FUNCTIONS OF DEMONSTRATIVE PRONOUNS
IN CONTEMPORARY SOIKKOLA INGRIAN

Elena Markus, Fedor Rozhanskiy
University of Tartu, EE
elmarkus@ut.ee, handarey@yahoo.com

Abstract. This article discusses demonstrative pronouns in the Soikkola dialect of the Ingrian language. The material for the research comes from a collection of spontaneous speech samples recorded in the 21st century. Most examples are being published for the first time. The article presents mainly a qualitative analysis and aims to provide an overview of the basic functions of the Ingrian demonstrative pronouns *tämä and še. It is shown that in contemporary Ingrian the demonstratives are used as deictic and anaphoric devices, and also as discourse markers. Individual speakers demonstrate differences in the use of these pronouns due to the contact influence from neighbouring languages.

Keywords: Ingrian, pronouns, demonstrative pronouns, deixis, anaphora, discourse markers, language contact

DOI: https://doi.org/10.12697/jeful.2023.14.2.03

1. Introduction

Four historical stems have been reconstructed as sources of contemporary Finnic demonstratives: *tämä, *še, *too and *taa (Larjavaara 1986: 75). The distribution of these stems across the Finnic languages is patchy (see the maps in Tuomi & Suhonen 2004: 424–428); the stem *taa is found only in South Estonian and in some North Finnic varieties (Larjavaara 1986: 95; Tuomi & Suhonen 2004: 427; Pajusalu 2015: 170–171). Larjavaara (1986: 75) mentions that the Proto-Finnic pronominal system, which had three pronouns and later acquired the fourth stem, significantly changed over time in particular varieties. Sometimes development was due to divergence, and sometimes the differences decreased because of convergent processes that led to levelling.

In contemporary Finnic languages, the system of demonstratives can be unipartite, bipartite or tripartite. In the latter case, proximal, medial,
and distal demonstratives are opposed. The functions of particular pronouns in demonstrative systems are not the same across the varieties. For example, the stem *se (which is found in all Finnic languages) functions as medial in Finnish and Karelian (and means ‘close to the listener’), but in South Estonian it has changed from medial to proximal, and in Livonian and North Estonian it has taken over all spatial spheres (Laanest 1982: 196–199; Pajusalu 1996: 148, 2009: 123; Tomingas 2018; Yurayong 2020: 31–34). In contemporary Votic and Veps, se functions as distal, while the proximal demonstratives are derived from the same stem *se combined with deictic intensifiers (Kettunen 1943: 403; Yurayong 2020: 34). This latter development was likely induced through contact with Russian that has a similar model (Yurayong 2020: 34, 205–206). The composition of demonstrative systems in contemporary Finnic languages is therefore highly heterogeneous.

Ingrian apparently had a tripartite system of demonstrative pronouns: proximal tämä ~ tää, medial se, and distal too, as attested in the oldest published description of Ingrian (Porkka 1885: 83–85). The dictionary by Nirvi (1971: 616, 514, 594) also contains all three pronouns but it gives only five Soikkola examples for the distal pronoun too. In contrast, there are five times more examples for tämä ~ tää and even more examples for se in this dictionary. This possibly indicates that too was much rarer in speech than the other two pronouns. Later publications (e.g. Laanest 1978; Saar 2017) do not mention the distal too at all, and it is also absent from published Soikkola Ingrian texts (Porkka 1885; Ariste 1960; Laanest 1966a) as well as from recent recordings. Contemporary Soikkola Ingrian has a two-way opposition between tämä, which functions as proximal ‘this’, and še, which is the unmarked pronoun, and which can read both as ‘this’ and ‘that’¹ and is far more frequent than tämä (Schwarz & Rozhanskiy 2022: 167).

The distribution and function of different demonstratives from syntactic, semantic and pragmatic points of view has been widely discussed for the major Finnic languages, Finnish and Estonian (cf. Laury 1996, 1997; Priiki 2017; Seppänen 1998; Pajusalu 2006, 2009, 2015; Nahkola et al. 2020; Reile et al. 2020; Taremaa et al. 2021). For other Finnic languages, considerably less research has been done so far. Karelian and Veps demonstratives were studied by Larjavaara (1986); Livonian

¹ Below we always gloss še as ‘that’.
Demonstrative proadjectives were described by Tomingas (2018). The doctoral thesis by Yurayong (2020) investigates demonstratives in most Finnic languages and also North Russian dialects, the primary focus being on postposed demonstratives. This in-depth study is in fact the only work where Ingrian demonstratives are considered (alongside other languages) from the point of view of their syntactic and pragmatic functions, including referential and information-structural uses.

When used as referential devices, Finnic demonstratives are usually classified as being either deictic or anaphoric (Laury 1996; Pajusalu 2006, 2009). Yurayong (2020: 153–160) offers a more detailed classification based on the chapter by König (2018) and the doctoral thesis by Becker (see 2021, the revised version), and distinguishes seven types of reference relations:

a) A **deictic** referent is new in the discourse, and it is usually a concrete visible entity that can be pointed to with a gesture.

b) A **recognitional** referent is also new information, but not directly visible; it is identified by the interlocutors through common or shared knowledge or experience.

c) An **absolutely unique** referent points to a unique entity, for example the Sun or God. Pajusalu (2009: 132) notes that such reference is not possible in Estonian; in the expression *see päike* ‘this sun’ the demonstrative conveys the physical presence of the sun but not its uniqueness.

d) An **anaphoric** reference is made to an entity that was already explicitly mentioned in the previous discourse. The gap between the antecedent and the demonstrative is usually short, but anaphora can also be used to reactivate referents mentioned 50–200 words earlier.

e) A **bridging** reference is similar to an anaphoric one, but it was not directly mentioned before; rather it associates with some previously mentioned entity through inalienable relations. For instance, ‘the footprints’ can be a bridging reference to the walking forest angel mentioned earlier in the text (example from southern Lude in Yurayong (2020: 158)).

f) A **situationally unique** reference also associates with some previously mentioned entity but does not have inalienable relations with it. The link between the reference and its antecedent is provided by the context.
g) Finally, an **establishing** reference is new information in the discourse that is immediately followed with some clarifying details about the referent.

For Ingrian postposed demonstratives, Yurayong (2020: 161) only observed deictic and anaphoric uses.

As concerns the distribution of demonstrative vs personal pronouns in their referential uses, a distinction is usually made between animate vs inanimate or human vs non-human referents. Pajusalu (2015: 169–170) notes that these distinctions are not strict in the Finnic family, and demonstratives can refer to both animate and inanimate referents. She mentions, however, that in Finnish and Estonian written discourse, probably under the influence of Indo-European languages, human/animate beings are more typically referred to by personal pronouns. Since contemporary Ingrian does not have a written variety, and was not subject to normalizing, we expect that here the human antecedents can be encoded with demonstratives.

Syntactically, a demonstrative can function as an independent constituent or as a determiner in a noun phrase (NP). Within an NP, demonstratives can both precede or follow a noun. In Ingrian, preposed demonstratives clearly prevail, but postposed usages also occur. Yurayong (2020: 186–187) notes that “deictic use is more frequently found in the use of preposed demonstratives, whereas anaphoric and other referential uses are more common for postposed demonstratives”.

