

FEDOR ROZHANSKIY & ELENA MARKUS
TARTU TARTU

Syncretism in Soikkola Ingrian verbal and nominal paradigms¹

1. Introduction

This paper discusses instances of syncretism across inflectional paradigms in Soikkola Ingrian. The topic of syncretism has been rather extensively discussed in the literature (Carstairs 1987; Blevins 1995; Stump 2001; Baerman et al. 2005; Trosterud 2006; Corbett 2007; Grünthal 2010a, and many more). Since our work is the first attempt to investigate syncretism in Ingrian, we choose to understand this notion in a maximally broad way. All occasions where we observe an identical element (not smaller than a word) in two cells of an inflectional paradigm are considered examples of syncretism. A more fine-grained analysis of Ingrian syncretism with possible distinctions between proper syncretism and homonymy (as expressed with various terms in some works, e.g., Trnka 1958; Blevins 1995; Stump 2001) remains the topic of further theoretical work. The broad understanding of syncretism is also convenient for practical implementation in language technology tools that perform morphological analysis.

Different types of syncretism have been distinguished in previous research. The approaches vary a lot, and the use of terminology is not consistent across the papers (including the term “syncretism” itself; see, for example, the alternative terms suggested in Haspelmath 2023: 4). An overview of various classifications of syncretism is beyond the scope of our study (many references and comments can be found, e.g., in Arkad’ev 2003; Baerman et al. 2005; Trosterud

1. The work by Fedor Rozhanskiy was supported by the University of Tartu, grant PHVEE18904. The work by Elena Markus was supported by the Estonian Research Council grant PRG1290.

2006; Hein & Weissman 2023). In this article, we propose several distinctions that are relevant for describing the syncretism that we observe in the Ingrian data.

a) Absolute vs. conditioned syncretism

Absolute syncretism is observed between several cells in a paradigm for all members of a certain word class (e.g., nouns), while conditioned syncretism only appears in a certain subgroup of words within a class. Usually, this subgroup is defined by some morphophonological parameters, i.e., it consists of particular paradigmatic types or of words with particular characteristics of the stem. For example, the syncretism of the nominative and accusative cases in the Russian noun *письмо* ‘letter.NOM.SG/ACC.SG’ only appears in certain paradigmatic types (defined by the grammatical gender, stem structure, and patterns of stress). Other nouns keep these two cases clearly distinct, e.g., Rus. *книга* ‘book.NOM.SG’ vs. *книгу* ‘book.ACC.SG’ (both examples are discussed by Baerman et al. 2005: 1–2). We call such syncretism *conditioned* (cf. also a similar opposition between symmetrical vs. asymmetrical syncretism introduced by Grünthal 2010a).

The proposed distinction between absolute and conditioned syncretism is formal. It does not require a deeper analysis of how syncretism has emerged, unlike for example the opposition between systematic and accidental homonymy (that partially intersects with the opposition under discussion) as described by Carstairs (1987: 93–102).

b) Complete vs. partial syncretism

With this division we distinguish the instances when two cells in an inflectional paradigm are fully identical from the instances when one cell is identical with an element (usually, a word) of an analytical form. It seems that the latter case is no less frequent than the former. *Partial* syncretism can be illustrated with the German form *arbeiten* when it is part of an analytical future tense, cf. *wir arbeiten* ‘we work’ and *wir werden arbeiten* ‘we will work’. At the same time, *arbeiten* can illustrate *complete* syncretism, as in *arbeiten* work.INF/PRS.PL ‘work, we work, you (pl) work, they work’.

Note that under partial syncretism we only consider the identity of full words, not the identity of morphemes or some other parts smaller than words.

c) Homonymic vs. polysemic syncretism

There are two main reasons why there can be identical forms in two different cells of a paradigm. In one scenario, a form can undergo phonetic changes for whatever morphophonological reasons and become identical with another form, although they were originally different. We call such instances *homonymic* syncretism. For example, synchronically the Russian *книги* is both the genitive singular and nominative plural of ‘book’, but historically it was only the genitive. The nominative plural used to be *книгы*, but it changed to *книги* when the velar consonants were palatalized.

In another scenario, identical forms which appear in several cells have the same origin. This situation can also be interpreted as one form that has several functions in the paradigm. For example, the contemporary Russian syncretic past tense form *видел* see.PST.1/2/3SG ‘I saw, you saw, he saw’ is historically a participle that used to combine into an analytical form with an auxiliary expressing the person. After the auxiliary was dropped, the forms of all three persons became identical. This kind of syncretism we label as *polysemic*.²

d) Variational syncretism

A language may have varying forms that occupy the same cell in a paradigm. One of the variants may happen to be syncretic with another cell in a paradigm, while the other variant is not. This is the case, for example, with the Luuditsa Votic illative singular. In our field materials, there are variants with and without the marker *-se/-se* (the unmarked variants are more frequent), e.g., *liiva ~ liivase* ‘sand.ILL’. The first variant is formally identical with the partitive and genitive singular forms of the same word: *liiva* ‘sand.GEN/PART/ILL’. Here we cannot say that the illative cell is fully identical with the partitive and genitive cells because there is also the second variant *liivase*. We call such instances *variational syncretism*. As this type of syncretism is relatively rare, we do not introduce a special label for the rest of the cases when there is no variation in a paradigm cell.



The rest of the article is organized as follows. Section 2 describes the data used for this research. In Section 3, we briefly present paradigmatic classes that we propose for Ingrian verbal and nominal inflection. This classification is relevant for the division into absolute vs. conditioned syncretism as described in the Introduction. Section 4 addresses the instances of verbal syncretism in Soikkola Ingrian while Section 5 discusses nominal syncretism. Section 6 summarizes the findings.

2. Data

The data used in this study come from recordings made during fieldwork with Soikkola Ingrian speakers in 2006–2023. The recordings include a vast corpus of elicited materials (about 700 hours) by more than 60 native speakers, and a significant part of these elicitations consists of verbal and nominal paradigms. For the research on syncretism, we used about 460 verbal and 1050 nominal paradigms (available in the morphological dictionary of Ingrian, see Rozhanskiy & Markus 2023). As a rule, a paradigm was recorded by several speakers. The main method of collecting the paradigms was the translation of simple sentences from Russian into Ingrian. The syntactic context of the sentence was designed to elicit the target morphological form.

