does not allow us to reach a definitive conclusion but it is likely that here we have a combination of idiolectal and free variation. Junus (1936: 99) gives *senen* as the genitive form and *sen* as the accusative, but we did not find any evidence supporting this distribution.

4.2.4. Illative singular of 'that'

Porkka (1885: 84) gives a long vowel in the illative, inessive and elative forms of *se* 'that' – *šiihe*, *šiin*, *šiind* respectively. In all other forms of this pronoun, the first vowel is short. The same distribution of long and short vowels in the paradigm of 'that' is found in Nirvi (1971: 514) and Laanest (1978: 252). The illative form *siihe* is also given by Saar (2017: 127).

We checked the length of vowels in the paradigm of 'that' and found that in the illative form the first vowel is usually short. The difference between *šiihe* and *šihe* is easily measurable – for *šiihe* we expect a V1/V2 ratio of 1.5–2, while *šihe* has the CVCV structure that requires an automatic prolongation of the second vowel and the estimated V1/V2 ratio is 0.5–0.9.

Of the 22 examples recorded by 8 native speakers, 17 examples have a ratio from 0.5–0.9, four examples have a ratio between 1.1 and 1.2, and one example has a ratio of 1.6. The average duration of V1 is 79 ms, which corresponds to the average duration of a short vowel (for a long vowel we would expect a duration of 120–150 ms). Thus, there is no doubt that in our data the dominant variant of the illative singular is *šihe*, while *šiihe* is a minor variant that appears only occasionally. Since there are so few examples with a long vowel, we cannot define the type of variation here. However, it likely represents idiolectal variation.

4.2.5. Nominative plural of 'that'

Porkka (1985: 84) and Nirvi (1971: 338) give the nominative plural of 'that' as *ne*, and Laanest (1978: 252) considers this the main variant, but in our data *need ~ ned* definitely dominates. Only one of the native speakers uses *nee ~ ne* while others use *need ~ ned*. In the Spontaneous Speech Corpus, *nee ~ ne* occurs in the samples from this one speaker and occasionally in the speech of another speaker (possibly as an allegro variant). Thus, this variation is idiolectal, but also partially linked to

register. Since the variant with the final consonant *neet* is proposed in Junus (1936: 99), it should not be ruled out that studying of the Ingrian language at school may have contributed to the spread of this variant.

In our data, the duration of the vowel in this pronoun mostly corresponds to a long vowel, although examples with a short vowel are also attested in the speech of several native speakers. We consider the long vowel in this pronoun to be an innovation. First, existing grammars give the variant with a short vowel as the main or only one possible, and second, there is no raising of the long mid vowel in *nee*, same as in *šee* (see 4.2.2). As the distribution of *ne(d)* vs *nee(d)* is speaker-dependent but not strict, we consider the variation of the vowel duration to be idiolectal with elements of free variation.

5.   Reflexive and reciprocal pronouns

In this section, three pronouns are analysed: the reflexive-possessive pronoun *oma* 'own', the reflexive pronoun *itse* 'oneself' and the reciprocal pronoun *toin-toišt* 'each other'. The paradigms of these pronouns display fewer irregularities than the personal or demonstrative pronouns, and none of the sources (Porkka 1885, Junus 1936, Laanest 1978, Saar 2017) offers the complete paradigms of these pronouns. Table 3 presents these paradigms as attested in our data. The illative and ablative forms of *itse* 'oneself' do not occur in our materials, so they are hypothetical.

Table 3. Paradigms of reflexive and reciprocal pronouns.

Number / Case	'own'		'oneself'	'each other'
	Singular	Plural	Singular	Singular
NOM	oma	omad	itse	(toin-toišt)
GEN	oman	oṁmiin	itsen	toin-toižen
PART	oṁmaa	oṁmia	itsiä ~ itsiädä ~ itsedä	toin-toišt
ILL	oṁmaa	oṁmii	ʾitsee	toin-toiššee
INE	omaž	omiž	itseež	toin-toižeež
ELA	omašt	omišt	itseešt	toin-toižeešt
ALL	omalle	omille	itselle	toin-toiželle
ADE	omal	omil	itseel	toin-toižeel
ABL	omalt	omild	ʾitseelt	toin-toižeeld
TRANS	omakš	omikš	itseekš	toin-toižeekš

5.1. General description

The reflexive–possessive pronoun *oma* 'own' belongs to one of the main nominal paradigm types (cf. *kana* 'chicken', GEN *kanan*, PART/ILL *kaňnaa*) and does not exhibit variation in its forms. Unlike *itse* 'oneself' and *toin-toišt* 'each other', *oma* has plural forms.

