

Saami Linguistics in Uppsala

The SAALS 4 Symposium

ROGIER BLOKLAND, MICHAEL RIEßLER
& TORBJÖRN SÖDER (EDS)



Studia Uralica Upsaliensia 42



UPPSALA
UNIVERSITET

Acta Universitatis Upsaliensis
Studia Uralica Upsaliensia 42

Saami Linguistics in Uppsala

The SAALS 4 Symposium

Rogier Blokland, Michael Rießler
& Torbjörn Söder (eds)



UPPSALA
UNIVERSITET

Uppsala 2025

© The authors

This work is licensed under a Creative Commons Attribution
NonCommercial-NoDerivatives 4.0 International License (CC BY-NC-ND)

ISSN 1101-7430

ISBN 978-91-513-2708-2 (print)

ISBN 978-91-513-2747-1 (electronic)

<http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-573352>

Distributed by Uppsala University Library
Box 510, SE-751 20 Uppsala, Sweden
www.ub.uu.se, acta@ub.uu.se

Printed by Kph Print AB, Uppsala 2025

Contents

Preface	5
MARKUS JUUTINEN & JUSSI YLIKOSKI: The fourth-person verb forms in Skolt Saami.....	9
OLLE KEJONEN: Verb inflection in Forest Lule Saami.....	37
RICHARD KOWALIK: South Saami finite verb morphology	99
JUHA KUOKKALA: Inchoative verbs in Saami: Derivational types and their variation	119
JOHANNA NICHOLS: Morphological complexity in Saami: A historical-typological view	159
KSENIA SHAGAL & ALENA BLINOVA: Bivalent verb classes in Skolt Saami: A pilot study	181
TORBJÖRN SÖDER: On the origin of South Saami <i>edjedh</i> ‘shall, will’	207
RIITTA-LIISA VALIJÄRVI & LILY KAHN: A comparative study of user attitudes to Sámi orthographies.....	235
TAARNA VALTONEN: Areal variation in structure types of Saami place names	253
JOSHUA WILBUR: <i>Bisumsáme tjállemvuoge birra</i> - A review of the official Pite Saami orthography	285

Preface

Saami languages have been taught and studied as an academic subject at Uppsala University since the late 19th century. Given this longstanding and celebrated tradition of research on Saami linguistics, we were especially honoured and pleased to host the SAALS 4 conference as a contribution to contemporary research on the Saami languages.

‘SAALS 4’ refers to the fourth conference in a series of international conferences on Saami linguistics. Tromsø hosted the inaugural conference in 2006 as well as the second one in 2009. The following conference took place in Freiburg in 2017 with the 2019 Uppsala conference resulting in the present volume. The last two took place in Tartu in 2022 and Helsinki in 2024. The next conference is planned for 2026 and will be organized by Nord University/Noerhte universitete/Nuortta universitietta in Norway.

Availing itself of the same name as the first two events, a third SAALS conference evolved into the current conference series after almost a decade. At the time, Michael Rießler and Joshua Wilbur founded a small working group in Saami linguistics at the Department of Scandinavian Studies at the University of Freiburg in Germany. Their intent was to re-establish the conference under its original name with a similar concept. The original organisers in Tromsø were glad to “hand over” the established name and gave their blessings to the new event. Since that time, it has become a tradition to decide the next location at the present conference, so organisers can begin preparations.

Another tradition involves inviting international guest and keynote speakers. This custom began at SAALS 1 and SAALS 2 in Tromsø, where Anna-Riitta Lindgren and Trond Trosterud from Norway gave keynote addresses as well as Zita McRobbie-Utasi and Ida Toivonen from Canada, and Pekka Sammallahti from Finland. The speakers covered broad perspectives ranging from the sociology of language to formal approaches of language description.

To further enrich the next conference, organisers invited a keynote speaker from outside of Saami studies to explore topics in general linguistics with potential Saami relevance. With this new tradition, Matthew Baerman of the University of Surrey delivered a talk on three-way length alternations in

morphology at the SAALS 3 Freiburg conference, and Johanna Nichols of the University of California, Berkeley, discussed morphological complexity in Saami at the 2019 SAALS 4 Uppsala conference. Maria Polinsky of the University of Maryland presented a keynote address on structural parallels between endangered and heritage languages at the 2022 SAALS 5 Tartu conference.

SAALS 6 returned to the original format with more than one invited speaker: Jussi Ylikoski of the University of Turku and Minerva Piha of Nord University. A recent and visible change in the SAALS 6 program features papers presented in various Saami languages alongside English.

Simultaneously, as traditions evolved, each SAALS conference came with innovations; SAALS 4 became the first to publish conference proceedings. Resulting from a long delay due to unforeseen (and foreseen) circumstances the decision was made to list only the authors' current affiliations in the present volume.

At present, academic research and pedagogy in Saami studies are relatively small; consequently, meetings are generally restricted to centres of Uralic studies overlapping with only a few universities worldwide. As such, the working atmosphere is small and collegial with participants already familiar on both personal and professional terms. Thus, the *Saami Linguistics Symposium* has never been formally constituted as an institution with members and a board.

Since the first meeting of SAALS, the purpose is to grant researchers the opportunity to present current findings to their colleagues, participate in formal and informal discussions on relevant topics, and explore potential opportunities for collaborative work. As always, the call for papers encourages younger researchers to present their work at the conference.

Concerning the present volume: each paper has been reviewed by two anonymous peer-reviewers, and with regard to language names we let each author use their preferred form, so there has been no harmonization from our side with regard to spellings of Sámi, Saami, Sami, etc.

Saami Linguistics Symposia (SAALS)

The following chronicle includes a summary of all SAALS meetings up to the present date.

SAALS 1 Tromsø

Since the 1st Saami Linguistics Symposium, held on 23–24 November 2006 at the University of Tromsø, already included an ordinal number in its official

title¹, the organisers Patrik Bye, Kjell Kemi, and Ove Lorentz seem to have planned this symposium as the initiation of a series of meetings.

This meeting had three invited speakers: Pekka Sammallahti ('How to describe quantity in North Saami'), Ida Toivonen ('A phonetic study of consonant length in Inari Saami'), and Trond Trosterud ('Language technology for endangered languages: Sámi as a case study').

SAALS 2 Tromsø

The second conference was held on 19–21 November 2009, again at the University of Tromsø. Organisers: Kjell Kemi, Ove Lorentz, Bruce Morén-Duolljá, Laila S. Oskarsson, Mikael Svonni and Trond Trosterud.

This meeting had three invited speakers: Anna-Riitta Lindgren ('Urban Saamis and Use of Saami language'), Zita McRobbie-Utasi ('An experimental phonetic approach to the analysis of the Skolt Saami vowel system'), and Pekka Sammallahti ('How to describe quantity in North Saami').²

SAALS 3 Freiburg

The conference was held on 19–20 October 2017 at the Freiburg Institute for Advanced Studies (University of Freiburg). Organisers: Michael Rießler, Joshua Wilbur, Tim Feist, and Flore Picard.

The conference invited Matthew Baerman as keynote speaker ('Inflectional morphology in a cross-linguistic perspective').

SAALS 4 Uppsala

The conference was held on 14–15 November 2019 at Uppsala University. Organisers: Rogier Blokland, Olle Kejonen, Richard Kowalik, Michael Rießler, Torbjörn Söder, and Riitta-Liisa Valijärvi.

The conference invited Joanna Nichols as keynote speaker ('Grammatical complexity in Saami: A historical-typological view').

SAALS 5 Tartu

The conference was held on 22–23 September 2022 at the University of Tartu. Organisers: Joshua Wilbur, Helen Türk, and Merit Niinemägi.

The conference invited Maria Polinsky as keynote speaker ('Structural parallels between endangered and heritage languages').

¹ A version of the original webpage (archived on 29 May 2010), including the call for papers, can be found under <https://web.archive.org/web/20100529174317/https://www.hum.uit.no/laponic/> (7.1.2023). However, the complete conference program could not be reconstructed.

² Similarly to SAALS 1 we were not able to reconstruct the complete program.

SAALS 6 Helsinki

The conference was held on 6–7 November 2024 at the University of Helsinki. Organisers: Ilona Kivinen, Ulla-Maija Forsberg, Riho Grünthal, Sampsa Holopainen, Mariann Bernhard, Rigina Ajanki, Pirjo Virtanen, Juha Kuokkala and Hanna Guttorm.

The conference invited Jussi Ylikoski (‘Sámegielaide sajádat urálalaš giellabearrašis ja -dutkanis’) and Minerva Piha (‘Saami language history then and now: research historical perspectives’) as keynote speakers.

The present open access publication has been financed by the Department of Modern Languages of Uppsala University and the Editorial Committee for *Acta Universitatis Upsaliensis*, for which the editors would like to express their gratitude. We also thank Brady King for technical assistance as well as Olle Kejonen, Richard Kowalik and Riitta-Liisa Valijärvi for having organized the conference together with us.

Rogier Blokland, Michael Rießler and Torbjörn Söder

18.12.2025, Uppsala

The fourth-person verb forms in Skolt Saami

Markus Juutinen, University of Oulu & Jussi Ylikoski, University of Turku

1. Introduction

This paper provides a description of a morphosyntactic phenomenon that is little known outside – or even inside – Saami linguistics: what is known as the impersonal, or indefinite, fourth-person verb forms in Skolt Saami.

Within Uralic linguistics, probably the best-known feature of Saami verbs is the dual vs. plural opposition, which appears to relate Saami morphosyntax to Khanty, Mansi, Samoyed and ultimately to Proto-Uralic. However, the easternmost Saami vernaculars that have survived to our day do not have dual except for Aanaar (Inari) Saami (and Skolt Saami personal pronouns), but their finite verbs do not always consist of only three persons in two numbers either. Instead, especially Skolt Saami has a seventh type of person category, occasionally labeled as an “impersonal”, “indefinite” or a “fourth” person. In practice, we are dealing with an inflectional category that is very similar to the one in Finnic, most commonly known as the “passive” in Finnish grammatical tradition, or the “impersonal” in Estonian grammar. Moreover, it is instructive to approach Skolt Saami also by comparing it with Finnish as well as other Saami languages under Finnish influence, as the Saami forms are not only historically related to those of Finnic, but Finnish is also the majority language that has had the most overwhelming effect on the development of Skolt Saami during the past century.

From the perspective of verb inflection, the three Saami languages of Finland can be divided in two “western” languages and one “eastern” despite the fact that the major dividing line between Western Saami and Eastern Saami is commonly drawn between North and Aanaar Saami (e.g., Sammallahti 1998b: 6–7). However, the traditional view is mostly based on phonological and morphological arguments, whereas recent lexical studies see Aanaar Saami as a relatively western Saami language (Rydving 2012; Tillinger 2014; however, see Juutinen 2023: 110). As regards the main theme of this study, it is noteworthy that especially the Skolt Saami paradigm is

analogous to that of Finnish instead of North and Aanaar Saami with identical person–number categories.

Tables 1–4. The indicative present tense forms for ‘go’ in North, Aanaar and Skolt Saami, and Finnish.

1. North Saami				3. Skolt Saami		
	SG	DU	PL		SG	PL
1	<i>manan</i>	<i>manne</i>	<i>mannat</i>	1	<i>mõõnam</i>	<i>mõõnnâp</i>
2	<i>manat</i>	<i>mannabeahtti</i>	<i>mannabehtet</i>	2	<i>mõõnak</i>	<i>mõõnnve'ted</i>
3	<i>manná</i>	<i>mannaba</i>	<i>mannet</i>	3	<i>mâânn</i>	<i>má'nne</i>
				4	<i>mõõnât</i>	

2. Aanaar Saami				4. Finnish		
	SG	DU	PL		SG	PL
1	<i>moonâm</i>	<i>manneen</i>	<i>moonnâp</i>	1	<i>menen</i>	<i>menemme</i>
2	<i>moonah</i>	<i>monâvettee</i>	<i>monâvetted</i>	2	<i>menet</i>	<i>menette</i>
3	<i>mana</i>	<i>moonnâv</i>	<i>maneh</i>	3	<i>menee</i>	<i>menevät</i>
				4	<i>mennään</i>	

In general, the fourth person in Skolt Saami is syntactically on a par with the first, second and third persons in singular and plural. Fourth-person forms in other moods are somewhat rare, but the present tense form in *-t* is paralleled by the past tense form *-š*. From a functional perspective, the Skolt Saami fourth person greatly resembles the general Finnic impersonal verb forms traditionally labeled as “passives” or impersonals, and the Finnic impersonals or passives are also actually regarded as the material origins of the Saami forms (E. Itkonen 1957: 4; Korhonen 1967: 346–348).

The fourth person has been well-known among the handful of specialists on Skolt Saami, but next to unknown outside Saami linguistics. The few earlier descriptions have mostly focused on the historical morphology of the verb forms in question, but very little has been said about the ways and reasons the forms are actually used; almost all details of its syntax and semantics have been left undescribed even in the most detailed accounts of Skolt Saami grammar (cf. Moshnikoff et al. 2020: 82; Feist 2015: 200, 234–235; Lehtinen 2018: 11–12, 95–96).

The structure of the paper is as follows: Section 2 is a brief introduction to the history of the Skolt Saami fourth person and to earlier research. Section 3 provides a concise description of the phenomenon as a specialized impersonal form, or a non-promotional passive, and its relation to other functionally related expressions in the language. As all speakers of Skolt Saami are

bilingual in the majority languages Finnish and Russian, the language is greatly affected by interference from the respective majority languages on both sides of the border. Section 4 presents further observations on the most recent developments of the fourth-person forms that show signs of extending their functions to those of third-person plural forms – and vice versa – as well as using the fourth person as a kind of promotional passive whose only core argument may be in the nominative instead of the accusative.

The study is based on virtually all relevant materials, from the first recordings of Skolt Saami folklore (T. I. Itkonen 1931) to later recordings (Giellagas Corpus) and ultimately the contemporary written language, largely consisting of official translations from Finnish (e.g., SIKOR corpus of about 213,000 words). However, the main result of fieldwork among present-day speakers is that many native speakers regard the fourth-person forms as foreign to their own idiolects or even the language in general (see Sections 3.1 and 4.1). Although fourth-person forms can also be found in Akkala, Kildin and Ter Saami, this study focuses mainly on Skolt Saami, but also takes into account some features of the fourth person in other eastern Saami languages.

2. Background

2.1. An ancient morphological loan from Finnic

The fourth-person forms are commonly regarded as an ancient morphological loan from Finnic, more precisely from the predecessors of present-day Karelian and Finnish. According to E. Itkonen (1957: 4) and Korhonen (1967: 346–348), both present and past tense suffixes have been borrowed from Karelian to Saami at the time when the Finnic impersonal forms had not taken over the functions of the plural third-person forms in Karelian. The impersonal forms appear to have replaced the plural third-person forms in Karelian centuries ago, but in Finnish and Veps the distinction between the two categories has been mostly preserved (Kettunen 1943: 57–61, 427; Nirvi 1947; Laanest 1982: 231).

Another sign of the long history of the fourth person is the fact that is attested in the very first documentation of a Saami language by Stephen Borough in 1557 (Korhonen 1967: 346). While Borough's data comes from an idiom that can be labeled as (Old) Kildin or Ter Saami, the easternmost Saami language, the range of this category has extended to the westernmost Eastern Saami, the now extinct Kemi Saami as used by Olaus Sirma in the 17th century (Bergsland 1984: B 35; Sammallahhti 1984: 148; 1998a: 84). As it happens, the only Eastern Saami language without a trace of the fourth person

is Aanaar Saami, the most Finnicized Saami language spoken today (Valtonen et al. 2022: 188-189).

From a functional perspective, the fourth person greatly resembles the general Finnic impersonal verb forms traditionally labeled as “passives” in Finnish grammar. From a morphological perspective, the similarity is not that obvious, but close enough for T. I. Itkonen (1942: 55) to have proposed that the origin of the Skolt Saami present-tense fourth-person form *jeä'le-t* (live-4) ‘one lives’ lies in Finnish *ele-tään* (live-PASS) id. and Karelian *ele-täh* (live-3PL) ‘they live; people live; one lives’ (Table 5).

Table 5. The indicative present and past-tense fourth-person forms for ‘live’ in Skolt Saami, and their cognates in Finnish and Karelian.

		PRESENT	PAST
Skolt Saami	4 (4th person)	<i>jeä'let</i>	<i>jie'lleš</i>
Finnish	PASS	<i>eletään</i>	<i>elettiin</i>
Karelian	3PL	<i>eletäh</i>	<i>elettih</i>

This view is supported by E. Itkonen (1957: 4) and Korhonen (1967: 346–348) who add that the past-tense fourth-person suffix *-š* probably goes back to the word-final *-h* in the Karelian past-tense suffix *-ttih*; substitution of Finnic *-h* with Saami *-š* appears to be a strategy of phonological-cum-etymological nativization, comparable to words like Skolt Saami *kaadâš* ‘jealous’; cf. Karelian *kajeh* id.

Tables 6–7. The negative indicative present tense forms for ‘go’ in Skolt Saami and Finnish.

6. Skolt Saami			7. Finnish		
	SG	PL		SG	PL
1	<i>jiðm mððñ</i>	<i>jeä'p mððñ</i>	1	<i>en mene</i>	<i>emme mene</i>
2	<i>jiðk mððñ</i>	<i>jeä'ped mððñ</i>	2	<i>et mene</i>	<i>ette mene</i>
3	<i>ij mððñ</i>	<i>jie mððñ</i>	3	<i>ei mene</i>	<i>eivät mene</i>
4	<i>jeüt mðñnu</i>		PASS	<i>ei mennä</i>	

As already shown by Tables 1–4, Skolt Saami verb inflection resembles that of Finnish much more than those of Aanaar and North Saami, for example. The poorly attested fourth-person forms for conditional (*jeälčeš* ‘one would live’) and potential (*jie'lžet* ‘one might live’) moods do not have obvious cognates in Finnic. From a morphological point of view, the most remarkable difference from Finnic is that Skolt Saami has a specialized fourth-person form even for the negative auxiliary (Miestamo & Koponen 2015: 355): *jeä-t*

mõnnu (NEG-4 go.CNG2) ‘one doesn’t go; people don’t go’. However, the lexical verb is in a connegative form otherwise used in certain prohibitive constructions (e.g., *jeállap mõnnu* (neg.IMP.1PL go.CNG2) ‘let’s not go’; whereas in Finnic, special passive (indicative) connegative forms (*mennä* go.PASS.CNG) are used. The functional yet not formal equivalents of Skolt and Finnish negative present tense constructions are seen in Tables 6–7.

2.2. History of research

As mentioned above, fourth-person forms have been attested already in the 16th and 17th centuries. The first linguist to have documented these forms appears to have been D. E. D. Europaeus (*lendshit* be.POT.4) in his records of Ter Saami in 1856 (Korhonen 1967: 347). Linguistic description of the fourth person appears to have begun in T. I. Itkonen’s (1942: 55) and E. Itkonen’s (1957: 4) brief comments on the origins of the forms, but the actual description is limited to seven detached verb forms (from Skolt, Kildin and Ter Saami), translated with Finnish passive verb forms. The first example sentences are given by Korhonen (1967: 346–348), but his eyes are also on historical morphology, the topic of his doctoral dissertation. However, in his subsequent paper on the expression of the indefinite subject in Saami languages, Korhonen (1970: 144) is apparently the first one to view the fourth person in a decidedly semantic context, albeit very briefly. In the first comprehensive grammar sketch of Skolt Saami, Korhonen (1973: 67ff.) is the first one to present the category in question in complete inflectional paradigms on a par with first-, second-, and third-person singular and plural. Moreover, this source also presents fourth-person forms for not only present and past indicatives, but also for conditional and potential moods (Korhonen 1973: 67ff.; see also Korhonen 1977: 81 for a brief mention of the phenomenon). These forms are also presented in the inflectional paradigms by Sammallahhti and Mosnikoff (1991: 169–179).

Zajkov (1987: 145–147; [Zaikov] 1996: 141) has discussed the development of the fourth-person forms in Akkala Saami. Sammallahhti (1998b: 29) mentions these forms as one of the two main criteria for distinguishing between Akkala Saami and Skolt Saami proper: in Akkala Saami, according to him, the fourth-person forms have entirely displaced the original third-person plural forms, thus resulting in the loss of the fourth person as an independent category. Although Kildin Saami is better studied than Akkala Saami, only indicative affirmative fourth-person verb forms have been described (see E. Itkonen 1957: 4; Korhonen 1967: 347; Kert 1971: 178, 549–550; but see Rießler 2022: 228–230 for the conditional mood).

The most recent descriptions of the fourth person in Skolt Saami are presented in the grammar by Moshnikoff et al. (2020) and the reference grammars by Feist (2010, 2015), and the phenomenon is also discussed in Lehtinen's (2018) master's thesis on Skolt Saami passive derivatives.

To briefly summarize the semantic characterizations given to the verb forms labeled as fourth person here, the following observations can be made: Most scholars of Skolt Saami have described the forms in question either in Finnish or at least with reference to the so-called passive in Finnish – either from a synchronic or diachronic perspective. T. I. Itkonen (1942: 55) calls it an impersonal (Finnish *persoonaton*) verb form, E. Itkonen (1957: 4) speaks of an impersonal passive form. The label “fourth person” is adopted from Korhonen (1967: 346–348; 1970: 144) who appears to have adopted it from Harms' (1962: 57–58 *et passim*) grammar of Estonian, even though the Estonian grammatical tradition refers to Harms' “fourth person” as an impersonal (*impersonaal* or *umbisikuline tegumood* ‘impersonal voice’) (see also Zajkov 1987: 145). Despite Korhonen's (1967, 1970) label “fourth person”, he characterizes these forms as verb forms for indefinite agents or subjects. Soon afterwards, Korhonen (1973: 67) stated that in addition to the first, second and third persons in singular and plural, Skolt Saami also has “a seventh personal form” or the “indefinite person” that can most often be translated with the Finnish passive, occasionally also with the third-person plural. Moreover, Bergsland (1984: B 35) refers to Kemi Saami forms as impersonal forms or “the ‘passive’ of the Finnish type”. However, the most influential modern Skolt Saami grammars by Moshnikoff et al. (2020: 80) and Feist (2010: 115; 2015: 200) call the phenomenon “a fourth, indefinite person”, thus following Korhonen's (1967, 1970) early wordings.

In principle, there is a significant difference whether a verb form is characterized as a personal form, be it a fourth, seventh or an indefinite person, or whether it is called impersonal, the way these forms were originally characterized (T. I. Itkonen 1942; E. Itkonen 1957), and the way its Finnic counterparts are often characterized. However, as the mainstream term “fourth person” is adopted in the present paper, the term in itself must be seen as a term only, and the nature of this category still remains open for competing interpretations.

Although the fourth-person forms in Skolt Saami and the neighboring languages have duly been mentioned by many grammarians and other scholars, the descriptions have remained quite scanty. Even Moshnikoff et al. (2020: 80–107) and Feist (2010: 115–136; 2015: 200–232) are content with providing just a few example sentences without truly analyzing their structure and meaning. This is all the more surprising in light of the fact, laconically

remarked by Schlachter (1970: 152–153) in his comment to Korhonen (1970), that the only truly indefinite verb forms in the entire Saami branch of Uralic languages are the fourth-person forms in East Saami. Although indefiniteness can be expressed by many kinds of clauses with personal verb forms, according to him, the fourth person is the only truly grammatical category for indefiniteness in Saami languages.

2.3. Skolt Saami conjugation within Saami languages

As already mentioned in the introduction, verb conjugation in Saami languages is most often known for the existence of the dual vs. plural distinction, but in our days this distinction is found only in the languages spoken west of Skolt Saami – the languages with no traces of the fourth person characteristic of the easternmost Saami languages instead. Even though it is in itself interesting that this major isogloss goes along the language border between Aanaar and Skolt Saami, it is even more interesting to note the dividing line can be observed inside the Skolt Saami, as the fourth person has not been attested in the Paččjokk–Peäccam and Njauddâm dialects, which have the dual vs. plural distinction instead. Sammallahti (1998b: 30–31) considers the dual vs. plural distinction one of the most important criteria for drawing the line between the northern (Njauddâm, 1 in Map 1; Paččjokk, 2; Peäccam, 3 and Mue’tkĕ, 4) and southern (Suõ’nn’jel, 5; Njuõttjäu’rr, 6; Sâârvesjäu’rr, 7) dialects of the language. This state of affairs is depicted in Table 8, which also shows the situation in Akkala, Kildin and Ter Saami in the east. The person-number categories of the western type are unanimously regarded as direct descendants of the Proto-Saami conjugation (e.g., Korhonen 1967; Sammallahti 1998b: 212–221).

Table 8. Person categories in the Saami languages, the impersonal fourth person included.

South, Ume, Pite, Lule, North and Aanaar Saami	Paččjokk–Peäccam Skolt Saami (as well as the extinct Njauddâm dialect)	Suõ’nn’jel and Njuõttjäu’rr Skolt Saami	Akkala	Kildin and Ter
1SG	1SG	1SG	1SG	1SG
1DU	1DU	1PL	1PL	1PL
1PL	1PL			
2SG	2SG	2SG	2SG	2SG

2DU	2DU	2PL	2PL	2PL
2PL	2PL			
3SG	3SG	3SG	3SG	3SG
3DU	3DU	3PL	3PL	3PL
3PL	3PL			
		4		4



Map 1. Traditional Skolt Saami area. Map by Timo Rantanen et al. 2022; modified by Ali Ylikoski.

The right-hand columns in Table 8 remind us of the fact that the person categories in Skolt, Kildin and Ter Saami are in principle identical. However, as regards the fate of the fourth person in the intermediate Akkala Saami, it is most interesting to note that even though the impersonal fourth-person in itself has been lost in the language, the loss of the category has not meant the loss

of the verb forms as such, but instead, the cognates of the Skolt, Kildin and Ter Saami fourth-person forms have replaced the earlier third-person plural forms (Zajkov 1987: 144ff.; [Zaikov] 1996: 141; Sammallahti 1998b: 29). What is more, this development is fully in accord with the development of the corresponding Finnic forms – the presumed source of the fourth-person forms in Saami – in Karelian (Nirvi 1947). However, while the merging of the third-person plural and the impersonal passive (i.e., the fourth person) appears to have taken place centuries ago in Karelian, the situation in Akkala Saami seems to be relatively new, as suggested by the remnants of the original third-person forms in the pioneering records of the language by Jens Andreas Friis (1867), e.g., *leije* (be.3PL), *läjji* (be.PST.3PL), *lenče* (be.COND.3PL), *jelläk* (be.NEG.3PL), *jellemenč* (be.NEG.PST.3PL), *je lenče* (NEG.3PL be.COND.CNG).¹ In any case, it is notable that the pattern that emerges here is that Akkala Saami and Karelian in the south have developed analogously, whereas Skolt Saami and Finnish in the west and Kildin and Ter Saami in the east have preserved the more original paradigms better (Tables 3–4, 9–10; but see Sections 4.1 and 4.2).

Tables 3–4 (repeated here) and 9–10. The indicative present tense forms for ‘go’ in Skolt Saami, Finnish, Akkala Saami and Karelian.

3. Skolt Saami			9. Akkala Saami		
	SG	PL		SG	PL
1	<i>mõõnam</i>	<i>mõõnnâp</i>	1	<i>mõõnam</i>	<i>mõõnnâp</i>
2	<i>mõõnak</i>	<i>mõõnnve'ted</i>	2	<i>mõõnak</i>	<i>mõõnnve'tte</i>
3	<i>mâânn</i>	<i>mâ'nne</i>	3	<i>mâânn</i>	<i>mõõnât</i>
4	<i>mõõnât</i>				

4. Finnish			10. Karelian		
	SG	PL		SG	PL
1	<i>menen</i>	<i>menemme</i>	1	<i>mänen</i>	<i>mänemmä</i>
2	<i>menet</i>	<i>menette</i>	2	<i>mänet</i>	<i>mänettä</i>
3	<i>menee</i>	<i>menevät</i>	3	<i>mänöy</i>	<i>männäh</i>
PASS	<i>mennään</i>				

It goes without saying that in the context of multifaceted language contacts among the languages in question – other neighboring Saami languages and Russian included – Table 3 must be understood as a rough generalization of the state of affairs in Skolt Saami. This is the starting point of the more

¹ Friis (1867) has also recorded third-person dual forms otherwise unattested in Akkala Saami. These have also been lost and replaced by fourth-person forms.

nuanced description of the nature of the fourth person in Skolt Saami in the following sections.

3. The main features of the fourth person in Skolt Saami

This section describes the main features of the fourth person in the language that can be characterized as “classical Skolt Saami” – the Skolt Saami spoken in its traditional territory prior to the Second World War as well as in the resettled areas in Inari, Finland, for some time before the abrupt wholesale language shift to Finnish. For the history of Skolt Saami speaking areas and history of language contacts, see Linkola & Sammallahti (1995: 46–55) and Juutinen (2023: 29–56). The sentences presented here also contain examples of the “classical” use of the fourth-person in our times.

3.1. An impersonal passive

As already seen above, the verb forms labeled as fourth-person forms have been characterized as “impersonal verb forms”, “impersonal passives”, “a seventh personal form”, “‘passive’ of the Finnish type” or “a fourth, indefinite person”. Although the use of all these labels can be justified, we wish to choose the attribute “impersonal” over “indefinite”. Not unlike in many other grammatical traditions, the latter term has other functions in the realm of (indefinite) pronouns, for example, but the term “impersonal” is not overloaded to the same extent. In any case, the term “impersonal” is to be understood as a label for the inflectional category that does not explicitly refer to first, second or third-person agents, but instead, the identity of the agent(s) is simply left unspecified (or “indefinite”).

Moreover, fourth-person forms can also be characterized as (impersonal) passive forms, as was already done by E. Itkonen (1957: 4) and Bergsland (1984: B 35), and less explicitly by Korhonen (1973: 67) who noted that the Skolt Saami fourth-person forms can usually be translated with the Finnish passive. We acknowledge the fact the label “passive” has been debated among Finnish grammarians who sometimes tend to equate the label with the personal passives characteristic of many Germanic passive constructions in which the expression of the agent is not always fully deleted in the sentence but rather demoted to an oblique position (see, e.g., Shore 1988 and Helasvuo 2006). However, following a less categorical approach represented by scholars such as Givón (2001: 127ff.), Siewierska (2008, 2010) and many other typologists, we do not see obstacles in characterizing most occurrences

of the fourth-person forms as (impersonal) passive clauses in which the agent is fully absent.

- (1) *Kǎddtǎǎ'lv* *tǎid* *nue'rrsiǎrid*
spring.winter.GEN that.PL.ACC string.game.PL.ACC
siǎ'rreš.
play.PST.4
‘In the early spring that string game (*nue'rrsiǎrr*) was played.’
(Če'vetjǎu'rr², Kotus 3319_1az: 55:48)
- (2) *Tǎin* *vǎnnsin* *mi'jjid* *uiddeeš*
that.COM boat.COM 1PL.ACC take.away.PST.4
tǎ'lvv-si'jdde.
winter-village.ILL
‘With that boat they took us away to the winter village.’ (Njeǎ'llem³, Kotus 12744_1a: 02:09)

In addition to transitive verbs (1–2), this category can be applied to virtually all kinds of verbs, including intransitives (3–4) and the copula (4). However, it is notable that quite like the impersonal verb forms in Finnic languages, the Skolt Saami fourth-person appears to be formed and almost always used with reference to actions and states in which the absent subject arguments (agents and themes) are human (but see Section 4.2):

- (3) *Tǎ'lvvpǎikka* *mǎ'nneš* *suukkǎmvǎnnsin*
winter.place.ILL go.PST.4 row.boat.COM
pǎ'rǧǧmannu *looppǎst* *da* *kie'sspǎikka*
August.GEN end.LOC and summer.place.ILL
puǎ'tteš *mǎǎusat* *vue'ssmannust.*
come.PST.4 back May.LOC
‘People moved to the winter village by rowboat at the end of August, and returned to the summer village in May.’ (SIKOR)
- (4) *To'ben* *nuǎ'tteš* *da* *le'jješ*
there fish.with.a.seine.PST.4 and be.PST.4
ǎǎutsǎǎ'jest *tollpeǎllsest* *kaa'fstǎǎlee'l,*
together next.to.fire have.a.coffee.break.CVB.INS

² The Skolt Saami spoken in Če'vetjǎu'rr (Finland) is the successor of the dialect spoken in Suǎ'nn'jel.

³ Unlike most speakers in Njeǎ'llem (Finland), this speaker does not speak the successor of the Paččjokk–Peǎccam dialect but that of Suǎ'nn'jel.

<i>bliinid,</i>	<i>kue'llkääkkaid</i>	<i>da</i>
pancake.PL.ACC	fish.cake.PL.ACC	and
<i>mää'rfid</i>	<i>pää'stee'l.</i>	
sausage.PL.ACC	bake.CVB.INS	

‘People fished with a seine and were together, having coffee at campfire, and making pancakes, fishcakes and sausages.’ (Kolttsaamelaiset 2013)

In addition to indicative fourth-person forms, grammars (e.g., Korhonen 1973: 67; Feist 2015: 204; Moshnikoff et al. 2020: 94–103) provide examples of analogous forms in the conditional and potential moods. Unlike the indicative fourth-person markers *-t* and *-š*, the conditional fourth-person forms in *-češ* and the potential in *-žet* cannot go materially back to their Finnic equivalents, but forms like *mõõnčeš* (go.COND.4) and *mõõnžet* (go.POT.4) appear to have developed by analogy to the Skolt Saami indicative paradigm. However, such forms are not attested in any of the corpora of spoken or written language (more than 500,000 words), but a couple of examples are provided by the Skolt Saami grammar by Moshnikoff et al. (2020: 101–103), as seen in (5). In other words, the conditional and potential forms must be regarded highly marginal, whether the ultimate reason is the unbalanced nature of the available corpora or the possible artificiality of the forms in question (cf. the introduction of the passive potential forms such as *mentäneen* ~ *mentänee* in the standard Finnish in the end of the 19th century; see Tunkelo 1934).⁴

- (5) *Kue'llšee'llmõ'sše* *taarbše'žet* *kuddnallšem*
 fish.catch.NMLZ.ILL need.POT.4 good.kind.ADV
säai'mid *da* *nuõ'ttid.*
 net.PL.ACC and seine.PL.ACC
 ‘One needs proper nets and seines for fishing.’ (Moshnikoff et al. 2020: 101)

As mentioned in Section 2.1, Skolt Saami differs from Finnic in having a special fourth-person form even for the negative verb (Miestamo & Koponen 2015: 355):

- (6) *Päällain* *siõrât,* *mutta* *jeät*
 ball.COM play.PRS.4 but NEG.4
nu'tt *siõrru* *ko* *mä'htt* *te'l*
 so play.CNG as then how

⁴ The use of the potential mood in a non-subordinate clause in (5) also bears a flavor of Finnish literary influence (cf. Bartens 1980).

siõ'rreš.

play.PST.4

‘One (still) plays with a ball, but one doesn’t play now the way it was played then.’ (Če’vetjäu’rr, Kotus 9832_2az: 30:14)

- (7) *Ooccmõðžžid* *jeüt* *maacctuku.*
search.NMLZ.PL.ACC NEG.4 return.CNG
‘Applications will not be returned.’ (SIKOR)

One of the reasons for the limited description of the fourth person may be the fact that these forms were quite infrequent in spoken language already in the 1970s (Korhonen 1973: 67) and many native speakers do not even recognize such forms anymore (Markus Juutinen, personal fieldwork notes). However, their use is quite common in contemporary written Skolt Saami.⁵ For example, the corpus of about 213,000 word forms of predominantly literary language (SIKOR) contains more than two thousand fourth-person forms, whereas the corpus of 40 hours and 300,000 word forms of spoken Skolt Saami (Giellagas Corpus) contains only about one hundred such forms. Lehtinen (2018: 95–96) reports only 36 fourth-person forms in a corpus of 17 hours of Skolt Saami. The increased use of these forms may be largely explained by direct (translational) or indirect influence from Finnish, the majority language with an overwhelming effect on virtually all writers and translators of Skolt Saami (cf. Juutinen 2023: 97–100).