2. See Baerman (2007: 546–548) on the diachronic aspects of syncretism.

3. Ingrian paradigmatic classes

Many patterns discussed in this article represent the morphophonologically conditioned syncretism that only appears in certain paradigmatic classes, both verbal and nominal. In this section, we briefly describe our system of paradigmatic classes for Soikkola Ingrian. The classification of verbal forms is discussed in detail in Rožanskij & Markus (2020); for nouns, it has not been previously described.

The system of verbal paradigmatic classes³ (henceforth verbal classes) is based on three parameters:

- a) whether a verb is one-stem (has only a vowel stem) or two-stem (has a vowel and consonant stem);⁴
- b) whether a verb has the past tense marker *-i* or *-iži*;
- c) what the morphophonological alternations in the main consonant (cluster) are. This parameter has four possible values: 1) no alternations, 2) consonant gradation (e.g., *temba-jaa*⁵ ‘pull-PRS.3SG’ – *temma-da* ‘pull-INF’, *kelba-jaa* ‘fit-PRS.3SG’ – *kelva-da* ‘fit-INF’), 3) gemination (e.g., *algu-maa* ‘begin-SUP’ – *alkku-a* ‘begin-INF’, *köhi-mää* ‘cough-SUP’ – *köhhi-ä*⁶ ‘cough-INF’), 4) both consonant gradation and gemination.

3. We would like to emphasize that our principles of classification and the way we divide words into morphemes are based on synchronic data. For the sake of elegance and consistency of the description, we sometimes ignore the underlying historical structure of the contemporary forms. This concerns, for example, our decision to distinguish the past tense marker *-iži* for all verbs where it precedes the personal endings, including those where historically *ž* was part of the stem due to the change **t > s* before *i*, e.g., *makka-iži-n* ‘sleep-PST-1SG’. Similarly, we distinguish the 3rd person present tense marker *-jaa/-jää*, even though historically *ja/jä* was part of the stem, e.g., *maga-jaa* ‘sleep-PRS.3SG’.

4. The coexistence of the so-called vowel stems and consonant stems in the same paradigm is common in Finnic languages (Laakso 2022: 243).

5. The transcription of Ingrian data in this article mostly follows the same principles as those implemented in *The Oxford Guide to the Uralic Languages*, see Markus & Rozhanskiy (2022: 312). The stressed long mid vowels, which can be pronounced by contemporary speakers as mid, mid-high, or high, are transcribed as mid-high.

6. This and other words where a short geminate is followed by two vowels are transcribed in Nirvi (1971) as having a long second vowel and a syllable boundary before the third vowel: *köhhiä* ‘cough.INF’, *käppiä* ‘beautiful’, *käššiä* ‘hand.PL.PART’, etc. In our data, the length of the second vowel in such forms varies a lot, and the syllable boundary is often lost, so the two last vowels are contracted into a diphthong and the words become disyllabic, e.g., *käppiä* ‘beautiful’. The question of how well the archaic

By the *main consonant/cluster*, we mean a consonant or a consonant cluster at the border between the last and penultimate syllable in the infinitive stem, e.g., *eḥti-ä* ‘have.time-INF’ (the main cluster is *ht*, cf. *eḥi-n* ‘have.time.PRS-1SG’), *töriš-šä* ‘chat-INF’ (the main cluster is *r*, cf. *törr̥iše-n* ‘chat.PRS-1SG’). For those two-stem verbs that have the infinitive stem consisting of only one syllable, the stem-final consonant is considered the main consonant cluster, e.g., *noiš-ša* ‘begin-INF’ (the main cluster is *š*,⁷ cf. *noiššo* ‘begin.PRS.3SG’).

Possible combinations of the binary parameters (a) and (b) give four options that we label with capital letters:

- A — one-stem verbs with the past tense marker *-i*;
- B — one-stem verbs with the past tense marker *-iži*;
- C — two-stem verbs with the past tense marker *-i*;
- D — two-stem verbs with the past tense marker *-iži*.

The four possible values of the parameter (c) in combination with the capital letters encoding the values of the parameters (a) and (b) constitute the indexes of the main paradigmatic classes in our classification (A₁, A₂, ..., D₁). For example, the index B₄ corresponds to the class of one-stem verbs that have the past tense marker *-iži*, and have both consonant gradation and gemination, like *luva-da* ‘promise-INF’ (cf. *luba-jaa* ‘promise-PRS.3SG’, *luṗpa-iž* ‘promise-PST.3SG’).

For some of the main classes, it is necessary to introduce subclasses of the second level, which we label as “d”, “g”, “m”, and “s”.

- d — verbs with a double set of alternating consonants. The second set appears due to the change of *d* into *ž* before the past tense marker *-i* (only happens in the main classes A₂ and A₄), e.g., *kün̄ttä-ä* ‘plow-INF’, *kündä-mää* ‘plow-SUP’, *künnä-n* ‘plow.PRS-1SG’, but *künž-i* ‘plow-PST.3SG’, *künšš-ii-d* ‘plow-PST-3PL’ with the alternating sets of clusters *n̄tt/nd/nn* and *nšš/nž*.
- g — verbs that have gemination in the consonant preceding the main consonant/cluster (happens in A₁, A₂, and C₁), e.g., *unohta-a* ‘forget-INF’, but *un̄nooha-n* ‘forget.PRS-1SG’ with the gemination of *n̄n/n* accompanying the gradation *ht/h*.

pronunciation of these words as trisyllabic has been preserved in the contemporary language requires further research. We choose the disyllabic variants with diphthongs as the main variants of transcription for such words.

7. Obviously, a geminate *šš* in *noišša* cannot be divided into two parts, but in morphophonological representation it is the consonant stem *noiš* plus the infinitive marker *-da* (*noiš-da*), so we distinguish *š* as the stem-final consonant.

- m — verbs with a monosyllabic vowel stem (only exist in A₁). Such verbs have certain specific markers different from regular A₁, e.g., *jää-vvä* ‘stay-INF’, *jää-kkää* ‘stay-IMP.2PL’, *jää-ttii* ‘stay-IPS.PST’.
- s — verbs that have an additional syllable in the supine and in many other forms as compared to the infinitive (happens in all A classes, also in C₁ and C₃), e.g., *kange-da* ‘harden-INF’, but *kangeno-maa* ‘harden-SUP’ and *kankkene-n* ‘harden.PRS-1SG’ with an additional syllable *-no-/-ne-* absent from the infinitive form.