In Ingrian, the possessive suffixes existed in the 19th century (e.g. *ättii-n* father-P1SG 'my father', *veljä-hä-d* brother-P3SG-PL 'his brothers' (Porkka 1885: 131–132, Rožanskij & Markus 2012: 473, 488)) but in the 20th century they became rare (Laanest 1966b: 107). However, our data contradict Nikolaev (2001: 450) who claims that by the end of the 20th century even the reflexive pronoun *itse* was used without possessive suffixes. According to our data, it is the only word in Ingrian that can still take possessive suffixes. Although the variants without the possessive suffixes indicated in Table 3 are much more frequent, some speakers also use the possessive forms of *itse.* The variation between the forms with and without the possessive suffixes is discussed in Section 5.2.1.

The reflexive pronoun *toin-toišt* does not have a nominative form. It never occurs in subject position. In the position of total object which requires the nominative form of a noun, we find only the partitive form *toin-toišt*. The plural forms of the reciprocal pronoun were not attested in our material except one example with *toin-toiššia* that we consider to have been built by analogy, cf. *toiššia* – a plural partitive form of *toin* 'other'.

There are no unambiguous examples of the essive, excessive and abessive forms of these pronouns in our material, therefore these are not presented in Table 3.

5.2. Variation

5.2.1. Forms of itse *'oneself' with and without possessive suffixes*

Porkka (1885: 83) notes that *itse* is typically used with possessive suffixes. In our material, the forms of *itse* without possessive suffixes are more frequent. However, a number of native speakers use forms with possessive suffixes in their speech. The following forms were attested in our material:

– with 1Sg possessive suffix: *itsiä-n* 'oneself.ᴘᴀʀᴛ-ᴘ1ꜱɢ', *itse-lle-en*[20] 'oneself-ᴀʟʟ-ᴘ1ꜱɢ';
– with 2Sg possessive suffix: *itsiä-äž ~ itsedä-äž ~ itsiättä-äž* 'oneself. ᴘᴀʀᴛ-ᴘ2ꜱɢ' (see Section 5.2.2 on the variation in the partitive singular forms), *itse-lle-ež* 'oneself-ᴀʟʟ-ᴘ2ꜱɢ';
– with 3Sg possessive suffix: *itse-he* 'oneself-ᴘ3ꜱɢ', *itse-he-n* 'oneself-ᴘ3ꜱɢ-ɢᴇɴ', *itsiättää ~ itseťtää* 'oneself.ᴘᴀʀᴛ.ᴘ3ꜱɢ', *itse-štää* 'oneself-ᴇʟᴀ.ᴘ3ꜱɢ', *itse-llää* 'oneself-ᴀᴅᴇ.ᴘ3ꜱɢ';
– with 1Pl possessive suffix: *itse-mme* 'oneself-ᴘ1ᴘʟ';
– with 3Pl possessive suffix: *itsiä-nne* 'oneself.ᴘᴀʀᴛ-ᴘ3ᴘʟ'.

All speakers who use forms with possessive markers also use forms without these markers. We were not able to define the features that condition the choice of forms. The small number of occurrences where the possessive markers are used does not allow us to describe the distribution of these forms across native speakers. Most likely, this is a case of free variation with elements of idiolectal variation. It should also be noted that the possessive forms are "atavistic" and sometimes used incorrectly, e.g. the 3Sg form can be used in a context requiring the 1Sg form.

5.2.2. Partitive singular of *itse* 'oneself'

There are several variants that appear in the context requiring the partitive singular form of *itse*. These variants are *itsiä ~ itsiädä ~ itsiedä* without the possessive suffix and *itsiättää ~ itseťtää* with the 3Sg possessive suffix. The emergence and distribution of these variants is not clear enough for us and definitely requires further detailed research. Here we discuss our observations and the main hypotheses.