It is also important to note that while earlier descriptions of the fourth person have not paid attention to its geographical distribution within Skolt Saami, it has been attested in the most vital southern dialects (Suõ’nn’jel and Njuõttjäu’rr) only, but not in the northernmost dialects of the language (see Table 8 in Section 2.3).

3.2. A non-promotional impersonal passive

In addition to being an impersonal passive without overt agents, the fourth person also differs from the best-known Western European passives by being non-promotional. This means that unlike in promotional passives which prototypically promote the patient argument from object to subject position, the Skolt Saami impersonal passive excludes – instead of simply demoting – the nominative-marked agent argument. However, the accusative-marked patient argument remains an accusative object and is thus not promoted to

⁵ A reviewer of this paper has pointed out that fourth-person forms are quite common in contemporary spoken Kildin Saami too.

nominative subject. The main rule is that fourth-person forms are non-promotional impersonals without subjects, and with accusative objects just like in other persons (see also Examples 1, 2, 5 and 7 above):

- (8) *Tõ'st* *mi'jjid* *pi'jješ* *au'tte*.
 that.LOC 1PL.ACC put.PST.4 car.ILL
 ‘There we were put in a car.’ (“There one put us in a car.”) (Če'vetjäu'rr, Kotus 17462_1c: 08:52)
- (9) *Kā'skkvue'žžid* *ko'skkēš* *di* *kā'skk-kue'lid*.
 dry.meat.PL.ACC dry.PST.4 and dry-fish.PL.ACC
 ‘Meat and fish were dried.’ (Če'vetjäu'rr, Kotus 11722_1a: 19:38)

As the only evident syntactic difference to the first, second and third-person clauses is the absence of subject, this is indeed the explanation for the fact that the non-promotional impersonal passive can also be applied to intransitive and copula verbs as seen above in (3–4) and (6). As pointed out by Givón (2001: 127), many formalist grammarians have preferred to limit the concept of passive to promotional passives, but this view is “an unfortunate by-product of nonfunctional, non-typological approaches to syntax” (see also Siewierska & Bakker 2013). Indeed, it appears that it is useful and informative to interpret all of the above examples of the fourth person as instances of non-promotional impersonal passive clauses. Reasons for this will become more evident in the following sections that show examples of recent innovations in which the fourth-person forms are used in a way that must be regarded as non-impersonal and non-passive instead (Section 4.3), as well as in sentences that must rather be described as promotional passive clauses (Section 4.4).

4. Related phenomena and recent developments

In the following, we present further observations on the most recent developments of the fourth-person forms and especially their relationship to third-person plural forms, both as passive and active predicates. The space does not allow us to extend the discussion to derived passive verbs that are clearly in the realm of derivation and not inflection, both as regards their morphology and their syntax. For the most common passive verbs in *-je-* and *-õðvvâ-* (e.g., *kāddjed* ‘be killed’ ← *kā'dded* ‘kill’; *valmštõðvvâd* ‘be prepared’ ← *valmšted* ‘prepare’), see Feist (2015: 119) and especially Lehtinen (2018; 2021).

4.1. Third-person plural as impersonal

It is well known that languages around the globe commonly use third-person plural forms for impersonal predicates, and Uralic is not an exception (e.g., Siewierska 2008: 11ff.; 2010; Siewierska & Papastathi 2011; F. Gulyás 2016; Klumpp & Skribnik 2022). Not unlike all other Saami languages, even Skolt Saami – in spite of the existence of the special fourth-person forms – employs its third-person plural forms (Moshnikoff et al. 2020: 82) for seemingly identical functions (10–12). Occasionally, the two alternatives seem to be in free variation even in a single sentence (11–12):

- (10) *Pue'rmõs* *puõccid* *kue'dde*
 good.SUP reindeer.PL.ACC leave.PST.3PL⁶
ja *hue'nmõõzzid* *ko'dde.*
 and bad.SUP.PL.ACC kill.PST.3PL
 ‘The best reindeer were left alive, and the weakest were killed.’
 (Če'vetjäu'rr, Kotus 631_1a: 03:23)

- (11) *Nä'de* *kue'zzid* *kåčča* *di* *poorât*
 then guest.PL.ACC invite.PRS.3PL and eat.PRS.4
di *mân̄ŋa* *õs* *vuüddje.*
 and then PTC sleep.INCH.PRS.3PL
Nu'bb *peeiv* *siõrât* *di* *nue'r-nue'r*
 second day.GEN play.PRS.4 and string-string.ACC
siõ'rre.
 play.PST.3PL
 ‘And then they invite guests and eat and after that they go to sleep. The other day they play and played the string game.’ (Če'vetjäu'rr, Kotus 11723_1a: 03:19)

- (12) *No* *tie'tteš=han* *kâ'l,* *mutta* *ko*
 well know.PST.4=DPT DPT but as
jiâ *ni* *huõllâm* *te'l.*
 NEG.3PL even care.PST.PTCP then
 ‘Well, one surely knew [that there were doctors], but one just didn't care about them at that time.’ (Če'vetjäu'rr, Kotus 11722_1a: 18:38)

⁶ As the perspective adopted in this article is partly historical-morphological, we approach the fourth-person as well as the third-person plural forms from this perspective, and gloss them consistently as 4 and 3pl according to their most original, canonical functions, even when the former are used as personal third-person forms and the latter as impersonals.

In fact, even in the southern dialects with fourth-person forms, the use of the third-person plural forms has increased at the expense of the special impersonal form. Although the fourth-person forms are commonly used in literary texts that are mostly created – written and often translated – by the most language-conscious educated members of the community, many laymen regard these forms as foreign or entirely unknown, and do not consider themselves as users of such forms (Markus Juutinen, personal fieldwork notes). From this perspective, it is all the more interesting that impersonal use of the third-person plural is nevertheless quite uncommon in the written language.

In light of the fact that the fourth-person forms are traditionally known in the southern dialects of Suõ'nn'jel, Njuõttjäu'rr and Sââ'rvesjäu'rr only, it is quite understandable that in the northernmost dialects of Paččjokk and Peäccam as well as in the extinct dialect of Njauddâm the same functions are usually covered by the third-person plural forms. The only formal difference to ordinary active clauses is the absence of the nominative subject.

- (13) *Pällsiõr* *kuõi'tnalla* *siõ'rre*.
 ball.game.ACC in.two.ways play.PRS.3PL
 ‘The ball game is played in two ways.’ (Paččjokk, T. I. Itkonen 1931: 141)

- (14) *De* *go* *ä'lǧǧe* *siõrrád,* *de*
 then when begin.PRS.3PL play.INF then
kuõi'tpeälla *ä'lǧǧe* *lee'd* *õõut*
 two.GEN.side.ILL must.PRS.3PL be.INF one.GEN
vee'rd *oummu*.
 amount person.PL
 ‘When one begins to play there have to be the same number of people on both sides.’ (Paččjokk, T. I. Itkonen 1931: 139)

It goes without saying that in many contexts it may be impossible to discern impersonal clauses from clauses where the absence of an overt subject could be explained as an instance of pro-drop. As noted by Dryer (2013), in many languages such as Finnish, first- and second-person pronouns may be often absent, whereas third-person pronouns are normally obligatory. Although verbal morphology and the entire syntax of Skolt Saami are quite analogous to those of Finnish, it is interesting to note that while the person and number marking in Skolt Saami finite verbs make it easy to omit the pronoun subject, according to Feist (2010: 252) this is not very common and even less so in the first- and second-person clauses. Somewhat surprisingly, Feist states that

third-person pronouns are omitted more often than other personal pronouns. However, as he does not differentiate between the singular and plural forms and does not focus his attention to impersonal expressions either, it is possible that many of the clauses that he considers instances of pro-drop could have been analyzed as impersonal passives instead.

While this paper does not attempt to reconstruct full verb paradigms of earlier phases of Skolt Saami dialects, it appears quite natural to think that the predecessors of the northern dialects have also possessed specialized fourth-person forms. The absence of these forms can be partly explained with reference to the long-standing contacts between Sea Saami dialects of North Saami and Njauddâm Skolt Saami in particular (Juutinen 2019; 2023: 36–39). Similar circumstances and especially resulting lexical influence from North Saami to Skolt Saami are also known for the Paččjokk area (Juutinen 2023: 39). On the other hand, it is notable that the fourth person is also absent in the westernmost living Eastern Saami language, Aanaar Saami, but then again, much more important evidence is recorded in the extinct Kemi Saami written by Olaus Sirma in the 17th century: Sammallahti (1984: 148; 1998a: 84) considers Sirma’s verb forms such as *pieiatte* (put.PRS.4) ‘is put’ and *sarnäte* (speak.PRS.4) ‘is spoken’ ‘is spoken’ as a sign of the archaic nature of the westernmost Kemi Saami. As these forms can be compared to present-day Skolt Saami *pijjât* and *säärnat* id., such forms most likely have existed in the northernmost dialects as well.

The use of third-person plural forms for impersonal predicates is not only universally common, but it is also remarkable that this is also the strategy employed by Russian, the majority language with the most long-standing influence on the easternmost Saami languages such as Skolt and Kildin Saami, which nevertheless have been able to maintain the fourth person as an independent morphosyntactic means to code impersonal passive sentences. However, as third-person plural forms are used for similar functions also in most, if not all, western Saami languages (Korhonen 1970), it is conceivable that these forms have always have at least some impersonal use in the easternmost Saami languages as well.

4.2. Impersonal as third-person plural

Not only are third-person plural verb forms used in subjectless impersonal clauses, but the fourth-person impersonal forms are used in contexts where they actually are not impersonal predicates. Instead, they seem to have become new third-person plural forms; cf. the use of Finnish *mennään* go.PASS as a new first-person plural form in *me mennään* ‘we go’.

To be sure, in a situation where third-person forms are being used like fourth-person forms, and vice versa, the first logical interpretation would be that the distinction between the two has been lost, and we are witnessing more or less free variation, which will presumably level out sooner or later. However, these developments have taken place in different parts of the language community. As mentioned above, the use of third-person forms in impersonal clauses is most common in the northernmost dialects without attested fourth-person forms. Perhaps not surprisingly, the opposite development is most common in the southernmost dialects.

As mentioned in Section 2, one of Sammallahti's (1998b: 29) main criteria for distinguishing between Akkala Saami and Skolt Saami proper is that in the former, the impersonal forms have entirely displaced the original third-person plural forms, resulting in the loss of the fourth person as an independent category. However, Sammallahti also notes that a partly analogous development can be observed in the southernmost Sââ'rvesjäu'rr dialect of Skolt Saami – the one with closest contacts with Akkala Saami. In Sââ'rvesjäu'rr Skolt Saami, it is possible to come across sentences such as (15) where the fourth-person form *jie'lleš* actually functions as a third-person form that has an overt nominative subject, the personal pronoun *sij* 'they'.

- (15) *Sij* *čâhčča* *to'b* *jie'lleš*.
 3PL at.autumn there live.PST.4
 'They (relatives) lived there in autumn.' (Yona, Kert & Zajkov 1988: 61⁷)

Quite obviously, this kind of use of fourth-person forms is due to the influence from the neighboring Akkala Saami where the fourth-person forms have entirely replaced the original third-person plural forms. More precisely, this may have happened after 1937–1938, when the Sââ'rvesjäu'rr Skolt Saami and a majority of the Akkala Saami were forced to move to Yona kolkhoz located between their old winter villages (Linkola & Sammallahti 1995: 53; Kert & Zajkov 1988: 3–4).

The use of the fourth person as third-person forms in Sââ'rvesjäu'rr Skolt Saami appears quite new also because of the fact that unlike in Akkala Saami, the fourth person has not fully replaced the old third-person plural, which is still used for the negation verb.⁸ In (16), *le'jješ* (be.PST.4) occurs in a context

⁷ Although Kert & Zajkov (1988) have labeled their texts as Akkala and Ter Saami, data about their informants reveals that part of them represent Skolt Saami dialects of Sââ'rvesjäu'rr and even Njuöttjäu'rr.

⁸ On the other hand, *jeät* [NEG.4] is also absent in Korhonen's (1973: 95–97) grammar sketch in spite of its presence in Korhonen's (1967: 91) historical and dialectological description of the Saami conjugation. See also Zajkov (1987: 157–159) for similar variation in Akkala Saami.

clearly atypical for impersonals, namely having an inanimate subject referring to iron ovens.

- (16) *Ru'vddkiuggân* *le'jješ* --
 iron.oven.PL be.PST.4
to'b lie *ru'vddkiuggân.*
 there be.PRS.3PL iron.oven.PL
 'There were iron ovens -- there are iron ovens.' (Yona, Kert & Zajkov 1988: 51)

- (17) *Päärnaid* *di* *ää'kk*
 child.PL.ACC and old.woman.ACC
jie lue'st *ni* *koozz*
 NEG.3PL leave.CNG NPM where.ILL
 'The children and the old woman are not allowed to go anywhere.' (Yona, Kert & Zajkov 1988: 75)

However, our understanding of the history of Sââ'rvesjäu'rr Skolt Saami remains limited, as the dialect was not properly documented before 1960s, and the only published materials consist of the texts published by Zajkov (1987) and Kert & Zaikov (1988).

Fourth-person forms are also attested as third-person forms also among those Skolt Saami speakers of Russia (formerly the Soviet Union) who did not live collectively with Akkala Saami. Another concentration for the Soviet Skolt Saami was in Pâ'jj-Tuállâm, where both the Njuöttjäu'rr and (to lesser extent) Mue'tkk dialects were spoken. Here, the use of the fourth person instead of original third-person forms appears more recent than in the Sââ'rvesjäu'rr dialect, as these innovations are very rare in the oldest records of the Njuöttjäu'rr dialect (e.g., T. I. Itkonen 1931).⁹ Moreover, there is much idiolectal variation among those Skolt Saami who have lived in Pâ'jj-Tuállâm.

Due to the paucity of language material from a handful of informants, it is difficult to present clear patterns of variation among the speakers of the Njuöttjäu'rr dialect. However, it is interesting to note that some speakers use fourth-person forms in third-person functions (with overt subjects) in the past tense only, but the original third-person plural forms have been retained in the

⁹ The only such sentence attested from the old Njuöttjäu'rr dialect is (i):

- (i) *Oummu* *äärat* *to'ben,* *ju'rddes,* *što*
 person.PL be.PRS.4 there think.PST.4 that
jie *māka* *kaaun* *mi'jjid.*
 NEG.3PL as.it.were find.CNG 1PL.ACC
 'People are there and think that [the villains] won't find them.' (T. I. Itkonen 1931: 232)

present tense. For some speakers, the old and new forms occur in free variation, as seen in the following examples from Pâ'jj-Tuållam:

(18)

<i>kue'ss</i>	<i>puä'tte</i> (01:45)	vs.	<i>kue'ss</i>	<i>puõ'tteš</i>	(01:40)
guest.PL	come.PRS.3PL		guest.PL	come.PST.4	
'Guests come.'			'Guests came.' (Kotus 12423_2az)		
<i>sij</i>	<i>ko'lle</i> (02:10)	vs.	<i>sij</i>	<i>ku'lleš</i>	(03:45)
3PL	hear.PRS.3PL		3PL	hear.PST.4	
'They hear.'			'They heard.' (Kotus 12423_1az)		
<i>sij</i>	<i>puä'tte</i> (23:15)	vs.	<i>sij</i>	<i>puõ'tteš</i> – <i>puõ'tte</i>	(23:54–24:00)
3PL	come.PRS.3PL		3PL	come.PST.4	come.PST.3PL
'They come.'			'They came.' (Kotus 12423_1az)		

Another speaker has adopted a somewhat different system, in which the original third-person plural forms are used for the copula and negation verb:

(19) *Ruõšš le'jje occanj.*

Russian be.PST.3PL few

'There were few Russians.' (Pâ'jj-Tuållam, Kotus 16620_1a 11:47)

(20) *A jälsteš sij puârast, sij*
 but live.PST.4 3PL well 3PL

õ'nne leehmid.

keep.PST.3PL cow.PL.ACC

'They were living well, they kept cows.' (Pâ'jj-Tuållam, Kotus 16620_1a 05:18)

Of course, the fourth-person form (*sij*) *jälsteš* 'they lived' above is not an indefinite, impersonal form anymore. Although Sammallahti (1998b: 29) considers this feature as a criterion for distinguishing between Akkala Saami and Skolt Saami, the above examples show that the phenomenon is not foreign to Skolt Saami either.

While the use of third-person plural instead of the impersonal fourth person can be easily understood in light of typological unmarkedness as well as

possible interference from Russian¹⁰, the opposite development calls for explanation. Why would a specialized category of impersonals replace a fully functional third-person plural? As examples above suggest, the Skolt Saami third-person plural forms are not as functional as one might expect. As certain classes of Skolt Saami verbs have homonymous third-person plural forms in the present and past tense, it appears some speakers use third-person forms for lexical verbs without such homonymy (e.g., *â'nne* [keep.PRS.3PL] vs. *õ'nne* [keep.PST.3PL]), but for verbs with ambiguous third-person plurals in common, classical Skolt Saami, fourth-person forms have replaced the old third-person forms and thus restored unambiguous tense-marking (for verb classes in Skolt Saami, see Moshnikoff et al. 2020: 306–340 and Koponen et al. 2022: 208–210). While it is true that *õ'nne* (keep.PST.3PL) in (20) has the unambiguous meaning ‘they kept’, but the corresponding forms for verbs like *jälsted* ‘live’ and members of many inflectional classes are potentially ambiguous: Although most Skolt Saami idioms seem to contain the expected form *jälste* (live.PST.3PL) with which *jälsteš* (live.PST.4) above could be replaced, the former is not only the past tense form, but homonymous with the corresponding present tense form (PRS.3PL). In other words, the use of *jälsteš* can be seen as a less ambiguous and more informative choice, especially when accompanied by the subject *sij* ‘they’. The past tense form *õ'nne* (keep.PST.3PL), in turn, does not need to be replaced by *õ'nneš* (keep.PST.4), as the distinction between the past (*õ'nne*) and present (*â'nne*) third-person plural forms has been retained. According to our preliminary observations, fourth-person forms are used in third-person plural functions predominantly in contexts where use of original forms would result in unwanted ambiguity as regards tense.

4.3. From non-promotional to promotional?

The above sections have described the Skolt Saami fourth person in what seems to be its original nature as regards argument marking: It is a non-promotional impersonal passive, where the agent is not expressed at all, and

¹⁰ For impersonal functions of third-person plural forms in Russian, see, e.g., Siewierska & Papastathi (2011) and the following examples:

- (ii) *Они* *покупают* *хлеб* *в* *магазине.*
 3PL buy.IPFV.3PL bread.ACC in store.PREP
 ‘They buy bread in the grocery store.’ (constructed example)
- (iii) *Хлеб* *покупают* *в* *магазине.*
 bread.ACC buy.IPFV.3PL in store.PREP
 ‘Bread is bought in the grocery store.’ (constructed example)

patient-marking is non-promotional, which in Skolt Saami means that fourth-person verb forms take accusative objects just like the finite verb forms in first, second and third person. However, as seen in the previous section, the fourth person has also acquired non-passive third-person functions among the easternmost Skolt Saami dialects spoken in Russia. Moreover, it is possible to observe another kind of deviation from what has been characterized as non-promotional impersonal passive above. To be specific, in contemporary literary texts written – and most often translated – by bilinguals in Finnish and in speech of Skolt Saami new speakers, the patient argument is occasionally in the unmarked nominative case instead of the accusative. However, it is important to note that the nominative seems to be used only in contexts where the corresponding argument in Finnish would be the nominative (see Juutinen 2023: 97–100):

- (21a) *Veä'kk* *ti'lljet* *pâi*
 help.NOM order.PRS.4 always
jie'ttnââmrest *112.*
 emergency.number.LOC 112
 ‘Help is always available via the emergency number 112.’
 (‘‘Help is ordered always from the emergency number 112.’’) (SIKOR)

- (21b) *Apu* *tilataan* *aina*
 help.NOM order.PRS.PASS always
hätänumerosta *112.*
 emergency.number.ELA 112
 (translation into Finnish)

- (22a) *Vaalin* *vaalšet* *ouddooumaž* *po'dde*
 election.PL.LOC choose.PRS.4 trustee.NOM period.ILL
1.1.2015–31.12.2017.
1.1.2015–31.12.2017
 ‘The [Skolt Saami] Trustee for the period 1.1.2015–31.12.2017 will be elected through an election.’ (YLE 2017a)

- (22b) *Vaaleilla* *valitaan* *luottamusmies*
 election.PL.ADE choose.PRS.PASS trustee.NOM
kaudelle *1.1.2015–31.12.2017.*
 period.ALL 1.1.2015–31.12.2017
 (translation into Finnish)

- (23a) *Če'vetjääu'rest* *riâššât* *veâl*
Če'vetjäu'rr.LOC *organize.PRS.4* *yet*
vaalsaggstõõllmõš *peeiv* *ouddâl*
election.discuss.NMLZ.NOM *day.GEN* *before*
saa'mi *ouddoummuvaalid,* *sue'vet*
Skolt.Saami.PL.GEN *trustee.election.PL.ACC* *Saturday.GEN*
4.10.
4.10.
‘An election debate will be held in Če'vetjäu'rr as late as one day before the Skolt Saami Trustee election, on Saturday the 4th of October.’ (YLE 2017b)

- (23b) *Sevettijärvellä* *järjestetään* *vielä*
Sevettijärvi.ADE *organize.PRS.PASS* *yet*
vaaliväittely *päivää* *ennen*
election.debate.NOM *day.PART* *before*
kolttien *luottamusmiesvaaleja,* *lauantaina*
Skolt.Saami.PL.GEN *trustee.election.PL.PART* *Saturday.ESS*
4.10.
4.10.
(translation into Finnish)

- (24a) *Cistt* *vuei'tet* *u'vdded* *õ'hhte*
prize.NOM *can.PRS.4* *give.INF* *one.ILL*
vuõittja, *le'be* *tõt* *vuei'tet*
winner.ILL *or* *that.NOM* *can.PRS.4*
jue'kked *kuei't* *vuõi'ti* *kõõsk.*
divide.INF *two.GEN* *winner.GEN* *between.*
‘The prize can be awarded to one winner, or it can be divided between two winners.’ (SIKOR)

- (24b) *Palkinto* *voidaan* *antaa* *yhdelle*
prize.NOM *can.PRS.PASS* *give.INF* *one.ALL*
voittajalle *tai* *se* *voidaan*
winner.ALL *or* *it.NOM* *can.PRS.PASS*
jakaa *kahden* *voittajan* *kesken.*
divide.INF *two.GEN* *winner.GEN* *between*
(translation into Finnish)

Although the criteria for choosing between the nominative-, genitive- and partitive-marking for subjects and objects are notoriously complex in Finnish (see, e.g., T. Itkonen 1979), it seems obvious that the nominative-marked patients of ordering (help), choosing (a trustee), organizing (a debate) and awarding (a prize) are in the nominative because the nominative would be used in the corresponding impersonal passive clauses in Finnish. However, while in Finnish the nominative can be considered an expected case for the object in such clauses, in Skolt Saami the expected objects would be the accusatives *vie'kk̃* (help.ACC), *ouddoummu* (trustee.ACC), *vaalsaggstõðllmõžž* (election.discuss.NMLZ.ACC), *ciist* (prize.ACC) and *tõn* (that.ACC) instead of the nominatives *veä'kk̃*, *ouddooumaž*, *vaalsaggstõðllmõž*, *cistt* and *tõt* seen above.

What is most interesting here is that although nominatives such as these can almost always be explained by interference from Finnish, a nominative argument of a Skolt Saami transitive verb is, in principle, always the subject of a clause. Therefore, in the context of the Skolt Saami grammar itself, the nominative arguments seen in (21–24) may be better analyzed as nominative subjects instead of nominative objects. In other words, in sentences like these fourth-person forms no longer function as non-promotional passives but instead, the patient looks like having been promoted to subject position; on the other hand, this analysis would entail that the verb no longer agrees with the subject in the way it has traditionally agreed. While nominative subjects in Skolt Saami impersonals are still rather exceptional deviations from the norm, it seems that the phenomenon is gradually gaining foothold. If continued, this kind of development could lead to a wholesale reanalysis of the alignment system. Alternatively, of course, it is possible to regard nominative arguments like these as objects, but as the nominative is not otherwise a (traditional) case for objects in Skolt Saami, it is somewhat awkward to think that this were the case only in sentences with passive functions and no overt subjects whatsoever. Yet again, if the nominative-marked patient arguments were regarded as objects, the fact that those alter with the more traditional accusative could be considered an instance of differential object marking, the details of which certainly call for further research. One of the future tasks in the field of Skolt Saami syntax is to provide an up-to-date picture of the gradual restructuring of the entire grammar as a result of all-embracing Finnish influence akin to that in the neighboring Aanaar Saami, as already described by Mettovaara (Mettovaara 2023; see also Juutinen 2023: 93–94, 111–112).

At the same time, it must be remembered that while the phenomenon seen here appears to be limited to the Skolt Saami written and spoken in Finland,

the use of (originally) fourth-person forms as third-person plural predicates of active clauses (Section 4.2) is limited to the Skolt Saami spoken in Russia, certainly without similar interference from Finnish. As a consequence, there appear to be no dialects, idiolects or concrete, attested sentences in which fourth-person forms would have both agents and patients marked in the nominative.

5. Conclusion

The preceding sections have aimed to provide a concise description of the so-called fourth-person verb forms in Skolt Saami. They can be functionally defined as an impersonal, non-promotional passive, a set of verb forms that make up the seventh “person form” within Skolt Saami inflection, morphologically and syntactically on a par with the first-, second- and third-person singulars and plurals. Special attention has been directed to the multifaceted relationships of the fourth-person and the third-person plural forms, as data from various dialects and the closest sister languages of Skolt Saami show that impersonal verb forms may acquire functions as personal verb forms, and vice versa, not unlike what has happened in Finnic languages.

Although the fourth person in Skolt Saami is hopefully better known now, observations on its past and present also provide new, interesting perspectives to future studies on the interplay between impersonal and personal verb forms (Sections 4.1 and 4.2). Moreover, it appears obvious that the future of the language itself will be dependent on the interrelations of Skolt Saami between Finnish and Russian. In a sense, these majority languages have been actively pulling the Skolt Saami language in two quite opposite directions for over a century, and the future of Skolt Saami is most likely that of Finnish–Saami bilinguals. From this point of view, it remains to be seen how the complex argument marking in Finnish will affect the originally straightforward argument marking in Skolt Saami, for example (cf. Section 4.3).

All in all, it appears that although Skolt Saami and other easternmost Saami languages are not the key languages in reconstructing the Proto-Saami conjugation, for example, they do offer important perspectives on the study of specialized impersonal verb forms and on contact linguistics in general. Although they lack the dual, the most emblematic inflectional category in Saami, languages like Skolt Saami may have something else interesting to offer: the Finnish “passive” and Estonian “impersonal” have attracted the attention of scholars working on passives or Uralic grammars in general for a long time, but similar – although not identical – phenomena can be found in Saami, too. It is to be hoped that the findings of the present paper prove that

individual Saami languages must be understood and described on their own premises and are in that way able to provide us with important and novel information about phenomena that are foreign to even their better-known sister languages.

Abbreviations

4—fourth person; ALL—allative; CNG—connegative; DPT—discourse particle; ESS—essive; ILL—illative; NPM—negative polarity marker; PART—partitive; POT—potential; SUP—superlative.

Data sources

Giellagas Corpus = The Giellagas Corpus of Spoken Saami languages. [Skolt Saami corpus under development.] The Saami Culture Archive, Giellagas Institute for Saami Studies, University of Oulu.

Itkonen T. I. 1931. *Koltan- ja kuolanlappalaisia satuja. I–II. Kolttalaisia ja kildiniläisiä satuja.* (Mémoires de la Société Finno-Ougrienne LX). Helsinki: Suomalais-Ugrilainen Seura.

Kert, Georgij M. & Zajkov Pekka M. 1988. *Obrazcy saamskoj reči.* Petrozavodsk: Karel'skij filial AN SSSR.

Kolttsaamelaiset 2013 = <https://www.kolttsaamelaiset.fi/se/saa'mpoort-riddu-puo'tte-jiannai-oummu-aalda-da-ku'k'ken-kie'zzloopptem-praa'znka-pa'rggmannust-2013/>

SIKOR. UiT The Arctic University of Norway and the Norwegian Saami Parliament's Saami text collection, Version 06.11.2018, URL: <http://gtweb.uit.no/korp/>

YLE 2017a = https://yle.fi/uutiset/osasto/sapmi/saami_ouddoummuvaalneattel_algg_passpeeiv_jionstet/7497221

YLE 2017b = https://yle.fi/uutiset/osasto/sapmi/cevetjaarest_vaalsastoollmos_veal_ouddal_ouddoummuvaalid/7505512

References

Bartens, Hans-Hermann. 1980. *Die Verwendung von Potential und Konditional im Lappischen* (Mémoires de la Société Finno-Ougrienne 177). Helsinki: Suomalais-Ugrilainen Seura.

Bergsland, Knut. 1984. *Eldre samiske tekster: i utvalg til forelesninger.* Tromsø: Universitetet i Tromsø, Institutt for språk og litteratur.

Dryer, Matthew S. 2013. Expression of pronominal subjects. In: Matthew S. Dryer & Martin Haspelmath (eds), *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (<http://wals.info/chapter/101>) (Accessed 2020-04-04.)

Feist, Timothy. 2010. *A Grammar of Skolt Saami.* Manchester: University of Manchester, Faculty of Humanities. (Doctoral dissertation.)

Feist, Timothy. 2015. *A Grammar of Skolt Saami* (Mémoires de la Société Finno-Ougrienne 273). Helsinki: Suomalais-Ugrilainen Seura.

F. Gulyás, Nikolett. 2016. *Személytelen szerkezetek finnugor nyelvekben. Funkcionális és tipológiai megközelítés.* Budapest: Eötvös Loránd Tudományegyetem. (Doctoral dissertation.) (DOI: 10.15476/ELTE.2016.165)

- Friis, Jens Andreas. 1867. *Russisk-lappiske Sprogprøver og Optegnelser om de russisk-lappiske Dialekters Grammatik*.
(<https://www.nb.no/items/3b83f5cc3d5e6bf91fc0d9dc7b4aab7e>) (Accessed 2024-06-25.)
- Givón, Talmy. 2001. *Syntax. An introduction*, vol. 2. Amsterdam: Benjamins.
- Harms, Robert Thomas. 1962. *Estonian grammar* (Indiana University publications, Uralic and Altaic series, vol. 12). Bloomington: Indiana University.
- Helasvuo, Marja-Liisa. 2006. Passive – personal or impersonal? A Finnish perspective. In Marja-Liisa Helasvuo & Lyle Campbell (eds), *Grammar from the Human Perspective: Case, space and person in Finnish*, 233–255. (Amsterdam studies in the theory and history of linguistic science, Series 4, Current issues in linguistic theory). Amsterdam: John Benjamins.
- Itkonen, Erkki. 1957. Suomalais-ugrilaisen kantakielen äänne- ja muotorakenteesta. *Virittäjä* 61(1). 1–23.
- Itkonen, Terho. 1979. Subject and object marking in Finnish: An inverted ergative system and an “ideal” ergative sub-system, 79–102. In Frans Plank (ed.), *Ergativity: Towards a Theory of Grammatical Relations*. London: Academic Press.
- Itkonen, T. I. 1942. Karjalaiset ja Kuolan-Lappi. *Kalevalaseuran vuosikirja* 22. 40–55.
- Juutinen, Markus. 2019. Koltansaamen Näätämon murteen ja merisaamen välinen koodien sekoittuminen vuonna 1920. *Puhe ja kieli* 39(2). 165–185.
- Juutinen, Markus. 2023. *Koltansaamen kielikontaktit: vähemmistökieli muuttuvassa kieliympäristössä*. Oulu: Oulun yliopisto. (Doctoral dissertation.)
<https://urn.fi/URN:ISBN:9789526238357>
- Kert, Georgij M. 1971. *Saamskij jazyk (kil'dinskij dialekt). Fonetika. Morfologija. Sintaksis*. Leningrad: Nauka.
- Kettunen, Lauri. 1943. *Vepsän murteiden lauseopillinen tutkimus* (Mémoires de la Société Finno-Ougrienne 86). Helsinki: Suomalais-Ugrilainen Seura.
- Klump, Gerson & Elena Skribnik. 2022. Information structuring. In Marianne Bakró-Nagy, Johanna Laakso & Elena Skribnik (eds), *The Oxford guide to the Uralic languages*, 1018–1036. Oxford: Oxford University Press.
- Koponen, Eino, Miestamo, Matti & Juutinen, Markus. 2022. Skolt Saami. In Marianne Bakró-Nagy, Johanna Laakso & Elena Skribnik (eds), *The Oxford guide to the Uralic languages*, 196–218. Oxford: Oxford University Press.
- Korhonen, Mikko. 1967. *Die Konjugation im Lappischen: Morphologisch-historische Untersuchung I. Die finiten Formkategorien* (Mémoires de la Société Finno-Ougrienne 143). Helsinki: Suomalais-Ugrilainen Seura.
- Korhonen, Mikko. 1970. Der Ausdruck des indefiniten Subjekts im aktivischen Satz des Lappischen. In Heikki Leskinen (ed.), *Symposium über Syntax der uralischen Sprachen 15.–18. Juli 1969 in Reinhausen bei Göttingen*, 142–149. Göttingen: Vandenhoeck & Ruprecht.
- Korhonen, Mikko. 1973. Kieliopin pääkohdat. In Mikko Korhonen, Jouni Moshnikoff & Pekka Sammallahhti. *Koltansaamen opas*, 9–97. *Castrenianumin toimitteita* 4.
- Korhonen, Mikko. 1977. Karjalaisten ja saamelaisten yhteydet kielentutkimuksen valossa. In Veijo Saloheimo (ed.), *Karjalan synty. Symposio 30.6.–2.7.1976 Joensuu: alustukset*, 75–82. Monistesarja A 24a. Joensuu: Joensuun korkeakoulu.

- Laanest, Arvo. 1982. *Einführung in die ostseefinnischen Sprachen*. Hamburg: Helmut Buske.
- Lehtinen, Miika. 2018. *Sää'mkiöl ve'rbbjäädtõõgg -jed da -õõvvâd*. Oulu: Oulu universite'tt. (Master's thesis.)
- Lehtinen, Miika. 2021. Skolt Saami passive verbs. *Finnisch-Ugrische Mitteilungen* 45. 15–52.
- Linkola, Martti & Pekka Sammallahti. 1995. Koltanmaa, osa Saamenmaata. In Tuija Saarinen & Seppo Suhonen (eds), *Koltat, karjalaiset ja setukaiset – Pienet kansat maailmojen rajoilla*, 39–57. Snellman-instituutin A-sarja 19. Kuopio: Snellman-instituutti.
- Mettovaara, Jukka. 2023. Syntax in transition: emergence of differential argument marking in Aanaar Saami. *Finnish Journal of Linguistics* 36. 36. 93–129.
- Miestamo, Matti & Eino Koponen. 2015. Negation in Skolt Saami. In Matti Miestamo, Anne Tamm & Beáta Wagner-Nagy (eds), *Negation in Uralic languages*, 353–375. Amsterdam: Benjamins.
- Moshnikoff, Satu, Jouni Moshnikoff, Eino Koponen & Miika Lehtinen. 2020. *Koltansaamen kieliooppi. Sää'mkiöl kiöllvue'ppes*. Aanaar: Sää'mte'gğ.
- Nirvi, R. E. 1947. *Passiivimuotojen aktiivistumisesta* (Suomi 104). Helsinki: Suomalaisen Kirjallisuuden Seura.
- Rantanen, Timo, Harri Tolvanen, Meeli Roose, Jussi Ylikoski & Outi Vesakoski. 2022. Best practices for spatial language data harmonization, sharing and map creation: A case study of Uralic. *PLoS ONE* 17(6). e0269648. <https://doi.org/10.1371/journal.pone.0269648>
- Rießler, Michael. 2022. Kildin Saami. In Marianne Bakró-Nagy, Johanna Laakso & Elena Skribnik (eds), *The Oxford guide to the Uralic languages*, 219–239. Oxford: Oxford University Press.
- Rydving, Håkan. 2013. *Words and varieties: Lexical variation in Saami* (Mémoires de la Société Finno-Ougrienne 269). Helsinki: Finno-Ugrian Society.
- Sammallahti, Pekka. 1984. Saamelaiden esihistoriallinen tausta kielitieteen valossa. In Jarl Gallén (ed.) *Suomen väestön esihistorialliset juuret* (Bidrag till kännedom av Finlands natur och folk 131), 137–156. Helsinki: Societas scientiarum Fennica.
- Sammallahti, Pekka. 1998a. Olaus Sirma: Kulnasatz niråsam. In Vesa Guttorm (ed.), *Giellačållosat* III, 84–95 (*Diedut* 1998(1)). Guovdageaidnu: Sámi Instituhtta.
- Sammallahti, Pekka. 1998b. *The Saami languages: An introduction*. Karasjok: Davvi Girji
- Schlachter, Wolfgang. 1970. Korreferat B [Comment to Korhonen [1970]]. In Heikki Leskinen (ed.), *Symposium über Syntax der uralischen Sprachen 15.–18. Juli 1969 in Reinhausen bei Göttingen*, 151–153. Göttingen: Vandenhoeck & Ruprecht.
- Shore, Susanna. 1988. On the so-called Finnish passive. *Word* 39. 151–176.
- Siewierska, Anna. 2008. Introduction: Impersonalization from a subject-centred vs. Agent-centred perspective. *Transactions of the Philological Society* 106(2). 1–23.
- Siewierska, Anna. 2010. From 3PL to passive: Incipient, emergent and established passives. *Diachronica* 27(1). 73–109.