There are also minor inflectional differences that appear due to the specific morphophonological structure of the stems in particular verbs. These details are explained in Rožanskij & Markus (2020: 118–119), but for further discussion of syncretism they are irrelevant.

The described system of verbal classes is summarized in Table 1.

Table 1: Verbal classes in Soikkola Ingrian (by Rožanskij & Markus 2020: 117)

	A	B	C	D
1	<i>lüpšää</i> ‘milk’ g <i>haravoj(j)a</i> ‘rake’ ^a m <i>šaa(v)a</i> ‘get’ s <i>valida</i> ‘choose’	<i>arvada</i> ‘under- stand’	<i>kuulla</i> ‘hear’ g <i>pilahušša</i> ‘get spoiled’ s <i>nuuhella</i> ‘smell’	<i>mättiššä</i> ‘climb’
2	<i>jättää</i> ‘leave’ d <i>hämmendää</i> ‘stir’ g <i>mörähtää</i> ‘cry’ s <i>habada</i> ‘get sour’	<i>leigada</i> ‘cut’		
3	<i>ujjua</i> ‘swim’ s <i>kangeda</i> ‘harden’	<i>šuvada</i> ‘love’	<i>šeišša</i> ‘stand’ s <i>ommella</i> ‘sew’	
4	<i>anĭtaa</i> ‘give’ d <i>tuntĭtia</i> ‘recognize’ s <i>paeda</i> ‘escape’	<i>keredä</i> ‘have time’	<i>tehä</i> ‘do’	

^a The problematic cases of variation in the length of *j* (as in *haravoj(j)a* ‘rake’) and epenthetic *v* (as in *šaa(v)a* ‘get’) are discussed in Rožanskij & Markus (2020: 104–105, 117).

The nominal classes are distinguished according to similar principles (Table 2). The numbers 1 to 4 in the paradigmatic index refer to the same four variants of the morphophonological alternations in the main consonant/cluster.⁸ The capital letters in the index refer to the nouns with the following characteristics:

8. In most cases, the main consonant/cluster is on the border between the last and penultimate syllable in the genitive stem.

- A — one-stem nouns with direct alternations (or without alternations);
 B — one-stem nouns with reverse alternations;
 C — two-stem nouns with direct alternations (or without alternations);
 D — two-stem nouns with reverse alternations.

By the direct vs. reverse alternations, we mean whether the genitive form of the noun has a weak or a strong grade of the alternating consonant/cluster. If it is weak, we call the alternation direct, as in, e.g., *matka* ‘trip’ – *mada-n* ‘trip-GEN’ – *matka-a* ‘trip-PART’. If the grade of the consonant/cluster in the genitive form is strong or if it has a stem with a short geminate, we call the alternation reverse, as in, e.g., *käppiä* ‘beautiful’ – *käppiä-n* ‘beautiful-GEN’ – *käbiä-ä* ‘beautiful-PART’.

Nominal subclasses “d”, “m”, and “g” refer to similar processes as in verbs:

- d — nouns with a double set of alternating consonants/clusters (**d* changes into *ž* before the plural marker *-i*), e.g., *käe-n* ‘hand-GEN’, *kättee* ‘hand.ILL’, but *käži* ‘hand’, *käšš-i-ä* ‘hand-PL-PART’.
 m — nouns with a monosyllabic vowel stem and with specific partitive and illative markers, e.g., *maa-da* ‘land-PART’, *maa-ha* ‘land-ILL’.
 g — nouns that have gemination in the consonant/cluster preceding the main consonant/cluster, e.g., *šeniikka* ‘groom’, but *šenniikka-n* ‘groom-GEN’.
 g’ — unlike in verbs, in nouns there are also instances of gemination in the consonant following the main consonant/cluster, e.g., *viigade* ‘scythe’, *viigattehe-n* ‘scythe-GEN’, but *viigattehh-i-a* ‘scythe-PL-PART’.

In the class C2.dg, two subclasses combine: there is a second set of alternations and also gemination in the preceding syllable, cf. *olue-šše* ‘beer-ILL’, *öllue-n* ‘beer-GEN’, and *oluiž-i-a* ‘beer-PL-PART’.

Table 2: Nominal classes in Soikkola Ingrian

	A	B	C	D
1	<i>lehmä</i> ‘cow’ <i>m puu</i> ‘tree’ <i>g naappuri</i> ‘neighbor’		<i>korluž</i> ‘shawl’	
2	<i>aitta</i> ‘granary’ <i>g šeniikka</i> ‘groom’		<i>pakkain</i> ‘frost’ <i>g’ vaabukkain</i> ‘raspberry’ <i>dg olud</i> ‘beer’	
3	<i>aiža</i> ‘shaft’	<i>käppiä</i> ‘beautiful’	<i>nain</i> ‘woman’	<i>avvain</i> ‘key’
4	<i>leibä</i> ‘bread’		– <i>d herž</i> ‘log’	<i>lammaž</i> ‘sheep’

4. Syncretism in the verbal paradigm

In Ingrian verbal paradigms, the following sets of syncretic forms are found:

1. The infinitive and 3Sg present tense form;
2. The infinitive and the impersonal present tense connegative;
3. The infinitive and the lexical verb in the negative 2Pl imperative;
4. The present and past tense forms;
5. The active past participle and the lexical verb in several analytical forms: perfect, pluperfect, and negative past;
6. The passive past participle and the lexical verb in the impersonal forms (perfect, pluperfect, and negative past);
7. The present tense connegative and 2Sg imperative (affirmative and negative);
8. The 3Sg conditional and conditional connegative;
9. The impersonal conditional and the lexical verb in the negative impersonal conditional.

Items (5) to (9) are characteristic of the Finnic languages in general, so we mention them only briefly.