The variant *itsiä* looks like a typical partitive form with the marker *-ä*, cf. *atti-ä* 'father.ᴘᴀʀᴛ' or *kiv̆vi-ä* 'stone.ᴘᴀʀᴛ'. However, such partitive forms are typical for nouns with a final *-i* in the nominative (cf. *ätti* 'father, *kivi* 'stone') while nouns with a final *-e* in the nominative

---

20 The small number of occurrences prevents us from making accurate measurements that could define the length of the vowel at the interface of the case and possessive suffixes. When this vowel is *ä*, the lack of qualitative reduction is evidence that this vowel is long. In other cases, we follow Laanest (1978: 251) who spells such forms with a long vowel.

typically derive the partitive from the consonant stem, e.g. *herneeh-t* 'peas.PART' from *herne* 'peas' (exceptions are very rare, e.g. *nukkia* 'doll.PART' from *nukke* 'doll'). Porkka (1885: 83) gives *itsiä-* as the partitive singular stem. Several occurrences of *itsiä* can be found in Nirvi (1971: 59, 77, 613), however the variant *itsi₍ā̆* with a long final *ā̆* is more frequent (Nirvi 1971: 97, 314, 412, 528, 590). In our material, we do not observe the opposition of *itsiä* vs *itsiää* (where the last variant is supposed to have the 3Sg possessive marker). The variant *itsiä* was attested for different persons of the subject, e.g. *hö̆ö̆ ombeliid itsiä vard* 'They sewed for themselves', *miulle pĭttää răĥhaa itsiä važ* 'I need money for myself'.

The variant *itsiädä* is attested in the Ingrian dictionary (*oppī itsi₍ädä vard* 'Study for yourself!' (Nirvi 1971: 97)) as well as in our material. If we assume that *itsiä-* is a root rather than a root plus the partitive marker, the emergence of *dä* looks reasonable: it is the common partitive singular suffix that is attached after a long vowel or diphthong, cf. *ö̆ö̆-dä* 'night-PART', *täi-dä* 'louse-PART'. However, we do not have any evidence that Ingrian speakers reinterpreted *itsiä* as a root.

It is not clear what the relation between *itsiädä* and *itsiä̆ttää* is. Theoretically, *itsiä̆ttää* could result from the gemination of *d* before a long vowel of the 3Sg possessive marker. However, Porkka never mentions *itsiädä*, and notes that in the partitive singular form that ends in a long vowel, the partitive ending *ta*[21] appears before the possessive suffix, e.g. *varis tuli küzümää **poikaataa*** 'A crow came to ask his/her son' (Porkka 1885: 79).

The general picture is also blurred because historically long vowels shortened in contemporary Soikkola Ingrian and their duration became even shorter than the duration of the prolonged second vowel in a (C)VCV foot (Markus 2011: 109–110). Nirvi (1971: 97) puts a syllable marker before -*ädä* in *itsiädä*, and such parsing corresponds to our observations (one would expect a strong qualitative reduction of the final vowel in *itsiädä* if *iä* belonged to one syllable, but this is not the case).

The forms *itsiädä* and *itsiä̆ttää* are distributed across the native speakers and we did not notice any variation of them in the speech

---

21 Porkka (1885) represents a short geminate plosive with a single letter (e.g. *t*) while we write it with a double letter with a breve (e.g. *t̆t*).

of any given speaker (although the variation *itsiä ~ itsiädä* seems to be rather typical). The form *itsiädä* appears with different persons of the subject (e.g. *höö itsiädä eväd hoidaneend* 'They didn't take care of themselves') while *itsiättää* appears in contexts appropriate for the 3Sg form.[22]

The variants *itsedä* and *itsettää* with the second vowel *e* are likely the result of a phonetic reduction in fluent speech. First, it is typical to have a variation *itsiädä ~ itsedä* or *itsiättää ~ itsettää* in the speech of a given speaker, and the variants with *iä* are more typical of careful speech. Second, the shortening of the diphthong *iä* into *e* does not look strange. For example, in some Lower Luga Ingrian varieties, the change *iä > e* became regular (Laanest 1966c: 83), e.g. *rätte* 'kerchief.PART' (< *rätti-ä*), *reppe* 'tear.INF' (< *reppi-ä*).