- Siewierska, Anna & Dik Bakker. 2013. Passive agents: prototypical vs. canonical passives. In Dunstan Brown, Maria Chumakina & Greville G. Corbett (eds), *Canonical morphology and syntax*, 151–189. Oxford: Oxford University Press.
- Siewierska, Anna & Maria Papastathi. 2011. Towards a typology of third person plural impersonals. *Linguistics* 49(3).
- Tillinger, Gábor. 2014. *Samiska ord för ord: att mäta lexikalt avstånd mellan språk* (Studia Uralica Upsaliensia 39). Uppsala: Acta Universitatis Upsaliensis.
- Tunkelo, E. A. 1934. Passiivin potentsiaali. *Virittäjä* 38. 89–92.
- Valtonen, Taarna, Jussi Ylikoski & Luobbal Sámmol Sámmol Ánte (Ante Aikio). 2022. Aanaar (Inari) Saami. In Marianne Bakró-Nagy, Johanna Laakso & Elena Skribnik (eds), *The Oxford guide to the Uralic languages*, 178–195. Oxford: Oxford University Press.
- Zaikov, Pekka M. 1996. Sprachliche Besonderheiten des Akkalalappischen. In Lars-Gunnar Larsson (ed.), *Laponica et Uralica: 100 Jahre finnisch-ugrischer Unterricht an der Universität Uppsala: Vorträge am Jubiläumssymposium 20–23. April 1994* (Studia Uralica Upsaliensia 26), 137–142. Uppsala: Uppsala universitet.
- Zajkov, Pekka M. 1987. *Babinskij dialekt saamskogo jazyka: fonologo-morfologičeskoe issledovanie*. Petrozavodsk: Karelija.

Verb inflection in Forest Lule Saami

Olle Kejonen, Uppsala University

1. Introduction

Lule Saami is a Uralic language of the Saamic branch, likely spoken by fewer than one thousand people belonging to the indigenous Saami minority of Norrbotten County, Sweden, and Nordland County, Norway. The situation of the Lule Saami language is in many ways paradoxical. On the one hand, it is an officially recognized language in the heart of Northern Europe, with an established literary standard and a relatively strong legal status. On the other hand, it is severely endangered, and there has been little description of spoken Lule Saami published since the mid-1900s. Even recent descriptions of the language largely build on the endeavors of scholars such as K. B. Wiklund (1868–1934), Harald Grundström (1885–1960) and Björn Collinder (1894–1983). To name but one example, the pedagogical grammar of Spiik (1989, first edition 1977) relies heavily on the textbook of Wiklund (1915, first edition 1901), which in turn derives from linguistic documentation carried out in the late 1800s.

While acknowledging the accomplishments of previous scholars, it is also important to describe the Lule Saami language as it is spoken today. The goal of this paper is thus twofold: first, to give a descriptive account of contemporary Lule Saami verb inflection, and second, to present original data on the critically endangered Lule Saami Forest dialect. Although the phonetics of this variety are the best described of any Lule Saami dialect—on account of Collinder’s 1938 monograph on the topic—its morphology and syntax have gone virtually undescribed to this day. This study takes a first step towards filling this knowledge gap.

The paper has the following structure. In Section 2, the background and data of the study are presented. Section 3 outlines the research history of the Forest dialect and Section 4 provides dialectological notes on the variety. In Section 5, which constitutes the main body of the paper, the inflectional verb morphology of the Forest dialect is described. Particularly noteworthy in this

section are the merged forms of the negative auxiliary and the verb ‘be’ presented in 5.10, some of which have not been previously described. Section 6 features a discussion on inflectional forms that are known from previous descriptions of Lule Saami, but that are either absent or only poorly attested in the present corpus. Most notably, these forms include verbs in the dual number and the potential mood. Section 7 contains a brief summary and discussion of the main findings of the study.

2. Background and data

This section outlines the background and data of the study, starting with a delimitation of the research area (2.1). This delimitation is followed by a description of the corpus on which the paper is based (2.2) and the transcription conventions used (2.3).

2.1 Area and speakers

The Lule Saami Forest dialect is traditionally spoken in the reindeer herding district of Jiellevárre (Sw. *Gällivare skogssameby*) in Norrbotten County, Sweden. This district ranges from Jiellevárre municipality in the northwest to the Baltic Sea in the southeast (see Map 1). It is also customary to include the varieties traditionally spoken in the districts of Sláhkka (Sw. *Slakka*, previously *Sierre*, Sw. *Serri*) and Udtjá (Sw. *Udtja*) in Jáhkâmáhkke (Sw. *Jokkmokk*) municipality under the heading of the Forest dialect (see Sammallahti 1998: 21–22 for a brief classification of Lule Saami dialects).¹ Data from the latter districts are however scarce, and according to Lule Saami scholar Susanna Angéus Kuoljok (2002a: 10, 2003: 9), there were presumably no speakers of the Forest dialect left in Sláhkka or Udtjá already at the beginning of this century (cf. Hedström 2011: 153). Consequently, the present account of the Forest dialect focuses solely on the variety spoken in the Jiellevárre district.

Unlike more southern varieties of the language, this dialect has been in close contact with the Finnish variety now known as Meänkieli. It has also been in contact with local Swedish dialects, including the varieties of Råne and Överkalix, in addition to Standard Swedish (see, e.g., Pihl 1924: 3, Collinder 1926, Källskog 1989, Winsa 1991: 32–38). As a result of forced

¹ For older classifications, see Wiklund (1890: III; 1891: I–II; 1915: 4), Qvigstad (1925: 19) and Grundström (1937: III–VII); cf. Rydving (2013: 42–43). See also the discussion on dialect levelling in Angéus Kuoljok (2002a: 10; 2003: 8).

relocations of reindeer herders during the early 1900s, there are also a few speakers of North Saami in the area (see Labba 2020: 145–153).

In 1930, Collinder stated that the Forest dialect was spoken by “half a thousand settled individuals in some twenty villages” in southeastern Jiellevárre (1930: 7). A few years later (1938: 5), he was more cautious in his estimation of the speaker population, stating that only a fourth of the 800 inhabitants of the parish spoke Saami. Today, this number has diminished significantly—I am only aware of a handful of active speakers of the dialect, all of whom are above 80 years of age and belong to two related families from the same village.

2.2 Data of the study

This paper is based on data gathered by the author through the small-scale project *Documenting the Lule Saami Forest dialect*, conducted under the auspices of the Uppsala Language Documentation Group (UPPLADOC) at Uppsala University, Sweden. The primary goals of the project, lasting from 2019 to 2020, were to assess the current situation of the presumably moribund Lule Saami Forest dialect and, if possible, to record spontaneous speech with one or several speakers of the variety. Secondary goals included elicitation of dialectal features and place names. Both the primary and secondary goals were met beyond expectation: through the project, some nine hours of spontaneous and elicited speech were recorded with two speakers of the dialect.

Data from the documentation project have also been included in a study on the development of some kinship terms in North Saami, Lule Saami and South Saami (Kejonen 2020a). An additional outcome of the project is the publication of two popular scientific papers presenting older folklore materials from the area in Lule Saami and Swedish respectively (Kejonen 2021a, second edition 2024; 2021b), as well as a brief description of Collinder’s work with the Forest dialect, also aimed at a general audience (Kejonen 2022).² A Meänkieli borrowing from the Forest dialect is treated in Kejonen (2024a: 138–139).

The two speakers interviewed for the project are Sören Andersson (SA), born in 1930, and Lars Eriksson (LE), born in 1944.³ Both speakers are from Sádinvárre (lit. ‘trap mountain’, Swedish *Flakaberg*, *Råne Flakaberg* or

² Preliminary findings of the project were also presented at the *Workshop on language documentation: multilingual settings and technological advances* (Uppsala, 24–25 October 2019) and at the *4th Saami Linguistics Symposium, SAALS 4* (Uppsala, 14–15 November 2019).

³ Here it might be of interest to mention that the speakers interviewed for the present study are closely related to the two speakers who provided most of Collinder’s data (see 3.3), one of them being Andersson’s father and the other being his grandfather’s brother. Andersson, in turn, is the half-uncle (SA *tjiehtjebielle*) of Eriksson.

Flakabergsåsen), also known as Vårre (lit. ‘mountain’), a small settlement close to the Arctic Circle in Boden municipality, Norrbotten County, Sweden.⁴ The majority language Swedish dominates the day-to-day life of both speakers, although they both value their Lule Saami mother tongue and speak it when they have the opportunity, for instance with their respective siblings.⁵



Map 1. Location of Sádinvårre and the Jiellevårre reindeer herding district (shaded) in Norrbotten County, Sweden. Map: Olle Kejonen, based on data from Google and Sámedigge (the Saami Parliament of Sweden).

⁴ Collinder (1938: 48) has noted the name Sádinvårre both with a second syllable *ε* and *a*: *sātemvårre* ~ *sātamvårre*, cf. SA *sádam*, SG.ACC *sádamav* ‘trap (for birds)’ (Sw. *flake*). Also noted by Collinder is the variant *vårre* (ULMA 15925: 5, cf. Kejonen 2021b: 35). Note that this settlement should not be confused with the nearby locality Sádek (Sw. *Flakaberg*, *Gällivare Flakaberg* or *Lill-Flakaberg*), as is done by Rydving (2013: 42). The town of Boden is referred to in Lule Saami by both speakers as *Båvdda* (see Section 4.1 on the lack of labial vowel harmony in this name), by SA also as *Suddás* (originally the oblique stem of *Suttas*, by which name the town is also known).

⁵ It is my impression that today, the Forest dialect is highly regarded by speakers of closely related varieties: on several occasions, I have heard speakers of both North Saami and Lule Saami making favorable comments about the dialect. This stands in stark contrast to the situation of the early 1900s, when Collinder noted that most Forest Saami regarded the neighboring mountain varieties of Lule Saami as a kind of *Hochsprache* (1938: 7), which was “better” or “more correct” than the Forest dialect (1930: 12–13).



Figure 1. Official Swedish and unofficial Lule Saami road sign, Sádinvárre, August 1, 2019. Note that the Lule Saami sign does not adhere to the standard orthography. Photo: Olle Kejonen.



Figure 2. Lars Eriksson, August 1, 2019. Photo: Olle Kejonen.



Figure 3. Sören Andersson, January 4, 2020. Photo: Olle Kejonen.

The interview with Eriksson was done at his home in Sádinvárre on August 1, 2019 and consists of two hours and forty minutes of audio and video. A first interview with Andersson was carried out at his home in Boden on January 4, 2020 and amounts to four hours and seven minutes of audio and video. A second interview with Andersson, scheduled for early 2020, had to be postponed due to the outbreak of the coronavirus (COVID-19) pandemic the same spring. Finally, on August 12, 2021, this interview could be carried out, amounting to two hours and twenty minutes of audio and video; note that this second interview with Andersson was carried out after the formal conclusion of the documentation project.⁶

The bulk of these over nine hours of recordings consists of semi-structured interviews and conversations on everyday topics such as family, weather and reindeer herding. Other topics range from traditional Saami foods to celestial names, e.g., SA *Tjuollda* ‘Pole star’ (lit. ‘stake’) and *Stálo bädnaga* ‘Orion’s belt’ (lit. ‘dogs of the *stállo*’, a mythological ogre). The interviews with Andersson also include two different recordings of a folktale about the origin of the hazel grouse. In addition to spontaneous speech, a sizeable share of the recordings consists of elicitation, focusing primarily on dialectologically interesting features of the variety. The interview with Eriksson also contains

⁶ Although the corpus strictly speaking consists of two idiolects, I take the language of these speakers to be representative of the present-day Forest dialect—not least due to the fact that they represent both families in which there are active speakers of the variety.

some elicitation on local place names, whereas the second interview with Andersson includes elicitation on personal pronouns. Additionally, a number of verb paradigms were elicited during the first interview with Andersson, making up the backbone of the present paper.

With the exception of elicitation settings, all contact with the interviewed speakers was carried out in Saami, with myself speaking Čohkkiras (*Jukkasjärvi*) North Saami, a variety closely related to Lule Saami. In elicitation, Swedish was used as the primary meta language. The first two interviews were recorded using a HD video camera and a wireless set with a lapel microphone; audio recordings were also made using a separate audio recorder. The third interview was recorded with a HD video camera and a wired lapel microphone. All interviews were also recorded with a mobile phone for backup purposes. It should be pointed out that the entire corpus has not been fully transcribed, nor archived, at the time of writing.

In addition to data on the Forest dialect, the present paper also contains some notes on the Lule Saami variety of the neighboring Unna Tjerusj district (Sw. previously *Sörkaitum*). Unless stated otherwise, data from this dialect stem from a small corpus (two hours from two male speakers, born in 1929 and 1933), gathered by the author in 2016 through the Uppsala University project *Linguistic Continua: a View from the North*.⁷ Comparative data on Čohkkiras North Saami stem from an ongoing documentation and description project on this variety (see Kejonen 2020b).

2.3 Transcription conventions

The gathered data are presented in a slightly adapted version of the Lule Saami standard orthography. While a full description of the correspondences between the Lule Saami orthography and phonology is beyond the scope of this paper, the following writing conventions can be noted specifically. The vowels /a/, o, eæ/ are represented by ⟨á, â, ä⟩ and the consonants /ŋ, ɲ, ʃ, tʃ, ts/ are represented by ⟨ŋ, ɲ, sj, tj, ts⟩ (geminated ⟨ŋŋ, ɲɲ, ssj, ttj, tts⟩).⁸ The voiced affricates /dʒ, dz/ are represented by the trigraphs ⟨dtj, dts⟩. For a concise description of Lule Saami phonology, see Wickman (1975); cf. Angéus Kuoljok (1997: 40–54) and Ylikoski (2022: 132–135).

The following adaptations to the standard orthography are made in this paper: ⟨ä̃⟩ is used for the long counterpart of /o/ and ⟨ë̃⟩ is used for short /e/ [ɛ] (in the standard orthography not distinguished from short ⟨â⟩ and long ⟨e⟩

⁷ Through this project, a total of 14 speakers of Lule Saami and North Saami were interviewed in Sweden and Norway, resulting in more than twelve hours of audio recordings.

⁸ In Lule Saami materials produced in Norway the grapheme ⟨æ̃⟩ is used instead of ⟨ä̃⟩.

respectively).⁹ Following the tradition in Saami linguistics, as well as pedagogical materials on Lule Saami (e.g., Spiik 1989), long geminates are indicated by the sign <ː> between two identical consonants, e.g., LE *bies'se* /s:s/ ‘birch bark’ (cf. *biesse* /ss/ ‘nest’). Additionally, minor adaptations of the orthography have been made in order to better reflect the pronunciation of the Forest dialect (see Section 4).

In Lule Saami, stress is fixed: the first syllable carries main stress and each following non-final odd syllable carries secondary stress. There are however a few exceptions to this rule—mainly in compounds and loanwords—so for the sake of clarity, I occasionally use the IPA signs <ˈ> and <ˌ> to indicate primary and secondary stress respectively, e.g., LE *'Gårkkåbies, áhpe* (place name, lit. ‘ant hill bog’, cf. *gårkkåbiesse* ‘ant hill’, *áhpe* ‘large bog’).

Where necessary, a broad phonetic transcription using IPA is included within square brackets. The speaker who has provided a certain form is indicated with his respective initials (SA, LE) where relevant, and elicited clauses are indicated as such. Materials quoted by other scholars are given in their original form, most commonly the Uralic Phonetic Alphabet (UPA, also known as the Finno-Ugric transcription, FUT).¹⁰ Glosses and quotes in other languages have been translated into English. For better comparability, hyphens and equals signs (indicating morpheme and clitic boundaries respectively) have sometimes been added also when quoting other scholars.

3. Previous research

The present description of the Forest dialect is concerned with the contemporary spoken language. Nevertheless, it could also be described as “philologically informed”, in the sense that earlier materials on the variety—both published and unpublished—provide an important reference point throughout the study. In this section, these legacy materials are presented, starting with the work of linguist K. B. Wiklund in the late 1800s and early

⁹ While Wickman (1975) does not include short /e/ in his phonemic analysis of Lule Saami, this vowel can be argued to have phonemic status at least in the present-day Forest dialect on the basis of minimal pairs such as SA, LE *dě* ‘then’ vs. *da* DEM.PL.NOM and SA *guoktĕ* ‘two’ vs. *guokti* ‘twice’. Note however the variation of second and third syllable *ě* and *a* documented by Collinder (1938: 48) for different subdialects of the Forest variety, as well as the variation in vocalism described for some verb suffixes in Section 5.3.

¹⁰ Wiklund’s earliest works on Lule Saami (1888, 1890, 1891) utilize a version of the Swedish Dialect Alphabet (Sw. *landsmålsalfabetet*), many graphemes of which are not found in the Unicode standard. When quoting these materials, I use the most similar characters available in Unicode, e.g., *d* with a curl <ḑ> for Wiklund’s “mouilliertes *d*” (1890: IV; 1891: 9). The inclusion of characters from the Swedish Dialect Alphabet in Unicode would be highly desirable for Saami linguistics, seeing that not only Wiklund, but also other scholars have used it in transcribing Saamic languages (see, e.g., Siegl 2017: 284, Note 47).

1900s (3.1), followed by the folklore texts noted down by clergyman Otto Lindgren in the 1910s (3.2) and the extensive materials of linguist Björn Collinder, primarily gathered in the 1920s (3.3). Finally, some notes are made on other materials on the dialect (3.4).

3.1 K. B. Wiklund

The first materials on the Forest dialect were gathered by K. B. Wiklund (1868–1934) in 1895 and 1906. During these trips, Wiklund noted down a couple hundred words from a handful of speakers, using a detailed phonetic transcription (cf. Collinder 1938: 7). Wiklund also documented a number of place names in the area. While some of Wiklund’s materials on the Forest dialect have been included in the Lule Saami dictionary of Grundström (1946–54), his notebooks from the area (KBW 1895 I–II, 1906) also contain unpublished data, primarily from a speaker born in Bålggem (Sw. *Pålkem*, previously *Vitträsk*). When quoting Wiklund’s materials on the Forest dialect, all data stem from this speaker unless explicitly stated otherwise.¹¹

3.2 Otto Lindgren

Clergyman Otto Lindgren (1884–1952) noted down a number of folklore texts in Stuur Sájjvva (Sw. *Storsaivis*) in 1912, amounting to more than 1,300 words (OL 1912; see also Kejonen 2021a). He also noted down a story in the Forest dialect in Suobbat (OL 1913), consisting of some 200 words. Lindgren knew Lule Saami well (Korhonen 2016: 777, Stoor 2016: 727), and although his texts use a practical orthography that does not take into account some phonological oppositions, his materials are of a high quality. Lindgren’s texts can be said to complement the materials of Wiklund and Collinder, as the latter two scholars almost exclusively documented single words, rather than connected speech.¹²

¹¹ It should however be noted that according to Collinder (1938: 7), the speech of this individual is not entirely representative of the Forest dialect; see also Note 22.

¹² Note that Lindgren’s materials from Stuur Sájjvva include a total of nine texts—not eight, as erroneously stated in Kejonen (2021a: 24). Parts of Lindgren’s folklore texts from the Forest dialect area have also been published in Swedish (Lindgren 1928, 1930, Kejonen 2021b). On behalf of Wiklund, Lindgren also collected linguistic questionnaires from two speakers of Lule Saami from Svenskudden (*Junkär* in his Lule Saami transcription), south of the present research area (cf. Kejonen 2021b: 35). These speakers agree in some respects with the Forest dialect (such as the vocalism in *pesti* ‘spoon’, see Section 4.1), but they also show some peculiar traits, including cases of labial umlaut, e.g., *kånokit* ‘to stop’ (OL 1923), cf. SA *gēnugit*, Standard Lule Saami *ganugit*. In this respect, this variety seems to agree with that of Gåjkkul (Sw. *Kåikul*) as described by Wiklund (1891: 48; cf. Collinder 1938: 66–69 and Wickman 1965: 424–425, see also Note 20). Lindgren further acquired a number of cultural items in the Forest dialect area, which are deposited at the museum of Västernorrland (*Västernorrlands museum*).

3.3 Björn Collinder

Most of the materials available on the Forest dialect were gathered by Wiklund's successor Björn Collinder (1894–1983), who carried out extensive linguistic documentation in the area in 1923, supplemented by shorter trips in 1925 and 1938 (Collinder 1938: 7). Collinder's work was focused on the sound system of the dialect, and resulted in the 1938 monograph *Lautlehre des Waldlappischen Dialektes von Gällivare*.¹³ This monograph is primarily based on data from two speakers (or “language masters”, Ger. *Sprachmeister*, Sw. *språkmästare*) from Sádinvárre: Erik Persson (1854–1949) and Paulus Andersson (1882–1958). In total, however, Collinder gathered materials from more than two dozen speakers from some fifteen villages in the area, all in detailed phonetic transcription. In this survey work, Collinder used a linguistic questionnaire originally constructed by Wiklund, consisting of more than sixty questions mostly concerning phonology.¹⁴

In addition to his 1938 monograph, parts of Collinder's data on the Forest dialect have also been published in the abovementioned dictionary of Grundström (1946–54). Sporadic notes on the variety are further found in other works by Collinder, including a few sociolinguistically oriented texts (1926, 1929, 1930) and an introduction to general linguistics (1959, with several later editions and a 1978 German translation).

Notebooks from Collinder's work on the Forest dialect are archived at the Swedish Institute for Language and Folklore (*Institutet för språk och folkminnen*) in Uppsala (ULMA 15924–36). These notebooks contain some unpublished data, including a small number of partial verb paradigms. Collinder also gathered place names in the Forest dialect area, which are available at the place name register of the Institute for Language and Folklore (*Ortnamnsregistret*). For a popular scientific description of Collinder's work with the Forest dialect, see Kejonen (2022).¹⁵

¹³ Collinder's monograph further contains a short text sample (1938: 170–172). This text, a folktale on the origin of the woodpecker, was however not originally composed in the Forest dialect, but is based on a recitation of a text from Wiklund's anonymously published Lule Saami text collection *Tálotj suptsasah ja áttá* ‘Old tales and new’ (1916: 3–4). While Collinder duly notes this fact, Wiklund, in turn, makes no mention of the original source of the text. It can however be noted that it corresponds almost word for word to a Pite Saami version of the same story, published by Ignác Halász (1893: 7–8) some twenty years earlier, as pointed out in Kejonen (2021a: 27, Note 4); cf. Kejonen (2024b).

¹⁴ While certainly suitable for its intended purpose, the historical-comparative focus of this questionnaire is apparent—for instance, there are no less than four questions concerning the second person dual of the imperative mood (investigating the *Überdehnung*, or “extra strong grade” of consonants, cf. Collinder 1938: 149), but not a single question concerning dual forms in the indicative.

¹⁵ See also Korhonen (2008: 11) for a recounting of an episode where Collinder lost his way skiing to Sádinvárre, only to be found by the speaker he was to interview.

3.4 Other materials

In addition to written materials, there are also audio tapes of the Forest dialect at the Institute for Language and Folklore, including recordings made by linguists Israel Ruong (1903–1986) and Olavi Korhonen (1938–).¹⁶ The Saami sections of the Swedish national broadcasting corporations (*Sameradion* and *SVT Sápmi*) have made a few interviews with speakers of the Forest dialect, including SA and LE, at least some of which are deposited at the archive of the Swedish national radio (*Sveriges Radio*).¹⁷ The abovementioned *Ortnamnsregistret* also contains data on local Saami place names, including materials gathered by scholar Edvin Brännström (1902–1979) and by Fahle Alrik Isberg (1909–1977).¹⁸

4. Dialectological notes

In terms of Saami dialectology, the Forest variety has been described as a “typical central dialect” of Lule Saami (Collinder 1938: 3), closely related to the dialects of the neighboring mountain Saami in southern Jiellevárre and northern Jáhkâmáhkke, while also agreeing in some respects with more southern varieties (Collinder 1926: 126, cf. 1953: 64).¹⁹ Considering the dearth of modern data available on the Forest dialect, this section describes some of the most salient features of the variety, focusing on traits that differ—or could be expected to differ—from the Lule Saami literary standard language as regards vocalism (4.1), consonantism (4.2), morphology (4.3), morphosyntax (4.4) and lexicon (4.5).

4.1 Vocalism

The most distinctive feature of the Forest dialect is undoubtedly its *ě*-inducing umlaut, a phonological process causing the counterpart of a Standard Lule

¹⁶ I have not yet had the opportunity to make any detailed studies of these recordings, but see notes 32 and 37.

¹⁷ In 2013, I was employed by the Swedish radio with the task of digitizing their Saami archive materials. While these recordings contain spontaneous speech in all Saamic languages spoken in Sweden, most of them are not publically available and have, to the best of my knowledge, not been utilized for the purpose of linguistic research (cf. Siegl 2017: 254, Note 3); see however Note 29.

¹⁸ Despite briefly doing work for the Swedish place name committee (*Ortnamnskommittén*), Isberg did not pursue a career in academia, but worked at the newspaper *Dagens Nyheter* from the 1940s until his retirement. During the Second World War, he took an active role in supporting the Norwegian resistance movement (see the obituary in *Dagens Nyheter*, May 20, 1977).

¹⁹ Collinder’s use of the term “central” should not be confused with the Central dialect of Lule Saami, as later defined by Sammallahiti (1998: 21–22).

Saami *a* in a stressed syllable to be realized as *ě* if the following nucleus is either *i* or *u*. This process affects simplex lexemes, e.g., SA, LE *běssti* ‘spoon’, *běktu* ‘via’ (standard *bassti*, *baktu*), as well as derived words, e.g., *ěkti* ‘once’ (standard *akti*) < *akta* ‘one’. The raising is also present in the inflection, e.g., SA, LE *běssiv* fry.PST.1SG (standard *bassiv*), cf. *basset* ‘to fry’; *běrgguj* work.SG.ILL (standard *bargguj*), cf. *barggo* ‘work’.²⁰ For a further description of this umlaut in the Forest dialect, see Collinder (1938: 60–64); for a historical-comparative study of the phenomenon, based on the materials of Wiklund and Collinder, see Wickman (1965).²¹

It should be pointed out that on two accounts, Sammallahti’s (1998: 21–22) classification of the Forest dialect is at odds both with my own data and with Collinder’s description of the variety. First, Sammallahti states that a second syllable long *á* is not subject to labial vowel harmony, i.e., not being realized as long *ā* when the nucleus of the preceding syllable is a short *á*. This is not the case in my data (e.g., SA, LE *álmāj* ‘man’, LE *gādāj* corral.SG.ILL, SA *bāhtjām* milk.PST.PTCP), nor in Collinder’s materials (1938: 57–58).²²

Second, Sammallahti claims that the Forest dialect has retained the opposition between *ä* and *ie* before a second syllable *e*, unlike in more southern dialects of Lule Saami where the two vowels have collapsed into *ie* in this position. On the contrary, the Forest dialect is also subject to this development, e.g., SA, LE *giergge* ‘stone’, *tjiehtje* ‘father’s brother’ (cf. Collinder 1938: 58–59).²³

²⁰ Notably, I have heard this umlaut even in the foreign place name SA SG.ELA *Gárasěvvunis* (North Saami *Gárasavvon*, Sw. *Karesuando*). The only exception to this rule in my data is the form SA *šaljijuj* yard.SG.ILL, instead of expected **šěljijuj*. In the texts of Lindgren, there is a form *vatti* give.PST.3SG (OL 1912: 3) in addition to forms with expected umlaut (cf. SA, LE *věttij* id.); see also Kejonen (2021a: 27, Note 3). Wiklund’s notebooks contain a few forms without umlaut, e.g., *älūtš* ‘western’, *ävūctis* (KBW 1895 II), PL.NOM *avuotisā* ‘reindeer harness’ (KBW 1895 I, speaker from Hiergáhpe); see also the exceptions discussed by Collinder (1938: 61–64). Note additionally that at least three words that have a first syllable *a* in Standard Lule Saami have *á* in my data, thus being exempt of umlaut: SA, LE *áktu* ‘alone’, *bákkoj* : *bággu-* ‘hazel grouse’, SA *rábo* : *ráhpuga-* ‘glove’ (standard *aktu*, *bakkoj* : *baggu-*, *rabo* : *rahpuga-*), see further Collinder (1938: 66–69).

²¹ In addition to the comparisons with other Saamic varieties made by Wickman, it can be noted that a strikingly similar vowel alternation is attested in eastern North Saami as well (Sammallahti 1998: 188).

²² One exception to this rule is the loanword LE *sámá*, ATTR *sámás* ‘nice’ (cf. Meänkieli *soma* id.). For the sake of completeness, it should also be pointed out that for one of the speakers interviewed by Wiklund, an etymologically long *á* is not subject to labial vowel harmony, e.g., *kóhtšām* awaken.PST.PTCP, *kóčtšājiv* awaken.PST.1SG, *pórrā* eat.PRS.3SG. An etymologically short *a*, on the other hand, has the quality [o] regardless of whether it is short or long, e.g., *pórrót* eat.INF, *pōrō* eat.CNG (KBW 1895 II). One token of lacking labial vowel harmony is also attested in the texts of Lindgren (Kejonen 2021a: 27, Note 8).

²³ This development is additionally seen even in the place names LE ‘*Gårkkábies*, *áhpe* (lit. ‘ant hill bog’) and LE ‘*Hierg*, *áhpe*, SA SG.ELA ‘*Hierg*, *ábies* (lit. ‘draft reindeer bog’, Sw. *Härkmyran*) in which the first elements *gårkkábiesse* ‘ant hill’ and *hiergge* ‘draft reindeer’ have lost their final syllables (cf. Spiik 1989: 102). Note further that Collinder (1938: 59) states that

Traits that differentiate the Forest dialect from the written standard language include a lack of labial vowel harmony in words where the consonant center is a cluster beginning with *v*, e.g., SA *bávnná* ‘tussock’, *ávtav* one.SG.ACC, *ávddal*, LE *ávddala* ‘before’, SA, LE *skávlla* ‘school’ (standard *bávnná*, *ávtáv*, *ávddál*, *ávddála*, *skávllá*).²⁴

Further unlike the standard language, the diphthong *oa* is subject to umlaut also when preceding a consonant center in Grade III, e.g., SA, LE *ádtjum* receive.PST.PTCP, SA *áhppiv* learn.PST.1SG (standard *oadtjum*, *oahppiv*). For the front vowel *ä*, I have noted the umlauted forms SA *eddnú* : *eddnú* ‘mother’s brother’, *bejsstum* get.ruined.PST.PTCP and LE *hehttu* must.PRS.2SG before a Grade III consonant (standard *änoj* : *äddnu*-, *bäjsstum*, *hähttu*), but also SA *gähttijiv* look.PST.1SG (standard *id.*), without expected umlaut.

The Forest dialect has retained *ä* in monosyllabic stems and before a second syllable *a*, unlike in the standard language where there is a long *e* instead (except before a consonant center in Grade III), e.g., SA, LE *lä*- ‘be’, *bäna* ‘dog’, LE *gärmaj* ‘snake’ (standard *le*-, *bena*, *germaj*). As in the standard language, an older *oa* can be pronounced more or less as a monophthong *ä* in the corresponding environment, e.g., SA, LE *māj* ‘we (two)’ (cf. Collinder 1938: 71). On occasion, I have, however, also noted a slight diphthong *oa*, e.g., SA *soahpap* agree.PRS.1PL (standard *sāhpap*).

4.2 Consonantism

The Forest dialect has lost the affricates *ts* /ts/ and *dtš* /dz/ through merger with *tj* /tʃ/ and *dtj* /dʒ/ respectively, e.g., SA, LE *tjiellá* bark.PRS.3SG, LE *boatjoj* ‘reindeer’ (standard *tsiellá*, *boatsoj*), SA *vádtjet* ‘to walk’, LE *vádtjiv* walk.PST.1SG (standard *vádtset*, *vádtšiv*); note, however, the geographical variation documented by Collinder (1938: 49–50, cf. Kejonen 2021a: 25). These mergers have also taken place in some other varieties of Lule Saami (Angéus Kuoljok 1997: 25, Korhonen 1979: 544; cf. Halász 1892: 258).²⁵

forms with retained *ä* are only sporadically found in the neighboring Unna Tjerusj dialect. In my material on that dialect, I have noted forms with *ä* for both interviewed speakers, e.g., *äddne* ‘mother’, *bävde* ‘table’, *gässeskávlla* ‘summer school’ (see also Angéus Kuoljok 1997: 45 and Sammallahti 1998: 21). In the word for ‘stone’, however, I have noted different forms for the two speakers: *gärgge* (with retained *ä*) and *giergge* (*ie* < *ä*).

²⁴ In Lindgren’s texts from Stuor Sávvva (OL 1912) there are forms such as *áutalatja* ‘a little bit before’ and *jáula* ‘Christmas’, but also *áuftáv* one.SG.ACC and *áuftá=k* one.SG.ACC=NPM. Collinder (1938: 57) has noted the place name *póuno-sīta* (Sw. *Bönlräsk*) as an exception to this rule. I have tentatively noted the name of this village as SA *Bávnnu*, LE SG.INE *Bávnon* (cf. Meänkieli *Pounu* id.) and the nearby river as SA, LE *Bávnoädnö*; cf. also Wiklund’s form *Pöunö-jaure* ‘Bönlräsk’ (KBW 1895 I).