The small number of published speech samples is a serious obstacle for detailed research on Ingrian demonstratives. There is currently no published corpus of the Ingrian language, while the number and size of the published texts is very limited. For Soikkola Ingrian, there are two tales published by Porkka (1885), ca. 2000 words; three tales published by Sovijärvi (1944), ca. 3000 words; texts and folklore pieces recorded mainly by one speaker (Ariste 1960), ca. 7500 words; and texts recorded by one speaker (Laanest 1966a), ca. 5000 words. The Ingrian material used in the study by Yurayong (2020) comes from the text collection by

---

2 A written variety and school teaching of Ingrian was introduced in the beginning of the 1930s (Musaev 2004: 248) but was banned by 1938.

3 In Lude and Veps, postposed demonstratives can also attach to non-nominal parts of speech (Yurayong 2020: 140–142).
Laanest (1966a); apart from the Soikkola dialect, the Heva data from the same publication are also considered.

For the current study, we use our own field recordings of Soikkola Ingrian spontaneous speech (see details in section 2). The size of this field corpus is bigger than in any of the published sources and contains data from many native speakers, which gives us the possibility to analyse the Ingrian demonstratives in more detail. This article is the first step of the analysis and presents preliminary research of a qualitative rather than a quantitative nature. Our aim is to give an overview of the functions of demonstratives in contemporary\(^4\) Soikkola Ingrian\(^5\) and also to formulate preliminary hypotheses for further study.

The paper is organized as follows. Section 2 provides an overview of the data used in our research, section 3 gives the basic information on the forms of Ingrian demonstrative pronouns, section 4 presents the analysis of their functions, section 5 offers some observations on the variation between the native speakers and in section 6 we formulate the conclusions and further hypotheses.

2. Data

The material for the current research comprises a collection of spontaneous speech samples that we recorded, together with colleagues, on linguistic field trips to the Soikkola peninsula. We have about 4 hours of transcribed texts in Soikkola Ingrian, consisting of approximately 20,000 words. These speech samples are mostly narratives and a few dialogues recorded by 21 speakers between 2006 and 2013. The genre of the texts is mostly life stories. The data were collected in different dialectal zones, namely in the northern, central (transitional) and

---

4 By “contemporary Ingrian” we mean the language used by Ingrian native speakers in the 21st century.

5 In this article, we do not consider demonstratives in Lower Luga Ingrian for two reasons. First, it is not just an Ingrian dialect in a strict sense, but a convergent language based on Ingrian and Votic with some influence of Ingrian Finnish and Estonian (Rožanskij & Markus 2013). The analysis of Lower Luga Ingrian material would be most effectively done in comparison with all these languages, but such a task is beyond the scope of this article. Second, we do not have a representative corpus of transcribed text samples in Lower Luga Ingrian.
southern parts of the Soikkola peninsula. The recordings were processed and transcribed in ELAN (2020), an audiovisual annotation tool, during the course of the project “Documentation of Ingrian: collecting and analysing fieldwork data and digitizing legacy materials” financed by the Endangered Languages Documentation Programme (2011–2013).

For each example, we provide a code to indicate the text which the example is taken from. The code combines the title of the text and the index of the native speaker. The list of Soikkola Ingrian speakers that are mentioned in this article is given in Table 1. The table specifies the year and place of birth, the place of recording, and the dialectal zone for each speaker. Villages mentioned in Table 1 are plotted on the map in Figure 1.

Table 1. List of native speakers.

<table>
<thead>
<tr>
<th>Index of speaker</th>
<th>Year of birth</th>
<th>Place of birth</th>
<th>Place of recording</th>
<th>Dialectal zone</th>
</tr>
</thead>
<tbody>
<tr>
<td>AG</td>
<td>1936</td>
<td>Reppola</td>
<td>Voloītsa</td>
<td>Northern</td>
</tr>
<tr>
<td>AI</td>
<td>1926</td>
<td>Viistina</td>
<td>Viistina</td>
<td>Transitional</td>
</tr>
<tr>
<td>EI</td>
<td>1929</td>
<td>Viistina</td>
<td>Otsave</td>
<td>Southern</td>
</tr>
<tr>
<td>EN</td>
<td>1932</td>
<td>Venakontsa</td>
<td>Venakontsa</td>
<td>Southern</td>
</tr>
<tr>
<td>GI</td>
<td>1936</td>
<td>Viistina</td>
<td>Viistina</td>
<td>Transitional</td>
</tr>
<tr>
<td>KV</td>
<td>1939</td>
<td>Säädinä</td>
<td>Säädinä</td>
<td>Southern</td>
</tr>
<tr>
<td>LK</td>
<td>1949</td>
<td>Saarove</td>
<td>Saarove</td>
<td>Southern</td>
</tr>
<tr>
<td>MM</td>
<td>1920</td>
<td>Viistina</td>
<td>Viistina</td>
<td>Transitional</td>
</tr>
<tr>
<td>OM</td>
<td>1931</td>
<td>Viistina</td>
<td>Viistina</td>
<td>Transitional</td>
</tr>
<tr>
<td>RP</td>
<td>1927</td>
<td>Saarove</td>
<td>Saarove</td>
<td>Southern</td>
</tr>
<tr>
<td>ST</td>
<td>1924</td>
<td>Mättüizi</td>
<td>Savimää</td>
<td>Northern</td>
</tr>
</tbody>
</table>

6 See Kuznetsova (2009: 19) on the dialectal zones of the Soikkola peninsula.
7 See details of the project in Rozhanskiy & Markus (2019).
8 The village names in Table 1 and Figure 1 correspond to the way they are attested in our field recordings but spelled according to the Finnish tradition with ū rendered as y and Š/ž rendered as s/z. In the case of several variants of a name being attested (e.g. Tammi- gondu ~ Tammikonttu ~ Tammikondo), we chose the variant that was confirmed by most speakers and corresponds most closely to Ingrian phonetics. The name Hovinmaa (in Russian, Krasnaja Gorka) does not occur in our records; it was reported by Aleksei Kriukov who recorded it during his fieldwork with Soikkola Ingrians.
For spelling the examples we use a standardized transcription which does not reflect specific phonetic features of individual speakers. In particular:

a) The historically long mid vowels oo, öö, ee in the stressed syllable that can be pronounced as mid, mid-high or high (e.g. [šoo ~ šöö ~ šuu] ‘swamp’) are always transcribed as mid-high (šöö);

b) The occasional durational shortening and qualitative reduction of the vowel at the stem–case marker boundary is not indicated, e.g. metsä-ž ‘forest-INE’ [metsäž ~ metsäž ~ metsež];

c) Short geminates are written as double characters with a breve (põ, įn, etc.) in all phonetic contexts;

---

d) Single \(p, t, k, \) and \(\dot{s}\) are spelled as voiced \(b, d, g, \ddot{z}\) except word-initially, in clusters with \(p, t, k, \dot{s}, f, \) or \(h,\) and in non-adapted borrowings, e.g. *haabä* ‘aspén’, *nüblä* ‘button’, *joob* ‘drink.PRS.3SG’, but *pada* ‘pot’, *läipšää* ‘milk.INF’.