The variant *itsiä* seems to be the most frequent across the native speakers if compared with *itsiädä* (*itsedä*) and *itsiättää* (*itsettää*) but the difference is not very big. For instance, in translations of the sentence *He loves only himself* recorded by many different speakers, we observe 12 occurrences of *itsiä*, 8 of *itsiädä* (*itsedä*), and 10 of *itsiättää* (*itsettää*).

We consider the distribution of *itsiä* vs *itsiädä* vs *itsedä* as speaker-dependent with elements of register variation (*itsiädä* vs *itsedä*) and free variation (*itsiä* vs *itsiädä/itsedä*).

6.   Conclusions

Table 4 summarizes all the cases of variation discussed above. The variation types are labeled in the table as D – dialectal, I – idiolectal, C – context, R – register, and F – free variation. The plus sign marks the attested types of variation. If only elements of some type are observed, we give the plus sign in parentheses "(+)"; if the picture is blurred and we do not have strong evidence but suspect that the type is possible, we use "?".

---

22   We treat the one exception that we encountered as an incorrect choice of form.

Table 4. Types of pronominal forms variation.

	D	I	C	R	F
Personal pronouns					
1st and 2nd person singular illative forms (*miuha ~ miuhe ~ miuhu*)	?	+			
3rd person singular pronoun: nominative (*hää ~ hään*)		+			(+)
3rd person singular pronoun: partitive (*händ ~ händä*)		+	+		+
Nominative of plural forms: quality of the vowel (*mÿÿ ~ möö ~ müü*)		+			
Nominative of the 3rd plural pronoun: the final consonant (*hÿÿ ~ hÿÿd*)		+		(+)	
Partitive of the plural pronouns: the final vowel (*meidä ~ meid*)		+			+
Allative forms (*miule ~ miulle*)		+		+	(+)
Demonstrative pronouns					
Nominative plural of 'this' (*nämäd ~ nämä*)		+			(+)
Nominative singular of 'that' (*še ~ šee*)		+			(+)
Genitive singular of 'that' (*šen ~ šenen*)		+			+
Illative singular of 'that' (*šiihe ~ šihe*)		?			
Nominative plural of 'that' (*need ~ nee ~ ned ~ ne*)		+		(+)	(+)
Reflexive pronouns					
Forms of *itse* 'oneself' with and without possessive suffixes		(+)			+
Partitive singular of *itse* 'oneself' (*itsiä ~ itsedä ~ itsiädä*)		+		(+)	(+)

The most unexpected conclusion we can draw is that a dialectal distribution of forms was not attested (only for the illative singular forms *miuha ~ miuhe ~ miuhu* we can hypothesize it). We can suggest several possible explanations for this fact. First, the Soikkola peninsula is a rather homogeneous zone from the point of view of linguistic diversity. Opposite the neighbouring Lower Luga region, where several Finnic varieties (Lower Luga Ingrian, Vaipooli Votic, Ingrian Finnish and Estonian) were in close contact with each other and a motley distribution of linguistic features emerged, the Ingrian language of the Soikkola peninsula does not demonstrate many noticeable traces of its recent contact with other Finnic languages. Second, some dialectal features of Soikkola Ingrian varieties have become blurred in recent decades due

to the natural everyday language environment having been lost, which resulted in speakers only communicating with a small number of his/her relatives or friends but not with the wider linguistic community in his/her settlement. Third, in the 21st century, the number of speakers with "simple" linguistic biographies (i.e. when a speaker lives in the same place where (s)he was born and previously his/her parents were born) reduced significantly: many people moved to different places or spent a significant part of their lives outside of their native villages. However, the latter explanation can only account for the lack of variation between villages but not between dialectal zones (as was mentioned in section 2, in most cases native speakers in the northern and southern dialectal zones did not mix).