²⁵ I have for instance noted the same innovation in Unna Tjerusj (e.g., *báhtjám* remain.PST.PTCP, standard *báhtsám*). This trait is further found in Čohkkiras North Saami; see also the

A conservative trait of both interviewed speakers is the retention of word-final *-tj*, e.g., in the reflexive pronoun SA *ietj* ‘oneself’ and in the nouns *sānjátj* ‘grainy snow’, *gieratj* (bird species, Sw. *spillkråka* ‘black woodpecker’) and SA, LE *niejdagutj* ‘little girl’ (standard *iesj*, *sānjásj*, *gierasj*, *niejdagusj*). Especially in connected speech, forms with a final fricative are however also attested, e.g., SA, LE *iesj*, *niejdagusj*, *ulmusj* : *ulmmu-* ‘person’ (standard *ulmusj* : *ulmutja-*); see also Collinder (1938: 55).²⁶

Both speakers have retained the opposition between consonant centers of the kind *jgg* [j^hk] and *jkk* [j^hhk], e.g., SA, LE *lájgge* ‘yarn’ vs. *lájkke* ‘lazy’ (standard id.), unlike some other varieties of Lule Saami, and unlike one of the speakers interviewed by Collinder (1930: 7; 1938: 48–49, cf. Angéus Kuoljok 1997: 53, Lagercrantz 1941: 306, Qvigstad 1925: 19).²⁷

There is also a retained opposition between consonant centers such as *hpp* and *pp* in cases where some other varieties have neutralized this opposition (Sammallahti 1998: 22, Angéus Kuoljok 1997: 53), e.g., SA *oahppá* learn.PRS.3SG vs. *oappáv* sister.SG.ACC (standard id.). Similarly, the Forest dialect has retained the opposition between consonant centers of the kind *tt* and *ht* (cf. Sammallahti 1998: 22), e.g., LE *Vuottas* (place name) vs. *muohta* ‘snow’ (standard id.).

As in some other varieties (and, optionally, in the standard language), the clusters *ppt* and *pptj* (including older *ppts*) have developed into *bdd* [b^ht] and *btj* [b^htj] in words such as SA *såbddå* ‘foam’, *vuobdda* ‘hair’, *gábdde* (a traditional Saami garment), LE *sjubttjasijt* story.PL.ACC (standard *supptsasijt* ~ *subttsasijt*).²⁸ An exception to this rule is however the adverb SA *ruopptot* ‘back’ (see also the variation documented by Collinder 1938: 131–132).²⁹

Another noteworthy feature of the Forest dialect is the presence of an epenthetic vowel in the strong grade of the consonant center written *llj* in the

corresponding developments in varieties of Ume Saami (Larsson 2012: 118, Hansegård 1988: 318) and South Saami (e.g., Magga & Mattsson Magga 2012: 17).

²⁶ The latter word is also used as a generic pronoun ‘one’, likely influenced by Swedish *man* ‘man; one’ (Juutinen & Mettovaara 2021: 115–116), e.g., SA *ulmusj má vajálduhtá* [person.SG.NOM DPT forget.PRS.3SG] ‘one forgets’.

²⁷ In one word, I have noted the consonant center *jdd* instead of expected *jtt*: SA *skájdde* (bird species, Sw. *fjällabb* ‘long-tailed jaeger’), cf. Standard Lule Saami *skájtti* id. (in northern Jiellevárre also documented with a final *-e*, Grundström 1946–54 s.v. *skai’hti*). This pronunciation can be attributed to a folk etymological connection with the word *skájdde* ‘land where two rivers meet’—according to SA, the tail of the bird in question looks like two rivers meeting (cf. Ryd 2017: 29).

²⁸ Collinder has noted this word with an initial *s-*, just as in the standard language: *suptjas*, PL.NOM *subtjas* ‘story’ (1938: 114). Both of the interviewed speakers however have an initial *sj-* instead: SA, LE *sjubtjas*, likely due to distant assimilation to the postalveolar affricate *tj*.

²⁹ This word is not attested in the corpus, but occurs in a radio interview with SA concerning the sacrificial site of Unna Sájjva, Sw. *Lillsaivis* (Heikki 2021).

standard language, e.g., SA, LE *vielja* [vielʲja] ‘brother’ (standard *viellja*, without an epenthetic vowel).³⁰

A few sound changes documented by Collinder in part of the Forest dialect area are not attested in my data. These changes include the loss of *sj* /ʃ/ through merger with *s* /s/ (also documented by Wiklund) and the loss of *nj* /ɲ/ through merger with *n* /n/ (Collinder 1938: 50–51).³¹ Rather, both *sj* and *nj* are retained, e.g., LE *rássjodit* ‘to rain’, *njunnje* ‘nose’ (standard *id.*), SA *njuovtjam gæssjku* [tongue.SG.NOM stiffen.PRS.3SG] ‘the tongue stiffens (when one does not speak Saami regularly)’ (standard *njuovtjav gassjku*); see however 4.3 on the illative singular of personal pronouns.

Collinder also reports a uvular pronunciation of *r* (UPA *ɣ*) for several speakers of the Forest dialect (1938: 13; 1930: 9). Both of the interviewed speakers have alveolar [r], but when asked about uvular rhotics, LE remembered that people from the nearby village Suobbat had this trait both when speaking Saami and Swedish (cf. Källskog 1992: 119).³²

4.3 Morphology

Questions of verb morphology will be treated in more detail below, but a few notes can be made here on the inflection and derivation of other word classes.

In the singular of personal pronouns, the Forest dialect has inessive forms such as the first person LE *muvn*, SA *muvn* ~ *muvně* [mu:n ~ mu:nɛ] and the

³⁰ The presence of an epenthetic vowel between *l* and *j* clearly shows that *lj* cannot be regarded a single segment in this variety. This is contrary to the analysis of Angéus Kuoljok (1997: 50), who posits a palatal lateral for (Standard) Lule Saami. It should further be noted that this feature is not restricted to the Forest dialect—I have recorded the pronunciation [vielʲja] ‘brother’ also in Unna Tjerusj (as well as in Eastern Čohkkiras North Saami); see also Wiklund (1891: 14), who noted a variation *pelje* ~ *pelʲje* ‘ear’, however without providing any provenance of the example. While the pronunciation with an epenthetic vowel between *l* and *j* was a *fait accompli* already when Wiklund first documented the Forest dialect in the late 1800s (e.g., *kälʲjē* ‘wide’, KBW 1895 II), the presence of this vowel in Unna Tjerusj is a fairly recent development—older records show forms without such a vowel, e.g., *vèllja* ‘brother’ (Grundström 1946–54 s.v. *vielʲja*). Potentially, the pronunciation with an epenthetic vowel could be due to influence from Meänkieli, where an epenthetic vowel can be found in the corresponding cluster (see, e.g., Airila 1912: 42). On the other hand, forms with an epenthetic vowel are also attested far south of the Meänkieli speaking area, e.g., Pite Saami *věäʲlʲja* ‘brother’ (Lagercrantz 1939, §8677:3). For a detailed study of the epenthetic vowel in Lule Saami, see Larsson (1991).

³¹ Both of these developments may be due to language contact—Meänkieli only distinguishes one sibilant, *s* /s/, in native vocabulary, and none of the co-territorial languages has a palatal nasal phoneme. The merger of *š* /ʃ/ with *s* /s/ in Duortnus (Torne) North Saami may also have been induced by contact with Meänkieli (cf. Lagercrantz 1941: 283; 1958: 167, Hansegård 1967: 96; 1988: 231, 321, Bergsland 1992: 13).

³² In a recording of a speaker of Lule Saami and Meänkieli from Suobbat, made in 1985 (DAUM BD 5070), I have noted uvular [ɣ] for both languages. See also Grundström (1946–54: 1696) on this trait in Lule Saami; cf. Halász (1886: 4; 1891: V), Hasselbrink (1981–85: 57; 1944: 119) and Kowalik (2023: 46) on South Saami and Nielsen (1926: 21) on North Saami, as well as Pohjanen (2022: 42) on Meänkieli.

second person SA *duvně* [tu:nɛ] (standard *mujna*, *dujna*), as well as elative forms such as the first person SA, LE *musstě* and the second person SA *dusstě* (standard *mujsta*, *dujsta*). In the illative, there is an *n* corresponding to the palatal *nj* of the standard language: first person SA, LE *muni*, second person SA *duni* (standard *munji*, *dunji*).³³

As in the standard language, the illative plural of parisyllabic nouns has the suffix *-jda*, e.g., SA *gâvi-jda* picture-PL.ILL (on the vocalism of this word, see Section 4.5), LE *sěri-jda* blueberry-PL.ILL (cf. Unna Tjerusj *máná-jdi* child-PL.ILL). Older materials on the Forest dialect show a vacillation between the suffixes *-jn* and *-jna* for the comitative singular of parisyllabic nouns (cf. Wiklund 1915: 27; 1891: 113). The present corpus only features forms in *-jn*, as in the modern standard language, e.g., SA, LE *běstijn* spoon.SG.COM.

In comparative and superlative forms derived from imparisyllabic adjectives, I have noted an *ä* in the penultimate nucleus (corresponding to the short *a* of the standard language), e.g., SA, LE *vuorrasäbbo* ‘older’, *vuorrasämos* : *vuorrasebmusa*- ‘oldest’ (standard *vuorrasabbo*, *vuorrasamos* : *vuorrasabmusa*-). Similarly, there is an *ie* (< *ä*) in the penultimate nucleus of actor nouns derived from imparisyllabic verbs (corresponding to the short *i* of the standard language), e.g., SA *lågäddieddjje*, SG.ILL *lågäddiddjáj* ‘teacher’ (standard *lågäddiddje*, *lågäddiddjáj*); see also Section 6.5. I have heard one token of the caritive derivational suffix *-diebme* ‘-less’ (standard *-dibme*), SA *tjalmediebme* ‘blind’ (lit. ‘eyeless’).

Noteworthy are also the numerals 11–19, which are formed by means of the element *-lâgeně* (lit. ‘on ten’, standard *-lâgenan*), e.g., SA, LE *aktalâgeně* ‘eleven’ < *akta* ‘one’, *guoktělâgeně* ‘twelve’ < *guoktě* ‘two’, cf. SA *něnně* ~ *nanně*, LE *nanně* ‘on’.

4.4 Morphosyntax

While attributive adjectives are generally indeclinable in Lule Saami, one of the interviewed speakers inflects the word *buorre* ‘good’ in agreement with the noun it modifies. For the other speaker, this word has an indeclinable attributive form instead (Table 1). The former speaker agrees in this respect with more northern varieties, whereas the other speaker agrees with more southern dialects, as well as with the standard language (Grundström 1946–54: 1843).

³³ To the best of my knowledge, forms like *muni* and *duni* are not attested in other dialects of Lule Saami; see however the variation *muññe* ~ *munne* 1SG.ILL noted by Lagercrantz (1939, §3688:2) in the North Saami Forest dialect of Juŋgá (Meänkieli *Parkalompolo*). For a discussion on the history of the palatal nasal in the illative singular of personal pronouns, see Wiklund (1896: 282–283), Korhonen (1981: 219–220) and Sammallahti (1998: 76).

Table 1. Partial paradigm of *buorre bāna* ‘good dog’.

	SA	LE
SG.NOM	buorre <i>bāna</i>	buorre <i>bāna</i>
SG.ACC	buorev <i>bādnagav</i>	buorre <i>bādnagav</i>
SG.ELA	buoret <i>bādnagis</i>	buorre <i>bādnagis</i>
ABE	buore <i>bādnagahtá</i>	N/A

This variation in agreement is not restricted to case, but applies to number as well: SA *buore gábmaga* [good.PL.NOM shoe.PL.NOM] ‘good shoes’ vs. LE *buorre bādnaga* [good.ATTR dog.PL.NOM] ‘good dogs’.³⁴

In the Lule Saami standard language, the verb *liĵku-* ‘like’ governs the illative case (Spiik 1989: 97, Wiklund 1915: 110). For both interviewed speakers, however, *liĵku-* generally takes an object in the accusative, as in (1). Interestingly, the corpus also contains one token of this verb governing the elative case (2), agreeing with Meänkieli, where the corresponding verb *tykätä* ‘to like’ similarly governs the elative (see further Kittilä & Ylikoski 2018 on the morphosyntax of verbs meaning ‘like’ in Saamic).

- (1) *bāhtju* *dě* *liĵkuji* **guobbarijt**
 reindeer.PL.NOM DPT like.PRS.3PL mushroom.PL.ACC
 ‘Reindeer like mushrooms.’ (LE)

- (2) *mān* *liĵkuv* **dat** **namás**
 1SG.NOM like.PRS.1SG DEM.SG.ELA name.SG.ELA
 ‘I like that name.’ (LE, elicited)

As seen in (3), the elative can also be used to encode the agent of passive verbs in *-dalla-* (such as *bārādalla-* ‘be bitten’), unlike in the standard language, where the illative is used in this function (Spiik 1989: 98, Wiklund 1915: 110; cf. however Spiik 1989: 99 and Wiklund 1915: 108 on other passives).

³⁴ In the texts of Lindgren, *buorre* ‘good’ is not inflected attributively: *puorre päddait* [good.ATTR bed.PL.ACC] (Stuor Sájvva, OL 1912: 4). Wiklund’s and Collinder’s materials contain no token of this word used attributively. Variation regarding the attributive inflection of *buorre* ‘good’ is also attested in North Saami (Nielsen 1926: 89) and Pite Saami (Lehtiranta 1992: 122); note additionally that there is some variation concerning the inflection of attributive adjectives documented in older descriptions of Lule Saami (Wiklund 1891: 203; 1915: 37–38). See further Rauhala (2013) regarding the history of the attributive inflection of this word (cf. also Stolz 2015: 294).

- (3) *mån* *bädnagis* *bårädëlliv*
 1SG.NOM dog.SG.ELA bite.PASS.PST.1SG
 ‘I was bitten by the dog.’ (SA, elicited)

4.5 Lexicon

The corpus includes a number of names for flora and fauna that appear to be unattested in other varieties of the language: SA *gådtjårásse* (plant species, lit. ‘piss grass’, Sw. *svinmålla* ‘white goosefoot’), *garssatjuojkka* (mosquito species, lit. ‘frozen ground mosquito’, Sw. *tjälmygga* id.), *dihátáj* (bird species, Sw. *sparvuggla* ‘pygmy owl’).³⁵ Noteworthy is also SA, LE *dåmmbå*, SA SG.ACC *dåmbåv* ‘horse’³⁶ and SA *båvna*, SG.ACC *båvnagav* ‘bear’ (derived from *båvnna* ‘tussock’ and used alongside SA *biernna* id.).³⁷ A verb that to my knowledge is not attested in other varieties of Lule Saami is LE *gien'niji* calve.PRS.3PL (cf. Collinder 1938: 59).

An unexpected final vowel is found in the word SA, LE *gávve*, SA PL.ACC *gávvi*, SG.ILL *gávvi* ‘picture’, which has a final *-å* (< *-a*) in most other Lule Saami dialects, but which is attested with *-e* also in Unna Tjerusj (Grundström 1946–54 s.v. *kåvvå*, cf. Aikio 2007: 36–37). Similarly unexpected is the final *-u* of SA *áddju*, SG.GEN *ádj* ‘grandfather (maternal or paternal)’, a parallel form of *áddjá* id.³⁸ A final *-u* is also found in SA *skiru*, SG.ACC *skiruv*, SG.ILL

³⁵ According to SA, the latter name is likely onomatopoeic (cf. also Standard Lule Saami *dihttá* ‘merlin’).

³⁶ Both speakers consider this word more correct than the loanword *hässta* id. (cf. Sw. *häst*). SA further states that he has seen this word in “an old book”, likely referring to the history textbook of Odhner in the so-called “South Lappish Book Language” (“sydlapska bokspråket”), where forms such as *tåmp* ‘horse’, SG.GEN *tåmpen* are found (1878: 97); cf. South Saami *dåâmhpē*, *dåâmpe* etc. and Ume Saami *tõm* ‘pə, PL.NOM *tõm* ‘p id. (ULMA 22480). Forms with a stem-final *-å* are attested in the dictionary of Lindahl & Öhrling (1780), e.g., SG.ACC *tåmpåb* (s.v. *swintet*); note also the form *dåmpå*, documented in the speech of the so-called “parish Lapps” in the 1770s (Larsson 2018: 99). In this context it may also be mentioned that the synonym given as *sioms* ‘horse’ by Larsson should read *siomb* instead (see Ihre 105, §1041), cf. e.g., *sjåmp* ‘horse’ (Lindahl & Öhrling 1780: 406); see also the derived word *siombäck* ‘shoe from leg skin of horse’ in the same word list (Larsson 2018: 82).

³⁷ In the dictionary of Grundström, the corresponding word is given as an adjective ‘tussocky’ (1946–54 s.v. *påunåk*). The meaning ‘bear’ is however also attested in older sources on Lule Saami (Wiklund 1890 s.v. *pounok*; Drake 1918: 327). In older records of the Forest dialect, the name of the mythological ogre *stållo* is also attested in the meaning ‘bear’: *stállō* ‘bear’ (Sádimvárre, ULMA 15927: 76, 79), *stållo* id. (Suobbat, DAUM BD 5070); cf. Drake (1918: 327) and Grundström (1946–54 s.v. *stállō*). Here a parallel can be noted between Lule Saami on the one hand, and the eastern Saamic languages Skolt Saami, Akkala Saami and Kildin Saami on the other: in the latter languages, a word corresponding to Lule Saami *stållo* is similarly attested in the meaning ‘bear’ (Itkonen 1958: 519–520, 1069 s.v. *stállō*). It should however be added that neither SA nor LE knows the word *stållo* in the meaning ‘bear’, but only in its mythological sense.

³⁸ Although SA says *áddju* himself, he also describes this form as “children’s language”, the regular form being *áddjá*. Here it can also be noted that Collinder translates the corresponding

skiruj ‘magpie’ (standard *skiri* ~ *skirri*). Further slightly different from the standard language is SA, LE *rávnnje*, SA SG.ACC *rávnjev*, SG.ILL *rávnnjáj* ‘rowan’ (standard *rávnnno*) and SA, LE *Svieri*, LE SG.ILL *Svierij* ‘Sweden’ (standard *Svierik*, see also the synonym *Ruoha* below). An unetymological cluster *pps* is seen in the word SA *idedisguoppsun* dawn.SG.INE, standard *-guoksun* (cf. Wiklund 1891: 144, Qvigstad 1925: 15).

A couple of words explicitly noted by Collinder (ULMA 15927: 83–84) as missing in Sádinvárre are attested in my data, including *vinndek* ‘window’ (SA SG.ELA *vinndegis*) and *gárránis* ‘raven’ (SA SG.INE *gárránisán*).³⁹ For one word, I have noted a different meaning as compared to the standard language: SA *bierre* ‘cousin’ (standard ‘relative’).

The lexicon of the Forest dialect is further characterized by a number of loanwords from Finnish/Meänkieli that to my knowledge are not attested further south, including LE *gis’sē*, SG.ACC *gissev* ‘cat’; *bihe*, SG.INE *bihen*, SG.ILL *biháj* ‘yard (of a house)’; SA *buvrrō*, SG.ACC *buvrov* ‘oatmeal’; *gáhto*, SG.GEN *gádo* ‘roof’; *sádne*, PL.NOM *sáne* ‘word’ (but cf. Lagercrantz 1939, §1893:4) and *vielín*, SG.ACC *vielínáv* ‘apron’ (cf. Meänkieli *kissa*, *piha*, *puuro*, *katto*, *sana*, *vyölinä* id.). Noteworthy in this context is also LE *Ruoha*, SG.INE *Ruohan*, SG.ILL *Ruohaj* ‘Sweden’ (cf. dialectal Finnish *ruoŕŕi*, Korhonen 1981: 161), which judging from its consonantism may have been mediated via Čohkkiras North Saami *Ruoha* id. (see Collinder 1938: 136–137 on the development of **ŕ* [θ] in the Forest dialect; other words in the data reflecting this sound are SA, LE *rahte*, SA SG.ACC *rahtev* ‘road’ and SA, LE *muodá*, SG.ACC *muodáv* ‘mother’s sister’, cf. Kejonen 2020: 48).⁴⁰

Further, the Forest dialect shows several borrowings from local Swedish varieties, including SA, LE *fielistit*, PST.PTCP *fielistam* ‘travel’, cf. Överkalix

form *áđju* simply as ‘grandfather’ (1938: 65, 102–103), whereas Wiklund translates *áđ’d’u* as ‘grandfather (children’s language)’ in his Forest dialect materials (Grundström 1946–54 s.v. *áđju*). The form in *-u* has probably had a wider geographical distribution in the past—while Grundström notes this form as unknown in northern and central Jämkämähkke, Wiklund (1890: 1) has documented it also in northern Jämkämähkke: *áđu* ‘grandfather (children’s language)’, cf. *áđjo* ‘great grandfather’ (“farfars far”) (Drake 1918: 221). Note additionally that the form *áđđjá* also has the meaning SA, LE ‘thunder’ (e.g., LE *áđđjá manná* [thunder.SG.NOM go.PRS.3SG] ‘it thunders’), as well as SA ‘large person (man or woman)’.

³⁹ As told by SA this bird has a white lucky feather which, if obtained, can aid its owner in finding lost objects (cf. Bartens 2022: 36–37 and references therein, see also Grundström 1929: 72, Labba 1969: 90–91 and Ryd 2007: 246–248; 2017: 170). In other varieties of Lule Saami, this feather is known as *lievedälge* (Grundström 1946–54 s.v. *lievē-(lāvē-)tāl’kē*). From my father, born in Buornna (Meänkieli *Purnu*) in 1944, I have heard the Meänkieli term *noitakynä* in the same meaning (“if you have that feather, you do well in court”). Neither of these terms was however recognized by SA. Note, finally, that while some dictionaries translate *gárránis* as ‘crow’ (Korhonen 2005, Kintel 2012 s.v.), the latter bird is referred to by SA as *vuoratjīs*.

⁴⁰ Also worth mentioning is SA, LE *těľěfoarnna* ‘telephone’, with an unexpected consonant cluster *rnn* : *rn* [r̥n : ɾ̥n] (cf. Standard Swedish *telefon*, Standard Lule Saami *těľěfävnnä*). A corresponding form is however attested in Meänkieli (OMSM s.v. *telefoorni*).

variety *feiLes* id. (ÖKM s.v. *färdas*, Pihl 1948: 54–55) and *flihttjit*, PST.1SG *flihttijiv* ‘drive logs’, cf. Överkalix variety *fLiTi* : *fLittje* ‘move; drive logs’ (ÖKM s.v. *flytta*, Pihl 1948: 54–55).⁴¹ Also belonging to this category is the word SA *iru*, SG.INE *irun*, SG.ILL *iruj* ‘backwater’, cf. Överkalix variety *iro* id. (ÖKM s.v. *eda*, Pihl 1924: 90; see also Kejonen 2022: 121, 123).

As in other Saamic varieties spoken in Sweden, there are also recent Standard Swedish loanwords in the Forest dialect, e.g., SA *sjyllta* ‘(road) sign’, *te,vea* ‘television’ (cf. Swedish *skylt*, *tv* id.). It should however be noted that both interviewed speakers express negative attitudes towards loanwords and appear to make a conscious effort to avoid obvious loans. For instance, SA explicitly shunned the loanwords *kujnna* ‘woman’ and *’fun,dierit* ‘to ponder’ (cf. Swedish *kvinna*, *fundera* id.), favoring the synonyms *nissun* and *ájádallat* instead.⁴²

5. Inflectional verb morphology

In this section, the inflectional verb morphology of the Lule Saami Forest dialect is described, starting with the identification of inflectional stems (5.1) and inflectional classes (5.2). Then, the inflectional categories of finite verbs (person, number, tense, mood) and the suffixes of these forms are presented (5.3). Section 5.4 describes the non-finite forms attested in the corpus, as well as their respective suffixes. The next three sections (5.5–5.7) describe the inflection of lexical verbs: parasyllabic verbs, imparisyllabic verbs and contracted verbs (on these terms, see 5.2). Sections 5.8 and 5.9 describe the inflection of non-lexical verbs: *lä-* ‘be’ and the negative auxiliary. Finally, in Section 5.10, the highly irregular negation of *lä-* ‘be’ is presented. Before proceeding to the finer details of the verb inflection in this variety, it might however be useful to give examples of the morphological and morphophonological processes relevant for the remainder of the paper.

In Lule Saami, inflectional forms can be encoded both by means of suffixes and by stem-internal processes. These processes are exemplified below, where

⁴¹ Considering its vocalism, the former word may be a borrowing separate from southern Lule Saami and Pite Saami *fielastit* id. (Korhonen 2005, Wilbur 2016 s.v., Ruong 1943: 191), Ume Saami *fielastidh* id. (Bergsland & Hasselbrink 1957: 72) etc.

⁴² Ironically, *ájádallat* ‘to think, ponder’ is also a loanword (cf. Meänkieli *ajatella* id.). On several occasions, I was told off by SA for using loanwords when speaking Saami. For instance, he did not approve of me using the (North Saami) words *hoahppu* ‘hurry’ and *ipmirdit* ‘to understand’ (cf. Meänkieli *hoppu*, *ymmärtää* id.). I was also rebuked for using the word *čielgi* ‘back’, which he perceived as a loan from Meänkieli *selkä* id. (for the record, it should however be pointed out that *čielgi* is not a borrowing, but a cognate of *selkä*). Instead of these—actual or presumed—loanwords, he preferred the words *ráhito* ‘hurry’, *dádjadit* ‘to understand’ and *tjavelk* ‘back’, which he regarded as more correct.

the verb *boahte-* ‘come’ is inflected by the use of a suffix (4a), consonant gradation (4b) and stem-final vowel alternation (4c) respectively; the latter process may also induce umlaut in the penultimate nucleus of the stem, as shown in (4d). Several of these processes may further be applied simultaneously, as in (4e).

(4a) <i>boahte-p</i> come-PRS.1PL ‘we come’	(4b) <i>boade</i> come.IMP.2SG ‘come!’
(4c) <i>boahtá</i> come.PRS.3SG ‘s/he comes’	(4d) <i>bǎhti</i> come.PRS.3PL ‘they come’
(4e) <i>bǎdi-jma</i> come-PST.1PL ‘we came’	

While the resulting system yields a significant number of stem allomorphs, it should be pointed out that all lexical verbs in Lule Saami have regular inflection—there are, for instance, no cases of suppletion for lexical verbs.

5.1 Inflectional stem

Lexical verbs in Lule Saami are divided into three inflectional classes: parisyllabic verbs, imparisyllabic verbs and contracted verbs (see further 5.2). The inflectional stem of parisyllabic and imparisyllabic verbs is retrieved by removing the suffix *-t* from the infinitive (which also functions as the citation form of verbs), e.g., *goarro-t* sew-INF ‘to sew’, *vuojadi-t* swim-INF ‘to swim’. Contracted verbs have two inflectional stems. The shorter of these two stems is the one preceding the infinitive marker, e.g., *lijkku-t* like-INF ‘to like’. The longer stem is characterized by an element *ji-*, which appears in certain inflectional forms, e.g., *lijjkuji-n* like-PST.3PL ‘they liked’. The longer stem is identical to the third person plural present, as well as to the second person singular past, e.g., *lijjkuji* like.PRS.3PL ‘they like’ (= like.PST.2SG ‘you liked’). The copula verb, which also functions as an auxiliary, has several inflectional stems (see 5.8); in this paper, I refer to this verb by its present indicative stem *lä-* ‘be’ (cf. Wiklund 1891: 268, Grundström 1946–54: 1887). The negative auxiliary verb similarly has several stem allomorphs (see 5.9), but lacks a convenient citation form.

A number of alternative analyses regarding the segmentation of Lule Saami verbs could potentially be put forward. For instance, it would from a synchronic perspective be possible to regard the *ji*-element of *lijkkujin* ‘they liked’ as belonging to the inflectional suffixes rather than to the stem (i.e., segmenting *lijkkuji-n* like-PST.3PL as *lijkku-jin* instead); this approach is favored in some pedagogical materials on Lule Saami (see, e.g., Spiik 1989: 66). In the present study, however, I have opted for that analysis which renders the system of suffixes as economical as possible—while the abovementioned analysis would reduce the stem allomorphy slightly, it would greatly increase the number of allomorphs for inflectional suffixes instead.

5.2 Inflectional classes

The vast majority of Lule Saami verbs belong to the large and open group of lexical verbs, all of which have a regular inflection. In addition to these verbs, there are also two non-lexical verbs, both of which have irregular inflection (including suppletion): *lä-* ‘be’ and the negative auxiliary. These two non-lexical verbs differ morphologically from lexical verbs in several respects, in addition to their irregular inflection. Whereas all lexical verbs are polysyllabic, the non-lexical verbs have a monosyllabic stem in several inflectional forms. The verb *lä-* ‘be’ is further unique in being inflected for the conditional mood (but see 6.3), and the negative auxiliary is unique in lacking non-finite forms altogether. The morphological criteria used to distinguish lexical and non-lexical verbs are summarized in Table 2.

Table 2. Features of lexical and non-lexical verbs.

	Lexical verbs	<i>lä-</i> ‘be’	NEG
Monosyllabic stem	–	+	+
Irregular inflection	–	+	+
Conditional forms	–	+	–
Non-finite forms	+	+	–

Turning then to the lexical verbs, three main inflectional classes can be discerned using morphological criteria. In the following, I use the traditional labels for these classes, that is, “parisyllabic verbs” (with an even number of syllables), “imparisyllabic verbs” (with an odd number of syllables) and “contracted verbs” (with an alternating number of syllables); see, e.g., Spiik (1989: 65) and the verb paradigms in Grundström (1946–54: 1847–1918). Note that the terms “parisyllabic” and “imparisyllabic” refer to the final stress group of the inflectional stem, and not necessarily to the total number of

syllables of a verb (on the stress group in Saami, see Sammallahti 1998: 39).⁴³ The term “contracted verb” refers to the fact that the longer stem of these verbs (see 5.1, 5.7) is the original from a historical point of view, whereas the shorter (or “contracted”) stem is a later development.⁴⁴ In the following, a brief characterization of the three inflectional classes is given.

- **Parisyllabic verbs** are both the most common and the morphologically most complex, employing not only suffixes, but also consonant gradation and stem-final vowel alternations—often with ensuing umlaut in the penultimate nucleus—as a means of encoding different inflectional forms. Most parisyllabic verbs in the corpus are disyllabic (e.g., *goarro-* ‘sew’), but there are also some tetrasyllabic verbs belonging to this class (e.g., the derived *bārādalla-* ‘be bitten; be eaten’ < *bārrå-* ‘eat’ and *biebbmagoahte-* ‘begin to feed’ < *biebbma-* ‘feed’). Note that only the final two syllables of a tetrasyllabic verb, i.e., the final stress group, is relevant for inflection. For instance, *bārādalla-* inflects identically to the disyllabic verb *balla-* ‘be afraid’ as regards consonant gradation, stem-final vowel alternations, umlaut and suffixes. Similarly, *biebbmagoahte-* inflects identically to disyllabic *boahte-* ‘come’.⁴⁵ Parisyllabic verbs are further divided into three subclasses on the basis of their respective final nucleus: *a*-stem verbs (including *å*-stem verbs, see below), *e*-stem verbs and *o*-stem verbs. The subclasses are subject to different stem-final vowel alternations.
- **Imparisyllabic verbs** lack consonant gradation, but have a small degree of stem-final vowel alternations. Unlike parisyllabic verbs, however, the vowel alternations of this class do not cause umlaut. All imparisyllabic verbs in the corpus have a trisyllabic stem. Many imparisyllabic verbs are derived, e.g., *sámásti-* ‘speak Saami’ < *sábme* ‘Saami’.
- **Contracted verbs** lack consonant gradation, but have stem-final vowel alternations in the third person singular present in this

⁴³ For instance, despite having five syllables, the verb *‘rássjodasj.goahte-* ‘begin to rain’ (Korhonen 2005 s.v.) inflects as a parisyllabic verb, as its final stress group is disyllabic.

⁴⁴ For the parisyllabic verbs *goarro-* ‘sew’ and *oadtjo-* ‘receive’, Wiklund (1891: 235, 260, 263) lists forms identical to those of contracted verbs alongside the expected disyllabic forms, e.g., *kōrun ~ kōruyn* sew.PST.3PL, *ožuuv ~ ožuyiv* receive.PST.1SG. Notably, however, such forms are absent in later descriptions of the language, including the textbook of Wiklund himself (1901, 1915).

⁴⁵ In fact, Kuokkala (2019) convincingly argues that the inchoative suffix *-goahte-* stems from (the predecessor of) the verb *boahte-* ‘come’ (cf. Ylikoski 2009: 127).

variety. Verbs of this class are characterized by having two inflectional stems: one disyllabic and one trisyllabic, e.g., *lieggi-* : *lieggiji-* ‘heat’. The final nucleus of a contracted verb is *á*, *i* or *u* (on contracted verbs in *ǣ̃*-, see below). Contracted verbs are not as common as parisyllabic or imparisyllabic verbs, and many verbs belonging to this class are derived, e.g., *dål’li-* ‘make fire’ < *dállå* ‘fire’, *gávnnu-* ‘exist’ < *gávnnu-* ‘find’.

The criteria used to distinguish the different inflectional classes of lexical verbs are shown in Table 3.

Table 3. Features of lexical verb classes.

	Pari-syllabic	Impari-syllabic	Contracted
Syllables in final stress group	two	three	two : three
Stem-final vowel	<i>a</i> (<i>ǣ̃</i>), <i>e</i> , <i>o</i>	<i>i</i>	<i>á</i> , <i>i</i> , <i>u</i>
Consonant gradation	+	–	–
Examples	<i>manna-</i> ‘go’ <i>áddidalla-</i> ‘imitate’ <i>bárrå-</i> ‘eat’ <i>boahte-</i> ‘come’ <i>vuodjgoahte-</i> ‘begin to drive’ <i>goarro-</i> ‘sew’	<i>gěnnugi-</i> ‘stop’ <i>ságasti-</i> ‘talk’	<i>ǣl’lá-</i> : <i>ǣl’láji-</i> ‘regain vigor’ <i>lieggi-</i> : <i>lieggiji-</i> ‘heat’ <i>gávnnu-</i> : <i>gávnnuji-</i> ‘exist’

Note that in Lule Saami, there is limited labial vowel harmony. If the vowel of a stressed syllable is a short *ǣ̃*, then an older *a* in the following syllable is also realized as *ǣ̃*, e.g., *lǣhkǣ̃-* ‘read’ (cf. North Saami *lohka-* id.). This phonological process is completely predictable, and verbs with a stem ending in *ǣ̃-* are consequently treated here as belonging to the subclass of *a*-stem verbs. Similarly, an older *á* is realized as *ǣ̃* if the nucleus of the preceding syllable is *ǣ̃* (cf. Section 4.1). While this phonological rule implies the existence of contracted verbs in *ǣ̃*-, reflecting an older *á*-, no clear instance of such a verb is attested in the corpus. Further, when presented with inflected forms of *ǣ̃*-stem verbs (*jǣhttǣ̃jin* begin.to.migrate.PST.3PL, *gǣhttǣ̃jiv* awaken.PST.1SG), SA did not seem to recognize them.⁴⁶ Consequently, more

⁴⁶ For the meaning ‘awaken’, SA preferred either of the verbs *mǣredi-* or *ábbudi-*. I have not been able to find any instance of an *ǣ̃*-stem verb in the Forest materials of Wiklund, Lindgren or Collinder; see also the forms without labial vowel harmony documented by Wiklund, given

data are needed in order to confirm the existence of $\tilde{a}$ -stem verbs in this variety.

5.3 Finite forms and their suffixes

Finite forms of lexical verbs agree in person (first, second, third) and number (singular, plural) with the subject of a clause, and they are further inflected for tense (present, past) and mood (indicative, imperative); the verb *lä*- ‘be’ is also inflected for the conditional mood (see 5.8).⁴⁷ Not all inflectional categories are mutually compatible: tense is only distinguished in the indicative mood. Further, only second person imperative forms are attested in the corpus. A number of inflectional forms known from previous descriptions of Lule Saami are not attested in the data, including verbs in the dual number, the potential mood and the so-called “imperative II” mood. For a further discussion on these forms, see Section 6.

Table 4. Inflectional suffixes of finite verbs.

		SG	PL
PRS	1	-v	-p
	2	—	-bit, -hpit
	3	—	—
PST	1	-v	-jma, -jmě, -jme
	2	—	-jda, -jdě, -jde
	3	-j	-n
IMP	2	—	-t ⁴⁸

in Note 22. It should be pointed out that $\tilde{a}$ -stem verbs—all which are inchoative derivations—are uncommon even in the varieties of Lule Saami where they do occur. In addition to the above-mentioned *jáhttä*- and *gáhttjä*- (standard *gáhttsä*-), I am aware of only three such forms: *bår'rä*- ‘begin to eat’ (Spiik 1989: 86), *däbdddä*- ‘begin to know’ etc. (Grundström 1946–54 s.v. *täb’tät*) and *lähkkä*- ‘begin to read’ (Kintel 2012 s.v. *lähkkät*).