### 3. Forms of the demonstrative pronouns in Soikkola Ingrian

Paradigms of Ingrian demonstrative pronouns were first listed by Porkka (1885: 83–84). In more recent studies, the full paradigms are not given; several basic forms are mentioned in Nirvi (1971: 338, 355, 514, 616), Laanest (1966b: 108, 1978: 252, 1986: 120), and Saar (2017: 127).

The full paradigms of demonstrative pronouns in Table 2 (cited from Schwarz & Rozhanskiy 2022) represent Soikkola Ingrian as spoken in the 21st century (the data are taken both from spontaneous speech and from elicitation sessions). The essive and excessive forms were not attested in our recordings, apart from the corresponding historical forms *šiin* and *šiind* that function as the inessive and elative forms in the contemporary language, and the historical essive *tänä* that has been lexicalized and is used exclusively in a temporal sense.

The short forms of the nominative (*tää*), genitive (*tään*), and nominative plural (*nää*) from *tämä* ‘this’, listed by Porkka (1885: 83) and Laanest (1966b: 108), are not attested in the contemporary materials.

---

10 Full paradigms of demonstrative pronouns are given also in Junus (1936: 99). However, we do not consider them, because this grammar is prescriptive, not descriptive. The forms from Junus (1936) do not always match our data. For instance, Junus lists *šenen* and *šen* as the genitive and accusative forms, respectively. In our materials, these two forms both function as genitive (the distribution is idiolectal), while there is no morphologically independent accusative form at all. Also, it is not clear whether the essive forms *tämännä* ‘this.ess’, *senennä* ‘that.ess’, *näinnä* ‘this.pl.ess’, and *niinnä* ‘that.pl.ess’ given by Junus have ever existed.
Table 2. Paradigms of demonstrative pronouns in contemporary Soikkola Ingrian.

<table>
<thead>
<tr>
<th>Case</th>
<th>‘this’</th>
<th>‘that’</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Singular Plural</td>
<td>Singular Plural</td>
</tr>
<tr>
<td>NOM</td>
<td>tämä nämäd ~ nämä</td>
<td>še ~ šee need ~ nee ~ ned ~ ne</td>
</tr>
<tr>
<td>GEN</td>
<td>tämän näijen</td>
<td>šenen ~ šen niijen</td>
</tr>
<tr>
<td>PART</td>
<td>tädä naidä</td>
<td>šidä niidä</td>
</tr>
<tr>
<td>ILL</td>
<td>tähä nähhe</td>
<td>šihe ~ šiihe niihe</td>
</tr>
<tr>
<td>INE</td>
<td>täž näž</td>
<td>šiin niiž</td>
</tr>
<tr>
<td>ELA</td>
<td>täšt näšt</td>
<td>šiind niišt</td>
</tr>
<tr>
<td>ALL</td>
<td>tälle näille</td>
<td>šille niiille</td>
</tr>
<tr>
<td>ADE</td>
<td>täl näil</td>
<td>šil niil</td>
</tr>
<tr>
<td>ABL</td>
<td>täld näild</td>
<td>šild niild</td>
</tr>
<tr>
<td>TRANS</td>
<td>täkš näikš</td>
<td>šikš niikš</td>
</tr>
</tbody>
</table>

In addition to the locative forms of pronouns, there are also pronominal adverbs that either have an identical form (e.g. täž ‘here’ and täž ‘this.INE’) or a very close form with the pronouns (e.g. tääl ‘here’ and täl ‘this.ADE’; it is rather typical for such forms to be pronounced similarly in fast speech). This homonymy sometimes poses a problem for analysing the corresponding forms. In (1), the adverb täšt ‘from here’ is syntactically dependent on the predicate pääššä poiž ‘get out’. In (2), the pronoun täšt ‘this.ELA’, which agrees with the head noun in case and number, hardly allows an adverbial interpretation ‘from here’ – there is no preceding context that says something about ‘here’. However, in (3), täšt can be easily interpreted both as an agreeing pronoun (‘from this house’) and as an adverb that depends on the predicate (‘were driven away from here from the house’).

(1) Kala_ja_metsä_EN
    miä e-n pääš-t täšt poiž
    1SG NEG-1SG get.in-PRTACT from.here away

‘I could not get out from here.’
Pedro_OM

\( tāšt \quad nūgūnāizet\text{-}nū \quad e\text{-}lō\text{-}št \quad mīā \quad hūvā\text{-}ā ē \quad e\text{-}n \quad nāā \quad mītāā ē \quad \text{NEG-1SG} \quad \text{see.CNG} \quad \text{what.NEG.PART} \)

‘I do not see anything good in this modern life.’

Šoda_AG

\( i \quad tāmā \quad mēidā \quad tāšt \quad koi\text{-}št \quad \text{and this 1PL.PART} \quad \text{this.ELA/from.here} \quad \text{house-ELA} \quad taaž \quad aje\text{-}ttii \quad pōiž \quad \text{again} \quad \text{drive-IMPRS.PST} \quad \text{away} \)

‘And this, we were again driven away from this house / from here from the house.’

Similarly, in examples (4) and (5), \( tāž \) and \( tāhā \) can be interpreted in two ways.

Elo(B)_AI

\( a \quad nūd \quad tāž \quad mēijen \quad otsaa\text{-}ž \quad ēllāā ē \quad \text{but} \quad \text{now} \quad \text{this.INE/her} \quad \text{1PL.GEN} \quad \text{edge-INE} \quad \text{live.PRS.3SG} \quad kāig \quad vāgi \quad noor\text{-}d \quad vākki\text{-}ā \quad ei\text{-}oo\text{-}gaa \quad \text{all old} \quad \text{people} \quad \text{young-PART} \quad \text{people-PART be.NEG.3SG=PTCL} \)

‘And now in this part of ours / here in our part only the old people live; there are no young people.’

Püüdämääž_GI

\( kūmmen \quad vai \quad kākšōišt \quad hēnki\text{-}ā \quad aïn \quad mēil \quad \text{ten or twelve} \quad \text{person-PART} \quad \text{always} \quad \text{1PL.ADE} \quad jōga \quad hōamuž \quad korjahu\text{-}i\text{-}d \quad tāž \quad <\ldots> \quad \text{every morning} \quad \text{get.together-PST-3PL} \quad \text{here} \quad aïn \quad tāhā \quad ūhTEE \quad pākkāa \quad korjahu\text{-}i\text{-}d \quad \text{always} \quad \text{here/this.ILL} \quad \text{one.ILL} \quad \text{place.ILL} \quad \text{get.together-PST-3PL} \quad \text{here} \quad \text{Ten or twelve people would always gather here at our place each morning. (They) always gathered in this one place / here, in one place.’}

This kind of homonymy can be a real challenge for the analysis of demonstrative pronouns. However, it is not very relevant for the current

---

11 It should be pointed out that the semantic opposition of the spatial vs deictic usage cannot be considered a reliable criterion for distinguishing between adverbs and pronouns. E.g. Etelämäki (2009: 40–43) notes that the Finnish demonstrative tāmā ‘this’ may be used as a means of organizing the interaction, for example marking “a new phase within the telling”, and concludes that the spatial use of demonstratives requires further investigation.
paper. None of the illustrative examples below contain homonymic forms that can be considered as adverbs. In section 5, where the ratio of different demonstratives is calculated, only the nominative forms are considered. The only place where calculations could be affected by this homonymy is Table 3 in section 4, which lists all forms of pronouns. In this calculation, we excluded all occurrences where the pronominal interpretation looked unnatural and artificial, and the adverbial interpretation looked definitively better. All ambiguous occurrences, where both the adverbial and pronominal interpretations are equally possible, were counted separately.