As a result, the most common type of variation observed in our study is idiolectal. This means that the variation seen in the Soikkola pronominal system does not depend much on macroparameters and therefore is poorly predicted. Even the speech of siblings does not display uniformity: very often siblings use different variants.

In quite a few cases, we observed free variation. This means that a single speaker may use different variants and there are no evident causes as to why a particular variant is preferred. However, we do not have instances of "pure" free variation in our data where two or more variants can be used by all speakers; usually free variation is combined with idiolectal variation and often is of minor importance.

Context and register variation are rare types and they are always combined with idiolectal, and sometimes other types of, variation.

In general, the pronominal system of Soikkola Ingrian has not changed much since the 19th century although a number of deviations from the previous descriptions are attested, e.g. the emergence of the variant *sihe* 'that.ILL' with a short first vowel and a decrease in the use of forms of *itse* 'oneself' with possessive suffixes.

Acknowledgements

Abbreviations

ABE – abessive, ABL – ablative, ACC – accusative, ADE – adessive, ALL – allative, ELA – elative, ESS – essive, GEN – genitive, ILL – illative, INE – inessive, INF – infinitive, NOM – nominative, PART – partitive, PL – plural, PRS – present tense, P1SG, P2SG, … P3PL – possessive markers of the corresponding person (1, 2 or 3) and number (singular or plural), SG – singular, SUP – supine, TRANS – translative.

References

Ariste, Paul. 1960. Isuri keelenäiteid. *Keel ja kirjanduse instituudi uurimused* 5, 7–116.

Bender, Byron W. 1998. Markedness and iconicity. Some questions. In Anna Siewierska & Jae Jung Song (eds.), *Case, Typology and Grammar: In honor of Barry J. Blake*, 7–70. Amsterdam/Philadelphia: John Benjamins. https://doi.org/10.1075/tsl.38.07ben.

Blokland, Rogier. 2012. Borrowability of pronouns: evidence from Uralic. *Finnisch-Ugrische Mitteilungen* 35, 1–34.

Boersma, Paul & David Weenink. 2020. Praat: doing phonetics by computer. Version 6.1.37, http://www.praat.org.

Cysouw, Michael. 2009. *The Paradigmatic Structure of Person Marking*. Oxford: Oxford University Press.

Diessel, Holger. 1999. *Demonstratives: Form, Function, and Grammaticalization*. Amsterdam/Philadelphia: John Benjamins. http://dx.doi.org/10.1075/tsl.42.

Drager, Katie K. 2015. *Linguistic variation, identity construction and cognition*. Berlin: Language Science Press. http://dx.doi.org/10.26530/OAPEN_603352.

Dufter, Andreas, Jürg Fleischer & Guido Seiler (eds.). 2009. *Describing and Modeling Variation in Grammar*. Berlin/New York: De Gruyter. http://dx.doi.org/10.1515/9783110216097.

Erelt, Mati. 2003. Syntax. In Mati Erelt (ed.), *Estonian Language*, 93–129. Tallinn: Estonian Academy of Sciences.

Heinsoo, Heinike & Margit Kuusk. 2002. Jõgõperä vadjalaste keelekasutus 20. sajandi lõpus. *Fenno-Ugristica* 24, 95–118.

Junus, Väino I. 1936. *Izoran keelen grammatikka: morfologia opettaijaa vart*. Leningrad/Moskova: Ucpedgiz.

Kiesling, Scott F. 2011. *Linguistic Variation and Change*. Edinburgh: Edinburgh University Press. https://doi.org/10.1515/9780748637638.

Kuznetsova, Natalia V. 2009. *Fonologičeskie sistemy ižorskix dialektov.* Unpublished doctoral dissertation. Sankt-Peterburg: Institut lingvističeskix issledovanij RAN.

Laanest, Arvo. 1966a. *Isuri murdetekste*. Tallinn: Eesti NSV Teaduste Akadeemia, Keele ja Kirjanduse Instituut.

Laanest, Arvo. 1966b. Ižorskij jazyk. In *Jazyki narodov SSSR. Vol. 3. Finno-ugorskie i samodijskie jazyki*, 102–117. Moskva: Nauka.