4.1. Infinitive and 3Sg present

The syncretism between these two forms appears in most classes of one-stem verbs with the past tense marker *-i* (class A), if their vowel stem ends in *a* or *ä*: e.g., *kaivaa* ‘dig.INF/PRS.3SG’ (A1), *kaštaa* ‘moisten.INF/PRS.3SG’ (A2), *vähendää* ‘lessen.INF/PRS.3SG’ (A2.d), *unohtaa* ‘forget.INF/PRS.3SG’ (A2.g), *ellää* ‘live.INF/PRS.3SG’ (A3), *anttaa* ‘give.INF/PRS.3SG’ (A4), *keļļttää* ‘forbid.INF/PRS.3SG’ (A4.d).

The merger of the two forms happens due to different phonological mechanisms that give the same result. When the infinitive marker *-a* or *-ä* is added to the stem with the same final vowel, the result is a long final vowel of the form, e.g., *kaiva-* ‘dig’ + *-a* ‘INF’ > *kaivaa* ‘dig.INF’. As for the 3Sg form, in the listed classes it is built through the lengthening of the stem-vowel, e.g., *kaiva-* ‘dig’ > *kaivaa* ‘dig.PRS.3SG’.

There is no syncretism, if the verb:

- has a stem-final vowel other than the short *a* or *ä*, cf. *hääülü-ä* ‘wander-INF’ vs. *hääülüü* ‘wander.PRS.3SG’ (A1), *šaa-vva* ‘get-INF’ vs. *šaa-b* ‘get-PRS.3SG’ (A1.m);
- uses the infinitive marker *-da* or *-dä*, e.g., *korja-da* ‘collect-INF’ vs. *korja-jaa* ‘collect-PRS.3SG’ (B1), including cases with assimilation, e.g., *haigoittas-ša* ‘yawn-INF’ vs. *haigoitta-jaa* ‘yawn-PRS.3SG’ (D1).

Since this syncretism is obviously restricted to certain stem-final vowels, it is classified as *conditioned* in our system. It is also *complete*, since for both syncretic forms there are no other elements in the corresponding cells. The origin of each form is different as shown above, so this syncretism is *homonymic*.

4.2. Infinitive and impersonal connegative

The negative present tense impersonal form is analytical: it consists of the negative auxiliary verb *ei* 'NEG.3SG' and the impersonal connegative of the lexical verb, e.g., *ei eledä* NEG.3SG live-IPS.CNG 'one does not live'. This connegative is identical with the infinitive in the subclasses "s" and "m" of the main class A and in all subclasses of the main classes B, C, D: e.g., *vali-da* 'choose-INF/IPS.CNG' (A1.s), *mušše-da* 'darken-INF/IPS.CNG' (A2.s), *kange-da* 'harden-INF/IPS.CNG' (A3.s), *pae-da* 'escape-INF/IPS.CNG' (A4.s), *lõõ-vvä* 'hit-INF/IPS.CNG' (A1.m), *arva-da* 'understand-INF/IPS.CNG' (B1), *leiga-da* 'cut-INF/IPS.CNG' (B2), *šuva-da* 'love-INF/IPS.CNG' (B3), *lää-dä* 'talk-INF/IPS.CNG' (B4), *jõõš-ša* 'run-INF/IPS.CNG' (C1), *nuuhel-la* 'smell-INF/IPS.CNG' (C1.s), *avahuš-ša* 'open-INF/IPS.CNG' (C1.g), *peš-šä* 'wash-INF/IPS.CNG' (C3), *ommel-la* 'sew-INF/IPS.CNG' (C3.s), *teh-ä* 'do-INF/IPS.CNG' (C4), *istuš-ša* 'sit.down-INF/IPS.CNG' (D1).

The listed subclasses cover all verbs that have the infinitive in *-da/-dä*, monosyllabic verbs (A1.m), and the verbs *tehä* 'do', *nähä* 'see' that constitute subclass C4. Additionally, the impersonal connegative and infinitive are syncretic in the verbs *aštavoj(j)-a* 'harrow-INF/IPS.CNG' (A1) and *haravoj(j)-a* 'rake-INF/IPS.CNG' (A1.g).

In all the listed cases, the markers of the infinitive and impersonal connegative are the same, and also the grade of the consonant/cluster in the corresponding forms is the same. There is no syncretism in A-class verbs except those mentioned above because they have the infinitive marked with *-a/-ä*, and the impersonal connegative marked with *-da/-dä*, e.g., *ellää* 'live-INF' vs. *ele-dä* 'live-IPS.CNG' (A3), *tuntti-a* 'recognize-INF' vs. *tunne-da* 'recognize-IPS.CNG' (A4.d), *nibikkoi-a* 'tie-INF' vs. *nibigoi-da* 'tie-IPS.CNG' (A2).

The syncretism of the infinitive and impersonal connegative is *conditioned* because it is observed only in certain paradigmatic classes. This is an instance of *partial* syncretism because the impersonal connegative form is always part of the analytical construction with the negative verb. This syncretism is *homonymic* because the two forms have a different origin (see, for example, Laanest 1978: 292, 296), which is especially obvious in the paradigmatic classes where they do not merge.

4.3. Infinitive and negative 2Pl imperative

The negative imperative 2Pl constructions are built with the 2Pl prohibitive of the auxiliary verb (*elkkää*) in combination with the lexical verb. For all Soikkola Ingrian verbs, irrespective of the paradigmatic class and of any morphophonological conditions, the form of the lexical verb in the negative 2Pl imperative is syncretic with the infinitive, cf. *lүpšää* ‘milk-INF’ – *elkkää lүpšää* ‘Do not milk! (Pl)’, *kүššü-ä* ‘ask-INF’ – *elkkää күššüä* ‘Do not ask! (Pl)’, *lää-dä* ‘speak-INF’ – *elkkää läädä* ‘Do not speak! (Pl)’, *kuunnel-la* ‘listen-INF’ – *elkkää kuunnella* ‘Do not listen! (Pl)’, *ištuš-ša* ‘sit.down-INF’ – *elkkää ištušša* ‘Do not sit down! (Pl)’.

This is an example of *absolute* syncretism; it works for all verbs in the language. The syncretism is *partial* because the negative imperative 2Pl construction is analytical and only the lexical verb is syncretic with the infinitive. In most Finnic varieties, the lexical verb in the negative 2Pl imperative is the same as in the corresponding positive form. In Ingrian, it is different, and according to Laanest (1978: 295) it was generalized by analogy with forms like **laulakak* > *laulä(σ)*. This explanation shows the different origin of the two discussed forms, so this syncretism is *homonymic*.