4. Function of demonstrative pronouns

This section presents a functional analysis of the demonstratives in contemporary Soikkola Ingrian. First, we discuss which morphological forms occur in our corpus.

Table 3 presents the number of occurrences of each form. If no number is indicated, the corresponding form is not attested in our data. A number in parentheses refers to ambiguous cases when a locative form can be interpreted either as a pronoun or as an adverb (see section 3).

<table>
<thead>
<tr>
<th>Case</th>
<th>‘this’ Singular</th>
<th>‘that’ Singular</th>
<th>‘this’ Plural</th>
<th>‘that’ Plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>NOM</td>
<td>tämä 77</td>
<td>nämä(d) 23</td>
<td>še(e) 156</td>
<td>ne(e)(d) 64</td>
</tr>
<tr>
<td>GEN</td>
<td>tämän 9</td>
<td>näijen 32</td>
<td>šen(en) 32</td>
<td>nijen</td>
</tr>
<tr>
<td>PART</td>
<td>tädä 11</td>
<td>näidä 11</td>
<td>šidä 72</td>
<td>niidä 7</td>
</tr>
<tr>
<td>ILL</td>
<td>tähä 2 (+4)</td>
<td>näihe 1</td>
<td>ši(i)he 2</td>
<td>niuhe</td>
</tr>
<tr>
<td>INE</td>
<td>täž (11)</td>
<td>näiž 1 (+3)</td>
<td>šiin 1</td>
<td>niž 1</td>
</tr>
<tr>
<td>ELA</td>
<td>täšt 1 (+1)</td>
<td>näišt 1</td>
<td>šiind 1</td>
<td>ništ</td>
</tr>
<tr>
<td>ALL</td>
<td>tälle 2</td>
<td>näille 2</td>
<td>šille 4</td>
<td>niille</td>
</tr>
<tr>
<td>ADE</td>
<td>täl 7</td>
<td>nail 2</td>
<td>šil 2</td>
<td>niil 1</td>
</tr>
<tr>
<td>ABL</td>
<td>täld</td>
<td>näild</td>
<td>šild 1</td>
<td>niild</td>
</tr>
<tr>
<td>TRANS</td>
<td>täkš</td>
<td>näikš</td>
<td>šikš</td>
<td>niikš</td>
</tr>
</tbody>
</table>
As seen from Table 3, the “syntactic cases” (i.e. the nominative, genitive and partitive) are considerably more frequent than other case forms; also, the forms of še (at least the syntactic cases) are generally more frequent than forms of tämä. The latter observation suggests that še is the unmarked variant in the pair ‘this – that’ (unlike, for example, in English).

The locative forms of demonstrative pronouns are rather rare in our materials, so we cannot offer any generalizations about their typical functions. In the majority of examples, these forms occur within an NP and represent one of the deictic or anaphoric uses discussed in this section.

Below we investigate the functions of demonstrative pronouns (primarily, their nominative forms) in our corpus. We distinguish two main groups of functions: referential uses of demonstrative pronouns and discourse uses. In its referential use, a demonstrative marks or brings into focus a certain referent. In its discourse use, a demonstrative pronoun functions as a discourse marker, i.e. marks hesitation, repairing, code-switching, transition to a new topic, etc.

As mentioned in the Introduction, two main classes of referential uses of demonstratives are usually distinguished for Finnic languages, namely deictic and anaphoric. The deictic reference points to a non-linguistic object that exists in space, while the anaphoric reference is made to something mentioned in the preceding text (cf. Halliday & Hasan 1976; Fillmore 1997). For the Ingrian examples observed in the data we offer a more fine-grained classification that intersects to a certain extent with the one proposed by Yurayong (2020).12

Since the function of a particular demonstrative form may differ depending on whether it occurs independently or as part of an NP, we treat these cases separately. Table 4 summarizes the functions of demonstratives observed in our material. The numbers in the right part of the table refer to the sections where the corresponding functions are discussed in the paper.

---

12 Our classification partially combines the parameters distinguished by Yurayong (2020: 153–172) as referential vs information-structural uses. We combine these, because the information-structural uses also serve to bring a referent into focus.
Table 4. Functions of demonstratives in Soikkola Ingrian and the corresponding sections in the article.

<table>
<thead>
<tr>
<th>Type</th>
<th>Sub-type</th>
<th>Function</th>
<th>Syntactic status</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Within NP</td>
</tr>
<tr>
<td>Referral</td>
<td>Deictic</td>
<td>Deictic introducing</td>
<td>4.1.1</td>
</tr>
<tr>
<td></td>
<td>Deictic</td>
<td>Deictic indexical</td>
<td>4.1.2</td>
</tr>
<tr>
<td></td>
<td>Anaphoric</td>
<td>Anaphoric (pure)</td>
<td>4.1.3</td>
</tr>
<tr>
<td></td>
<td>Mixed</td>
<td>Inexact reference</td>
<td>4.1.4</td>
</tr>
<tr>
<td></td>
<td>Mixed</td>
<td>Choosing a referent</td>
<td>4.1.5</td>
</tr>
<tr>
<td></td>
<td>Anaphoric</td>
<td>Serving as a personal pronoun</td>
<td>4.2.5</td>
</tr>
<tr>
<td></td>
<td>Deictic</td>
<td>Deictic contrast</td>
<td>4.2.6</td>
</tr>
<tr>
<td>Discourse</td>
<td>Change in syntactic structure, searching for the correct word, transition to a new sub-topic</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

4.1. Referential uses within an NP

4.1.1. Deictic introducing

Yurayong (2020: 154–155) distinguishes a recognitional reference type that refers to something not mentioned in the preceding text but identifiable to the interlocutors through common or shared knowledge. We propose a more general deictic introducing type where a demonstrative provides a link to a referent from the field of the speaker’s knowledge. This knowledge may be shared by both interlocutors and in this case the demonstrative serves as a device that “synchronizes” the corresponding piece of the speaker’s and hearer’s knowledge, cf. (6). It is precisely this usage which is called recognitional by Yurayong (2020).

(6) Pahhain elo MM
(A reference is made to the folklore ensemble from the village of Logi.)

meidä jokka paikka-a vejdel-üti
1pl.part every place-ill carry.around-imprs.pst
ühellain ku nüttä nämäd logovee-n oma-d
same when now this.pl Logi-GEN own-pl.nom
katso-i-tta töö faklörnöi-dä nu vod
look-pst-2pl 2pl folklore-part well ptcl

‘We were taken to all places. The same way as these from the Logi village now. Did you see (the performances) of the folk ensemble? That’s it.’
However, in a very similar way the speaker may refer exclusively to her/his knowledge. Sometimes it is not possible to define whether the referent is known also to the listener. For instance, in (7), the listener might know about the Finnish researchers that come to visit the Ingrian villages or (s)he might not.