Laanest, Arvo. 1966c. *Ižorskie dialekty. Lingvogeografičeskoe issledovanie.* Tallin: Institut jazyka i literatury.

Laanest, Arvo. 1978. *Istoričeskaja fonetika i morfologija ižorskogo jazyka.* Unpublished doctoral dissertation. Tallin: Institut jazyka i literatury Akademii nauk Èstonskoj SSSR.

Laanest, Arvo. 1982. *Einführung in die ostseefinnischen Sprachen*. Hamburg: Helmut Buske Verlag.

Laanest, Arvo. 1986. *Isuri keele ajalooline foneetika ja morfoloogia.* Tallinn: Valgus.

Labov, William. 1972. *Language in the inner city: studies in black English vernacular.* Philadelphia: University of Pennsylvania Press.

Labov, William. 1979. Locating the frontier between social and psychological factors in linguistic variation. In Charles J. Fillmore, Daniel Kempler, & William S.-Y. Yang (eds.), *Individual Differences in Language Ability and Language Behavior*, 327–340. New York: Academic Press. http://dx.doi.org/10.1016/B978-0-12-255950-1.50025-1.

Leib, Hans-Heinrich. 1993. *Linguistic variables. Towards a unified theory of linguistic variation.* (Current Issues in Linguistic Theory 108). Amsterdam/Philadelphia: John Benjamins. https://doi.org/10.1075/cilt.108.

Markus, Elena. 2011. The phonetics and phonology of a disyllabic foot in Soikkola Ingrian. *Linguistica Uralica* 47(2), 103–119. http://dx.doi.org/10.3176/lu.2011.2.03.

Metslang, Helle, Külli Habicht & Tiit Hennoste. 2020. Introduction: What Varies when Language Varies? *Open Linguistics* 6, 690–692. https://doi.org/10.1515/opli-2020-0039.

Nikolaev, Ilya. 2001. Osnovnye tendencii v izmenenii grammatičeskoj sistemy ižorskogo jazyka v tečenie XX veka. In *Congressus Nonus Internationalis Fenno-Ugristarum 7.–13.8.2000 Tartu. Pars V. Dissertationes sectionum: Linguistica II*, 449–453. Tartu.

Nirvi, Ruben E. 1971. *Inkeroismurteiden sanakirja.* (Lexica societatis Fenno-Ugricae XVIII)*.* Helsinki: Suomalais-Ugrilainen Seura.

Porkka, Volmari. 1885. *Ueber den Ingrischen Dialekt: mit Berücksichtigung der übrigen finnisch-ingermanländischen Dialekte*. Helsingfors: J. C. Frenckell & Sohn.

Rožanskij, Fedor I. & Markus, Elena B. 2012. "Zolotaja ptica" (publikacija ižorskoj skazki, zapisannoj v XIX veke). *Acta linguistica petropolitana* 8(1), 448–503.

Reppen, Randi, Susan M. Fitzmaurice & Douglas Biber. 2002. *Using Corpora to Explore Linguistic Variation*. (Studies in Corpus Linguistics 9). Amsterdam/Philadelphia: John Benjamins. http://dx.doi.org/10.1075/scl.9.

Saar, Eva. 2017. *Isuri keele Soikkola murde sõnamuutmissüsteem.* Dissertationes philologiae uralicae Universitatis Tartuensis 19. Tartu: University of Tartu.

Siemund, Peter. 2011. Universals and variation: an introduction. In Peter Siemund (ed.), *Linguistic Universals and Language Variation*, 1–20. Berlin/New York: De Gruyter. http://dx.doi.org/10.1515/9783110238068.1.

Sovijärvi, Antti. 1944. *Foneettis-äännehistoriallinen tutkimus Soikkolan inkerois-murteesta* (Suomi-kirjan eripainoksia 5). Helsinki: Suomalaisen kirjallisuuden seura.

Tamminga, Meredith, Laurel MacKenzie & David Embick. 2016. The dynamics of variation in individuals. *Linguistic Variation* 16(2), 300–336. https://doi.org/10.1075/lv.16.2.06tam.