4.4. Present and past tense

The syncretism between the present and past tense forms appears if a verb:

- (1) has a vowel stem ending in a short *i* or in a diphthong with *i* as the final component (in our dataset, these diphthongs are *oi*, *öi*, and *ai*);
- (2) attaches the past tense marker *-i*.

These conditions are met in class A (except subclasses “s” and “d” where no verbs with the stem-final *i* were attested).

In such verbs, five personal forms are identical in the present and past, while the 3Sg form is different, see examples of A2, A3, and A4 classes in Table 3. Examples of 1Sg in other classes are: *kääri-n* ‘roll.PRS/PST-1SG’ (A1), *härraavoi-n* ‘rake.PRS/PST-1SG’ (A1.g), *nai-n* ‘marry.PRS/PST-1SG’ (A1.m), *märreehi-n* ‘ruminate.PRS/PST-1SG’ (A2.g), *karži-n* ‘scrape.PRS/PST-1SG’ (A3), *hүүdöi-n* ‘harden.PRS/PST-1SG’ (A3), *revi-n* ‘tear.PRS/PST-1SG’ (A4), *leivoi-n* ‘knead dough.PRS/PST-1SG’ (A4).

The syncretism appears due to the merger of the past tense marker *-i* with the stem-final *i*. The 3Sg present and past forms are not syncretic for several reasons: either there is gemination in one of the stems but not in the other (classes A1.g, A4 ending in a diphthong, cf. *jäättöjä* ‘get frozen’ in Table 3), or

Table 3: Examples of past and present tense syncretism

	<i>ehtii</i> ‘have time’ class A2		<i>köhhii</i> ‘cough’ class A3		<i>jäättöjä</i> ‘get frozen’ class A4	
	Present	Past	Present	Past	Present	Past
1Sg	<i>ehi-n</i>		<i>köhi-n</i>		<i>jääöi-n</i>	
2Sg	<i>ehi-d</i>		<i>köhi-d</i>		<i>jääöi-d</i>	
3Sg	<i>ehtii</i>	<i>ehi</i>	<i>köhhii</i>	<i>köhi</i>	<i>jäättöi</i>	<i>jäädöi</i>
1Pl	<i>ehi-mmä</i>		<i>köhi-mmä</i>		<i>jääöi-mmä</i>	
2Pl	<i>ehi-ttä</i>		<i>köhi-ttä</i>		<i>jääöi-ttä</i>	
3Pl	<i>ehtii-d</i>		<i>köhhii-d</i>		<i>jäättöi-d</i>	

there is a long final *ii* in the present but a short *i* in the past (classes A1, A2, A2.g ending in *i*, cf. *ehtii* ‘have time’ in Table 3), or both (classes A3, A4 ending in *i*, cf. *köhhii* ‘cough’ in Table 3). In class A1.m, there is a specific PRS.3SG marker *-b*, so there is no syncretism with the past tense: *nai-b* ‘marry.PRS-3SG’ vs. *nai* ‘marry.PST.3SG’.

Verbs of the classes A1 and A2 that have a vowel stem ending in a diphthong present a special case. In earlier sources, their 3SG forms are distinguished through the length of the final diphthong: *murjoi* ‘wrinkle.PRS.3SG’ vs. *murjoï* ‘wrinkle.PST.3SG’ (Nirvi 1971: 319), *hì`èstöi* ‘sweat.PRS.3SG’ vs. *hì`èstöï* ‘sweat.PST.3SG’ (Sovijärvi 1944: 22). In the experimental research (Rozhanskiy & Markus 2019), it is shown that the length contrast of the final diphthongs is preserved only by some contemporary speakers, but lost in the innovative idiolects, so for them such verbs are syncretic in all six personal forms: *aštavoi* ‘harrow.PRS/PST.3SG’ (A1), *leikkoi* ‘cut.PRS/PST.3SG’ (A2).

The syncretism addressed in this subsection is *conditioned* and *complete* in our classification. It is again *homonymic* because the morphological structure of the forms is different due to the presence/absence of the past tense marker.

4.5. Active past participle and lexical verb in analytical forms

In the following five subsections, we will briefly discuss several cases of syncretism that are common for most Finnic varieties, including Soikkola Ingrian. All these cases involve analytical forms, where the lexical verb is identical with some other form in the paradigm. As mentioned above, we aim at covering all cases of formal identity in Ingrian morphological paradigms. So, we list these forms although some other definitions of syncretism do not cover them.

The form identical with the past tense active participle is used in the negative past tense, positive and negative perfect tense (indicative and conditional), and positive and negative pluperfect tense. In Ingrian, the active participle distinguishes the singular and plural forms. Correspondingly, the singular form of the participle is found in all the mentioned tense/mood singular forms, while the plural form of the participle is found in plural forms.

Table 4 contains forms of the verb *anġtaa* ‘give’. All singular forms (illustrated with the 1Sg form) are partially syncretic with the active past participle *anġtaand* (Sg); all plural forms (illustrated with the 3Pl form) are partially syncretic with the active past participle *andaneend* (Pl).

Table 4: Personal analytical forms of the verb *anġtaa* ‘give’

Polarity	Mood	Tense	1Sg	3Pl
Positive	Indicative	Perfect	<i>oġn anġtaand</i>	<i>ovad andaneend</i>
		Pluperfect	<i>olin anġtaand</i>	<i>oġliid andaneend</i>
	Conditional	Perfect	<i>oġliiġin anġtaand</i>	<i>oliġivad andaneend</i>
		Negative	Indicative	Past
Negative	Indicative	Perfect	<i>en oġ anġtaand</i>	<i>eväd oġ andaneend</i>
		Pluperfect	<i>en old anġtaand</i>	<i>eväd olleend andaneend</i>
		Conditional	Perfect	<i>en oġliiġ anġtaand</i>

4.6. Passive past participle and the lexical verb in analytical forms

The form identical with the past tense passive participle is used in the impersonal forms with the same tense, mood, and polarity characteristics as those discussed in 4.5. Table 5 presents the impersonal forms of *anġtaa* ‘give’ that are partially syncretic with the passive past participle *annettu*.