(7) Elo(B)_AI
(There was no preceding talk mentioning the Finnish visitors.)

\[ a \ šiž \ vod \ nüd \ ku \ nämäd \ šıpмела-
\[ \text{ and then } \ ptcl \ now \ when \ this.pl \ Finn-pl.nom \]
\[ \text{ ałkko-i-d } \ tänne \ kävvä ... \]
\[ \text{ begin-pst-3pl } \ here \ go.1nf \]

‘And then, now when these Finns started to visit here …’

In other cases, the referent is known exclusively to the speaker. Thus, the function of the demonstrative pronoun is to mark the introduction of some new information from the speaker’s sphere of knowledge into the current discourse, as in (8) and (9).

(8) Püüdämääž_GI
(The preceding narrative described the process of fishing. There was no mentioning of the brigade.)

\[ \text{kaigin } \ kâ-i-vâd \ ŝe \ briga-
\[ \text{ all.together } \ go-pst-3pl \ that \ brigade \ also \]
\[ \text{ kakštoišt } \ henkki-ä \ ain \ kümmen \ vai \]
\[ \text{ twelve } \ person-part \ always \ ten \ or \]
\[ \text{ kakštoišt } \ henkki-ä \]
\[ \text{ twelve } \ person-part \]

‘Everybody used to go (fishing). That brigade also... (There were) always twelve people, ten or twelve people.’

(9) Konna_ST
(The story tells about frogs that the speaker very much dislikes. The nephew is introduced for the first time.)

\[ a \ ŝämâ \ plenmääniikk \ ku \ männöö \ šaunaa \]
\[ \text{ and this nephew } \ when \ go.prs.3sg \ sauna.i.ill \]
\[ \text{ nii } \ šeël \ ono \ taiťaa \ lä уме-mb \ talvee-l \]
\[ \text{ so } \ there \ be.prs.3sg \ probably \ warm-comp \ winter-ade \]
\[ \text{ nii } \ krookkaa-d \ toiže-n \ kerra-n \]
\[ \text{ so } \ croak.prs-3pl \ other-gen \ time-gen \]
\[ \text{ nii } \ hää \ tulloo \ ŝe \ šiun \ dovarišša-d \]
\[ \text{ so } \ 3sg \ come.prs.3sg \ ptcl \ 2sg.gen \ friend-pl.nom \]
Examples (6–9) demonstrate different variants of the distribution of knowledge between the speaker’s and hearer’s “territory of information” (in terms of Kamio 1997). Apparently, both tämä and še can be used for synchronizing the speaker’s and hearer’s knowledge, but it is not entirely clear which factors trigger the choice of a particular pronoun. Evans et al. (2018a: 123–134) have shown that, cross-linguistically, demonstrative systems are among the grammatical means used to encode differences in the accessibility of an entity or state of affairs to the speaker and addressee. The authors introduce a category of “engagement” to refer to such cases. Applied to Ingrian, it remains a question for further investigation to understand how the use of tämä vs še correlates with the asymmetry of knowledge between the interlocutors.

4.1.2. Deictic indexical

This type of reference is made to an object visible to the speaker and hearer (10, 11), so there is no need to synchronize the knowledge between the interlocutors as in the previous type. The demonstrative pronouns in this type of reference are often accompanied by a pointing gesture (cf. the deictic reference by Yurayong 2020).

(10) Hirvi_GI
(The speaker tells and shows how she was bitten by mosquitoes in her garden.)

\[
\begin{array}{lllll}
\text{jala-}d=\text{ki} & \text{kaig} & \text{pal-}l\text{o-}i-d & \text{höö} & \text{niin} \\
\text{leg-PL.NOM=PTCL} & \text{all} & \text{burn-PST-3PL} & \text{3PL} & \text{so} \\
\text{kovast} & \text{i} & \text{kae-d} & \text{nämäd} & \text{paiga-d} \\
\text{very} & \text{and} & \text{hand-PL.NOM} & \text{this.PL} & \text{place-PL.NOM} \\
i & \text{liitsa} & \text{niin} & \text{on} & \text{paljo} & \text{tänäh} & \text{v99-n} \\
\text{and} & \text{face} & \text{so} & \text{be.PRS.3SG} & \text{many} & \text{this}^{13} & \text{year-ESS}
\end{array}
\]

‘Even my legs were all burning. They (bit) so much, the hands, these places, and the face, there are so many (of them) this year.’

---

13 Since we treat the historical essive form of tänäh as lexicalized (see section 3), we do not mark it with a grammatical gloss.
4.1.3. Anaphoric (pure)

A pure anaphoric reference is made to a word mentioned in the preceding text (12, 13).

\[ \text{(12)} \]
\[ \text{(The cherry tree was mentioned several sentences before.)} \]
\[ \text{män-i-mmä a še meži-marja on} \]
\[ \text{go-pst-1pl and that cherry be.prs.3sg} \]
\[ \text{mokkooma šuur mokkooma polnoi kaig} \]
\[ \text{such big such full all} \]
\[ \text{‘We went, and that cherry(-tree) is so big all full (of berries).’} \]

\[ \text{(13)} \]
\[ \text{Püüdämääž GI} \]
\[ \text{minjä nüttä taaž kera} \]
\[ \text{daughter-in-law now again also} \]
\[ \text{tämä minjä noiz-i jo} \]
\[ \text{this daughter-in-law become-pst.3sg already} \]
\[ \text{maišto-i šao-i što hūvā rokka} \]
\[ \text{taste-pst.3sg say-pst.3sg that good cabbage.soup} \]
\[ \text{‘Now also the daughter-in-law, this daughter-in-law started (to eat the cabbage soup, she) already tried (and) said that (it is) a good cabbage soup.’} \]

4.1.4. Mixed anaphoric-deictic: inexact reference

We use this term to denote cases when a reference is made not to a particular word in the text, but to the object described in the preceding context. It may be a word that was intended to be pronounced but in fact was not (so one can only guess what word it is), or a referent that was described by a whole situation but not named (compare with the notion of a bridging referent in the classification by Yurayong (2020: 157–158)). This type of uses can therefore be interpreted both as anaphora and deixis.
(14) Keeled(B)_KV
(The speaker tells about belongings that were hidden before the deportation but then disappeared.)

\[ miä\ od-i-n\ maa-št\ no\ ku\ need \]
\[ 1SG\ take-PST-1SG\ land-ELA\ well\ when\ that.PL \]
\[ meijen\ midä\ meil\ ol-i\ hauva-ttu\]
\[ 1PL.GEN\ what.PART\ 1PL.ADE\ be-PST.3SG\ bury-PRTPASS \]

\[ vaattii-d\ ja\ kaigveez\ šinne\]
\[ clothes-PL.NOM\ and\ everything\ there \]
\[ koi-d\ polde-tti\]
\[ house-PL.NOM\ burn-IMPRS.PST \]

\[ a\ ked\ ked-lekkenää\ need\ havva-d \]
\[ but\ who.PL.NOM\ somebody.PL.NOM\ that.PL\ hole-PL.NOM \]
\[ kaig\ ott-ii-d\ ott-ii-d\]
\[ all\ take-PST-3PL\ take-PST-3PL \]

‘I took from the ground, well... When those our (things) that we had hidden, the clothes and everything there... The houses were burnt. And somebody (emptied) those pits, took everything.’