⁴⁷ Tense and mood can additionally be expressed with periphrastic constructions consisting of a finite form of *lä*- ‘be’ coupled with a non-finite verb. Such constructions include the present perfect (*läv gärrum* [be.PRS.1SG sew.PST.PTCP] ‘I have sewn’), the past perfect (*lědjiv dav vuojnnám* [be.PST.1SG DEM.SG.ACC see.PST.PTCP] ‘I had seen it’), the present progressive (*läv goarromin* [be.PRS.1SG sew.PROG] ‘I am sewing’), the past progressive (*lěj sjivnnjedime* [be.PST.3SG create.PROG] ‘was creating’) and the conditional (*lulun vuojnnet* [be.COND.3PL see.INF] ‘they would see’), all examples from SA. Also attested is a future construction consisting of *lä*- ‘be’ and the supine: SA *lä=k dân manátjit?* [be.PRS.2SG=Q 2SG.NOM go.SUP] ‘Are you going?’ (Sw. “Ska du till att fara?”); on this construction, see Ylikoski (2016: 221–226); cf. also Lehtiranta (1992: 103).

⁴⁸ In the Forest dialect, the ending of the second person plural imperative is -t for all inflectional classes. In the modern standard language, however, there is also a longer second person plural ending -htit in the (first) imperative of imparisyllabic and contracted verbs, e.g., *málesti-htit* cook-IMP.2PL, *lieggiji-htit* heat-IMP.2PL. While the latter suffix is used e.g. in the textbook of Tuolja & Angéus Kuoljok (2013: 141) and in the New Testament (*Ädä Testamentta*, 2003), it is noteworthy that this ending appears to be lacking in the earlier descriptions of the language

Finite verbs are formed both by adding suffixes and by stem-internal processes (consonant gradation and stem-final vowel alternations). As the non-linear morphology of finite verbs is largely dependent on the respective inflectional class of a verb, these processes are described in sections 5.5–5.9. The suffixes of finite verbs are portmanteau morphemes, which can simultaneously encode person, number, tense and mood. Several forms lack overt suffixes; these forms are indicated by a minus sign (–) in Table 4.⁴⁹

The first person singular suffix *-v* is identical in the present and past tenses; all other person and number suffixes differ in the present and past tenses, as well as in the imperative mood. There is no specific feature in common for the present tense forms, nor for the attested imperative forms. In several forms of the past tense there is an element *-j-*, which corresponds to the Proto-Saamic past tense marker **-j-* (Korhonen 1967: 194–197; 1981: 265–266); this element is however not a productive tense marker.

Most suffixes are identical for all three lexical verb classes. One exception is the suffix of the second person plural present, which has the allomorph *-bit* for parasyllabic and contracted verbs, but *-hpit* for imparisyllabic verbs, e.g., *manna-bit* go-PRS.2PL vs. *gǽnugi-hpit* stop-PRS.2PL. The ending *-bit* goes back to an older disyllabic suffix (Korhonen 1967: 317), cf. Standard Lule Saami *gulla-bihtit* hear-PRS.2PL. Such a disyllabic suffix is documented in the Forest dialect by Wiklund (*kulla-pēčtit* id., KBW 1895 II), as well as by Lindgren (*pivva-pettit* keep.warm-PRS.2PL, OL 1912: 4). In Collinder’s materials, however, I am only aware of forms with the shorter suffix, e.g., *vaddē-pit* give-PRS.2PL (Unna Sájvva, ULMA 15926: 106).

The past plural suffixes of the first and second person have several different allomorphs, ending either in *-ě*, *-e* or *-a*. One of the interviewed speakers consistently uses the allomorphs with a final *-a* (e.g., LE *maná-jma* go-PST.1PL, *sámásti-jma* speak.Saami-PST.1PL, *lě-jma* be-PST.1PL). The other speaker shows considerable variation: for imparisyllabic verbs and *lä-* ‘be’, he gives forms ending in *-e* (e.g., SA *gǽnugi-jme* stop-PST.1PL, *gǽnugi-jde* stop-PST.2PL, *lě-jme* be-PST.1PL, *lě-jde* be-PST.2PL), but for parasyllabic and contracted verbs, both forms with a final *-a* and with a final *-ě* are attested, see Table 5. It should however be noted that the forms with a final *-ě* in my data

made by Spiik (1989), Grundström (1946–54) and Wiklund (1915, 1891), as well as in the previous translation of the New Testament (*Åtå Testamenta*, 1903). Corresponding forms are however found in other Saamic languages, e.g., Pite Saami *málesti-htit* cook-IMP.2PL (Sjaggo 2015: 30, cf. Lehtiranta 1992: 153); see also Korhonen (1967: 324).

⁴⁹ In accordance with the Lule Saami orthography of the time, Lindgren’s texts have a final *-h* in the second person singular and in the third person plural present, as well as in the connegative, see, e.g., example (39) in 6.3.

are only attested in elicitation, and not in spontaneous speech. The conditional has a final *-a*: *lulu-jma* be.COND-1PL, *lulu-jda* be.COND-2PL (cf. Section 5.8).

Table 5. Intraspeaker variation in past tense suffixes (SA).

	<i>manna-</i> 'go'	<i>boahte-</i> 'come'	<i>goarro-</i> 'sew'	<i>dål'li-</i> 'make fire'	<i>lieggi-</i> 'heat'	<i>lijkku-</i> 'like'
PST.1PL	<i>maná-</i> <i>jma</i>	<i>bãdi-</i> <i>jma</i>	<i>goaro-</i> <i>jma</i>	<i>dål'li-</i> <i>jmě</i>	<i>lieggi-</i> <i>jma</i>	<i>lijkku-</i> <i>jma</i>
PST.2PL	<i>maná-</i> <i>jda</i>	<i>bãdi-</i> <i>jdě</i>	<i>goaro-</i> <i>jda</i>	<i>dål'li-</i> <i>jdě</i>	<i>lieggi-</i> <i>jda</i>	<i>lijkku-</i> <i>jda</i>

It is possible that this vacillation reflects a former dialectal variation: in Collinder's materials from Sádinvárre, for instance, imparisyllabic verbs have suffixes ending in *-e* and parisyllabic verbs have suffixes in *-a*, e.g., *a'tnalè-ìmè* hold-PST.1PL, *lokō-ìma* read-PST.1PL (1938: 37, 48). In Oajgek (Sw. *Rikti-Dockas*), on the other hand, Collinder has documented long *-e* for imparisyllabic verbs, but short *-ě* for parisyllabic and contracted verbs: *tār'pahè-ìmě* need-PST.1PL, *lokō-ime* read-PST.1PL, *àr^epi-jme* inherit-PST.1PL (ULMA 15926: 78–80).⁵⁰

5.4 Non-finite forms and their suffixes

In addition to finite forms, Lule Saami verbs are inflected for a number of non-finite forms. As non-finites, these forms are distinguished morphologically from their finite counterparts by their lack of inflection for person, number and tense. The following non-finite forms are attested for lexical verbs: the infinitive, the progressive, the past participle, the supine and the connegative.⁵¹ As regards the function of these forms, it can be noted that

⁵⁰ The system with *-e* for imparisyllabic verbs and *-a* for parisyllabic and contracted verbs is also found in Grundström's verb paradigms from Jåhkâmáhkke, whereas the system with *-e* for imparisyllabic verbs and *-ě* for parisyllabic and contracted verbs is found in the mountain dialects of Jiellevárre (Grundström 1946–54: 1847–1918). The latter system is also found, e.g., in the textbook of Wiklund (1915). In the modern standard language, all verb classes have final *-a*, e.g., *sámásti-jma* speak.Saami-PST.1PL, *maná-jma* go-PST.1PL, *lieggi-jma* heat-PST.1PL. Note additionally that for the imparisyllabic verbs quoted above, Collinder has noted a penultimate vowel *è* or *ě* where I have noted *i*, e.g., SA *rádedi-jme* discuss-PST.1PL. In the same position, Wiklund has noted short *ě*: *mälēstě-imē* cook-PST.1PL (KBW 1895 II).

⁵¹ In the research history of Lule Saami, the progressive has varyingly been referred to as the “state form” (Sw. *tillståndsformen*, Spiik 1989: 109), the “second gerund” (Sw. *gerundium II*, Wiklund 1915: 115) or the “bound gerund” (Ger. *Verbundenes Gerundium*, Grundström 1946–54: 1847–1918). Note also that confusingly enough, Lagercrantz (1939: 1207 et passim) refers to this form as the “first gerund” (Ger. *Gerund I*); see also Note 91.

unlike in the standard language, the connegative is the form of a negated verb only in the present tense and in the imperative. In the past tense, the past participle functions as a connegative instead (see Section 5.9). Noteworthy from a comparative perspective is the supine, perhaps most aptly described as a purposive converb (cf. Ylikoski 2016; 2022: 140). While frequent in Lule Saami, this form is largely unproductive in other Saamic languages.⁵²

Apart from being invariant as regards person, number and tense, non-finite forms are generally not inflected for mood. One exception to this rule is however the verb *lä-* ‘be’, which has dedicated connegative forms in the conditional and in the imperative (see Section 5.8).⁵³ The negative auxiliary verb lacks non-finite forms altogether.

Table 6. Suffixes of non-finite verbs.

	Suffix
infinitive (INF)	<i>-t</i>
progressive (PROG)	<i>-min, -me</i>
past participle (PST.PTCP)	<i>-m</i>
supine (SUP)	<i>-tjit, -ttjat, -ttjit</i>
connegative (CNG)	—
passive past participle (PASS.PST.PTCP)	<i>-dum, -um</i>

Non-finites are formed both by means of adding suffixes and by stem-internal processes; the non-linear morphology of these forms is described separately for each inflectional class (5.5–5.8). The suffixes of the infinitive (*-t*) and the past participle (*-m*) are identical for all three inflectional classes; the connegative lacks a suffix.⁵⁴ The progressive has the ending *-min* for parasyllabic and contracted verbs, and the ending *-me* for imparisyllabic verbs. Similarly, the supine has different allomorphs for different inflectional classes—for parasyllabic and contracted verbs, the ending is *-tjit*; for imparisyllabic verbs, the two interviewed speakers use different suffixes: SA

⁵² The supine also has other uses, e.g., functioning as the complement of an adjective in clauses such as SA *mánná=l buorak lágáttjit* [child.SG.NOM=be.PRS.3SG good.SG.NOM read.SUP] ‘the child is good at reading’ (cf. Tuolja & Angéus Kuoljok 2013: 147); see also the future construction mentioned in Note 47. In addition to Lule Saami, the supine is also productive in Western Čohkkiras North Saami, e.g., *viečča-žit ~ viečča-žik ~ viečča-sit* fetch-SUP. Although rare, it is also attested for parasyllabic verbs in modern Pite Saami, e.g. *stågá-tjit* play-SUP, *tjálá-tjit* write-SUP, *uru-tjit* live-SUP (Joshua Wilbur, p.c.); see also Lehtiranta (1992: 103) and Sjaggo (2015). The supine is very poorly attested in traditional Ume Saami (see Korhonen 1974: 198), but remarkably, it is found in the modern literary language (e.g. *njåmmielub váldá-tjit* [hare.SG.ACC take-SUP] ‘in order to take the hare’, Álgguogáhtie 2022).

⁵³ See also the “imperative II” connegative forms in examples (40) and (41), Section 6.4, documented by Wiklund.

⁵⁴ In Lindgren’s texts, the connegative has a final *-h*, see Note 49.

-*ttjat*, LE -*ttjit* (see further Korhonen 1974: 202 on this variation).⁵⁵ The suffixes of non-finite verbs are given in Table 6.

In addition to the abovementioned non-finite forms, it is also possible to posit a passive past participle, characterized by the suffix *-dum* for parisyllabic and contracted verbs, and *-um* for imparisyllabic verbs (e.g., LE *tjåle-dum* write-PASS.PST.PTCP ‘written’, SA *bidtji-dum* build-PASS.PST.PTCP ‘built’, *gatjost-um* hang-PASS.PST.PTCP ‘hung up’). This form can be used together with the verb *lä-* ‘be’ (5), as well as attributively (6).

- (5) *da* *lědjin* *gåttedum*
 DEM.PL.NOM be.PST.3PL kill.PASS.PST.PTCP
 ‘They (the bears) were killed.’ (SA)
- (6) *dat* *lä* *njålguk,* *basedum*
 DEM.SG.NOM be.PRS.3SG tasty.SG.NOM fry.PASS.PST.PTCP
 bierggo
 meat.SG.NOM
 ‘That’s tasty, fried meat.’ (SA, elicited)

In the Lule Saami standard language, the forms in *-(d)um* co-exist with longer forms, analyzable as past participles of passive verbs in *-(d)uvva-*, e.g., *tjåle-duvva-m* write-PASS-PST.PTCP ‘written’. In my data on the Forest dialect, however, only the synchronically unsegmentable forms in *-(d)um* are attested. These forms are further productive, as evidenced by their occurrence with recent loanwords, e.g., SA *jippsi-dum* put.in.cast-PASS.PST.PTCP ‘put in a cast (about a leg)’, cf. Sw. *gipsad* id.

Finally, a number of non-finites known from previous descriptions of Lule Saami are only poorly attested in my data. On these forms, see Section 6.5.

5.5 Inflection of parisyllabic verbs

Parisyllabic verbs are characterized by a large degree of non-linear morphology. Most notably, these verbs are subject to consonant gradation, a morphophonological process affecting the quantity and/or quality of word-internal consonants and consonant clusters.⁵⁶ This process affects the

⁵⁵ Lindgren has noted a form *låskala-dtjat* hit-SUP in Suobbat (OL 1913), implying a Grade III consonant in the supine suffix (cf. central Jåhkåmåhkke *målēsta-dt’šat* cook-SUP, Grundström 1946–54: 1886).

⁵⁶ Note that in the Forest dialect, the regular reflex of the Proto-Saamic consonant center **š* [ǫ] is a *d* /*t* that is exempt of consonant gradation (cf. Collinder 1938: 135). Consequently, there is a small number of verbs that lack consonant gradation in this variety, e.g., SA, LE *guode-* ‘leave’, *oade-* ‘sleep’, LE *gåde-* ‘knit’. This rule is however not confined to verbs, and it is

consonants between the two final nuclei of the verb stem—compare the forms *boahtá* come.PRS.3SG ‘comes’ and *boadá* come.PRS.2SG ‘you come’ or *ájádělliv* think.PST.1SG ‘I thought’ and *ájádaláv* think.PRS.1SG ‘I think’. Lule Saami consonants and consonant clusters alternate in three absolute grades, labeled Grade I–III (see, e.g., Spiik 1989: 22–29), but for the purposes of this paper, it is sufficient to distinguish between two relative grades: the strong grade and the weak grade.⁵⁷ In the examples above, the cluster *ht* in *boahtá* ‘comes’ is in the strong grade and the single consonant *d* in *boadá* ‘you come’ is in the weak grade; the geminate *ll* in *ájádělliv* ‘I thought’ is in the strong grade and the singleton *l* in *ájádaláv* ‘I think’ is in the weak grade. Table 7 shows the gradation patterns of parisyllabic verbs, indicating the weak grade (WK) and the strong grade (STR).

Table 7. Consonant gradation of parisyllabic verbs.

		SG	PL	Non-finite	
PRS	1	WK	STR	INF	STR
	2	WK	STR	PROG	STR
	3	STR	STR	PST.PTCP	STR
PST	1	STR	WK	SUP	WK
	2	STR	WK	CNG	WK
	3	WK	STR	PASS.PST.PTCP	WK
IMP	2	WK	STR		

Another characteristic feature of parisyllabic verbs is that in a number of forms, the final vowel of the verb stem changes. These stem-final vowel alternations can be quantitative (e.g., short *a* to long *á*), qualitative (e.g., short *a* to short *i*) or both (e.g., long *o* to short *u*). As mentioned above, parisyllabic verbs can be divided into three different subclasses, based on the final nucleus of the inflectional stem: *a*-stem verbs (including *á*-stem verbs), *e*-stem verbs and *o*-stem verbs. These stem-final nuclei are subject to different stem-final vowel alternations.

In many cases, the change of a stem-final vowel causes an umlaut in the preceding nucleus, e.g., SA *děbmuv* glue.PST.1SG ‘I glued’, cf. *damov* glue.PRS.1SG ‘I glue’, *dabmo-* ‘glue’ (for an overview of umlaut in Standard

therefore treated here as a regular part of morphophonology, rather than as an exception in the verb inflection.

⁵⁷ A third relative grade, called the “extra strong” or “overlong” grade, is not relevant for the inflectional forms treated here (see however Collinder 1938: 149). Note further that in Swedish- and Norwegian-language descriptions of Saami, there is a clearer terminological distinction between the absolute and relative grades, with the former being referred to as *grad* ‘grade’ and the latter referred to as *stadium* ‘stage’.

Lule Saami, see Angéus Kuoljok 1997: 45–47; for details concerning the Forest dialect, see Collinder 1938: 57–71). The stem-final vowel alternations of parisyllabic verbs are given in Table 8. Note that the vowel alternations $\bar{a} > \bar{a}$ and $e > \bar{a}$, given in parentheses, are only applicable to verbs in which the penultimate nucleus is $\bar{a}$ (see Section 5.2).

Table 8. Stem-final vowel alternations of parisyllabic verbs.

Stem-final vowel	Form(s)	Example(s)
$a > \acute{a}$ ($\bar{a} > \bar{a}$)	PRS.3SG	<i>manná</i> PRS.3SG < <i>manna</i> - ‘go’ <i>bárrá</i> PRS.3SG < <i>bárrá</i> - ‘eat’
$a > i$	PRS.3PL, PST.1SG, PST.2SG, PST.3PL, IMP.2PL	<i>měnni</i> PRS.3PL < <i>manna</i> - ‘go’
$e > \acute{a}$ ($e > \bar{a}$)	PRS.1SG, PRS.2SG, PRS.3SG, PST.PTCP, SUP	<i>boahtá</i> PRS.3SG < <i>boahte</i> - ‘come’ <i>gáddá</i> PRS.3SG < <i>gádde</i> - ‘kill’
$e > i$	PRS.3PL, all PST forms, IMP.2PL	<i>báhti</i> PRS.3PL < <i>boahte</i> - ‘come’
$o > u$	PRS.3SG, PRS.3PL, PST.1SG, PST.2SG, PST.3PL, IMP.2PL, PST.PTCP, SUP	<i>gárru</i> PRS.3SG < <i>goarro</i> - ‘sew’

In addition to these vowel alternations, there is also a phonological process that is characteristic of Lule Saami, namely the lengthening of a short a to long $\acute{a}$ when the preceding nucleus is stressed and when both the preceding consonant and nucleus are short, e.g., *maná* go.IMP.2SG ‘go!’ < *manna*- ‘go’ (cf. North Saami *mana* id., with a short second syllable a). In stems which are subject to labial vowel harmony (see Section 5.2), short $\bar{a}$ is lengthened to $\bar{a}$, e.g., *bárrá* eat.IMP.2SG ‘eat!’ < *bárrá*- ‘eat’.

In the following tables (9–11), sample paradigms are given for the a -stem verb *manna*- ‘go’, the e -stem verb *boahte*- ‘come’ and the o -stem verb *goarro*- ‘sew’. In those cases where a particular inflectional form is unattested for these three verbs, the corresponding form is given for another verb of the same inflectional subclass.

Table 9. Inflection of the parisyllabic a -stem verb *manna*- ‘go’ (SA).

	SG	PL	Non-finite	
PRS	1	<i>maná-v</i> <i>manna-p</i>	INF	<i>manna-t</i>
	2	<i>maná</i> <i>manna-bit</i>	PROG	<i>manna-min</i>
	3	<i>manná</i> <i>měnni</i>	PST.PTCP	<i>manna-m</i>
PST	1	<i>měnni-v</i> <i>maná-jma</i>	SUP	<i>maná-tjit</i>
	2	<i>měnni</i> <i>maná-jda</i>	CNG	<i>maná</i>

	3	<i>maná-j</i>	<i>měnni-n</i>	PASS.PST.PTCP	<i>dagá-dum</i> ⁵⁸
IMP	2	<i>maná</i>	<i>měnni-t</i>		

Table 10. Inflection of the parisyllabic *e*-stem verb *boahte-* ‘come’ (SA).

		SG	PL	Non-finite	
PRS	1	<i>boadá-v</i>	<i>boahte-p</i>	INF	<i>boahte-t</i>
	2	<i>boadá</i>	<i>boahte-bit</i>	PROG	<i>boahte-min</i>
	3	<i>boahtá</i>	<i>bāhti</i>	PST.PTCP	<i>boahtá-m</i>
PST	1	<i>bāhti-v</i>	<i>bādi-jma</i>	SUP	<i>tjálá-tjit</i> ⁵⁹
	2	<i>bāhti</i>	<i>bādi-jdē</i>	CNG	<i>boade</i>
	3	<i>bādi-j</i>	<i>bāhti-n</i>	PASS.PST.PTCP	<i>vuotje-dum</i> ⁶⁰
IMP	2	<i>boade</i> ⁶¹	<i>bāhti-t</i>		

Table 11. Inflection of the parisyllabic *o*-stem verb *goarro-* ‘sew’ (SA).

		SG	PL	Non-finite	
PRS	1	<i>goaro-v</i>	<i>goarro-p</i>	INF	<i>goarro-t</i>
	2	<i>goaro</i>	<i>goarro-bit</i>	PROG	<i>goarro-min</i>
	3	<i>gārru</i>	<i>gārru</i>	PST.PTCP	<i>gārru-m</i>
PST	1	<i>gārru-v</i>	<i>goaro-jma</i>	SUP	<i>gāru-tjit</i>
	2	<i>gārru</i>	<i>goaro-jda</i>	CNG	<i>goaro</i>
	3	<i>goaro-j</i>	<i>gārru-n</i>	PASS.PST.PTCP	<i>goaro-dum</i>
IMP	2	<i>goaro</i>	<i>gārru-t</i>		

5.6 Inflection of imparisyllabic verbs

Compared to parisyllabic verbs, the morphology of imparisyllabic verbs is to a large degree concatenative, and involves only two morphophonological processes. First, in some inflectional forms (PRS.1SG, PRS.2SG, IMP.2SG, PST.PTCP, CNG, SUP) the final *i* of the inflectional stem changes into *a*, e.g., *gēnugav* stop.PRS.1SG < *gēnugi-* ‘stop’. Note that the supine suffix *-ttjit* of one speaker causes the *a* of the penultimate nucleus to be realized as *ě* instead: LE *vuojaděttjit* swim.SUP (cf. SA *vuojadattjat* id.).

⁵⁸ *dahka-* ‘make’

⁵⁹ *tjálle-* ‘write’

⁶⁰ *vuohije-* ‘shoot’

⁶¹ In addition to the expected forms with word-initial stress, Collinder has also noted the second person singular imperative of this word with main stress on the second syllable: *poⁿtě-* ‘come!’ (1938: 43). The imperative forms in my data have word-initial stress, just like other inflectional forms, e.i., SA, LE ‘*boade* id.

Second, in the third person singular present, the inflectional stem loses its final vowel, e.g., SA *hállán* fly.away.PRS.3SG < *hálláni*- ‘fly away’. As a result of phonotactical constraints, the final consonant or consonants of the inflectional stem may also undergo change in this form (see Tuoljja & Angéus Kuoljok 2013: 50–51, Wiklund 1915: 74, Note 1), e.g., SA *tjuormas* hail.PRS.3SG < **tjuormast* < *tjuormasti*- ‘hail’,⁶² *dárbaj* need.PRS.3SG < **dárbah* < *dárbahi*- ‘need’.⁶³ For LE, I have however also noted a form in *-a*: *ráhkesta* love.PRS.3SG.

The final vowel of the stem is also lost in the passive past participle, in which the suffix *-um* is added to the final consonant of the inflectional stem, e.g., *gatjostum* hang.PASS.PST.PTCP < *gatjosti*- ‘hang (trans.)’. A sample paradigm of the imparisyllabic verb *gěnugi*- ‘stop’ is given in Table 12.

Table 12. Inflection of the imparisyllabic verb *gěnugi*- ‘stop’ (SA).

		SG	PL	Non-finite	
PRS	1	<i>gěnuga-v</i>	<i>gěnugi-p</i>	INF	<i>gěnugi-t</i>
	2	<i>gěnuga</i>	<i>gěnugi-hpit</i>	PROG	<i>gěnugi-me</i>
	3	<i>gěnuk</i> ⁶⁴	<i>gěnugi</i>	PST.PTCP	<i>gěnuga-m</i>
PST	1	<i>gěnugi-v</i>	<i>gěnugi-jme</i>	SUP	<i>gěnuga-ttjat</i>
	2	<i>gěnugi</i>	<i>gěnugi-jde</i>	CNG	<i>gěnuga</i>
	3	<i>gěnugi-j</i>	<i>gěnugi-n</i>	PASS.PST.PTCP	<i>gatjost-um</i> ⁶⁵
IMP	2	<i>gěnuga</i>	<i>gěnugi-t</i>		

5.7 Inflection of contracted verbs

Contracted verbs are characterized by having two stem allomorphs: one shorter stem, ending in either of the vowels *á*-, *i*- or *u*- (on contracted verbs in *ā̃*-, see 5.2), and one longer stem, ending in either of these vowels followed by an element *ji*-. From a historical point of view, the stem ending in *ji*- is the older one, from which the shorter (or contracted) stem has been derived. In

⁶² For LE, I have noted an initial *s*- in this word: *suormastij* hail.PST.3SG, *suormastam* hail.PST.PTCP. Note further that Collinder (1938: 54) has also documented a third person singular form *mālēst* ‘cooks’, with a retained word-final cluster.

⁶³ In some cases, the third person singular *-j* appears to have spread to other forms through analogy: SA *vállajav* lie.down.PRS.1SG, *vállajiv* lie.down.PST.1SG, *huomajiv* realize.PST.1SG, LE *huomaja* realize.CNG (standard *vellahav*, *vellahiv*, *huomahiv*, *huomaha*). For SA, I have noted the forms *dárbaja* need.CNG, *dárbajam* need.PST.PTCP (standard *dárbaha*, *dárbaham*) in spontaneous speech and *dárbaha* need.CNG, *dárbahav* need.1SG.PRS (standard id.) in elicitation.

⁶⁴ The difference between word-final *-k* and word-internal *-g*- is orthographical: both graphemes represent /k/. Similarly, both word-final *-t* and word-internal *-d*- represent /t/ in forms such as SA *vuosjvet* crave.PRS.3SG and *vuosjvedav* crave.PRS.1SG (standard *vuosvet*, *vuosvedav*).

⁶⁵ *gatjosti*- ‘hang (trans.)’.

the contemporary verb inflection, the distribution of the two stems is tied to specific cells in the inflectional paradigm (PRS.3PL, PST.1SG, PST.2SG, PST.3PL, IMP.2PL).

In previous descriptions of the Lule Saami Forest dialect, the third person singular present of *i*-stem verbs has been documented with a final long *-e*, e.g., *reštē* cross.PRS.3SG, cf. *reštit* cross.INF (Collinder 1938: 61), *tāddnē* tin.PRS.3SG, cf. *tēddnit* tin.INF (KBW 1895 II).⁶⁶ Similarly, *u*-stem verbs have been documented with a final long *-o* (UPA *o*) in the same form, e.g., *hāχtō* must.PRS.3SG (cf. *heχtūji* must.PRS.3PL, Collinder 1938: 62, 59).⁶⁷

Table 13. Inflection of the contracted verb *lieggi*- ‘heat’ (SA).

		SG	PL	Non-finite	
PRS	1	<i>lieggi-v</i>	<i>lieggi-p</i>	INF	<i>lieggi-t</i>
	2	<i>lieggi</i>	<i>lieggi-bit</i>	PROG	<i>lieggi-min</i>
	3	<i>liegge</i>	<i>lieggiji</i>	PST.PTCP	<i>lieggi-m</i>
PST	1	<i>lieggiji-v</i>	<i>lieggi-jma</i>	SUP	<i>lieggi-tjit</i>
	2	<i>lieggiji</i>	<i>lieggi-jda</i>	CNG	<i>lieggi</i>
	3	<i>lieggi-j</i>	<i>lieggiji-n</i>	PASS.PST.PTCP	<i>lieggi-dum</i>
IMP	2	<i>lieggi</i>	<i>lieggiji-t</i>		

In my data, the analysis of these forms is complicated by the fact that both speakers tend to pronounce second syllable *e* and *o* as fairly close monophthongs, making them difficult to distinguish qualitatively from their short counterparts *i* and *u* respectively.⁶⁸ Nevertheless, there is a clear opposition between forms such as LE *hāhtto* must.PRS.3SG and *hehttu* must.PRS.2SG, reflected in the quality of the word’s first vowel (cf. also SA *bājssto* get.ruined.PRS.3SG and *bejsstum* get.ruined.PST.PTCP). Tentatively, I have also noted an opposition between forms such as SA *rēsste* cross.PRS.3SG and *rēssti* cross.CNG, although this contrast is admittedly not as clear-cut as that of *u*-stem verbs.⁶⁹ A sample paradigm of the contracted verb *lieggi*- ‘heat’ is given in Table 13.

⁶⁶ Note that while the alternation of a second syllable *i* and *e* triggers different first syllable vowels for the speaker interviewed by Wiklund, Collinder’s data show variation in this respect (for examples, see Collinder 1938: 61).

⁶⁷ The form *hāχtō* (Collinder 1938: 59), with a long vowel in the first syllable, appears to be a misprint, cf. *hāχtō* (1938: 62), *hāxtō* id. (ULMA 15932: 104).

⁶⁸ See also the native-speaker spelling of Sádinvárre as <Satimvarri> in Figure 1. It should additionally be pointed out that the vowel denoted with the quality *e* in Collinder’s monograph on the dialect (e.g., *passēt* ‘to fry’, *vaddēt* ‘to give’, 1938: 170) is regularly corresponded by the more close *é* in his notebooks (e.g., *passēt*, *vaddēt*, ULMA 15935: 24).

⁶⁹ While third person singular present forms in *-e*, *-o* are also found in more northern varieties of the language (Grundström 1946–54: 1905, 1911), Standard Lule Saami has *-i*, *-u* in the corresponding forms, e.g., *hāhttu* must.PRS.3SG, *rassti* cross.PRS.3SG. In Collinder’s notes from

5.8 Inflection of *lä-* ‘be’

The verb *lä-* ‘be’, which functions both as a copula and as an auxiliary, is the most common verb in the corpus. While having suffixes identical to those of imparisyllabic verbs, this verb has a partially irregular inflection. Most notably, the infinitive, past participle and imperative forms are suppletive, using forms identical to those of the lexical verb *ârro-* ‘live, reside’.⁷⁰ The inflectional paradigm of *lä-* ‘be’ is given in Table 14; for the highly irregular negation of this verb, see Section 5.10.

Table 14. Inflection of *lä-* ‘be’ (SA).

		SG	PL	Non-finite	
PRS	1	<i>lä-v</i>	<i>lä-p</i>	INF	<i>ârro-t</i>
	2	<i>lä</i>	<i>le-hpit</i>	PROG	N/A
	3	<i>lä, =l</i>	<i>lä, =l</i>	PST.PTCP	<i>ârru-m</i>
PST	1	<i>lëdji-v</i>	<i>lë-jme</i>	SUP	N/A
	2	<i>lëdji</i>	<i>lë-jde</i>	CNG	N/A
	3	<i>lë-j</i>	<i>lëdji-n</i>	PASS.PST.PTCP	N/A
COND	1	<i>lulu-v</i>	<i>lulu-jma</i>	COND.CNG	<i>lulu</i>
	2	<i>lulu</i>	<i>lulu-jda</i>		
	3	<i>lulu-j</i>	<i>lulu-n</i>		
IMP	2	<i>âro</i>	<i>ârru-t</i>		

The element *-dj-* in the past tense of the first and second person singular, as well as the third person plural, goes back to a geminated form of the Proto-Saamic past tense marker **-j-* (Korhonen 1967: 195–196; 1981: 265). In addition to the forms given in Table 14, I have also noted a few parallel forms in *li-*: SA, LE *lip* be.PRS.1PL, SA *li* be.PRS.3PL, *lij* be.PST.3SG.⁷¹

Risuk (Sw. *Risappi*) there is a final *-j* in the third person singular present of contracted verbs, such as *tällit* ‘to pack shoes with hay’ and *smäxkkut* ‘to taste’, making these forms identical to the corresponding past tense forms: *tällij*, *smäxkkuj* PRS.3SG = PST.3SG (cf. *tälli*, *smäxkku* PRS.2SG, ULMA 15924). This trait is also attested in Pite Saami (Lehtiranta 1992: 93).

⁷⁰ Wiklund’s materials on the Forest dialect include an infinitive form *lëœk’ët* be.INF alongside the suppletive form *ôrrowt*, as well as imperative forms of the same stem: *lëœk’is* be.IMP.3SG, *lëœk’is* be.IMP.3SG (KBW 1895 II). Collinder, on the other hand, has noted that Wiklund’s form *lëœk’ët* “ought to be altogether missing in the Forest Lapp area” (“torde alldeles saknas i skogs-lappsområdet”, ULMA 15927: 30). In the Lule Saami standard language, both forms in *ârro-* and in *liehke-* are used, e.g., *ârrot* ~ *liehket* be.INF; forms in *liehke-* however appear to be more common in Norway than in Sweden (cf. Spiik 1989: 69).

⁷¹ In the Lule Saami standard language, the verb ‘be’ has two parallel sets that differ in their respective vocalism. One of these sets (e.g., *lev* ‘I am’, *ledjiv* ‘I was’) is typically described as occurring clause-initially, whereas the other set (e.g., *lav* ‘I am’, *lidjiv* ‘I was’) occurs in other positions (Spiik 1989: 68–71); note however that in some texts, such as the New Testament (*Adâ Testamentta*, 2003), forms in *le-* are used rather than forms in *la-* (cf. Angéus Kuoljok 2002b: 16). In the Forest dialect, I have not been able to confirm two separate sets of the verb

A characteristic feature of the verb *lä*- ‘be’ is that the third person singular present *lä* is often shortened to the non-syllabic clitic *=l* if the preceding word ends in a vowel (7). This shortening is however not obligatory, as shown in (8).

- (7) *muv vuorrasāmos vieljja=l lågev*
 1 SG.GEN old.SPRL.ATTR brother.SG.NOM=be.PRS.3SG ten.SG.NOM
jage musstě vuorrasäbbo
 year.SG.GEN 1 SG.ELA old.CMPR.SG.NOM
 ‘My oldest brother is ten years older than me.’ (LE)

- (8) *muvva vieljja lä årudagån*
 1 SG.GEN brother.SG.NOM be.PRS.3SG dwelling.SG.INE
 ‘My brother is at home.’ (SA, elicited)

The homophonous third person plural present *lä* can also be cliticized as *=l*, an example of which is given in (9). As in the singular, however, this shortening is not obligatory, as shown in (10).⁷²

- (9) *gusá=l val årrum*
 cow.PL.NOM=be.PRS.3PL just be.PST.PTCP
 ‘(We) have had cows (but not poultry).’ (SA)

- (10) *da lä vuorrasäppo gå mån*
 DEM.PL.NOM be.PRS.3PL old.CMPR.PL.NOM than 1 SG.NOM
 ‘They are older than me.’ (LE)

While the second person singular present *lä* is homophonous with the full forms of the third person singular and plural, this form is not attested as cliticized, but only in its full form (11). For an overview of the cliticization of the verb ‘be’ in the western Saamic languages, see Siegl (2019).

- (11) *gån bárnne lä dån?*
 who.SG.GEN son.SG.NOM be.PRS.2SG 2SG.NOM
 ‘Whose son are you?’ (LE)

lä- ‘be’ as used depending on position in the clause. See also Wiklund (1915: 77–78), who describes different realizations of this verb as governed by prosodic factors rather than by word order proper.