The same form is also used in the stative passive constructions, e.g., *metsä ei old kerdettü* ‘The forest was not touched’. It is not obvious whether such passive constructions should be included in the verbal paradigm, and we do not address this question in our paper.

Table 5: Impersonal analytical forms of the verb *anĭtaa* ‘give’

Polarity	Mood	Tense	Impersonal
Positive	Indicative	Perfect	<i>on annettu</i>
		Pluperfect	<i>oli annettu</i>
Negative	Conditional	Perfect	<i>ölliiž annettu</i>
	Indicative	Past	<i>ei annettu</i>
		Perfect	<i>ei-oo annettu</i>
		Pluperfect	<i>ei old annettu</i>
	Conditional	Perfect	<i>ei ölliiž annettu</i>

4.7. 2Sg imperative and present tense connegative

The 2Sg imperative form is identical with the present tense connegative, and also with the form of the lexical verb in negative 2Sg imperative in all Finnic varieties (Laakso 2022: 247–249), and also in Soikkola Ingrian, cf. *maĭkaa* ‘Sleep! (Sg)’, *elä maĭkaa* ‘Do not sleep! (Sg)’, *miä en maĭkaa* ‘I do not sleep’.

4.8. 3Sg conditional and conditional connegative

The 3Sg conditional is syncretic with the conditional connegative used in negative present conditional forms in all persons, cf. *jättäiž* ‘leave.COND.3SG’ and *en jättäiž* ‘NEG.1SG leave.COND.CNG’, *ed jättäiž* ‘NEG.2SG leave.COND.CNG’, *eväd jättäiž* ‘NEG.3PL leave.COND.CNG’, etc.

4.9. Impersonal present conditional (positive and negative)

The impersonal present conditional form is syncretic with the form of the lexical verb used in the negative impersonal present conditional forms, cf. *keide-ttäiž* ‘cook-IPS.COND’ and *ei keide-ttäiž* ‘NEG.3SG cook-IPS.COND’.

We classify all cases of syncretism discussed in Sections 4.5.–4.9. as *absolute* and *partial* because they apply to all verbs without any morphological restrictions and there are analytical constructions involved. They are also *polysemic* because no traces of different origins of the syncretic forms are attested.

5. Syncretism in the nominal paradigm

Unlike verbs, the nominal paradigm in Ingrian does not have analytical forms, therefore all the instances of syncretism among the nouns are complete. Analyzing nominal inflection in Finnic languages, Grünthal (2010a: 106–108) demonstrates that the number of syncretic forms in a paradigm is directly connected with the degree of fusion/agglutination in the morphology of a language. The agglutinative languages tend to preserve the one-form-one-meaning principle with little deviation, so they have considerably less syncretism compared to fusional languages. Among the Finnic languages, one can observe a geographical distribution of the discussed features. Languages of the southern group are more fusional and have a lot of syncretism, while the northern languages are more agglutinative, and syncretism is rather an exception in their inflectional paradigms.

Ingrian belongs to the northern group of Finnic, together with Finnish, Karelian, and Veps. Not surprisingly, we find only few instances of syncretism in the Ingrian nominal inflection. Most of them are morphophonologically conditioned and work only with certain stem-final vowels or in certain paradigmatic classes.

The following sets of syncretic forms are found in the nominal paradigms:

1. The partitive and illative singular;
2. The partitive singular and nominative plural;
3. The genitive and essive singular;
4. The genitive and essive plural.

5.1. Partitive and illative singular

This syncretism is structurally similar to the syncretism between the infinitive and 3Sg present verbal forms. The partitive marker has several variants, one of them is *-a/-ä*. The illative form can have a dedicated marker *-šše*, but more frequently the illative is built through the prolongation of the stem-final vowel. For one-stem nouns that end in a short *-a/-ä* or in a diphthong with *-a/-ä* as the second component, the partitive and illative singular forms coincide. These conditions are met in classes A (except A1.m) and B, e.g., *heinää* ‘hay.PART/ILL’ (A1), *kadajaa* ‘juniper.PART/ILL’ (A1.g), *aittaa* ‘granary.PART/ILL’ (A2), *haaþþaa* ‘aspen.PART/ILL’ (A3), *jalkkaa* ‘leg.PART/ILL’ (A4), *moržiaa* ‘bride.PART/ILL’ (B3).

This syncretism cannot appear in the classes C and D of our classification because in the two-stem nouns the partitive form is always built from

the consonant stem. The partitive is therefore always distinct from the illative forms built from the vowel stem, cf. *lunda* ‘snow.PART’ vs. *lum̃mee* ‘snow.ILL’ (C3), *hevoišť* ‘horse.PART’ vs. *heboižee* ‘horse.ILL’ (D4). The two forms are also not syncretic in one-stem nouns ending in a long vowel because they have specific markers both for the partitive and illative, cf. *kiukka-da* ‘stove-PART’ vs. *kiukka-šše* ‘stove-ILL’ (from *kiukkaa*, A1), *maa-da* ‘land-PART’ vs. *maa-ha* ‘land-ILL’ (from *maa*, A1.m).

The partitive and illative can only coincide in the singular. In plural, the two forms are always different because two different markers are added to the same plural stem, e.g., *aitto-j-a* ‘granary-PL-PART’ vs. *aitto-i* ‘granary-PL-ILL’ (A2).

This syncretism is *complete* and *conditioned* for the reasons just described. It is *homonymic* because the origin of the forms is not the same. The partitive singular form is built with the marker *-a/-ä* added to the stem ending in *a/ä*, so it results in a long final vowel, e.g., *heinä* ‘hay’ + *-ä* ‘PART’ > *heinää* ‘hay.PART’. In the illative form, the original marker was **-hV*, but the intervocalic *h* was lost, and the form is marked with the lengthening of the stem-final vowel (Laanest 1978: 216–217), e.g., *heinä* > *heinää* ‘hay.ILL’.

For the B3 class words, this syncretism is *variational*, because some speakers tend to inflect these words similarly to A1 words that end in diphthongs, e.g., *moržiaa* ~ *moržiada* ‘bride.PART’, *moržiaa* ~ *moržiašše* ‘bride.ILL’ (the second variants of the forms are built by analogy with A1 words like *pehmiä* ‘soft’, cf. *pehmiädä* ‘soft.PART’, *pehmiäšše* ‘soft.ILL’).