Although example (14) mentions the buried things, the word *hauda* ‘hole, pit’ does not occur previously. There is no pure anaphoric reference here, and the demonstrative can be analysed either as referring to a word that was omitted (although implied) or as referring to an object from the speaker’s sphere of knowledge, similar to the type ‘Deictic introducing’.

(15) Hirvi_GL

\[ mänänä\ voo-n\ miul\ šue-d\ šö-i-väd \]
\[ last\ year-ESS\ 1SG.ADE\ wolf-PL.NOM\ eat-PST-3PL \]
\[ koira-n\ a\ nüttä\ še\ šuži\ ammu-tti\]
\[ dog-GEN\ but\ now\ that\ wolf\ shoot-IMPRS.PST \]

‘Last year wolves ate my dog. And now that wolf was shot.’

At first, the speaker describes the situation with a general non-specific referent: ‘wolves’ in the plural. The plural form does not necessarily mean here that the dog was eaten by several wolves: in such a context, the plural form is semantically unmarked if compared with its singular counterpart and can refer to both a single wolf and several wolves. In the second sentence, a concrete wolf is reported to have been shot, and it is implied that it was the one that had eaten the dog. Again, the demonstrative may be analysed here as referring either to the wolf
from the text above (labelled with the generic plural “wolves”) or to the specific wolf from the speaker’s sphere of knowledge.

4.1.5. Mixed anaphoric-deictic: choosing a referent

A demonstrative pronoun marks a new referent that is being introduced into the discourse, when there are other potential referents in the same text.

(16) Püüdämääž_GI
a Še ton’a kumba ono smenkovaaž
but that Tonja which be.prs.3sg Smenkovo-ine
‘And Tonja, who is (= lives) in Smenkovo…’

In the preceding narrative, different sisters of the speaker were mentioned. Now she starts talking about another sister. It should be noted that in (16) the newly introduced referent is followed by a restrictive attributive phrase, which corresponds to the establishing type in the classification by Yurayong (2020). However, this is not always the case, cf. (17).

(17) Pahhain_elo_MM
voi-b ol-la Še vunukka tušoo
be.able-prs.3sg be-inf that grandchild come.prs.3sg
nii vęg-b miun kalmoi-lle
so carry.prs-3sg 1sg.gen graveyard-all
‘Maybe this grandson will come, then he will take me to the cemetery.’

In the preceding text, another grandson of the speaker was mentioned, so the demonstrative marks the change of the referent to a different grandson.

4.2. Referential uses of independent demonstratives

4.2.1. Deictic introducing

The deictic introducing type that provides a link to a referent from the field of the speaker’s knowledge cannot be expressed with independent demonstratives. If the referent is not visible (unlike in the deictic indexical type), and has not previously been mentioned, it can only be referred to with a full NP and not with an independent pronoun.
4.2.2. Deictic indexical

Unlike deictic introducing, deictic indexical reference, which is often accompanied by a gesture, is rather common for independent demonstratives (18, 19).

(18) Pahhain_eolo_MM
(The conversation takes place in the speaker’s house, to which the reference is made.)

```
kodi meil ol-i meil ol-i vanha kodi
house 1PL.ADE be-PST.3SG 1PL.ADE be-PST.3SG old house
```

tämä on teh-ü uuž meil
this be.PRS.3SG do-PRTPASS new 1PL.ADE

“We had a house, we had an old house, this (one) is the new (one) made.”

(19) Lastochkad_LK14
```
tul-kkaa katso-kkaa tul-kkaa katso-kkaa
come-IMP.2PL look-IMP.2PL come-IMP.2PL look-IMP.2PL
```

```
midä ono paljo lastoška-a
what.PART be.PRS.3SG many swallow-PART
```

```
štvo vod nämäd ol-laas nqrrikkaiš-t
that PTCL. this.PL be-IMPRS.PRS young-PL.NOM
```

```
nämäd ol-laas jo vanhekkaiš-t
this.PL be-IMPRS.PRS already old-PL.NOM
```

‘Go and look, how many swallows! These (ones) are young, and these (ones) are already older.’

4.2.3. Anaphoric (pure)

An anaphoric reference expressed with an independent pronoun is only possible when the antecedent is located closely in the preceding text (otherwise, it becomes unclear to which entity the demonstrative refers). Typically, such reference is made to a word within the same sentence (20, 21) or within the preceding sentence (22) but no further back in the discourse.

---

14 This speaker has some idiolectal features. E.g. in this example, we observe katsōkkaa ‘look.IMP.2PL’ instead of katsogaa, the impersonal form ollaa ‘be.IMPRS.PRS’ in the personal context, and the Russian borrowing lastoška ‘swallow’ instead of the regular Ingrian pääšköi.
4.2.4. **Mixed anaphoric-deictic: inexact reference**

This reference type is rather common for independent demonstratives. The pronoun refers to some situation described in the previous narrative for which the speaker did not use any specific cover term. In (23), še refers to the situation of speaking in Russian that is described in the preceding sentence. The utterance in (24) comes after the speaker tells about her mother knowing many songs in the Ingrian language that the children did not learn for some reason. The independent demonstrative še refers to the knowledge of Ingrian songs that has been lost.
4.2.5. Anaphoric: serving as a personal pronoun

Independent demonstratives can be used anaphorically in place of personal pronouns. This often happens when a human referent is distinguished among a group of similar people (25, 26) or a switch to a different referent takes place (27).

(25) Elo(B)_AI
(In the preceding narrative, the speaker mentioned other members of her family.)

a šiž tämä miun toīže-n šiar
but then this 1SG.GEN other-GEN sister

galina ivanna še ku kā-i tōō-ž
Galina Ivanovna that when go-PST.3SG work-INE
‘And then my other sister, Galina Ivanovna, when she went to work…’

(26) Püüdämääž_GI

a še ton’a kumba ono smenkovaazž
but that Tonja which be.PRS.3SG Smenkovo-INE

še ei kā̄-nd mere-l še teg-i
that NEG.3SG go-PRTACT sea-ADE that do-PST.3SG

tōō-dā narvaaž
work-PART Narva-INE
‘And that Tonja, who is (= lives) in Smenkovo, she did not go to the sea. She worked in Narva.’

15 This form is not attested either in our data or in other sources on Ingrian known to us. It is possibly a participle kelvand (from kelvada ‘to fit, to suit, to be worth’) that was built incorrectly. However, as pointed out by one of the anonymous reviewers, similar forms occur in Estonian dialects, especially in the island dialect, cf. a group of participles which is labelled “reduplicated forms” in Viikberg (2020: 279).

16 The speaker means that her sister did not work in a fishing crew.
The beginning of this example was quoted above as (16). Here again the reference is made to one of the sisters, while several other sisters were mentioned in the preceding story. After this sister (the new referent) is introduced, she is consistently referred to with the demonstrative še and not with the personal pronoun hää.