⁷² It is perhaps no coincidence that in example (10), the uncliticized form *lä* follows a monosyllabic word, whereas in (9), the cliticized *=l* attaches to a disyllabic form.

Historically, the conditional was a productive mood for all verb classes in Lule Saami but the negative auxiliary (see Section 6.3). In the present corpus, however, conditional forms are only attested for *lä-* ‘be’. In these forms, person and number suffixes identical to those of the past tense are added to a stem *lulu-*, e.g., *lulu-n* be.COND-3PL ‘they would’. Forms in *lulu-* can be coupled with a verb in the infinitive, such as *oasstet* ‘to buy’ in (12).

- (12) *mån* *luluv* *ådä* *vuojánáv*
 1SG.NOM be.COND.1SG new.ATTR vehicle.SG.ACC
oasstet *jus* *lulun* *rudá*
 buy.INF if be.COND.3PL money.PL.NOM
 ‘I would buy a new car (lit. vehicle) if I had money.’ (SA, elicited)

The conditional in *lulu-* does not need to precede the infinitive, but may occur after it instead, e.g., SA *gávnnat lulun* [find.INF be.COND.3PL] ‘they would find’. The only non-finite conditional form is the conditional connegative, which consists of the bare stem *lulu*.⁷³

Finally, a note on *lä-* ‘be’ and interrogatives is in order. In Lule Saami, polar questions are most commonly encoded by the second-position clitic =*gu* ~ =*gus*.⁷⁴ This interrogative clitic can be shortened to =*k* when attaching to a verb ending in a vowel, e.g., LE *jugá=k* drink.PRS.2SG=Q. In the standard language, however, =*k* does generally not attach to the verb ‘be’ (Tuolja & Angéus Kuoljok 2013: 20), nor to the negative auxiliary.⁷⁵ In the Forest dialect, on the other hand, there is no such restriction, as shown in examples (13) and (14). This trait is also found in the notebooks of Collinder, see example (15) from a speaker born in Suobbat.

⁷³ From a strictly synchronic perspective, it would also be possible to regard *lulu-* as a defective verb with the meaning ‘would’, rather than as an inflectional form of *lä-* ‘be’ (cf. Ylikoski 2022: 139). Note also that while finite forms of *lulu-* have suffixes identical to those of the past indicative, the conditional connegative lacks a suffix, agreeing in this respect with the present connegative of other verbs (i.e., *lulu* and not **lulum*). Here it can also be mentioned that in written Lule Saami, one may occasionally see *lulu-* inflected as a contracted verb, e.g., *lulujin* be.COND.3PL (Olsen & Utsi Vars 2012: 63).

⁷⁴ The host of this clitic is most often a verb, but it may also belong to another word class, e.g., SA *ieddne=gu lij?* [mother.SG.NOM=Q be.PST.3SG] ‘Mother, was it?’, *vidás=gu?* five.people.SG.NOM=Q ‘Are there five (of us, speakers of the Forest dialect)?’.

⁷⁵ I am aware of only two instances in the written standard language where the shortened clitic =*k* attaches to the verb ‘be’: *lettji=k* be.POT.3PL=Q (Tuolja & Angéus Kuoljok 2013: 152) and *lähppe=k* be.PRS.2DU=Q (Tuolja 1987: 38, 43). In both of these examples, the form to which =*k* attaches is disyllabic, suggesting that stress may be a decisive factor in the shortening of this clitic.

- (13) *lǎ=k* *dǎn* *vuojnnám* *riek* *stuor*
 be.PRS.2SG=Q 2SG.NOM see.PST.PTCP really large.ATTR
tjuormmasijt?
 hail.PL.ACC
 ‘Have you seen really large hail?’ (SA)

- (14) *lǎ=k* *suomagiellaj?*
 be.PRS.3SG=Q Finnish.SG.ILL
 ‘Is it in Finnish?’ (LE, about a road sign)

- (15) *lǎ=k* *ton* *kullam*
 be.PRS.2SG=Q 2SG.NOM hear.PST.PTCP
 ‘Have you heard?’ (Suobbat, ULMA 15926: 42)

The shortened interrogative clitic =*k* may also attach to *lǎ-* ‘be’ in indirect questions, as shown in (16).

- (16) *jus* *lǎ* *dě* *tjuosek* *jěli*
 if be.PRS.3SG DPT reindeer.with.a.blaze.SG.NOM or
lǎ=k *dě* *rijhmek*
 be.PRS.3SG=Q DPT dark.reindeer.SG.NOM
 ‘...if it’s a reindeer with a blaze, or if it’s a dark reindeer...’ (LE)

As with other verbs, the interrogative clitic may however remain unshortened also when attaching to *lǎ-* ‘be’, as seen in (17) and (18); see also Section 5.9 on the interrogative clitic =*k* and the negative auxiliary.

- (17) *lǎ=gus* *dǎn* *skibás?*
 be.PRS.2SG=Q 2SG.NOM sick.SG.NOM
 ‘Are you sick?’ (SA, elicited)

- (18) *lǎ=gu* *duv* *áhhttje* *gullam*
 be.PRS.3SG=Q 2SG.GEN father.SG.NOM hear.PST.PTCP
dav?
 DEM.SG.ACC
 ‘Has your father heard that?’ (SA)

5.9 Inflection of the negative auxiliary

In Lule Saami, standard negation is expressed by means of a negative auxiliary verb, the inflection of which is irregular (for a survey of negation in Lule

Saami, see Ylikoski & Kejonen, forthcoming). In the Forest dialect, this verb has the inflectional stem *i-* in the indicative singular, and *ä-* in the indicative plural. The third person singular *ij* has an irregular final *-j* (for a discussion on the origin of this form, see Korhonen 1967: 230–234; 1981: 272–273).

The forms of the imperative mood are consistent with a suppletive parisyllabic stem **alle-*, which is subject to both consonant gradation and stem-final vowel alternation (the latter in turn inducing an umlaut *a > ě* in the first syllable). It should however be noted that forms with an initial *ě-* are also found in dialects of Lule Saami which lack the characteristic *ě*-inducing umlaut of the Forest dialect, e.g., Unna Tjerusj *alē ~ elē* NEG.IMP.2SG, *allit ~ elliit* NEG.IMP.2PL (Grundström 1946–54: 1918).

Table 15. Inflection of the negative auxiliary (SA).

		SG	PL
PRS	1	<i>i-v</i>	<i>ä-p</i>
	2	<i>i</i>	<i>e-hpit</i>
	3	<i>ij</i>	<i>ä</i>
IMP	2	<i>ale</i>	<i>ělli-t</i>

In addition to the forms in Table 15, the corpus also contains a longer variant of the first personal singular form, SA *ivva*, in which the final consonant has been lengthened and a paragogic *-a* has been added (cf. the forms SA, LE *muṽ ~ muṽva* 1SG.GEN, SA *dav ~ davva* DEM.SG.ACC; see also Wiklund 1891: 117).⁷⁶

In the Lule Saami standard language, the negative auxiliary verb is inflected not only in person, number and mood, but also in tense. That is, in addition to the present tense forms, there is also a full set of past tense forms. Such forms are also attested for the Forest dialect: *iṭṛiw* NEG.PST.1SG, *iṭṛi* NEG.PST.2SG, *ēṭimē* NEG.PST.1PL, *ēṭitē* NEG.PST.2PL (Collinder 1938: 37, 104), *ēimē* NEG.PST.1PL, *ēitē* NEG.PST.2PL, *āṭjkā* NEG.PST.3DU (KBW 1895 II). In the present corpus, however, no such forms are produced spontaneously by the two interviewed speakers—not in free speech, nor in elicitation. Instead, present tense forms of the negative auxiliary are combined with the past

⁷⁶ I have noted corresponding disyllabic forms of the negative auxiliary verb in a conversation with a speaker of Pite Saami: *ivva* NEG.PRS.1SG, *ijja* NEG.PRS.3SG. In Bålggem, Collinder (1938: 19) has noted an emphatic third person singular of the negative auxiliary, *i:hi:j*, with an epenthetic *h*. Such an epenthetic *h* is also found elsewhere in the language; for instance, I have noted the accusative singular forms *āhav* ‘the letter *a*’ and *behav* ‘the letter *b*’ in Unna Tjerusj (see also Wiklund 1891: 8, 117–118, 185).

participle to form a past tense reading.⁷⁷ Compare the present tense clause in (19) with the past tense clause in (20).

- (19) *mån* *iv* *máhte* *vuodjat*
 1SG.NOM NEG.PRS.1SG can.CNG swim.INF
 ‘I can’t swim.’ (SA)

- (20) *mån* *iv* *máhttám* *mav=gě*
 1SG.NOM NEG.PRS.1SG can.PST.PTCP REL.SG.ACC=NPM
dárogielav
 Swedish.SG.ACC
 ‘I didn’t know any Swedish (when starting school).’ (SA)

However, when presented with the Standard Lule Saami third person singular past tense form *ittijj*, SA readily accepted it and produced the clause (21), in which this form is coupled with the past participle. This construction is also attested in Collinder’s materials from a speaker born in Unna Sájvva (22).

- (21) *muvva* *ieddne* *ittijj* *gǽrrum*
 1SG.GEN mother.SG.NOM NEG.PST.3SG sew.PST.PTCP
 ‘My mother didn’t sew.’ (SA, elicited)

- (22) *it’šiv* *p’öi’xkām*
 NEG.PST.1SG understand.PST.PTCP
 ‘I didn’t understand.’ (Unna Sájvva, ULMA 15936)

Note that this construction differs from the one prescribed in the written standard language, where past tense forms of the negative auxiliary are coupled with the indicative connegative instead, e.g., *ittijj goaro* [NEG.PST.3SG sew.CNG] ‘didn’t sew’. In Lindgren’s texts from the area, both this latter type of negation (23) and the type with present tense forms of the negative auxiliary (24) are attested.⁷⁸

⁷⁷ In my data from Unna Tjerusj, I have similarly only noted past negatives of this kind, e.g., *ij boahátám* [NEG.PRS.3SG come.PST.PTCP] ‘he did not come’ (cf. however Grundström 1946–54: 1917). In more northern and eastern Saamic varieties, this is the only way of forming a past negative; the loss of tense inflection for the negative auxiliary in these varieties may be due to influence of Finnic (cf. Bergsland 1946: 170, Korhonen 1967: 214; 1981: 310, Sammallahti 1998: 76–77; 224–225). It should however be noted that this “northern” type of negation is also found further south within Lule Saami. The same construction is also attested in Pite Saami, e.g., *i neita mannám* [NEG.PRS.3SG girl.SG.NOM go.PST.PTCP] ‘the girl did not go’ (ULMA 22200: 4).

⁷⁸ Collinder’s materials also feature the standard type of negation, albeit in a recited text (cf. Note 13), e.g., *i’i sitā* [NEG.PST.3SG want.CNG] ‘didn’t want’ (1938: 170). Although Wiklund’s

- (23) *itji* *tålåv* *åttjoh*
 NEG.PST.3SG fire.SG.ACC receive.CNG
 ‘...(she) didn’t get fire...’ (Stuor Sávvva, OL 1912: 3)

- (24) *I* *tat* *maksam.*
 NEG.PRS.3SG DEM.SG.NOM pay.off.PST.PTCP
 ‘It didn’t pay off.’ (Stuor Sávvva, OL 1912: 3)

Here it can also be mentioned that as in the case of *lä-* ‘be’, there is no restriction in this dialect against the combination of the negative auxiliary and the shortened interrogative clitic =*k* (cf. 5.8); an example is given in (25). The shortening of the question clitic is however not obligatory, as shown in (26).

- (25) *i=k* *dån* *jugá* *tjáhpav?*
 NEG.PRS.2SG=Q 2SG.NOM drink.CNG coffee.SG.ACC
 ‘Don’t you drink coffee (lit. black)?’ (LE, elicited)

- (26) *i=gus* *dån* *jugá* *tjáhpav?*
 NEG.PRS.2SG=Q 2SG.NOM drink.CNG coffee.SG.ACC
 ‘Don’t you drink coffee (lit. black)?’ (SA, elicited)

5.10 Negation of *lä-* ‘be’

The negation of *lä-* ‘be’ is highly irregular in the Lule Saami Forest dialect, and the following description should be seen as tentative. The obtained negated forms of *lä-* ‘be’ are given in Table 16.

Table 16. Negation of *lä-* ‘be’ (SA).

		SG	PL
PRS	1	<i>i-v lě</i>	<i>ällě-p</i>
	2	<i>illě</i>	<i>ällě-bit</i>
	3	<i>ile</i>	<i>ällě</i>
PST	1	<i>i-v li, illi-v</i>	<i>ělli-jma</i>
	2	<i>illi</i>	<i>ělli-jda</i>
	3	<i>illim</i>	<i>ělli-n</i>
IMP	2	<i>ale åro</i>	<i>ělli-t åro</i>

materials feature tokens of the negative auxiliary inflected for the past tense, these forms are only noted in isolation, and not together with a negated verb.

In most cells of the paradigm, the negative auxiliary and *lä-* ‘be’ have been amalgamated, e.g., *illě* ‘you are not’.⁷⁹ The third person singular present *ile* has an unexpected long *e* in the second syllable. Similar (albeit not identical) forms have also been documented in the Forest dialect by Collinder (*ilē*, ULMA 15936) and Wiklund (*ilēεε*, KBW 1895 II). The corresponding past tense form *illim* stems from an amalgamation of the negative auxiliary and a past connegative form of *lä-* ‘be’ (cf. Standard Lule Saami *ij lim*). Again, a similar (but not identical) form has been documented by Wiklund (*ilim*, KBW 1895 II). A form with a geminate consonant is attested in Lindgren’s materials from Suobbat (*illim*, OL 1913). In the first person singular of the past tense, I have noted both an amalgamated form *illiv* and a bipartite *iv li* ‘I was not’. The amalgamated form is also attested by Collinder (*illiw*, 1938: 45).

The plural forms are strikingly regular. In these forms, there is a stem *ällě-*, which takes the regular subject suffixes of parisyllabic verbs, e.g., *ällě-p* ‘we are not’, *ällě-bit* ‘you are not’. This stem reflects the plural *ä-* of the negative auxiliary having been amalgamated with *lä-* ‘be’. In the past tense forms, where the vowel of the second syllable is *i*, the first vowel is more close than in the corresponding present tense forms, e.g., *ěllin* ‘they were not’. To the best of my knowledge, forms like these have not been previously described for Lule Saami (cf. Standard Lule Saami *ep la* [NEG.PRS.1PL be.CNG] ‘we are not’, with the verb ‘be’ in the connegative and with subject marking on the negative auxiliary). Forms in *ällě-* can be used as an auxiliary, coupled with a verb in the past participle (27). They can also be used as a verb ‘be not’, as shown in (28) through (30).

- (27) *mij* *ällěp* *goas'sěk* *bårråm* *dáppe*
 1PL.NOM be.NEG.PRS.1PL ever eat.PST.PTCP here
 ‘We have never eaten (that) here.’
 (LE, answering the question “Do you eat mushrooms?”)

- (28) *dij* *ällěbit* *skihpasa*
 2PL.NOM be.NEG.PRS.2PL sick.PL.NOM
 ‘You aren’t sick.’ (SA, elicited)

⁷⁹ In the Lule Saami standard language, which features fewer amalgamated forms of the negative auxiliary and the verb ‘be’ than the Forest dialect, such forms have a consonant center in Grade III, e.g., PRS.2SG *il’la*, PRS.3PL *äl’la*. Based on the available data, I have not been able to ascertain whether the corresponding forms in Table 16 represent consonants of Grade II or Grade III—consequently, I have opted for the less marked notation of Grade II.

- (29) *da* *ěllin* *boatjojhoajdon*
 DEM.PL.NOM be.NEG.PST.3PL reindeer.herding.SG.INE
 ‘They weren’t in reindeer herding.’ (LE)
- (30) *ja* *dě* *ěllin* *tjuojka* *jěli*
 and then be.NEG.PST.3PL mosquito.PL.NOM or
muojva *jěli* *väppsá*
 blackfly.PL.NOM or wasp.PL.NOM
 ‘...and then there weren’t any mosquitos or blackflies or wasps...’ (LE)

In the imperative mood, the regular imperative forms of the negative auxiliary are coupled with the suppletive imperative connegative form *áro*, e.g., *ěllit áro* [NEG.IMP.2PL be.IMP.CNG]. In the conditional, the connegative form *lulu* occurs together with an indicative form of the negative auxiliary (31); cf. the corresponding positive clause (12) in Section 5.8.

- (31) *mãn* *iv* *lulu* *oasstet* *ådǎ*
 1SG.NOM NEG.PRS.1SG be.COND.CNG buy.INF new.ATTR
vuojánav *vájku* *lulun* *rudá*
 vehicle.SG.ACC despite be.COND.3PL money.PL.NOM
 ‘I wouldn’t buy a new car (lit. vehicle) even if I had money.’
 (SA, elicited)

Finally, it should be mentioned that a peculiar feature regarding the negation of *lǎ*- ‘be’ is found in older records of the Forest dialect. In the materials of Wiklund, there are first person singular forms of the negative auxiliary with a plosive *-b*, rather than *-v*, when coupled with the verb *lǎ*- ‘be’. Such forms are attested from two different speakers: *ib lǎ* ~ *ib^e lǎ* ‘I am not’, *ib^e lim* ‘I was not’ (KBW 1895 II), *ib^e lǎ* ‘I am not’ (KBW 1895 II, speaker from Muorkke, Sw. *Muorkafors*). When coupled with other verbs, however, only forms in *-v* are attested, e.g., *iv mālēstě_a* [NEG.PRS.1SG cook.CNG] ‘I don’t cook’.⁸⁰ Here it should be added that in the first interview with SA, I tentatively noted the form *ib lě* on one occasion—notably, this was before I was aware of the corresponding forms in Wiklund’s materials. However, when presented with this form during the second interview, SA did not accept it.

⁸⁰ Concerning the origin of the forms in *-b*, one is tempted to consider them an archaism, reflecting an older first person singular marker elsewhere corresponded by *-v* (both ultimately going back to Proto-Saamic **-m*). Here, a comparison could potentially be made with those varieties of Ume Saami that have *-b* for monosyllabic stems, but *-v* for polysyllabic stems (see, e.g., Siegl 2017: 270–271).

6. Other inflectional forms

This section discusses inflectional forms that are known from previous descriptions of Lule Saami, but that are either absent or only poorly attested in the present corpus. These forms include verbs inflected for the dual number (6.1), the potential mood (6.2), synthetic conditional forms (6.3), the “imperative II” mood and imperative forms in the first and third person (6.4), as well as a number of non-finite forms (6.5).

6.1 The dual number

In the Lule Saami standard language, finite verbs are as a rule inflected for dual number when the subject of a clause is two humans (see however Wiklund 1915: 102–103 and Tuolja & Angéus Kuoljok 2013: 42 on variation regarding the use of the dual number). Dual number is further obligatorily encoded for personal pronouns and possessive suffixes.⁸¹ In the present corpus, however, no dual forms of finite verbs are found. Instead, plural forms are used even when the subject of a clause is two individuals, as shown in (32) through (33). These plural verb forms can be accompanied by a personal pronoun in the plural, as in (34).

- (32) *běru* *mån* *ja* *muv* *áhkásj*
just 1SG.NOM and 1SG.GEN wife.SG.NOM
dánně *árrop*
DEM.PROX.SG.INE live.PRS.1PL
‘Just me and my wife live here.’ (LE)

- (32) *Anna* *ja* *Lisa* *lä* *oappátja*
Anna and Lisa be.PRS.3PL sister.CORR.PL.NOM
‘Anna and Lisa are sisters.’ (SA, elicited)

- (33) *iektu* *muvva* *vieljja* *ja* *ieddne*
yesterday 1SG.GEN brother.SG.NOM and mother.SG.NOM
vuolggin *Váhtjerij*⁸²

⁸¹ In addition to a third person singular *-s* in the reflexive pronoun SA *ietjas* REFL.GEN.3SG ‘his own’ and a third person plural *-sa* in the reciprocal adverb SA *gaskasa* RECP.3PL ‘with each other’, the present corpus contains only one token of a possessive suffix, the first person singular *-m* in the elicited kinship term SA *ednum* mother’s.brother.SG.ACC.POSS.1SG ‘my mother’s brother’.

⁸² The Swedish place name *Gällivare* was first translated as *Jiellevárre* (illative *Jiellevárráj*) by SA, but he then corrected himself to *Váhtjer* (illative *Váhtjerij*); both names refer to the same town. Both SA and LE further state that the name *Jiellevárre* is a misinterpretation of a form “*Jiellovárre*”, lit. ‘crack mountain’ (cf. Collinder 1964 s.v. *jielle*, Johansson 1989: 67).

go.PST.3PL Váhtjer.SG.ILL

‘Yesterday my brother and mother went to Váhtjer.’ (SA, elicited)

- (34) *dat* *lä* *danně* *věhi*
 DEM.SG.NOM be.PRS.3SG DEM.SG.INE little
nuortap *giella* *měn* *gal=ham*
 north.CMPR.ATTR language.SG.NOM but surely=indeed
mij *dě* *sābadip*
 1PL.NOM DPT agree.PRS.1PL
 ‘He’s got a little more northern language, but we (he and I) get along
 for sure.’
 (LE, about a speaker of North Saami in a nearby village)

Although not producing any dual verb forms himself, SA readily accepted the clause *māj mēnnin* [1DU.NOM go.PRS.1DU] when presented with it, and correctly translated it as Sw. *vi far* ‘we go’. Both speakers further produced clauses in which the first and second person dual pronouns *māj* and *dāj* are coupled with a finite verb in the plural, as shown in the elicited examples (35) through (37).⁸³

- (35) *māj* *áhkájn* *läp* *Suddásin*
 1DU.NOM wife.SG.COM be.PRS.1PL Boden.SG.INE
 ‘My wife and I are in Boden.’ (SA, elicited)

- (36) *māj* *dābddâp* *nuppijt*⁸⁴
 1DU.NOM know.PRS.1PL RECP.PL.ACC
 ‘We (two) know each other.’ (SA, elicited)

- (37) *dāj* *mannabit*
 2DU.NOM go.PRS.2PL
 ‘You (two) go.’ (LE, elicited)

⁸³ The corpus contains no token of a personal pronoun in the third person dual (standard *sāj*). In general, however, both speakers use the demonstrative *dat* ‘that, it’ rather than personal pronouns in the third person—even the singular and plural forms *sân* ‘s/he’ and *sij* ‘they’ are only represented by a few tokens in the corpus.

⁸⁴ While bipartite reciprocals such as the plural accusative *nuppe nuppijt* have been described for Lule Saami before (Grundström 1946–54: 1823–32), I am not aware of any previous mention of single plural forms of *nubbe* ‘second; other’ being used as a reciprocal pronoun, such as *nuppijt* in (36). I have however noted the same construction also in Unna Tjerusj: *mij dādjadip nuppijt* [1PL.NOM understand.PRS.1PL RECP.PL.ACC] ‘we understand each other’. Corresponding forms are further found in Duortnus North Saami (Kejonen 2017: 41, 49–50); see also Rasmus (2021: 139).

As in the indicative mood, only singular and plural verb forms are attested in the imperative. For instance, both speakers stated that the form *bāhtiit* come.IMP.2PL can be used when addressing two—as well as several—addressees.

Dual verb forms are found in Wiklund’s materials on the Forest dialect (e.g., *ēsstin* have.time.PRS.1DU, *lāpā* be.PRS.3DU, KBW 1895 II). For one speaker from Mörtberg,⁸⁵ however, Wiklund noted that “moreover, he was sometimes a bit uncertain of the dual forms and wanted to use plurals instead” (“han var för öfrigt ibland litet osäker på dualformerna och ville sätta pluraler i stället”, KBW 1906). Even so, Wiklund was still able to elicit dual forms from this speaker, e.g., *rāppē* open.IMP.2DU, *pōtika* come.PST.3DU. Lindgren’s texts from the area contain no dual forms.

Like Wiklund, Collinder was also able to obtain dual forms in the Forest dialect area, but he nevertheless noted an absence of dual forms for several speakers (1938: 149).⁸⁶ When noting down the nominative of the personal pronoun *mōḷi* 1DU.NOM, he further remarked that the oblique case forms *munnu* 1DU.GEN, *munnuw* 1DU.ACC and *munnuš* 1DU.ELA were hardly used (“föga brukliga”, ULMA 15925: 7). As in the present-day data, Collinder also noted a clause where a personal pronoun in the dual is coupled with a verb in the plural (38).

- (38) *lōxkit* *tōd'ja*
 read.IMP.2PL 2DU.NOM
 ‘Read you (two)!’ (Unna Sájvva, ULMA 15926: 61)

It can be pointed out that the use of dual personal pronouns with plural verb forms is a trait also attested in other Saamic languages, including Skolt Saami (see, e.g., Feist 2015: 200) and South Saami (Nilsson-Mankok 1970: 8, Ejerhed 1979: 188, Kowalik 2023: 211); see also Kejonen (2017: 32–33) for examples from North Saami.⁸⁷ This trait is further in line with a general

⁸⁵ For this speaker, Wiklund has noted the Lule Saami form *Sir^εkūl* (corrected from *Sēr^εkūl*) as corresponding to the Swedish place name *Mörtberg* (lit. ‘Roach mountain’); cf. *sir^εkā* ‘roach’ noted for the same speaker (Standard Lule Saami *siergga*). For a speaker from Unna Sájvva, Wiklund has noted a form *Sēr^εkūn* ‘Mörtberg’ (KBW 1906), and from a speaker born in Hiergáhpe, he has noted the form *Sēr^εkūn* id. (KBW 1895 I).

⁸⁶ Notably, one of the speakers interviewed by Collinder in Suobbat provided dual forms in the imperative mood, but not in the indicative, e.g., *qapā^x* open.IMP.2SG vs. *qax^εpē* open.IMP.2DU vs. *qēxpit* open.IMP.2PL, but *jau^εlap* say.PRS.1PL = say.PRS.1DU (ULMA 15926: 32–38). The number of documented verb forms from this speaker is however very small; see also Note 14 on the disproportionate number of imperative dual forms in Collinder’s materials.

⁸⁷ Note that the variety described by Ejerhed could potentially be classified as belonging to Ume Saami rather than to South Saami proper (cf. Kejonen 2017: 21). Most notably, Ejerhed (1979: 187) points out that a speaker interviewed by her had consonant gradation, a trait

tendency in the Saamic languages, according to which the dual number is retained longer for personal pronouns than for verbs, and longer for verbs than for possessive suffixes, i.e., an implicational hierarchy: personal pronouns > finite verbs > possessive suffixes (Kejonen 2017: 15–24, see also Rasmus 2021).

6.2 The potential mood

In the Lule Saami standard language there is a potential mood, which conveys a meaning of an action being possible or presumed (cf. Wiklund 1891: 278; 1915: 112, Spiik 1989: 105). The potential can also be used for future time reference (Ylikoski 2016; 2022: 139). A few potential forms have been documented in the Forest dialect by Wiklund (e.g., *ěñtšău* have.POT.1SG, KBW 1895 II) and by Collinder (*mālēstijav* cook.POT.1SG, *mālēstijap* cook.POT.1PL, 1938: 164).⁸⁸ Lindgren’s texts from the area do not contain any potential forms. Today, the potential does not seem to be productive in the Forest dialect—the present corpus contains no token of the potential mood in spontaneous speech, and when presented with the Standard Lule Saami forms *littjav* be.POT.1SG, *manátjav* go.POT.1SG and *manátjip* go.POT.1PL, SA did not seem to recognize them.

6.3 Synthetic conditional forms

Older descriptions of Lule Saami include the conditional as a productive mood. In these descriptions, synthetic conditionals can be formed by using a suffix *-lulu-* or the like, e.g., *kulā-lulu-v* hear-COND-1SG ‘I would hear’ (see the paradigms in Wiklund 1915: 65–82; 1891: 261–273 and Grundström 1946–54: 1847–1918; see also Wiklund 1891: 278–279; 1915: 113 on the use of the conditional). However, even these early descriptions note that it is possible to use a periphrastic construction instead, where only the auxiliary ‘be’ is inflected for the conditional, e.g., *lulu-v kulla-t* [be.COND-1SG hear-INF] ‘I would hear’ (Wiklund 1915: 81; 1891: 236). The latter system is the one prevailing in the Lule Saami standard language; later descriptions of the language only mention synthetic conditionals as “older forms” (see, for instance, Spiik 1989: 105).⁸⁹

traditionally seen as a defining feature of Ume Saami (interestingly, however, Ejerhed reports consonant gradation only for nouns, and not for verbs).

⁸⁸ In Collinder’s manuscript, the latter form is given as *mālēstid’šap* (ULMA 15934: 96), implying a Grade III consonant in the potential suffix (cf. Unna Tjerusj *mālēstid’t’šap* id., Grundström 1946–54: 1882).

⁸⁹ Occasionally, one may nevertheless come across synthetic conditional forms such as *galga-lulu* will-COND.CNG (Mark 13:14) also in written Lule Saami. Noteworthy in this context is also

As regards the Forest dialect, Wiklund has noted one token of a synthetic conditional: *küllŭ-lŭlŭ-i* hear.PASS-COND-3SG ‘it would be heard’ (“det skulle höras”, KBW 1895 II). In 1923, the nephew of the speaker who provided this form was interviewed by Collinder. When presented with the form, the nephew accepted it, but “could not pronounce it” himself (“Formen *kullulului* knäsatte Sandberg, men han kunde ej uttala den!”). Collinder further noted that “Moreover, there ought to be no other conditionals than of the kind *lului kullut* in the entire Forest Lapp area” (“Det torde f.ö. i hela skogslappsområdet ej finnas andra konditionaler än av typen *lului kullut*”, ULMA 15927: 29).

This appears to be the case also today; when presented with the synthetic conditional *oastá-lulu-v* buy-COND-1SG ‘I would buy’, neither SA nor LE accepted it. Both interviewed speakers however readily produced conditional clauses of the type *luluv oasstet* ‘I would buy’, an example of which is given in (12) in Section 5.8. It can also be added that Lindgren’s folklore texts from the area contain no synthetic conditionals, but only constructions of the same kind as in the present-day language (39).

- (39) *Jus mi tu tietet luluime,*
 if 1PL.NOM 2SG.ACC know.INF be.COND.1PL
ih tån luluh viessot.
 NEG.PRS.2SG 2SG.NOM be.COND.CNG live.INF
 ‘If we would have known about you, you wouldn’t be alive.’
 (Stuor Sájuvva, OL 1912: 5)

6.4 Imperative forms

In the Lule Saami standard language, imperative forms are not restricted to the second person, but are also found in the third person (singular, dual, plural) and in the first person (dual, plural).⁹⁰ In the Forest dialect, both Wiklund and Collinder have documented third person singular imperative forms: *lēōk’is* be.IMP.3SG (KBW 1895 II), *mennis* go.IMP.3SG, *pāxtjis* stay.IMP.3SG (Collinder 1938: 160). In the present corpus, however, only second person imperatives are attested.

The Lule Saami standard language further features a second imperative mood, the “imperative II”, which is commonly described as conveying a more

the peculiar form *sidálulut*, found in the dictionary of Spiik (1994 s.v. *skola*), which taken at face value appears to be an infinitive of a conditional form of *sihta-* ‘want’ (i.e., *sidá-lulu-t* want-COND-INF).

⁹⁰ Wiklund has described Lule Saami imperative forms also in the first person singular, albeit only in the “imperative II” mood, e.g., *lāhkuv* read.IMP.I.1SG, *alluv* NEG.IMP.I.1SG (1915: 65–82; cf. 1891: 241, 256–257, 264, 267, Grundström 1946–54: 1865, 1883).

polite or modest request than the imperative proper (Wiklund 1891: 279; 1915: 114, Grundström 1946–54: 1851, Spiik 1989: 103). Such forms have been documented in the Forest dialect by Wiklund: see examples (40) and (41), as well as the form *lēkūs* be.IMPII.3SG (KBW 1895 II). The present corpus however contains no such forms, and as far as I have been able to tell, neither Collinder’s nor Lindgren’s materials on the Forest dialect feature forms of this kind either.

- (40) *ēlluo* *tō* *nōu* *jēulūw!*
 NEG.IMPII.2SG DPT so say.IMPII.CNG
 ‘Don’t say that!’ (KBW 1895 II)

- (41) *ēllūw* *tēwkuw!*
 NEG.IMPII.2SG do.IMPII.CNG
 ‘Don’t do!’ (Mörtberg, KBW 1906)

It should be added that the elicitation of imperative forms was not exhaustive, and that the topic therefore would merit further investigation.

6.5 Non-finites

A number of forms traditionally described as part of the non-finite verb inflection of Lule Saami are only poorly attested in my data. Such forms include the “verb genitive” (or “converb of manner and concomitance”, Ylikoski 2022: 140), e.g., SA, LE *váttje* ‘by foot’ < *vádtje* ‘walk’, SA *tjuojga* ‘by means of skiing’ < *tjuojgga* ‘ski’ (only tokens). Also poorly attested is the “gerund” (or “converb of simultaneity”, Ylikoski 2022: 140), e.g., SA (*muwva*) *viesodijn* ‘during (my) life’ < *viesso* ‘live’ (only token).⁹¹

The form traditionally referred to as the “present participle” is only attested in the function of an actor noun, e.g., SA, LE *bargge* ‘worker’ < *bargga* ‘work’. As noted in Section 4.3, such actor nouns derived from imparisyllabic verbs have a vocalism that deviates from the Lule Saami standard language,

⁹¹ These two forms are relatively rare also in other varieties of spoken Lule Saami, but are found, e.g., in the New Testament (*Adā Testamennta*, 2003), cf. Angéus Kuoljok (2004: 35). The “verb genitive” is also known as the “verbal adverb” (Ger. *Verbaladverbium*, Wiklund 1928) or the “adverbial form of the verb” (Ger. *Adverbialform des Verbs*, Grundström 1946–54: 1847–1918). The “gerund” is also known as the “simultaneity form” (Sw. *samtidighetsformen*, Spiik 1989: 108), the “first gerund” (Sw. *gerundium I*, Wiklund 1915: 115) or the “absolute gerund” (Ger. *Absolutes Gerundium*, Grundström 1946–54: 1847–1918). Adding to the terminological confusion, Lagercrantz (1939: 1207 et passim) refers to this form as the “second gerund” or “gerundive” (Ger. *Gerund II*, *Gerundiv*); see also Note 51.

The “verbal noun” (or “action nominalization”, Ylikoski 2022: 140) is only attested as the first component of compound nouns, e.g., *SA vuohtjem-ruhta* ‘bounty’ lit. ‘shooting-money’ (< *vuohtje-* ‘shoot’, *ruhta* ‘money’), *tjåhkkåhim-sadje* ‘seat’ lit. ‘sitting-place’ (< *tjåhkkåhi-* ‘sit’, *sadje* ‘place’).⁹² It is possible that some—or all—of these forms are better treated as derivations than as non-finite verb forms in the Forest dialect; more data are however needed for a definitive analysis.

Earlier materials on the Forest dialect—as well as the standard language—include a form in *-mis*, *-mes* known as the “action elative” (or “infinitive II”, Ylikoski 2022: 140, cf. Spiik 1989: 108), which is governed by the verb *hiejtte* ‘stop, quit’ (for a list of other verbs known to govern this form, see Tuolja & Angéus Kuoljok 2013: 140). In the present corpus, no such forms are attested. Instead, the infinitive is used in this context (42).

- (42) *bäna=l* *häjttám* *hállat*
dog.SG.NOM=be.PRS.3SG stop.PST.PTCP bark.INF
'The dog has stopped barking.' (SA, elicited)

When presented with the clause in (43), noted by Collinder in Sádek, SA accepted it, but said that he himself would rather produce the clause given in (44), i.e., using the infinitive instead.