Van Olmen, Daniël, Tanja Mortelmans & Frank Brisard (eds.). 2019. *Aspects of Linguistic Variation*. (Trends in Linguistics. Studies and Monographs 324). Berlin/New York: De Gruyter. https://doi.org/10.1515/9783110607963.

Kokkuvõte. Anna Schwarz, Fedor Rozhanskiy: Isuri keele Soikkola murde pronoomenite süsteem: isikulised, demonstratiivsed, refleksiivsed ja retsi-prooksed pronoomenid ning nende varieerumine. Uurimus analüüsib isuri keele Soikkola murde pronoomenite süsteemi. Artikkel on olemuselt kirjeldav. Olemasolevad Soikkola murret puudutavad uurimused esitavad üsna pealiskaudse pronoomenite analüüsi ning puuduvad 21. sajandi andmetel põhinevad üksikasjalikud kirjeldused. Artikkel keskendub nii kõnelejasisesele kui ka kõnelejatevahelisele varieerumisele neljas isuri pronoomenite rühmas: personaal-, demonstratiiv-, refleksiiv- ja retsiprookpronoomenid. Autorite poolt 21. sajandil kogutud välitööde andmeid võrreldakse 19. ja 20. sajandil kirjutatud grammatikate andmetega. Uurimuses jõutakse järeldusele, et kuigi pronoomenite süsteem pole oluliselt muutunud, on isuri pronoomenite vormides ja nende varieerumises mitmeid uuendusi. Kõige sagedamini esineb varieerumist üksikute kõnelejate vahel, kuid mitte murde allrühmade vahel. Levinud on ka vaba varieerumine, samas kui muud tüüpi varieerumised on vähem olulised.

Märksõnad: isuri keel, pronoomenid, personaalpronoomenid, demonstratiiv-pronoomenid, refleksiivpronoomenid, retsiprookpronoomen, varieerumine, kirjeldus

Appendix 1. Vowel quality in plural personal pronouns

The research questions were:

1. What is the quality of the vowel *ö̈ö̈* in the 1st plural pronoun *mö̈ö̈* if compared with *üü* and *ö* in the first syllable of other words?

2. What type of variation between native speakers can we see based on the quality of *ö̈ö̈* in *mö̈ö̈*?

Data. Three sets of token words were analysed for twelve native speakers:

1. Word forms with *ö* in the first syllable (e.g. *köhä* 'cough', *vöglä* 'beetroot');

2. The personal pronoun *mö̈ö̈* 'we';

3. Word forms with *üü* in the first syllable (e.g. *püüťtää* 'to catch', *hüüťtöjä* 'to harden (about food)'). For 2 native speakers there were no examples with *üü*, therefore word forms with a short *ü* were used instead.

Every set contained from 10 to 15 token words.

The token words were taken from our corpus of elicitations, mostly from the questionnaire on vowel quality and quantity in monosyllabic words (this questionnaire was designed and recorded by Irina Brodskaya in 2014).

Experiment. We analysed F1 in the middle point[23] of the first vowel (F2 values of *öö* and *üü* do not show significant differences). Figure 2 presents our measurements: the average F1 for three vowels, as produced by twelve speakers. The X-axis contains labels for native speakers. Based on several features which usually define the dialectal characteristics of an idiolect[24] we ordered the native speakers starting from the representatives of the northernmost varieties (left) to the representatives of the southernmost varieties (right). The Y-axis shows F1 in Hz. In cases where the difference between the values of F1 for two vowels is not statistically significant ($p > 0.05$), these vowels are outlined with a solid line. A dashed line outlines the cases where the difference is questionable ($0.01 < p < 0.05$).

---

23 In fact, measurements were made for three points: one third, one half (middle) and two thirds of a vowel. As we did not observe any tendency towards diphthongization for long vowels, the middle point was chosen as the most representative.

24 These features are place of birth, place of residence, and place of birth of the speaker's parents.
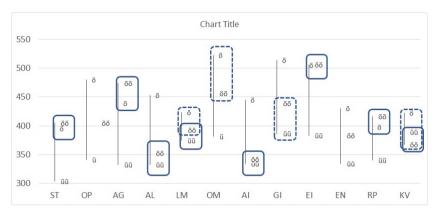
Figure 2. F1 (in Hz) of vowels in the first syllable.