(27) Maod2_LK
(In the preceding narrative, the nephew of the speaker was mentioned.)

\[
\begin{align*}
&\text{miä} & \text{katso-n} & \text{mado} & \text{pr’amo} & \text{koi-n} \\
&\text{ald} & \text{tälle} & \text{šao-n} & \text{što} & \text{tol’ko} \\
&\text{elä} & \text{mää} & \text{o} & \text{hiligaišta} & \text{hiligaišta} \\
\end{align*}
\]

\begin{align*}
&\text{miä} & \text{katso-n} & \text{mado} & \text{pr’amo} & \text{koi-n} \\
&\text{ald} & \text{tälle} & \text{šao-n} & \text{što} & \text{tol’ko} \\
&\text{elä} & \text{mää} & \text{o} & \text{hiligaišta} & \text{hiligaišta} \\
\end{align*}

‘I look: a snake (comes out) right from under the house. I tell him (the nephew) that “Just do not go (= move), stay still, still!”’

In (27), if the speaker had used the personal pronoun hänelle ‘to him’, that could have been addressed to the snake, so she used the demonstrative tälle ‘to this’ to mark the switch to the nephew as the preceding referent.17

4.2.6. Deictic contrast

In the data, we have examples of contrastive relations when an object is opposed or compared to another object. It is however not typical in Ingrian to use the pair of demonstratives tämä ‘this’ vs še ‘that’ for expressing this kind of contrast. More often a demonstrative is opposed to a pronoun from another category (28), or the contrast is expressed with two demonstratives of the same type (29, cf. also (19) above with two demonstratives nämäd both used for deictic indexical reference).

In (28), events happening on work days are referred to with the relative pronoun migä ‘what’, and they are compared to the events happening on weekends, expressed by the demonstrative še ‘that’.

---

17 The problem of logophoricity and the distribution of demonstrative and personal pronouns has been extensively discussed for Finnish (see, for example, Kaiser (2018) and the publications referred to therein).
(28) Pedro_OM
(In the preceding narrative, the speaker tells how people used to dress up on weekends when she was young.)

\[ a \, \text{nütä} \, \text{migä} \, \text{argi-pää-n} \, \text{še} \, \text{pühä-päivää-n} \]

but now what workday-ess that holiday-ess

‘But now \textbf{what} (a person wears) on a workday, \textbf{this} (same (s)he wears) on a holiday.’

Example (29) describes the situation where a person who went fishing himself for the collective farm got less fish than those who did not go fishing but had larger families. The demonstrative \textit{še} ‘that’ refers anaphorically to the relative pronoun \textit{ken} ‘who’ and is contrasted with the plural form \textit{need} of the same demonstrative.

(29) Kala_ja_metsä_EN

\begin{verbatim}
ken itse kā-i mere-l še välištää
who oneself go-pst.3sg sea-ade that occasionally
ša-i vejel vähe-mb kuin need kel
get-pst.3sg more small-comp as that.pl who.ade
öl-lii-d šuure-d pere-d
be-pst-3pl big-pl.nom family-pl.nom
\end{verbatim}

‘(The one) \textbf{who} used to go to the sea himself, \textbf{he (lit. \textit{that})} would occasionally get even less (fish) than \textbf{those} who had big families.’

Similarly, in the example from Heva Ingrian (Laanest 1966a: 49, quoted in Yurayong 2020: 165), the contrast of two types of bears is expressed via the same postposed demonstrative \textit{ne} ‘that.pl’: \textit{metsä karhu-d ne on suure-p. a kagra karhu-d ne on pikkarais-t. just nigu koira} ‘Forest bears are big. But bears in the oat field are small. Just like a dog’.

4.3. Discourse uses of demonstratives

Both \textit{tämä} ‘this’ and \textit{še} ‘that’ demonstratives can function as placeholders and mark certain complications in text production. It has been noted that demonstratives are a very common source of placeholders in the world’s languages, because the “pointing function of demonstrative

---

18 Here we observe a form levelling: the speaker pronounced \textit{pered} instead of \textit{peřreehed} ‘families’.
reference allows the speaker to draw the hearer’s attention to the yet-to-be-specified referent while attempting to articulate it” (Hayashi & Yoon 2010: 48). Demonstratives may function as placeholders also in Finnic languages, cf. Keevallik (2010) on the Estonian see ‘this’.

The most typical situations when demonstratives function as placeholders in Ingrian are the following:

a) The speaker fails to construct the initially intended syntactic structure of the clause, so she changes the structure. The point of change is marked with a demonstrative (30, 31).

(30) Püüdämääž GI
kaigelaže-d <pause> tämä noor vägi
different-PL.NOM this young people
ei ol-d miž tõõ-dää tehha
NEG.3SG be-PRTACT where work-PART do.INF
‘All kind of... well, young people, there was no place to work.’

(31) Lehmä EI
a möõ <e-e> tämä <pause> a meil
but 1PL this but 1PL.ADE
eińää lehmä-ä šiž ei ol-d
else cow-PART then NEG.3SG be-PRTACT
konž vėę-ittyi mejje-d Šoomee
when carry-IMPRS.PST 1PL-ACC Finland.ILL
‘But we... well, we did not have a cow anymore. When they brought us to Finland …’

b) The speaker fails to find the correct word (32, 33).

(32) Püüdämääž GI
liha ol-i oma nii kiukkaaż ain še
meat be-PST.3SG own so oven-INE always that
ol-i <pause> muńna-a ja lihha-a
be-PST.3SG potato-PART and meat-PART
pan-du kiukka-šše
put-PRTPASS oven-ILL
‘We had our own meat. So there was always in the stove … There were some potato and meat, put into the stove.’

---

19 The common infinitive form of this verb in contemporary Ingrian is tehä but this particular speaker uses the form with a geminated consonant.
At first the narrator cannot remember the name of the food made in the oven. She replaces it with *še* ‘that’, and after a pause she finds another way of describing the food that was prepared.

(33) Šoomeež OM

\[ mõõ \ hänən \ kera \ ühe-l \ <pause> \ še \]

\[ 1\text{pl} \ 3\text{sg.gen} \ \text{with} \ \text{one-ade} \ \text{that} \]

\[ emänaižeel \ ol-i-mma \]

mistress-ade be-pst-1pl

‘We worked (lit. were) together with her for the same mistress.’

The speaker hesitates while trying to recall the word for ‘mistress’. In Soikkola Ingrian, the usual term is *emändä*, but either the speaker did not remember it or for some reasons she chose a more peripheral word *emän(n)ain*.

A pause while looking for a word which is overtly marked with a demonstrative is especially common when the speaker is trying to render in Ingrian a certain concept taken from the Russian language (34).

(34) Keeled(B) KV

(The story tells about the problem of indicating one’s Ingrian nationality in passports.)

\[ peräšt \ šoa-n \ to \ ol-i \ kovašt \ stroga \]

after war-gen ptct be-pst.3sg very [rus] strict

\[ hervi-tii \ što \ <pause> \ hervi-tii \ što \]

be.afraid-imprs.pst that be.afraid-imprs.pst that

\[ tämü <pause> no \ migä-ikkeä \ vod \ <pause> \]

tämä well something ptcl

\[ vod \ migä-ikkeä \ špiöŋi \]

ptcl something [rus] spies

‘After the war it was very strict. (People) were afraid that... (people) were afraid that, well... well, (we were) some kind of spies.’

In (34), the speaker is searching for a word in Ingrian, fails to find it, inserts *tämä* as a placeholder, and finally opts to say the word in Russian.
In (35), the demonstrative tämä signals the code-switch to a Russian adjective.