- (43) *heixtē* *kačātimēs*⁹³
stop.IMP.2SG ask.ACTELA
'Stop asking!' (Sádek, ULMA 15926: 20)

- (44) *hiejte gatjádít!*
 stop.IMP.2SG ask.INF
 ‘Stop asking!’ (SA, elicited)

The “verb abessive” (or “negative converb”, Ylikoski 2022: 140), found in Standard Lule Saami (e.g., *tjáledak* ~ *tjálek* ‘without writing’) is not attested

⁹² Note that the form *tjahkkähim-* differs from the verbal noun proper (standard *tjahkkähibme* ‘the act of sitting’); forms of this kind have been referred to as “short” or “shortened” verbal nouns (Wiklund 1915: 82–83, Angéus Kuoljok 2013: 104). The so-called “agentive passive participle” is not attested in the Forest dialect (cf. *vuõïⁿnēm* ‘seen’ in the clause *munnu vuõïⁿnēm al̥ō* [1DU.GEN see.AG.PASS.PTCP reindeer.doe.SG.NOM] ‘the reindeer doe seen by us (two)’; Wiklund 1915: 116; see also Ylikoski 2009: 130–135 on the corresponding North Saami form).

⁹³ For a speaker from Allevárre (Sw. *Västerberget*), Collinder has noted the corresponding form with an additional final vowel *-ε*: *kačātimēse* (ULMA 15926: 6).

in my data, although admittedly, this form was not specifically sought after in elicitation. I am not aware any verb abessives in the Forest dialect materials of Wiklund, Lindgren or Collinder either.⁹⁴

7. Conclusions and discussion

In this paper, I have described the inflectional verb morphology of Lule Saami, using ample original data from the critically endangered Forest dialect. Here, lexical verbs belong to one of three inflectional classes, and finite forms are inflected for person (first, second, third), number (singular, plural), tense (present, past) and mood (indicative, imperative). There are also a number of non-finite verb forms, including the infinitive, progressive, past participle, supine, connegative and passive past participle. Two non-lexical verbs (the negative auxiliary and *lä-* ‘be’) have irregular inflection. Both finite and non-finite verbs are formed by means of adding suffixes, as well as by the stem-internal processes of consonant gradation and stem-final vowel alternations.

The negation of *lä-* ‘be’ is highly irregular in this variety and would merit further investigation. Particularly noteworthy are the plural forms, in which there is a stem *ällě-* ‘be not’ that takes regular subject suffixes (e.g., *ällěp* ‘we are not’, *ěllijma* ‘we were not’). These forms, which have not been previously described, stem from an amalgamation of the negative auxiliary and the verb ‘be’. Also in the singular there are amalgamated forms, such as *illiv* ‘I was not’. With the exception of the negation of *lä-* ‘be’, however, the verb forms attested in the corpus largely agree with the Lule Saami standard language. In this respect, the Forest dialect could indeed be seen as a “typical central dialect” of Lule Saami, as it has been described by Collinder (1938: 3).

Nevertheless, the description of the verb inflection presented here differs in several respects from the picture painted in pedagogical materials on Lule Saami, such as Spiik (1989) and Kejonen (2013). Most importantly, a number of inflectional forms found in previous descriptions of the language are not attested in the data. These forms include verbs in the dual number and the potential mood, as well as synthetic conditional forms. Imperative forms in the first and third person are also unattested in the corpus, as are forms of the “imperative II” mood. For the negative auxiliary verb, no past tense forms were produced spontaneously by the interviewed speakers.

Taken together, the verb inflection of the Lule Saami Forest dialect appears to be less complex than that of the Lule Saami standard language: while a

⁹⁴ See also Wiklund (1915: 115), who notes that verb abessives are “rather uncommon and are preferably replaced by circumlocutions” (“Dessa former äro emellertid tämligen ovanliga och ersättas helst med omskrifningar”).

lexical verb can be inflected in 43 finite forms in the standard language, the present corpus features no lexical verb inflected in more than 14 finite forms. Further, some non-finite forms known from previous descriptions of the language are also unattested in the—admittedly small—corpus. While a more comprehensive study, drawing from a larger corpus of elicited and non-elicited speech, would perhaps show some of these forms to be productive in the Forest dialect, the overall tendency in this variety seems to be that of reduction in the verb morphology.

In many cases, the loss of inflectional verb morphology can likely be attributed to language contact and obsolescence. This is perhaps most evident in the loss of the dual number and the potential mood, in which the Forest dialect both reduces its morphological complexity and adapts to the co-territorial languages Swedish and Meänkieli, which lack corresponding forms.⁹⁵

In sum, the present study highlights a potential discrepancy between existing descriptions of Lule Saami and varieties of the contemporary spoken language. As such, this paper should be seen as a call for further documentation and description of Lule Saami as it is spoken today.

Acknowledgments

I am profoundly grateful to Sören Andersson and Lars Eriksson for choosing to share a small portion of their traditional knowledge with me. *Gijtto!* I am also grateful to Anund Vitsak[†] and Lars Pittja[†] of the Unna Tjerusj district and to several speakers of Čohkkiras North Saami for comparative data on these varieties. Many thanks are further due to Torgny Hedström, whose keen eye—and ear—for detail has improved this study significantly, as well as to Adam Horn af Åminne, Richard Kowalik, Olof Lundgren, an anonymous reviewer and the editors of the present volume for helpful comments. I would additionally like to express my gratitude to the kind and helpful staff of the Institute for Language and Folklore in Uppsala for swift and diligent assistance in matters concerning older materials on the Forest dialect, including materials on place names provided by Björn Lundqvist. This study was partly funded by the Uppsala Language Documentation Group (UPPLADOC), which I gratefully acknowledge.

⁹⁵ While there is a potential mood in Standard Finnish, the potential is not productive in Meänkieli (Pohjanen 2022: 106, Hansegård 1967: 96; 1988: 224, Airila 1912: 17, 167, Note 1).

Abbreviations

Grammatical abbreviations follow those listed in the Leipzig Glossing Rules with the following additions: ABE—abessive; ACTELA—action elative; AG—agentive; ATTR—attributive; CMPR—comparative; CNG—connegative; CORR—correlative; DPT—discourse particle; ELA—elative; ILL—illative; INE—inessive; IMPII—imperative II; NPM—negative polarity marker; POT—potential; SPRL—superlative; STR—strong grade; SUP—supine; WK—weak grade. Other abbreviations: Ger.—German; LE—Lars Eriksson; N/A—not attested; SA—Sören Andersson; Sw.—Swedish.

Unpublished sources

- DAUM BD 5070 = Korhonen, Olavi. 1985. [Recording from Suobbat]. Institutet för språk och folkminnen, Uppsala.
- KBW 1895 I = Wiklund, K. B. Gellivare I. 1895. K. B. Wiklunds samling, 28, Cb 45:1, Uppsala universitetsbibliotek.
- KBW 1895 II = Wiklund, K. B. Gellivare II. 1895. K. B. Wiklunds samling, 28, Cb 45:2, Uppsala universitetsbibliotek.
- KBW 1906 = Wiklund, K. B. Jokkmokk 1906. K. B. Wiklunds samling, 28, Cb 47, Uppsala universitetsbibliotek.
- OL 1912 = Lindgren, Otto. Berättelser av skogslapphustrun Greta Andersson i Storsaivis, Gellivare socken vintern 1912, upptecknade av komminister Otto Lindgren. K. B. Wiklunds samling, 17, Cb 10, Uppsala universitetsbibliotek.
- OL 1913 = Lindgren, Otto. [Letter to K. B. Wiklund, November 29, 1913]. K. B. Wiklunds samling, A 2, Uppsala universitetsbibliotek.
- OL 1923 = Lindgren, Otto. Frågelista 2. besvarad af Anders Svensk. K. B. Wiklunds samling, 17, Cb 10, Uppsala universitetsbibliotek.
- ULMA 15924–36 = Collinder, Björn. Anteckningar om Lule lappmarks norra skogsdialekt. Sommaren 1923. Institutet för språk och folkminnen, Uppsala.
- ULMA 22200 = Johansson, Carl. Uppteckn. fr. Arjeplog samt efter Karesuando- o. Jukkasjärvilappar. Institutet för språk och folkminnen, Uppsala.
- ULMA 22480 = Calleberg, C. Axel. Uppteckningar på lapska. Institutet för språk och folkminnen, Uppsala.

References

- Ådå Testamentta*. 2003. 2nd ed. Uppsala: Svieriga ráamátsiebrre.
- Åtå Testamentta*. 1903. Stockholm: P. Palmquists Aktiebolag.
- Aikio, Ante. 2007. Etymological nativization of loanwords: a case study of Saami and Finnish. In Ida Toivonen & Diane Nelson (eds), *Saami linguistics* (Current Issues in Linguistic Theory, 288), 17–52. Amsterdam: John Benjamins.
- Airila, Martti. 1912. *Äännehistoriallinen tutkimus Tornion murteesta: murteen suhdetta Suomen muihin murteihin silmälläpitäen*. Helsinki: Suomalaisen Kirjallisuuden Seura.
- Álguogáhtie. 2022. *Ubmejesámien mánánappa*. <https://www.isof.se/nationella-minoritetssprak/laromedel/pedagogiska-material/ubmejesamien-mananappa> (last accessed November 8, 2024).
- Angéus Kuoljok, Susanna. 1997. *Nominalavledningar på ahka i lulesamiskan* (Studia Uralica Upsaliensia 28). Uppsala: Uppsala universitet.
- Angéus Kuoljok, Susanna. 2002a. Julevsámegiella. *Bårjås: diedalasj ájggetjála Árran - julevsáme dávvervuorkás*. 10–18.

- Angéus Kuoljok, Susanna. 2002b. Giella Ádå Testamenta ádå járggálimen. *Sámi diedalaš áigečála* 1/2002: 12–18.
- Angéus Kuoljok, Susanna. 2003. *Lulesamiska: ett hotat minoritetsspråk i Skandinavien* (Endangered languages and cultures 1). Uppsala: Uppsala universitet.
- Angéus Kuoljok, Susanna. 2004. Ádå Testamentnna ádåsis. *Sámi diedalaš áigečála* 1/2004: 32–42.
- Bartens, Hans-Hermann. 2022. Der Kuckuck in den Glaubensvorstellungen der Saamen. *Finnisch-Ugrische Forschungen* 67: 5–48.
- Bergsland, Knut. 1946. *Røros-lappisk grammatikk: et forsøk på strukturell språkbeskrivelse*. Oslo: Aschehoug.
- Bergsland, Knut. 1992. Language contacts between Southern Sami and Scandinavian. In Ernst Håkon Jahr (ed.). *Language contact: theoretical and empirical studies* (Trends in Linguistics. Studies and Monographs, 60), 5–15. Berlin: Mouton de Gruyter.
- Bergsland, Knut & Gustav Hasselbrink. 1957. *Sámien lukkeme-gärjá: sydlapsk läsebok med grammatik och ordlista* (svensk upplaga). Gardvik: P.J. Fjellström.
- Collinder, Björn. 1926. Från språkgränsen i Gällivara. *Norrboten*. 120–126.
- Collinder, Björn. 1929. Det mångtungade Norrbotten. *Svenska turistföreningens årsskrift*. 135–146.
- Collinder, Björn. 1930. Småbidrag till den sociologiska lingvistik. *Nysvenska studier* 10: 1–14.
- Collinder, Björn. 1938. *Lautlehre des Waldlappischen Dialektes von Gällivare* (Suomalais-Ugrilaisen Seuran Toimituksia 74). Helsinki: Suomalais-Ugrilainen Seura.
- Collinder, Björn. 1953. *Lapparna: en bok om samefolkets forntid och nutid*. Stockholm: Forum.
- Collinder, Björn. 1959. *Språket: inledning till det vetenskapliga språkstudiet*. Stockholm: Natur och kultur.
- Collinder, Björn. 1964. *Ordbok till Sveriges lapska ortnamn*. Uppsala: Kungl. Ortnamnskommissionen.
- Collinder, Björn. 1978. *Sprache und Sprachen: Einführung in die Sprachwissenschaft*. Hamburg: Buske.
- Drake, Sigrid. 1918. *Västerbottenslapparna under förra hälften av 1800-talet: etnografiska studier*. Uppsala: Uppsala universitet.
- Ejerhed, Eva. 1979. Strukturella aspekter på sydsamisk tvåspråkighet. In Astrid Stedje & Peter af Trampe (eds), *Tvåspråkighet: föredrag vid det Andra nordiska tvåspråkighetssymposiet, anordnat 18–19 maj 1978 av Språkvetenskapliga sektionen vid Stockholms universitet*, 186–194. Stockholm: Akademilitt.
- Feist, Timothy. 2015. *A grammar of Skolt Saami* (Suomalais-Ugrilaisen Seuran Toimituksia 273). Helsinki: Suomalais-Ugrilainen Seura.
- Grundström, Harald. 1929. Sägner och folktro bland lappar och nybyggare i Jokkmokk. *Norrboten*. 55–75.
- Grundström, Harald. 1937. Inledning. In Anta Pirak, *Jáhttee saamee viessoom*, III–XIV. Uppsala: Almqvist & Wiksell.
- Grundström, Harald. 1946–54. *Lulelappsk ordbok. Lulelappisches Wörterbuch* (Skrifter utgivna genom Dialekt- och folkminnesarkivet i Uppsala. Ser. C: 1). Uppsala: Lundequistska bokhandeln.

- Halász, Ignác. 1886. *Svéd-lapp nyelv. II. Jemtlandi lapp nyelvmutatványok*. Budapest: MTA.
- Halász, Ignác. 1891. *Svéd-lapp nyelv. IV. Déli-lapp szótár*. Budapest: MTA.
- Halász, Ignác. 1892. Harmadik lappföldi útam. *Budapesti szemle* 70: 253–284.
- Halász, Ignác. 1893. *Svéd-lapp nyelv. V. Népköltési gyűjtemény a Pite lappmark arjepluogi egyházkerületéből*. Budapest: MTA.
- Hansegård, Nils Erik. 1967. *Recent Finnish loanwords in Jukkasjärvi Lappish* (Studia Uralica et Altaica Upsaliensia 3). Uppsala: Uppsala universitet.
- Hansegård, Nils Erik. 1988. *Språken i det norrbottensfinska området*. Luleå: Höskolan i Luleå.
- Hasselbrink, Gustav. 1944. *Vilhelminalapskans ljudlära: med särskild hänsyn till första stavelsens vokaler*. Uppsala: Uppsala universitet.
- Hasselbrink, Gustav. 1981–85. *Südlappisches Wörterbuch. I–III* (Skrifter utgivna genom Dialekt- och folkminnesarkivet i Uppsala. Ser. C: 4). Uppsala: Lundequistska bokhandeln.
- Hedström, Torgny. 2011. Le gus boares lādna vaj substitusjávnná? Bondsk påverkan på samiska. In Torbjörn Söder (ed.). *Čieža čállosa. Gietjav tjállaga. Tjijhtje tjaalegh. Sju uppsatser i samiska* (Opiscula Uralica 10), 125–248. Uppsala: Institutionen för moderna språk, Uppsala universitet.
- Heikki, Jörgen. 2021. *Odda girji Unna Sávvva bálvvosbaikki birra*. Sameradion, August 9, 2021. <https://sverigesradio.se/artikel/odda-girji-unna-sajvva-balvvosbaikki-birra> (last accessed June 12, 2023).
- Ihre 105 = [Holmberger, Per]. *Lapparnes Björn wisa, Om deras öfriga, bekanta widskepelse, Vocabularium Linguae Lapponum mendicantium, Verborum Conjugatio*. Uppsala universitetsbibliotek. <http://urn.kb.se/resolve?urn=urn:nbn:se:alvin:portal:record-183592> (last accessed June 12, 2023).
- Itkonen, Toivo Immanuel. 1958. *Koltan- ja kuolanlapin sanakirja. Wörterbuch des Kolta- und Kolalappischen. I–II* (Lexica Societatis Fenno-Ugricae 15). Helsinki: Suomalais-Ugrilainen Seura.
- Johansson, Carl. 1989. *Mujto: minnen från jägar- och fiskartiden och den gamla renkonstens dagar* (Skrifter utgivna av Dialekt-, ortnamns- och folkminnesarkivet i Umeå. Ser. C: 5). Umeå: Dialekt-, ortnamns- och folkminnesarkivet.
- Juutinen, Markus & Jukka Mettovaara. 2021. Saamelaiskielten indefiniitti-pronominien jäljillä. In Mika Härmäläinen, Niko Partanen & Khalid Alnajjar (eds). *Multilingual Facilitation*, 104–127. Helsinki: University of Helsinki Library.
- Källskog, Margareta. 1989. ”Nog kan jag finska och lapska har jag lärt mig”: utskrift av en inspelning på överkalixmål från Storsaivis, Gällivare sn. In Lars Bleckert & Maj Reinhammar (eds). *Mål och minnen: vänskrift till Sven Söderström 1988 från arbetskamraterna på ULMA*, 50–64. Uppsala: Dialekt- och folkminnesarkivet.
- Källskog, Margareta. 1992. *Attityd, interferens, genitivsyntax: studier i nutida Överkalixmål* (Skrifter utgivna genom Dialekt- och folkminnesarkivet i Uppsala. Ser. A: 18). Uppsala: Dialekt- och folkminnesarkivet.
- Kejonen, Olle. 2013. *101 julevsáme verba: 101 lulesamiska verb*. Stockholm: Vulkan.

- Kejonen, Olle. 2017. *Dual number in the North Saami dialect of Ofoten and Sør-Troms*. MA thesis, Uppsala University.
<http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-323384> (last accessed June 12, 2023).
- Kejonen, Olle. 2020a. On the development of some Saami kinship terms. In Rogier Blokland & Riitta-Liisa Valijärvi (eds). *Där Östersjön är Västersjön – Seal, kus Läänemeri on Idameri: festskrift till Virve och Raimo Raag – pühendusteos Virve ja Raimo Raagile*, 43–53. Uppsala: Uppsala universitet.
- Kejonen, Olle. 2020b. North Saami, Čohkkiras variety (Sweden, Norway) – Language Snapshot. *Language Documentation and Description* 17: 178–185.
- Kejonen, Olle. 2021a. Om offer: en berättelse från 1912. In Kerstin Andersson (ed.). *Unna Saiva: en skändad samisk offerplats*, 24–27. Stockholm: Vulkan.
- Kejonen, Olle. 2021b. ”Tauta är ett levande väsen”: skogssamiska berättelser om sjukdom och smitta. *Samefolket* 102 (5): 34–35.
- Kejonen, Olle. 2022. Björn Collinder och den lulesamiska skogsdialekten. In Kerstin Andersson (ed.). *Vuovddesáme: skogssamer i Flakaberg*, 118–124. Stockholm: Vulkan.
- Kejonen, Olle. 2024a. Tre samiska lånord i meänkieli. *Svenska landsmål och svenskt folkliv* 2023 (146): 137–142.
- Kejonen, Olle. 2024b. *Pite Saami tales of terror*. Review of Halász, Ignác. 2023. *Alvos giehto. Förskräckliga berättelser från Arjeplog*. *ARV: Nordic Yearbook of Folklore* 80: 202–204.
- Kintel, Anders. 2012. *Julevsáme-dá[r]o báhkogirje*.
http://gtweb.uit.no/webdict/ak/smj2nob/s_smj2nob.html (last accessed June 12, 2023).
- Kittilä, Seppo & Jussi Ylikoski. 2018. Some like it transitive: remarks on verbs of liking and the like in the Saami languages. In Ilja A. Seržant & Alena Witzlack-Makarevich (eds). *Diachrony of differential argument marking* (Studies in Diversity Linguistics 19), 455–480. Berlin: Language Science Press.
- Korhonen, Mikko. 1967. *Die Konjugation im Lappischen: Morphologisch-historische Untersuchung. 1, Die finiten Formkategorien* (Suomalais-Ugrilainen Seuran Toimituksia 143). Helsinki: Suomalais-Ugrilainen Seura.
- Korhonen, Mikko. 1974. *Die Konjugation im Lappischen: Morphologisch-historische Untersuchung. 2, Die nominalen Formkategorien* (Suomalais-Ugrilainen Seuran Toimituksia 155). Helsinki: Suomalais-Ugrilainen Seura.
- Korhonen, Mikko. 1981. *Johdatus lapin kielen historiaan* (Suomalaisen Kirjallisuuden Seuran Toimituksia 370). Helsinki: Suomalaisen Kirjallisuuden Seura.
- Korhonen, Olavi. 1979. *Báhkogir'je: julevusámes dárrui, dáros julevusábmái. Lulesamisk svensk, svensk lulesamisk ordbok*. Luleå: Länsskolnämnden.
- Korhonen, Olavi. 2005. *Báhkogirje: julevsámes dárruj, dáros julevsábmáj. Ordbok: lulesamisk svensk, svensk lulesamisk*. Jokkmokk: Samernas utbildningscentrum.
- Korhonen, Olavi. 2008. Är lappologin död? *Oknytt* 3–4: 6–17.
- Korhonen, Olavi. 2016. Samiskan under fyra sekel i Svenska kyrkans arbete. In Daniel Lindmark & Olle Sundström (eds). *De historiska relationerna mellan Svenska kyrkan och samerna: en vetenskaplig antologi*, 735–795. Skellefteå: Artos.
- Kowalik, Richard. 2023. *Towards a grammar of spoken South Saami*. Stockholm: Stockholm University.

- Kuokkala, Juha. 2019. Saami -(š)goahti-inchoatives: their variation, history, and suggested cognates in Veps and Mordvin. *Suomalais-Ugrilaisen Seuran Aikakauskirja* 97: 153–181.
- Labba, Andreas. 1969. *Anta*. Stockholm: Bonniers.
- Labba, Elin Anna. 2020. *Herrarna satte oss hit: om tvångsförflyttningarna i Sverige*. Stockholm: Norstedts.
- Lagercrantz, Eliel. 1939. *Lappischer Wortschatz. I–II* (Lexica Societatis Fenno-Ugricae 6). Helsinki: Suomalais-Ugrilainen Seura.
- Lagercrantz, Eliel. 1941. *Synopsis des Lappischen*. Oslo: Oslos etnografiske museum.
- Lagercrantz, Eliel. 1958. *Lappische Volksdichtung II: Lyngnälappische, Nordwestlappische und Westfjordlappische Texte* (Suomalais-Ugrilaisen Seuran Toimituksia 115). Helsinki: Suomalais-Ugrilainen Seura.
- Larsson, Lars-Gunnar. 1991. Glidvokalen i lulesamiskan: en dialektgeografisk studie på grundval av Harald Grundströms ordbok. *Svenska landsmål och svenskt folkliv* 1990 (113): 167–199.
- Larsson, Lars-Gunnar. 2012. *Grenzen und Gruppierungen im Umesamischen* (Veröffentlichungen der Societas Uralo-Altaica 83). Wiesbaden: Harrassowitz.
- Larsson, Lars-Gunnar. 2018. *Per Holmberger och sockenlapparnas språk* (Acta Bibliothecae R. Universitatis Upsaliensis 52). Uppsala: Uppsala universitet.
- Lehtiranta, Juhani. 1992. *Arjeploginsaamen äänne- ja taivutusopin pääpiirteet* (Suomalais-Ugrilaisen Seuran Toimituksia 212). Helsinki: Suomalais-Ugrilainen Seura.
- Lindahl, Erik & Johan Öhrling. 1780. *Lexicon Lapponicum, cum interpretatione vocabulorum Sveco-Latina et indice Svecano Lapponico*. Stockholm.
- Lindgren, Otto. 1928. Lapska sagor och sägner från skogslappar i södra Gellivare. *Norrbottnen*. 170–173.
- Lindgren, Otto. 1930. Där lapskt och svenskt mötas. In Georg Bergfors (ed.). *Norrbottnen: läsebok för skola och hem. 1*, 357–359. Uppsala: Lindblad.
- Magga, Ole Henrik & Lajla Mattsson Magga. 2012. *Sørsamisk grammatikk*. Karasjok: Davvi girji.
- Nielsen, Konrad. 1926. *Lærebok i lappisk: utarbeidet på grunnlag av dialektene i Polmak, Karasjok og Kautokeino. 1, Grammatikk: lydlære, formlære, orddannelseslære og syntaks samt tillegg*. Oslo: Brøgger.
- Nilsson-Mankok, Erik. 1970. *Hur sydsamiskan försvinner*. Umeå: University of Umeå, Department of General Linguistics.
- Odhner, Clas Theodor. 1878. *C. T. Odhners lärobok i fäderneslandets historia: under författarens ledning bearbetad för folkskolan. C. T. Odhneren äppetuskirje aittiglandan suptsasikum: reitetum almogen skoli tet. Same kiälai puoktetum Johannes Mörtselest*. Stokkholman: P. A. Norstedt ja Pardneh.
- ÖKM = *ÖverköLisma:Le. Överkalixmålet*. 1999. Överkalix: Arbetarnas Bildningsförbund - Studieförbundet Vuxenskolan.
- Olsen, Kirsten Marit & Bodil Utsi Vars. 2012. *Biejvvegajjno: åhpaddidje viehkke- ja ressurssagirje 1–4. Åssko, iellemvuojnno ja etihkka. Kárásjjáhkå*: Davvi girji.
- OMSM = *Ordbok meänkieli-svenska-meänkieli*. Meän Akateemi – Academia Tornedaliensis. <https://språk.isof.se/meänkieli> (last accessed June 12, 2023).
- Ortnamnsregistret*. Uppsala: Institutet för språk och folkminnen. <https://ortnamnsregistret.isof.se> (last accessed June 12, 2023).
- Pihl, Carin. 1924. *Överkalixmålet. 1*. Uppsala: Uppsala universitet.

- Pihl, Carin. 1948. *Verben i överkalixmålet* (Skrifter utgivna genom Dialekt- och folkminnesarkivet i Uppsala. Ser. A: 5). Uppsala: Lundequistska bokhandeln.
- Pohjanen, Bengt. 2022. *Meänkieli: grammatik, lärobok, historik, texter*. Överkalix: Barents publisher.
- Qvigstad, Just Knud. 1925. *Die lappischen Dialekte in Norwegen* (Oslo Etnografiske Museums Skrifter 1–1). Oslo.
- Rasmus, Serge. 2021. Ohcejoga nuoraid duála geavaheapmi. *Sámi diedalaš áigečála* 2021: 123–154.
- Rauhala, Ilona. 2013. Attribuutin kongruenssi: adjektiivit, demonstratiivit ja *para. *Suomalais-Ugrilaisen Seuran Aikakauskirja* 94: 199–221.
- Ruong, Israel. 1943. *Lappische Verbalableitung dargestellt auf Grundlage des Pitelappischen*. Uppsala: Uppsala universitet.
- Ryd, Yngve. 2007. *Ren och varg: samer berättar*. Stockholm: Natur & kultur.
- Ryd, Yngve. 2017. *Doahkke – fågelflock: samiska fågelnamn och myter*. Skellefteå: Ord & visor.
- Rydvig, Håkan. 2013. *Words and varieties: lexical variation in Saami* (Suomalais-Ugrilaisen Seuran Toimituksia 269). Helsinki: Suomalais-Ugrilainen Seura.
- Sammallahti, Pekka. 1998. *The Saami languages: an introduction*. Kárášjohka: Davvi girji.
- Siegl, Florian. 2017. Tryggve Sköld och umesamiskan. Intryck från en oavslutad språkdokumentation. *Norsk Lingvistisk Tidsskrift* 35: 252–290.
- Siegl, Florian. 2019. Nebensächlichkeiten zur Kopula *leä-* im Umesaami unter Berücksichtigung der Verhältnisse in den anderen westsaamischen Sprachen. *Ural-Altaische Jahrbücher (Neue Folge)* 27: 93–127.
- Sjaggo, Ann-Charlotte. 2015. *Pitesamisk grammatik: en jämförande studie med lulesamiska* (Senter for samiske studier, Skriftserie 20). Tromsø: Tromsø universitet.
- Spiik, Nils Eric. 1977. *Lulesamisk grammatik*. Jokkmokk: Förf.
- Spiik, Nils Eric. 1989. *Lulesamisk grammatik*. 2nd ed. Jokkmokk: Sameskolstyrelsen.
- Spiik, Nils Eric. 1994. *Lulesamisk ordbok: svensk-samisk*. Jokkmokk: Sameskolstyrelsen.
- Stolz, Thomas. 2015. Adjective-noun agreement in language contact: loss, realignment, and innovation. In Francesco Gardani, Peter Arkadiev & Nino Amiridze (eds). *Borrowed morphology* (Language Contact and Bilingualism 8), 269–301. Berlin: De Gruyter Mouton.
- Stoor, Krister. 2016. Svenska kyrkan och jojken. In Daniel Lindmark & Olle Sundström (eds). *De historiska relationerna mellan Svenska kyrkan och samerna: en vetenskaplig antologi*, 711–734. Skellefteå: Artos.
- Tuolja, Karin & Susanna Angéus Kuoljok. 2013. *Giellaj hilá*. 2nd ed. Jokkmokk: Samernas utbildningscentrum.
- Tuolja, Lars Matto. 1987. *Tjaktjalasta*. Jokkmokk: Sámi girjjit.
- Wickman, Bo. 1965. Ein lappischer Umlautfall und seine Bedeutung für die lappische Vokalgeschichte. In Aleksandr Vasil'evič Isačenko, Wilhelm Wissmann & Hermann Strobach (eds), *Beiträge zur Sprachwissenschaft, Volkskunde und Literaturforschung Wolfgang Steinitz zum 60. Geburtstag am 28. Februar 1965* dargebracht, 424–431. Berlin: Akademie-Verlag.

- Wickman, Bo. 1975. Phonologische Analyse des Lulelappischen. In Karl Erich Heidolph (ed.). *Symposion Phonologische Analyse der uralischen Sprachen: Berlin, 17.–20. September 1974*, 1–5. Berlin: Akademie der Wissenschaften der DDR.
- Wiklund, K. B. 1888. Tillägg. [Addition to: Nordlinder, E. O. 1887. Förteckning öfver Lule-socknarnas person- och ortnamn]. *Svenska landsmål och svenskt folkliv* 6 (3): 29.
- Wiklund, K. B. 1890. *Lule-Lappisches Wörterbuch* (Suomalais-Ugrilaisen Seuran Toimituksia 1). Helsinki: Suomalais-Ugrilainen Seura.
- Wiklund, K. B. 1891. *Laut- und Formenlehre der Lule-lappischen Dialekte*. Stockholm: P. A. Norstedt & Söner.
- Wiklund, K. B. 1896. *Entwurf einer urlappischen Lautlehre I: Einleitung, Quantitätsgesetze, Accent, Geschichte der hauptbetonten Vokale* (Suomalais-Ugrilaisen Seuran Toimituksia 10). Helsinki: Suomalais-Ugrilainen Seura.
- Wiklund, K. B. 1901. *Lärobok i lappska språket*. Uppsala: Lundequistska bokhandeln.
- Wiklund, K. B. 1915. *Lärobok i lappska språket*. 2nd ed. Stockholm: Björck & Börjesson.
- Wiklund, K. B. 1916. *Tålotj suptsasah ja åtå*. Uppsala: Almqvist & Wiksell.
- Wiklund, K. B. 1928. Das lappische Verbaladverbium und einige andere Kasus des Verbalstammes. In: *Festskrift til rektor J. Qvigstad. 1853 4. april 1928*. (Tromsø museums skrifter 2), 316–353. Tromsø: Tromsø museum.
- Wilbur, Joshua (ed.). 2016. *Pitesamisk ordbok samt stavningsregler* (Samica 2). Freiburg: Department of Scandinavian Studies.
- Winsa, Birger. 1991. *Östligt eller västligt? Det äldsta ordförrådet i gällivarefinskan och tornedalsfinskan* (Studia Fennica Stockholmiensia 2). Stockholm: Almqvist & Wiksell.
- Ylikoski, Jussi. 2009. *Non-finites in North Saami* (Suomalais-Ugrilaisen Seuran Toimituksia 257). Helsinki: Suomalais-Ugrilainen Seura.
- Ylikoski, Jussi. 2016. Future time reference in Lule Saami, with some remarks on Finnish. *ESUKA – JEFUL* 7–2: 209–244.
- Ylikoski, Jussi. 2022. Lule Saami. In Marianne Bakró-Nagy, Johanna Laakso & Elena Skribnik (eds). *The Oxford guide to the Uralic languages*, 130–146. Oxford: Oxford University Press.
- Ylikoski, Jussi & Olle Kejonen. Forthcoming. Negation in Lule Saami. In Matti Miestamo, Ljuba Veselinova & Héloïse Calame (eds). *Negation in the languages of the world II: Eurasia*. Berlin: Language Science Press.

South Saami finite verb morphology

Richard Kowalik, Nord University, Bodø

1. Introduction

In this article, I argue for a classification of South Saami verbs into three main inflectional classes and present the finite verb forms in the language.

South Saami (glottonym *áarjelsaemien giele*; ISO 639-3: *sma*; Glottocode: *sout2674*) belongs to the Saamic branch of the Uralic language family and constitutes the southern end of the Saamic dialect continuum. South Saami is a small and endangered language spoken in Sweden and Norway. There is no official number of speakers but estimates around 300–500 speakers are usually accepted and agreed upon, see, e.g., Sammallahti (1998). The group of first language speakers is much smaller than that number; the number of second language users on the other hand might be increasing.

South Saami has a comparatively long research tradition that dates back into the 19th century, see, e.g., Halász (1886). However, comprehensive descriptive work has not been carried out since the 1940s. See Bergsland (1946), Hasselbrink (1944) and Lagercrantz (1923), and compare for instance Magga and Magga (2012), which is largely based on Bergsland (1982). This latter publication, in turn, is a standardized version that relies on Bergsland (1946). Furthermore, contemporary material for the language today reflects often a prescriptive practice of language. This article aims, thus, at filling a gap in the description of the language. Additionally, adding to our understanding of South Saami, from the perspective of language *use*, but also from the perspective of language *change*, given the fact that previous studies are based on a language that was in use almost a century ago.

This article is based on the verb chapter of my doctoral thesis, *A grammar of spoken South Saami* (Kowalik 2023), and focuses on morphology and the inflectional patterns of finite verb forms. The copula verb and the negative auxiliary are not included, as they behave differently. The article is organized as follows: In section 2, I present a brief background to the language of investigation and the data for the study. In section 3, I present the general

properties of the verb in South Saami, including the subject markers. The three inflectional classes are presented in section 4. Results are discussed in section 5 and a conclusion is given in section 6.

2. Background

2.1 The language situation

In order to give this study a context, I want to make a brief note on the language situation of South Saami. A way to depict the situation is to categorize the different varieties, or possibly “layers”, that are found in South Saami. Simplified, the language is used by three generations: the older generation (age 70+) of first language speakers; a middle generation of second language speakers, language activists and teachers; and a younger generation of second language speakers, whose language usually relies on teaching and revitalization efforts of the middle generation.¹ Teaching and revitalization follow a prescriptive language use, and speakers that reclaim their language usually depend on, or orientate themselves towards the standardized South Saami. Standardized South Saami, and the official orthography (adopted in 1978), is also used in the written language.

This study is based on spoken language as used by the older generation of speakers. This generation represents arguably the last generation of first language speakers of South Saami, born during the 1940s. The number of these speakers is naturally much lower than the total number of speakers. The older generation uses a language that presents many interesting features which have not yet been documented or studied in detail. This language differs in some respects from the language represented in e.g., Bergsland (1946), but also from the (written) standard language. Exceptions to this division exist; that is, there are fluent speakers, teachers and activists even in the younger generations. These are usually of high value for language revitalization and the speaker community.

The language of the older generation is perhaps the most innovative language of continuous use of South Saami. It reflects the South Saami as spoken by the generation of grandchildren of the speakers that former descriptions are based on, e.g., Hasselbrink (1944). The majority language, that is, Norwegian or Swedish, is usually the stronger language of the speakers. Language contact has contributed to the current structure of the

¹ Exceptions to this division exist; that is, there are fluent speakers, teachers and activists even in the younger generations. These are usually of high value for language revitalization and the speaker community.

language, and contact phenomena and loan words can easily be identified in South Saami. However, the observed features are usually consistent between different idiolects and dialects in the data. The language of study, despite its regional varieties, shares thus a common structure. This is a strong argument against processes that can be observed in connection with language death as described in for instance Dyirbal (Schmidt 1983).