Results. As follows from Figure 2, there are several classes of idiolects from the point of view of the vowel quality in the 1st plural pronoun.

Class 1. Four speakers have pronominal vowel *öö* that has the same quality as *ö* (ST, AG, EI, and RP).

Class 2. Four speakers have pronominal vowel *üü* (AL, LM, AI, and KV). Two of them (LM and KV) demonstrate a very subtle difference between *ö* and the pronominal vowel.

Class 3. Two speakers distinguish three vowels: *ö*, *ọ̈ọ̈* and *üü* (OP, EN).

Two speakers demonstrate marginal cases: OM belongs either to class (1) or to class (3), and GI belongs either to class (2) or to class (3).

It can clearly be seen that none of these classes tends to be to the left or to the right side of the scale (northernmost vs southernmost varieties). Thus, the distribution based on the vowel quality in the 1st plural pronoun is of an idiolectal rather than dialectal nature.[25]

---

25 Note that 3 of the speakers (AI, GI, EI) are sisters but their idiolects do not belong to the same class.

Appendix 2. Duration of the consonant in the allative marker of personal pronouns

The research question was whether there is variation between single-ton and geminate consonants in the allative marker -*(l)le* of the personal pronouns.

Data. The pronouns *miul(l)e* '1SG.ALL', *hänel(l)e* '3SG.ALL' and *meil(l)e* '1PL.ALL' were extracted from the recordings of the Basic Grammar Questionnaire and segmented in Praat. Usually there were 2–3 tokens for each speaker. Most of the token words were in clause-medial position but some of them (especially in the case of *miu(l)le*) were in sentence-final position. A preliminary analysis showed that the phrasal position was a factor of secondary importance from the point of view of the duration (excluding the final vowel). Thus, we did not distinguish pronunciations with different phrasal positions in our analysis.

Experiment. The average duration of every segment was calculated for three pronouns. As the speech tempo varied from speaker to speaker and influenced the durations, we also calculated the relative duration of each segment (as a percentage of the whole form). Table 5 presents these measurements.

Table 5. Average duration (in ms and %) of segments in the pronominal forms.

miule			m	iu	l	e	
Duration (ms)			**85**	**157**	**74**	**119**	35 speakers 87 tokens
StDev			21	49	23	76	
Duration (%)			**21**	**38**	**18**	**23**	
StDev			5	7	4	8	
hänele	h	ä	n	e	l	e	
Duration (ms)	**79**	**75**	**62**	**89**	**76**	**89**	34 speakers 77 tokens
StDev	26	24	15	23	23	43	
Duration (%)	**16**	**16**	**13**	**19**	**16**	**18**	
StDev	4	4	3	3	3	6	
meile			m	ei	l	e	
Duration (ms)			**90**	**169**	**71**	**88**	30 speakers 59 tokens
StDev			25	46	17	28	
Duration (%)			**22**	**40**	**17**	**21**	
StDev			5	7	3	5	

Results. Table 5 shows that the average duration of the consonant in the allative marker is shorter than the duration of all other segments in *miule* '1SG.ALL' and *meile* '1PL.ALL' and of the majority of segments in *hänele* '3SG.ALL'. It is also more than twice as short as the diphthongs in *miule* '1SG.ALL' and *meile* '1PL.ALL'. The relative duration corresponds to the expected duration of a singleton. The standard deviation of *l* is not higher than that of other segments (if some of the pronunciations had a singleton *l* and others had a geminate *ll* with a longer duration, we would expect the standard deviation to be high).

Two main phonetic features of a geminate are (a) a duration which is 2 to 3 times greater than that of a singleton (Sovijärvi 1944: 14, Markus 2011: 111), and (b) higher intensity. In our data, we do not notice any pronunciations where the consonant in the allative marker is unambiguously a geminate, although sometimes the analysed consonant sounds more intense than a typical singleton. This "shortened" geminate is attested in the samples from several speakers. Thus, we conclude that in contemporary Soikkola Ingrian the basic variant of the allative pronominal forms contains a singleton, and the variant with a geminate should be considered as secondary.