Sometimes the speaker starts with a wrong code and then repairs the phrase in the intended code; the moment of repair is marked with a demonstrative. In (36), the speaker starts to pronounce the date in Russian, stumbles, inserts tämä and then succeeds in expressing the date in Ingrian.

(35) Kazahstaniž_RP
a šiž šoda loppu-i šiž tämä
but then war finish-pst.3sg then this
v’elikij meild20 stalin
[RUS] great 1pl.abl [RUS] Stalin
‘And then the war ended, then, well, our great Stalin...’

(36) Pahhain_eło_MM
forty-ade three-ade year-ade 1pl.part
vee–tii poiz <d’ev’e...> tämä ühekšän-toižee–l
carry-imprs.pst away this nineteen-ade
tšislaa–l v d’ekabr’e
number-ade [rus] in December
‘They brought us away in the year forty-three. The nineteenth of December.’

c) A demonstrative pronoun can be used as a discourse marker to indicate a transition to a new sub-topic in the narrative (37, 38).

(37) Tapuuna_AI
šiž veli ol–i mokkooma tubli mežž
then brother be-pst.3sg such brave man
šiž tämä hää šañoo
then this 3sg say.prs.3sg
‘Then, my brother was a brave person. Then, well, he says...’

---

20 It is not clear why the ablative form is used. Probably the genitive form meijen ‘our’ was intended.
Although both tämä and še occur in the examples as discourse markers, our preliminary observations show that tämä is considerably more frequent in this function: around 45% of examples where tämä occurs in the nominative form (35 out of 77 occurrences) are discourse uses, while for the nominative še the percentage of discourse uses is around 12% (18 out of 156 occurrences, see Table 3). We can propose two possible explanations for this imbalance. On the one hand, it may reflect the Russian influence on the contemporary bilingual Ingrian speakers. In Russian, the proximal demonstrative pronoun èto ‘this’ is often used as a discourse marker (see, for example, Podlesskaya 2010: 12, 28) but its distal counterpart to ‘that’ is not. On the other hand, it might be the case that a proximal demonstrative is generally more preferable as a placeholder, because it correlates with the speaker’s sphere of knowledge and is therefore more natural for marking the speaker’s difficulties in text production (cf. again the notion of engagement in Evans et al. (2018a, b)). This question remains for further investigation.

5. Variation between the speakers and contact influence

In section 4, we observed the differences between the demonstrative pronouns tämä ‘this’ and še ‘that’. In this section, we look at whether the use of these demonstratives is homogeneous for all speakers or whether there are significant individual differences. Figure 2 plots the ratio between the frequency of occurrences of the nominative forms tämä and še for particular speakers. We use the data from 7 native speakers for whom the number of demonstrative pronouns is the biggest in our corpus. Speakers are ordered from the one with the lowest to the one with highest tämä – še ratio. The indexes of the speakers are indicated on the X-axis.

21 The narrator means the period of the German occupation during the Second World War.
Figure 2. The ratio of occurrences of täma to occurrences of še by different speakers.

It can be clearly seen that there is a difference between the speakers: Speaker LK has a ratio of 2.3, meaning that tämä is more than twice as frequent in her speech as še. For all other speakers, the ratio is below 1, so they use še more frequently than tämä. On the left side we see two speakers who do not use tämä at all (EN) or use it minimally (ST). The four speakers in the middle have more homogeneous ratios.

The differences between the speakers do not correlate with the dialectal zones, for instance EN and LK both represent the southern zone, while ST is from the northern zone. These differences can however be explained with the speakers’ language biographies. Both EN and ST have spent significant parts of their lives in Estonia. In Standard Estonian (based on the northern dialect), see ‘this’ is basically the only demonstrative pronoun (Pajusalu 2006), while tema ‘he/she’ functions as a third person singular pronoun. Thus, it would appear that the Estonian system influenced the use of Ingrian demonstratives in the speech of EN and ST, and they (almost) completely switched to a unipartite system.

Speaker LK is the youngest and her speech demonstrates more Russian elements and more code-switching into Russian compared to other speakers. We suggest that the predominance of the proximal tämä over še reflects the Russian influence in her case, because in Russian the unmarked and more frequent pronoun is the proximal ětot ‘this’ and not the distal tot ‘that’. Other speakers, GI, AI, OM, MM, lived their whole lives in Ingrian villages and communicated in their native language.

It therefore seems that the demonstrative pronominal system is very sensitive to contact influence.
6. Conclusions

Contemporary Ingrian has preserved only two components of the original tripartite system of demonstrative pronouns: proximal tämä and neutral še. The pronoun še is used as the unmarked variant in this pair. It is very untypical for Ingrian to express a contrast between referents by opposing these two demonstratives. Based on these observations, we hypothesize that the system of Soikkola Ingrian demonstratives was developing towards a unipartite system in a similar fashion to northern Estonian dialects (Pajusalu 2006). This shift was, however, slowed down due to the increase in discourse uses of tämä. In contemporary data, the demonstrative tämä is frequently found as a discourse marker. This may well be due to the influence of the Russian language, where the proximal demonstrative ètot is often used in discourse functions unlike its distal counterpart. The contemporary picture is far from being complementary distribution, but we nevertheless observe a slight tendency towards distributing the referential vs discourse uses between še and tämä, respectively. In a way, the preservation of the bipartite system in Soikkola Ingrian due to the Russian influence is reminiscent of the situation in Votic and Veps, where the bipartite system was restored due to the Russian influence (see Introduction).

The material that we analysed shows a wide range of functions for the demonstrative pronouns. In the texts recorded in the middle of the 20th century, Yurayong (2020) only found the deictic and anaphoric uses (in terms of his classification). In the contemporary data, we also have examples of other reference types, including recognitional and bridging references, as labelled in the classification by Yurayong (2020).

Yurayong’s (2020: 186) suggestion that “deictic use is more frequently found in the use of preposed demonstratives, whereas anaphoric and other referential uses are more common for postposed demonstratives” was not entirely confirmed by the contemporary Ingrian material. Postposed demonstratives are quite rare in our data, while preposed demonstratives have various functions including anaphoric ones.

We found many examples where the reference to human antecedents is encoded with demonstratives, and not with personal pronouns. This fact supports the observation expressed by Pajusalu (2015) about the difference between written and spoken languages. The predominance of third person pronouns referring to human/animate entities in written
Finnish and Estonian is believed to have been artificially introduced under the influence of Indo-European languages (Pajusalu 2015: 170). Ingrian, which had a written variety only for a very short time period, seems to tolerate the use of demonstratives for referring to human beings quite easily.

Although the functions of demonstrative pronouns are similar for most speakers of Soikkola Ingrian, we observed individual differences, particularly in speakers who had been exposed to another language (Estonian or Russian) for extended periods of time. The speakers who lived in Estonia for a long time have lost or considerably reduced their use of tämä. On the contrary, the speaker with the greatest Russian influence uses this pronoun more frequently than the others. Apparently, the demonstrative system is very sensitive to contact influence, so the language biographies of individual speakers should be taken into account when making general conclusions about its function.

Acknowledgements

We would like to thank two anonymous reviewers for many valuable comments on the article.

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

Abbreviations

References


Märksõnad: isuri keel, asesõnad, demonstratiivpronoomenid, deiksis, anafoorid, diskursusemarkerid, keelekontakt