5.2. Partitive singular and nominative plural

In Ingrian, all nouns that end in *-n* in the nominative and have the vowel stem in *-že* have two variants of the nominative plural: the expected regular variant with the marker *-d* added to the vowel stem (e.g., *mogomaiže-d* ‘such-PLNOM’, *ihmiže-d* ‘man-PLNOM’, *heppoize-d* ‘horse-PLNOM’) and a more frequent contracted variant with the final vowel dropped (e.g., *mogomaiš-t* ‘such-PLNOM’, *ihmiš-t* ‘man-PLNOM’, *heppoiš-t* ‘horse-PLNOM’).

All such words have the consonant stem in *š* and build the partitive singular forms with the marker *-t*. In classes C1 and C3, the partitive singular forms are identical with the contracted variants of the nominative plural, e.g., *hiljaiš-t* ‘quiet-PART/PLNOM’ (C1), *ihmiš-t* ‘man-PART/PLNOM’ (C1), *jogahiiš-t* ‘every-PART/PLNOM’ (C3), *roožvoiš-t* ‘pink-PART/PLNOM’ (C3).

There is no syncretism outside C1 and C3 because the main consonant is different in the two forms, cf. *idigaiš-t* ‘insect-PART’ vs. *idikkaiš-t* ‘insect-PLNOM’ (C2.g’), *pahaiš-t* ‘bad-PART’ vs. *pahhais-t* ‘bad-PLNOM’ (D3), *hevoiš-t* ‘horse-PART’ vs. *heppoiš-t* ‘horse-PLNOM’ (D4).

This syncretism is obviously *homonymic* and *conditioned*. It represents a *variational* case because only one variant of the nominative plural is syncretic with the partitive form.

5.3. Genitive and essive singular

Those nouns that have a vowel stem ending in a long vowel, have syncretic genitive and essive singular forms, e.g., *puu-n* ‘tree-GEN/ESS’ (A1.m), *lämmää-n* ‘warm-GEN/ESS’ (A1), *vaattii-n* ‘clothes-GEN/ESS’ (C2). For all other nouns, the genitive and essive are always different in the singular because of the stem-final vowel length⁹ and additionally the consonant grade or gemination (when applicable), e.g., *külmä-n* ‘cold-GEN’ vs. *külmää-n* ‘cold-ESS’, *käppiä-n* ‘beautiful-GEN’ vs. *käbiää-n* ‘beautiful-ESS’, *naiže-n* ‘woman-GEN’ vs. *naiššee-n* ‘woman-ESS’, *aida-n* ‘granary-GEN’ vs. *aittaa-n* ‘granary-ESS’.

This syncretism is *complete* and *conditioned*, as it is restricted to nouns of a specific morphophonological structure. It is *homonymic* because the essive marker *-n* was generalized in contemporary Soikkola Ingrian, but more variants of the marker were attested in earlier sources (Laanest 1978: 222).

5.4. Genitive and essive plural

The genitive and essive plural forms are identical for all nouns in Ingrian, irrespective of any morphophonological conditions, e.g., *lehm-ii-n* ‘cow-PL-GEN/ESS’, *naaburi-loi-n* ‘neighbor-PL-GEN/ESS’, *käbiö-i-n* ‘beautiful-PL-GEN/ESS’. This syncretism happens because both forms are built with the same marker *-n* added to the plural stem.

This syncretism is *absolute*. It developed due to the generalization of the essive marker *-n* as described in 5.3, so it is *homonymic*. This case represents another example of *variational* syncretism in Ingrian because for all nouns with the plural marker *-i* that have gradation or gemination in the vowel stem, there are two parallel variants of the genitive plural, e.g., *aitto-i-n* ~ *aido-i-n* ‘granary-PL-GEN’, *lamppah-i-n* ~ *lambah-ii-n* ‘sheep-PL-GEN’, *leipp-ii-n* ~ *leiv-ii-n* ‘bread-PL-GEN’. There is no such variation in the plural essive forms, but they are always syncretic with one of the variants of the genitive plural.

9. Speakers from the southern parts of the Soikkola peninsula tend to lose the distinction between long and short vowels in non-first syllables, so they might pronounce the genitive and essive singular forms in a similar way, see more details in Markus & Rozhanskiy (2017: 116–117).

6. Conclusions

Table 6 summarizes all the cases of syncretism discussed in this paper and their characteristics. A plus in parentheses means that this value is true only for a small part of the relevant cases (see 5.1. for the variational syncretism in nouns from the class B₃).

Table 6. Syncretism in Soikkola Ingrian

Type	Absolute (+) / Con- ditioned (-)	Complete (+) / Par- tial (-)	Homonymic (+) / Poly- semic (-)	Vari- ational
Inf and Prs.3Sg	-	+	+	
Inf and Ips.Cng	-	-	+	
Inf and Neg.Imp.2Pl	+	-	+	
Prs and Pst	-	+	+	
Active participle and analytical forms	+	-	-	
Passive participle and analytical forms	+	-	-	
Imp2Sg and Cng	+	-	-	
Cond.3Sg and Cond.Cng	+	-	-	
Ips.Cond and Neg.Ips.Cond	+	-	-	
Part and Ill	-	+	+	(+)
Part and PlNom	-	+	+	+
Gen and Ess	-	+	+	
Gen.Pl and Ess.Pl	+	+	+	+

Although the Ingrian language represents the northern group of Finnic, it displays quite a few instances of syncretic forms. However, as in Veps (cf. Grünthal 2010b), syncretism in Ingrian does not normally blur the distinctions between individual categories. Ingrian is a language with a highly complex system of morphophonological alternations, both qualitative and quantitative, and most syncretic non-analytical forms only occur in certain paradigmatic classes or with certain constraints on the stem structure. Additionally, the syntactic context usually helps to unambiguously distinguish between the syncretic forms. Unlike in the southern Finnic languages, syncretism in Ingrian does not occur in the grammatical cases of the nouns. The only exception is the merger of the partitive singular and nominative plural, but it only takes

place in a limited number of nouns with a certain stem structure, and the difference in number is an additional feature that helps to distinguish the forms in context.

Syncretism is more often found in the Ingrian verbal rather than nominal paradigms, which is not surprising considering that the verbal paradigm has considerably more forms, including analytical ones (however, if we exclude the analytical forms and look at complete syncretism only, the nouns demonstrate more instances of syncretism than the verbs). The most prominent blurring of morphosyntactic properties occurs between the present and past tense forms of *i*-stem verbs. When these occur in a sentence, they indeed require strong syntactic or lexical cues (for instance, temporal adverbials) in order to be set apart.

Abbreviations

ACC	accusative	NOM	nominative
CNG	connegative	PART	partitive
COND	conditional	PL	plural
ESS	essive	PLNOM	nominative plural
GEN	genitive	PRS	present tense
ILL	illative	PST	past tense
IMP	imperative	SG	singular
IPS	impersonal	SUP	supine
INF	infinitive	1, 2, 3	1st, 2nd, 3rd person
NEG	negative		

References

- Arkad'ev 2003 = Аркадьев П. М. 2003. *Типология надежного синкретизма*. Курсовая работа. Москва: Российский государственный гуманитарный университет.
- Baerman, Matthew & Brown, Dunstan & Corbett, Greville G. 2005. *The syntax-morphology interface: A study of syncretism* (Cambridge Studies in Linguistics 109). Cambridge: Cambridge University Press.
 <<https://doi.org/10.1017/CBO9780511486234>>
- Baerman, Matthew. 2007. Syncretism. *Language and Linguistics Compass* 1(5). 539–551. <<https://doi.org/10.1111/j.1749-818x.2007.00024.x>>
- Blevins, James P. 1995. Syncretism and paradigmatic opposition. *Linguistics and Philosophy* 18(2). 113–152. <<https://www.jstor.org/stable/25001583>>
- Carstairs, Andrew. 1987. *Allomorphy in inflection*. London: Croom Helm.

- Corbett, Greville G. 2007. Deponency, syncretism, and what lies between. In Baerman, Matthew & Corbett, Greville G. & Brown, Dunstan & Hippisley, Andrew (eds.), *Deponency and morphological mismatches* (Proceedings of the British Academy 145), 21–43. Oxford: Oxford University Press.
 <<https://doi.org/10.5871/bacad/9780197264102.003.0002>>
- Grünthal, Riho. 2010a. Sijasynkretismi morfologian koetinkivenä. *Journal of Estonian and Finno-Ugric Linguistics* 1:2. 91–113.
- Grünthal, Riho. 2010b. Syncretism in the central Veps local case system. *Linguistica Uralica* XLVI(4). 250–260.
 <<https://doi.org/10.3176/lu.2010.4.02>>
- Haspelmath, Martin. 2023. Coexpression and synexpression patterns across languages: Comparative concepts and possible explanations. *Frontiers in Psychology* 14: 1236853. <<https://doi.org/10.3389/fpsyg.2023.1236853>>
- Hein, Johannes & Weisser, Philipp. 2023. Syncretism: Recurring patterns. In Ackema, Peter & Bendjaballah, Sabrina & Bonet, Eulàlia & Antonio Fábregas (eds.), *The Wiley Blackwell companion to morphology*, volume 5, 2551–2576. Hoboken: Wiley-Blackwell.
- Laakso, Johanna. 2022. Finnic: General introduction. In Bakro-Nagy, Marianne & Laakso, Johanna & Skribnik, Elena (eds.), *The Oxford guide to the Uralic languages*, 240–253. Oxford: Oxford University Press.
 <<https://doi.org/10.1093/oso/9780198767664.001.0001>>
- Laanest 1978 = Лаанест, А. Х. 1978. *Историческая фонетика и морфология ижорского языка*. Диссертация на соискание ученой степени доктора филологических наук. Таллин: Институт языка и литературы АН ЭССР.
- Markus, Elena & Rozhanskiy, Fedor. 2017. The essive in Ingrian. In de Groot, Casper (ed.), *Uralic essive and the expression of impermanent state* (Typological Studies in Language 119), 113–129. Amsterdam: John Benjamins. <<https://doi.org/10.1075/tsl.119.05mar>>
- Markus, Elena & Rozhanskiy, Fedor. 2022. Ingrian. In Bakro-Nagy, Marianne & Laakso, Johanna & Skribnik, Elena (eds.), *The Oxford guide to the Uralic languages*, 308–329. Oxford: Oxford University Press.
 <<https://doi.org/10.1093/oso/9780198767664.003.0018>>
- Nirvi, Ruben E. 1971. *Inkeröismurteiden sanakirja* (Lexica Societatis Fenno-Ugricae XVIII). Helsinki: Suomalais-Ugrilainen Seura.
- Rozhanskiy, Fedor & Markus, Elena. 2019. On the quantitative contrast of diphthongs in Soikkola Ingrian. (Paper presented at the Linguistic Forum 2019: Indigenous languages of Russia and beyond, Moscow, April 4–6, 2019.)
- Rožanskij & Markus 2020 = Рожанский, Ф. И. & Маркус, Е. Б. 2020. Система парадигматических классов глагола в сойкинском диалекте ижорского языка. *Вопросы языкознания* 3. 101–134.
 <<https://doi.org/10.31857/S0373658X0009348-1>>

-
- Rozhanskiy, Fedor & Markus, Elena. 2023. *IngriDict 1.0: Morphological dictionary of Soikkola Ingrian*. (Computer program.) Retrieved 2023-05-29 from <<https://ingrian.org/ingrian-dictionary>>
- Sovijärvi, Antti. 1944. *Foneettis-äännehistoriallinen tutkimus Soikkolan inke-roismurteesta* (Suomi-kirjan eripainoksia 5). Helsinki: Suomalaisen Kirjallisuuden Seura.
- Stump, Gregory T. 2001. *Inflectional morphology: A theory of paradigm structure* (Cambridge Studies in Linguistics 93). Cambridge: Cambridge University Press. <<https://doi.org/10.1017/CBO9780511486333>>
- Trnka, Bohumil. 1958. On some problems of neutralization. In *Omagiu lui Iorgu Iordan cu prilejul împlinirii a 70 de ani*, 861–866. București: Academia Republicii Populare Romîne.
- Trosterud, Trond. 2006. *Homonymy in the Uralic two-argument agreement paradigms* (Suomalais-Ugrilaisen Seuran Toimituksia 251). Helsinki: Finno-Ugrian Society.