2.2 Earlier and current classifications

Different linguistic classifications and descriptions may serve different purposes, which may range from diachronic to didactic purposes.

Earlier descriptions of the verbal system in South Saami have often a high number of verb groups and distinguish between different “stem classes”, based on the number of syllables of verb stems. Lagercrantz (1923) divides verbs into two groups based on the syllable structure of the verb stem, with three respectively two subgroups. Hasselbrink’s (1981–1985) description has six verb groups, among these a group of “former contracted” verbs. Differences between these groups are restricted to the form of the stem, as suffixes do not differ considerably between the groups. In Bergsland (1946), verbs are divided into two main groups, where one has five subgroups. In Bergsland (1994), seven verb groups can be identified, although the means for this classification are not presented.

Table 1. Verb groups in South Saami according to Frändén et al. (2007)

Group	Disyllabic	Trisyllabic	Irregular	Contracted
1	<i>båtedh</i> ‘come’	<i>goltelidh</i> ‘listen’	<i>lea-</i> (copula)	<i>billijidh</i> ‘be scared’
2	<i>darjodh</i> ‘do’	<i>dåeriedidh</i> ‘follow’	NEG aux	
3	<i>guarkedh</i> ‘understand’	<i>aatskadidh</i> ‘try’		
4	<i>barkedh</i> ‘work’	<i>jåhttjidh</i> ‘move’	<i>edtjedh</i> ‘shall’	
5	<i>domtedh</i> ‘feel’	<i>daaroestidh</i> ‘speak Scand.’		
6	<i>veedtjedh</i> ‘get’			

Current descriptions, for instance Magga & Magga (2012) and learner-focused material such as the verb handbook *Verbh* (‘Verbs’) by Frändén et al. (2007),

operate with a division of verbs into two “main verb groups”. The division is based on the number of syllables of the stem (roughly, disyllabic vs. trisyllabic verbs), with six respectively five further “verb groups”, illustrated (with example verbs from Frändén et al. (2007)) in Table (1) above.

The system presented in these works reflects standardized South Saami. It is valid for verbs in prescriptive language use and sources and its purpose can be said to be mainly didactic.

The classification of verbs based on number of syllables has a long tradition and a historical motivation in Saamic linguistics. The purpose of the proposed classification in the present article, however, is purely synchronic. The corpus data collected for this study did not support a classification based on syllables from a pure synchronic point of view. Instead, a slightly different pattern emerges and I propose a grouping into three *inflectional classes*. I will use this term in contrast to *conjugational classes*, which I define as taking different finiteness markers, and which is not the case in South Saami.

2.3 Data and method

The data consists of around 35.5 hours recording with 12 different speakers of South Saami (most of whom are 75+ years of age today), plus complementary hand-written notes. The corpus material is collected by me (the author) since the year 2017 and ongoing. Of a total of 35.5 hours recording, around 12.5 hours are spontaneous speech, and the remaining 23.5 hours are elicitation. The recordings were processed in the annotation tool *ELAN* (Sloetjes & Wittenburg, 2008), transcribed into South Saami in an adapted orthography and translated into Swedish and English by the author, with the help of speakers if necessary. During data collection, focus is put on free (spontaneous) speech; however, paradigms presented below are usually complemented with elicited forms. The documentation project has furthermore a focus on the southern dialect of South Saami (today mainly spoken in the Swedish county of Härjedalen), but the data cover also central and northern varieties of the language.

The basis for the classification presented in this article are forms of ca. 300 different lexical verbs (types). Of these 300 verbs, a total of around 870 different (inflected) forms are attested. All verb forms were compiled and sorted according to their inflectional behaviour. The current data do not allow for a precise token-count.

3. Verbs in South Saami

3.1. General properties

Verbs in South Saami inflect for person and for the number categories singular and plural. Dual marking exists but is optional and not regularly used. Verbs do not inflect for gender.

South Saami has finite and non-finite verb forms. Verbal categories, such as various tenses, can be encoded in different ways, in bound and periphrastic expressions. Finite verb forms are used for the encoding of present and past. Non-finite verb forms are used in periphrastic expressions that encode perfect, pluperfect and progressive. The future is expressed with present tense or with modal verbs. However, verbal categories are not discussed in this article.

Verbal morphology is mostly concatenative in South Saami; verb forms are formed by attaching suffixes to a stem, and align with the following pattern (A):

(A) stem – (past tense marker) – subject suffix

The verb stem is defined as the form that does not undergo any changes when suffixes are added. The stem is thus not reflected in the infinitive form (which usually triggers reduction of the thematic vowel, e.g., *båetie*- ‘come-’ but *båetedh* ‘to come’). The thematic vowel is the vowel a stem ends in. All stems end in one of the thematic vowels *-e*, *-ie* or *-a*.

As scheme (A) illustrates, it is in general possible to segment and mark morpheme boundaries for verbs in South Saami. However, some forms in the paradigm reflect morpho-phonological alternations which make a segmentation less obvious.

Some scholars have chosen not to mark morpheme boundaries in South Saami, e.g., Ylikoski (2018) or Siegl (2017), whereas others mark boundaries, such as Blokland and Inaba (2015). I choose to mark morpheme boundaries for verbs (but not for the copula) in this article, as I believe that this adds transparency to the analysis.

The subject suffixes are in general the same for all verbs; two forms have allomorphs due to phonotactic restrictions. The stem vowels of the stems to which the suffixes attach, however, may change due to morpho-phonological alternations, and thematic vowels can change due to phonotactic restrictions. Based on this behaviour, verbs in the data are divided into three **inflectional classes**.

Class 1 contains verbs with the thematic vowel *-e* and is the largest verb class. Its inflection is linear but shows reduction of the thematic vowel in some forms. The group has two subgroups, one for disyllabic stems (1a) and one for trisyllabic stems (1b).

Class 2 contains verbs with the thematic vowel *-ie*.² All stems are disyllabic and have vowel alternations of the stem vowel in the present tense singular and in the entire past paradigm. The class is not large but contains verbs with a high frequency.

Class 3 contains verbs with the stem vowel *-a*, all stems are disyllabic. The inflection is linear and shows no alternation of the stem vowel. It is the smallest class but its verbs are also frequent.

A scheme for the thematic vowels, an example verb and the morphophonological characteristics (either no alternation of the stem, or umlaut/vowel alternation in parts of the paradigm) is presented in Table 2. Inflected forms of the example verbs for the first person singular and plural present tense and the first person singular past tense are given in Table 3.

Table 2. Scheme of alternations in the verb groups.

Class	Thematic vowel	Example verb	Morphophonological characteristic
1a	<i>-e</i>	<i>lohke-</i> ‘read’	stem unaltered (except 3SG)
1b	<i>-e</i>	<i>vaestede-</i> ‘answer’	stem unaltered
2	<i>-ie</i>	<i>båetie-</i> ‘come’	vowel altern. in PRS SG and PST
3	<i>-a</i>	<i>darja-</i> ‘do’	stem unaltered

Table 3. Example verbs for the three inflectional classes.

Class	Example verb	Present 1SG	Present 1PL	Past 1SG
1a	<i>lohke-</i> ‘read’	<i>lohke-m</i>	<i>lohke-be</i>	<i>lohk-i-m</i>
1b	<i>vaestede-</i> ‘answer’	<i>vaestede-m</i>	<i>vaestiedi-bie</i>	<i>vaested-i-m</i>
2	<i>båetie-</i> ‘come’	<i>båata-m</i>	<i>båete-be</i>	<i>bööť-i-m</i>
3	<i>darja-</i> ‘do’	<i>darja-m</i>	<i>darje-be</i>	<i>darja-ji-m</i>

The inflectional classes are presented in detail in section 4 below. By inflectional classes I mean different inflectional patterns that share the same finite markers (subject suffixes) but may show some differences in the behaviour of the stem vowel. I use this term in contrast to *conjugational classes*, which I understand as inflectional patterns with fundamental differences such as different person/number suffixes.

² I propose a long vowel phoneme /i:/ for the phonemic status of this vowel (Kowalik 2023: 81).

The verb type frequency³ differs between the classes. Class 1, comprising the sub-classes 1a and 1b, is the largest class and has the highest type frequency. Class 2 and 3 are considerably smaller, but their members have usually a high token frequency. The distribution of the different classes (the verb type frequency) based on a sample of 296 different verbs in the corpus is presented in Table 4:

Table 4. Distribution of verb classes in the data.

Class	Total number of verb types	Percentage
1a	134	44 %
1b	94	32 %
2	46	16 %
3	22	8 %
<i>total: 296</i>		

Based on the data, the following general remarks about verbs in South Saami can be made:

- About two thirds of all verb types in the data have a disyllabic stem, i.e., belong to class 1a, 2 and 3 (see Table 4).
- About one third of all verb types have a trisyllabic stem, i.e., belong to class 1b (see Table 4).
- Verbs with four, five and six syllables can be derived; however, these are less frequent in the data (see section 4 for examples).
- Only three verbs in South Saami are monosyllabic (both their stem and their finite forms): These are the copula (functioning also as auxiliary in periphrastic tenses), the negative auxiliary and a contracted form of the modal verb *galkedh* ‘shall’, i.e., *gam*, *gah*, *ga* (1/2/3SG present tense).
- Vowel alternation (umlaut) is important for one of the three verb classes (class 2, see section 4.1), which has fewer types but a high token frequency.
- Around eight percent of all verbs in the data are loan words (from Swedish/ Norwegian). All loan verbs are members of class 1 (mainly of 1a, some of 1b), which is also the largest class.
- Based on the accommodation of loan words, class 1 can be described as the only productive one.

³ I am aware of the limitations of these frequency counts. However, the figures may still be useful for a better understanding of the use of the verb classes in South Saami.

3.2 Subject suffixes

The subject suffixes (or person/number suffixes) are the same for all inflectional classes, with the exception of allomorphic forms for the first and second person plural suffixes (*-be/-bie* and *-de/-die*) due to phonotactic restrictions in the language, and two other allomorphs due to geographical variation. The suffixes for the present and past tense are presented in Table 5 below. The parentheses in the columns for the dual indicate that these suffixes have only been attested in elicitation.

Table 5. Subject suffixes for present and past.

	Present			Past		
	Singular	Dual	Plural	Singular	Dual	Plural
1	<i>-m</i>	<i>-n</i>	<i>-be, -bie; -o</i>	<i>-m</i>	<i>-men</i>	<i>-mh; -o, -bo</i>
2	<i>-h</i>	<i>(-den)</i>	<i>-de, -die</i>	<i>-h</i>	<i>-den</i>	<i>-dh</i>
3	<i>-Ø</i>	<i>(-gan)</i>	<i>-h</i>	<i>-Ø</i>	<i>(-gan)</i>	<i>-n</i>

The singular suffixes (*-m*, *-h*, *-Ø*) are identical for both tenses. The suffixes for the dual are different in the two tenses in the first person but are identical for the second and third person. Plural marking is different for all persons in the two tenses both in the standard and northern varieties: (*-be/-bie*, *-de/-die*, *-h* in present tense, *-mh*, *-dh*, *-n* in past tense). In the southern dialects, first person plural is identically marked in both tenses (*-o*). Variation and allomorphy are discussed in the following paragraphs.

3.2.1 Phonotactic allomorphy: *-be/-bie* and *-de/-die*

In Table 5 above, there are two suffixes listed for the first and second person plural present tense, *-be/-bie* and *-de/-die*. I will call these suffixes “short” (*-be*, *-de*) and “long” (*-bie*, *-die*) subject suffixes. As a rule, short suffixes attach to all disyllabic verb stems (that is, verbs belonging to class 1a, 2 and 3), long suffixes attach to trisyllabic verb stems (class 1b). These suffixes are thus allomorphs which occur in phonotactically different environments.

The allomorphy is triggered by word stress and phonotactic constraints in the language. Words in South Saami have a fixed word-initial stress. Words with more than three syllables receive a secondary stress on the penultimate syllable. Suffixes are always unstressed. A phonological word with the structure "CV.CV (two short open syllables with stress on the first one) is avoided. Instead, the final vowel (that is part of the suffix) is lengthened: *-be* into *-bie*, *-de* into *-die*. Thus, it is *loh.ke.be* ‘we read’ but *vae.stie.di.bie* ‘we answer’. See also the paradigms in section 4.1 below.

Some very few counterexamples to this rule exist in the data, in both directions (short suffixes used with trisyllabic verbs, e.g., *gavlesebe* ‘we hear each other’, and long suffixes used with disyllabic verbs). This could possibly indicate that the difference between the suffixes in general is less important to the speakers.

3.2.2 Dialectal allomorphy (1): final *-m* and *-b*

Another allomorphy is the dialectal variation between final *-m* and final *-b*. This is a geographical variation; southern and standard South Saami use final *-m*, northern dialects are attested using final *-b*. In the verbal system, this affects the first person singular (*-m*) as well as the first person plural past tense (*-mh*).

Examples are given below: (1) and (3) are examples from southern dialects (*eelki-m*, *doeki-mh*), and (2) and (4) (*eelki-b*, *lohki-bh*) from northern dialects:

- (1) *manne* ***eelk-i-m*** *skovle-m*
 1SG.NOM start-PST-1SG school-ACC.SG
desnie *Mihte-sne*
 DEM.LOC.SG Mittådalen-LOC.SG
 ‘I started school there in Mittådalen.’ [sma20170516a]
- (2) *goh* *manne* *lim* *tijhtje* *jaepie-n*
 when 1SG.NOM be.PST.1SG seven year-GEN.SG
båaeries dellie ***eelk-i-b*** *skovle-sne*
 old then start-PST-1SG school-LOC.SG
 ‘When I was seven years old then I started in school.’ [sma20170919a]
- (3) *dah* *ledtie-h* *mijjieh* ***doek-i-mh***
 DEM.NOM.PL bird-NOM.PL 1PL.NOM sell-PST-1PL
 ‘We sold these birds.’ [sma20170614d]
- (4) *jih* *dihte* *learare*
 and DEM.NOM.SG teacher
idtji *dihte* *guarka-h*
 NEG.AUX.PST.3SG 3SG.NOM understand-IRR
maam *mijjieh* ***lohk-i-bh***
 what.ACC.SG 1PL.NOM read-PST-1PL
 ‘And that teacher, he didn’t understand what we read.’ [sma20170926j]

The suffix *-b* as 1SG marker are attested with all speakers from the northern area in the data. However, speakers are not always consistent and may also use *-m*. This might be due to the influence from the standard language.

3.2.3. Dialectal allomorphy (2): The southern marker *-o*

In the southern dialects of South Saami (Härjedalen), a different marker for the first person plural past is attested. Whereas the standard marker is *-mh*, with its northern variant *-bh*, in the dialects of Härjedalen, *-o*, sometimes *-bo*, is also used. In the data, the suffix occurs more frequent in the past tense. The suffix *-o* is attached either to the verb stem (in the present) or to the past tense marker *-i* (in the past). The suffix *-bo* requires the “long” verb forms and attaches to the long past tense marker *-ji* or trisyllabic verbs. With the exception of trisyllabic verb stems, both variants can be used with the same verb. First person plural past tense forms can thus be either of the following (see examples in Table (6)), depending on the dialect:

Examples for these suffixes are given in (5), (6) and (7):

Table 6. First person plural past tense forms.

Verb stem	Standard	Northern	Southern (1)	Southern (2)
<i>juhtie-</i> ‘move’	<i>juht-i-mh</i>	<i>juht-i-bh</i>	<i>juht-i-o</i>	<i>juhtie-ji-bo</i>
<i>barke-</i> ‘work’	<i>bark-i-mh</i>	<i>bark-i-bh</i>	<i>bark-i-o</i>	<i>barka-ji-bo</i>
<i>höölteste-</i> ‘dwell’	<i>hööltest-i-mh</i>	<i>hööltest-i-bh</i>	<i>n/a</i>	<i>hööltest-i-bo</i>

- (5) *dellie monnah datnem sijht-o gihtjij*
 now 1DU.NOM 2SG.ACC want.PRS-1PL ask.INF
naan baaka-j bijre
 some word-GEN.PL around
 ‘Now we want to ask you about some words.’ [sma20200219e]

- (6) *mijjieh provhk-i-o tjoekij*
 1PL.NOM use.to.do-PST-1PL ski.INF
Haandskina-sse
 Handskinnsvålen-DAT.SG
 ‘We used to ski to Handskinnsvålen’ [sma20180614a]

- (7) *jih bijjen Rödfjäll-sne giedtie*
 and ontop Rödfjäll-LOC.SG reindeer.field
lij gusnie mijjieh miesie-h
 be.PST.3SG where 1PL.NOM calf-NOM.PL
mierhkiesj-i-bo
 mark-PST-1PL

‘And up on Rödfjäll was the field where we marked the reindeer calves.’
[sma20180605c]

However, the *-o/bo* suffix is not used consistently by all speakers of the southern dialect. This is probably due to the influence of the standard language.

3.2.4 Dialectal allomorphy (3): final *-n* instead of *-m*

Another variation of the first person singular suffix is attested in the data of two of the southern speakers. Instead of the first person marker *-m* (or *-b* in northern dialects), *-n* can also be used (8):

(8)	<i>manne</i>	<i>mujhtie-ji-n</i>	<i>gumhtie</i>
	1SG.NOM	remember-PST-1SG	how
	<i>maahte-me</i>	<i>orre-me</i>	
	can-PTCP	be-PTCP	
	‘I remembered how it could have been.’ [sma20181025k]		

In the data, this is restricted to two speakers from the southern dialects (Härjedalen).

The final *-n* is found in North Saami as well, where it is the standard first person singular marker. It reflects the change from the Proto-Saamic 1SG marker **-m* into *-n*.

3.2.5 Dual suffixes

The dual as a category requires a short remark of its own. Dual verb morphology, with the exception of the first person, is in general not used in free speech by the speakers I worked with. The dual suffixes in parentheses in Table 5 are only attested in elicitation, whereas the suffixes for first person present (*-n*) and first and second person past (*-men*, *-den*) are attested in free, spontaneous speech. That is, third person dual forms are not attested in spontaneous speech in the data. Already Bergsland (1946, p.171) noted that dual verb morphology is in general not used. Many native speakers have a meta-linguistic awareness for this, and I have recorded comments like “we don’t use these [dual] forms” several times (e.g., in recording [sma20170508d]). Loss of dual number is cross-linguistically common, see, e.g., Corbett (2000: 269). The use of dual pronouns, however, is still obligatory in South Saami; this can be observed in other languages as well, such as Slovene (Corbett 2000: 94).

In the data, two strategies for dual reference have emerged: a neutralization of *number* (DU=PL, i.e., plural verb morphology is used for dual reference,

combined with dual pronouns)⁴ and a neutralization of person (one dual form, historically the first person dual marker, is used for all persons in the dual). The first strategy is more common in the data. It is in line with cross-linguistic tendencies and found in other Saamic languages as well, see, e.g., Feist (2015: 200). The second strategy, in which the dual number feature is kept but without person distinction, has to the best of my knowledge previously not been described for South Saami. This topic, however, will not be discussed any further in the present article, but the general absence of dual marking means that less morphological forms and changes of stem vowels need to be taken into account in the system.

4. Inflectional classes of verbs

In this section, I present paradigms for the three inflectional classes and discuss their inflection in more detail. The classes are presented in a descending order according to their type frequency (their size).

4.1. Inflectional class 1: e-stems

Inflectional class 1 consists of verbs with a thematic vowel -e and is the largest class. It includes both disyllabic and trisyllabic verb stems, which each form a subclass, labelled (1a) and (1b). Common for all verbs in this class is that the stem does not show vowel alternation. An exception is the third person singular present tense of disyllabic verbs, which is marked with a morpho-phonological alternation of the stem vowel and the thematic vowel. Another feature common for verbs of class 1 is that the thematic vowel is dropped before the past tense marker -i-. The personal pronouns do not distinguish between third person dual and plural.

Example verbs for subclass (1a) are *lohke-* ‘read’ and *gööle-* ‘fish’; example verbs for subclass (1b) are *vaestede-* ‘answer’ and *tjihtede-* ‘push one’s way forward’. See the paradigms in Tables 7 and 8 for subgroup (1a):

Table 7. Inflectional class 1a: *lohke-* ‘read’.

	Singular	Dual	Plural
Present 1	<i>lohke-m</i>	<i>lohkie-n</i>	<i>lohke-be</i>
2	<i>lohk-h</i>		<i>lohke-de</i>
3	<i>lâhka-Ø</i>		<i>lohkie-h</i>
Past 1	<i>lohk-i-m</i>	<i>lohk-i-men</i>	<i>lohk-i-mh, -o</i>

⁴ The personal pronouns do not distinguish between third person dual and plural.

2	<i>lohk-i-h</i>	<i>lohk-i-dh</i>
3	<i>lohk-i-Ø</i>	<i>lohk-i-n</i>

The thematic vowel *-e* disappears in the second person singular present tense when it meets the suffix *-h* (i.e., *lohke+h* becomes *lohkh*).

Table 8. Inflectional class 1a: *gööle-* ‘fish’.

	Singular	Dual	Plural
Present 1	<i>gööle-m</i>	<i>göölie-n</i>	<i>lohke-be</i>
2	<i>gööl-h</i>		<i>lohke-de</i>
3	<i>göölie-Ø</i>		<i>lohkie-h</i>
Past 1	<i>gööl-i-m</i>	<i>gööl-i-men</i>	<i>gööl-i-mh</i> , <i>gööli-ji-bo</i>
2	<i>gööl-i-h</i>	<i>gööl-i-dh</i>	
3	<i>gööl-i-Ø</i>	<i>gööl-i-n</i>	

The third person plural ends either in *-ie* (e.g., *göölie*) or in *-a* (e.g., *lähka*). Synchronically, this can be described as follows: Verbs with a back vowel in the stem, as *lohke-*, are marked with a vowel alternation in the third person present tense and end on *-a* (*lähka*). Verbs with a front vowel in the stem, like *gööle-*, end on a long vowel *-ie* in the third person present tense (*göölie*). All verbs of this class show a long vowel *-ie* in the dual form.

Trisyllabic verbs (subgroup 1b) behave slightly different. While the stem remains unaltered, the vowels in the suffixes of trisyllabic verbs in the present tense plural are lengthened due to the phonotactic structure of a word in South Saami: 1PL *-be* into *-bie*, 2PL *-de* into *-die*. These suffixes are thus allomorphs which occur in different environments: *-be/-de* attach to disyllabic verb stems and follow thus an unstressed syllable (e.g., *loh.ke-be* ‘we read’), *-bie/die* attach to trisyllabic verb stems and follow thus a syllable with a secondary stress (e.g., *’tjih.te.di-bie* ‘we push forward’).

Paradigms for these two verbs are presented in Tables 9 and 10 below:

Table 9. Inflectional class 1b: *vaestede-* ‘answer’.

	Singular	Dual	Plural
Present 1	<i>vaestede-m</i>	<i>vaestiedie-n</i>	<i>vaestiedi-be</i>
2	<i>vaested-h</i>		<i>vaestiedi-de</i>
3	<i>vaestede-Ø</i>		<i>vaestied-h</i>
Past 1	<i>vaestied-i-m</i>	<i>vaestied-i-men</i>	<i>vaestied-i-mh</i>
2	<i>vaestied-i-h</i>	<i>vaestied-i-dh</i>	
3	<i>vaestied-i-Ø</i>	<i>vaestied-i-n</i>	

Table 10. Inflectional class 1b: *tjihtede-* ‘push forward’ (No dual forms attested).

	Singular	Plural
Present 1	<i>tjihtede-m</i>	<i>tjihtedi-bie</i>
2	<i>tjihted-h</i>	<i>tjihtedi-die</i>
3	<i>tjihtede-Ø</i> , <i>tjihtede-a</i>	<i>tjihted-h</i>
Past 1	<i>tjihted-i-m</i>	<i>tjihted-i-mh</i>
2	<i>tjihted-i-h</i>	<i>tjihted-i-dh</i>
3	<i>tjihted-i-Ø</i>	<i>tjihted-i-n</i>

Some verbs of this group with a long vowel in the stem, such as *vaestede-* ‘answer’, show a change of the second vowel (V2) *-e-* into *-ie-* (*vaestede-* vs. *vaestiede-*). The possible vowel alternations of V2 of verbs belonging to class 1b are schematized in Table 11 below:

Table 11. Vowel change of V2 in trisyllabic verbs (1b).

V2 vowel change	example verb stem	example verb form
<i>e = e</i>	<i>tjihtede-</i> ‘push’	<i>tjihtedim</i> (1SG past)
<i>e ⇒ i</i>	<i>vaestede-</i> ‘answer’	<i>vaestiedim</i> (1SG past)
<i>e ⇒ a</i>	<i>vuartesje-</i> ‘see’	<i>vuartasjim</i> (1SG past)
<i>a/o = a/oe</i>	<i>daaraste-/</i> <i>daaroste-</i> ‘talk Scand.	<i>daarastim</i> (1SG past) <i>daaroestim</i> (1SG past)

Whereas the first pattern (*e = e*, no vowel change) is the most frequent one in the data (see *tjihtede-* ‘push forward’ in Table 10), the changes *e* into *ie* and *e* into *a* are also common. The last pattern occurs least in the data.

Some verbs of subclass 1b can be shortened by syncope; usually, the vowel in the second syllable is omitted. For instance, the verbs *sop.tses.tidh* ‘talk’ or *hööl.tes.tidh* ‘live, dwell’ is contracted to *soptstidh* [sopts.tit] resp. *hööltstidh* [höölts.tit].

Verbs that originate from Swedish or Norwegian loans, as well as new loan-words, are usually members of this verb class (usually 1a). Example verbs are *provhke-* ‘use to’ (Swe/Nor *bruka/bruke*), *skeele-* ‘peel’ (Swe/Nor *skala/skrelle*), *bigke-* ‘build’ (Swe/Nor *bygga/bygge*), *tuhtje-* ‘think, consider’ (Swe *tycka*), *lyjhke-* ‘like’ (Nor *like*), *rååke-* ‘meet’ (Swe/Nor *råka/råke*). The other pattern, 1b, is less usual but occurs as well: *heannede-* ‘happen’, cf. Swe/Nor *hända/hende*.

4.2 Inflectional class 2: ie-stems

Inflectional class 2 consists of verbs with an even number of syllables in the stem that end in the vowel *-ie*, such as *båetie*- ‘come’, *tjælie*- ‘write’, *juhtie*- ‘move’ or *vedtie*- ‘give’. This class has vowel alternations in the stem in the present singular and in the past (as well as the non-finite forms infinitive and perfect participle). The past is formed by both vowel alternation and the past tense marker *-i*. The stem remains unaltered in the first person dual and third person plural present tense. Example paradigms for the class 2 verbs *båetie*- ‘come’ and *juhtie*- ‘move’ are given in Tables 12 and 13 below. Verb class 2 comprises about 20 percent of all verbs in the data and has thus a rather low type-frequency, but the token-frequency for these verbs is usually high.

Table 12. Inflectional class 2 verb *båetie*- ‘come’.

	Singular	Dual	Plural
Present 1	<i>båata-m</i>	<i>båetie-n</i>	<i>båete-be</i>
2	<i>båata-h</i>		<i>båete-de</i>
3	<i>båata-Ø</i>		<i>båetie-h</i>
Past 1	<i>bööť-i-m,</i> <i>båetie-ji-m</i>	<i>bööť-i-men</i>	<i>bööť-i-mh,</i> <i>bööť-i-o</i>
2	<i>bööť-i-h</i>	<i>bööť-i-dh</i>	
3	<i>bööť-i-Ø,</i> <i>båetie-j-a</i>	<i>bööť-i-n</i>	

Table 13. Inflectional class 2 verb *juhtie*- ‘move’.

	Singular	Dual	Plural
Present 1	<i>jahta-m</i>	<i>juhtie-n</i>	<i>juhte-be</i>
2	<i>jahta-h</i>		<i>juhte-de</i>
3	<i>jahta-Ø</i>		<i>juhtie-h</i>
Past 1	<i>juht-i-m</i>	<i>juht-i-men</i>	<i>juht-i-mh, -o, -ji-bo</i>
2	<i>juht-i-h</i>	<i>juht-i-dh</i>	
3	<i>juht-i-Ø, juhtie-ji-Ø</i>	<i>juht-i-n, juhtie-ji-n</i>	

There is another pattern for forming the past tense in this class attested in the data. In this pattern, the stem remains unaltered and the past tense marker *-i* is suffixed to the stem. To avoid hiatus, a semi-vowel *j* is inserted before the *-i*. The open syllable *-ie* is thus closed into *-ij*. Then, the subject markers are suffixed. This results in the so called “long” past tense forms of the verbs (cf. the past tense of inflectional class 3). Examples are given in the Tables 12 and

13 above: 1SG past *båetiejim*, 3SG past *båetija* and 3SG past *juhtieji*, 1PL past *juhtiejibo*, 3PL past *juhtiejin*.

For *juhtie*- ‘move’, the stem vowel *u* is only affected for the singular forms of the present tense (*u* into *a*); there is no vowel alternation in the past tense. Different forms are also attested in the past tense; both regular and “long” forms are attested in the data, exemplified in the table with the third person singular (*juht-i = juhtie-ji*) and third person plural (*juht-i-n = juhtie-ji-n*). This variation is characteristic for inflectional class 2 verbs in the data.

The third person singular past tense has yet another parallel form, ending in *-a* instead of *-Ø*: For the long forms *båetie-ji -Ø* come-PST-3SG, *juhtie-ji -Ø* move-PST-3SG the parallel forms *båetie-j-a* come-PST-3SG, *juhtie-j-a* move-PST-3SG are also attested. This occurs also with verbs of class 1a. It is a less frequent form in the data, but occurs regularly, especially in the data for the southern dialects. It shows a parallel to the third person singular present tense form of these verb classes as it also ends in *-a*.

Vowel alternations, and a marker for the singular present tense triggered by morphophonology are the defining characteristics of this inflectional class. These forms are difficult to motivate synchronically. However, these are frequent forms and the verbs of this class are frequent, so the forms may be explained from a usage perspective.

Interestingly, several tokens without vowel alternation (where one would expect such) are attested in the data. This suggests that vowel alternations, which reflect umlaut patterns, might not be of fundamental importance for inflection. Examples are: 3SG past of *vuejnedh* ‘see’, attested forms are *vööjni* (stem vowel alternation) and *vuajni* (no alternation in the stem) (cf. the long form: *vuejnieji*).

4.3 Inflectional class 3: a-stems

Inflectional class 3 is a rather small class and comprises about 10 percent of the verbs in the data. However, verbs of class 2 can also be inflected according to the past tense scheme of this class, which makes it an interesting inflectional pattern. Whereas verb type frequency is low, token frequency of its class members is often high. Inflectional class 3 consists of verbs with a disyllabic stem in *-a*. In the northern dialects, this vowel is usually *-oe*. Paradigms for the example verbs are given in Tables 14 and 15 for *dåaja*- ‘cut’ and *guarka*- ‘understand’. Other examples for verbs in this class are *tjåadtja*- ‘stand’ and *darja*- ‘do’.

Inflectional class 3 verbs have no vowel alternation in the stem. Exceptions can be, again, found in the third person singular present tense. Like verbs of class 1a, this form may also be marked with a morpho-phonological device

reflecting umlaut (e.g., *gårkå*, *dååjå*). Usually, the stem alone is used as 3SG form (as in *guarka*, *dåaja*).

Table 14. Inflectional class 3 verb *dåaja*- ‘cut’.

	Singular	Dual	Plural
Present 1	<i>dåaja-m</i>	<i>dåajie-n</i>	<i>dåaje-be</i>
2	<i>dåaja-h</i>		<i>dåaje-de</i>
3	<i>dåaja-Ø</i> , <i>dååjå-Ø</i>		<i>dåaja-h</i>
Past 1	<i>dåaja-ji-m</i>	<i>dåaja-ji-men</i>	<i>dåaja-ji-mh</i>
2	<i>dåaja-ji-h</i>		<i>dåaja-i-dh</i>
3	<i>dåaja-ji-Ø</i>		<i>dåaja-ji-n</i>

Table 15. Inflectional class 3 verb *guarka*- ‘understand’.

	Singular	Dual	Plural
Present 1	<i>guarka-m</i>	<i>guarkie-n</i>	<i>guarke-be</i>
2	<i>guarka-h</i>		<i>guarke-de</i>
3	<i>guarka-Ø</i> , <i>gårkå-Ø</i>		<i>guarka-h</i>
Past 1	<i>guarka-ji-m</i>	<i>guarki-i-men</i>	<i>guarka-ji-mh</i>
2	<i>guarka-ji-h</i>		<i>guarka-ji-dh</i>
3	<i>guarka-ji-Ø</i>		<i>guarka-ji-n</i>

The past tense is formed by adding the past tense maker *-i-* to the stem (i.e., *dåaja+i* ‘s/he cut’), resulting in the extra syllable *-ji-* (*dåaja-ji-*). The class shows syncretism in 2SG and 3PL present tense. In the southern dialects, an alternative 3PL form on *-ieh* (e.g., *dåajieh* ‘they cut’) is attested.

This group reflects a difference between the data and the standard language: whereas members of this class in the data end on *-edh* in the infinitive (like *darjedh* ‘make’, *arredh* ‘be, live’, *saarnedh* ‘tell’), they end in *-odh* in the standard language (*darjodh*, *årrodh*, *saarnodh*).

5. Discussion

In the data, three main inflectional patterns emerge, which are here described as three inflectional classes of verbs. The classes do not constitute different conjugational classes but reflect different inflectional patterns. The classification is primarily based on the thematic vowel of verb stems (either *-e*, *-ie* or *-a*) and this vowel’s behaviour in inflection, instead of being based on the distinction of syllables of a verb. However, phonotactics in the language are still relevant in the system, and the subgroup of class 1 are partly based on

the number of syllables. The verbal system in South Saami is usually known to be complex, but the morphology is generally concatenative. Morpho-phonological vowel alternations in the stem are restricted to one, albeit frequent, inflectional class. Especially the third person singular present tense is marked by this device. This type of marking is synchronically unmotivated, but occurs in frequent forms and may therefore be explained from a usage perspective. In these frequent forms, umlaut patterns may be preserved. These verb forms in the paradigm can be seen as the result of frequent use and diachronic changes.

The third person singular is in general zero-marked and has thus no explicit subject suffix. As has been shown, the third person plural is expressed with the bare verb stem in most classes, and the stem plus past tense marker in the past tense. However, there seems to be a certain association with *-a* as a third person marker. While some verbs show a stem vowel alternation and a third person marker *-a/-ǫ*, in the past tense long forms, an *-a* is added to the past tense marker (e.g., *juhtie-j-a* move-PST-3G). Whether this is an analogy to the morpho-phonological marking (as explained above) or not remains to be investigated. Another possibility is that *-ja* is an allomorph of the past tense marker (*i = ji = ja*).

Stem vowel alternation is a characteristic feature of South Saami. However, it is not a predominant feature and occurs mainly within one verb class (class 2). This class has a low type frequency, but its members have a high token frequency. The majority of verbs, as well as loan words, are inflected according to a linear, concatenative pattern without vowel alternations (class 1). Variations in verb forms, that is, forms without vowel alternations (such as *vuajni* instead of *vööjni*) as well as alternations where no such are expected (*gārka* instead of *guarka*) are attested. This could indicate that South Saami is undergoing change regarding vowel alternations, as these alternations are not entirely obligatory.

In this article, focus was put on finite verb forms. Space did not allow to present non-finite forms. However, these forms and their segmentation aligns with the proposed classification.

6. Conclusions

In this article, I presented a classification of South Saami verbs into inflectional classes based on empirical data of spoken language. The purely synchronic classification is an attempt to present the verbal system as coherent and consistently as possible, and delete any redundancies in the description. The verbal system as described here differs from other contemporary

descriptions mainly in the fact that there are fewer inflectional classes. The difference between “odd-syllabic” (trisyllabic) and “even-syllabic” (disyllabic) verbs is reduced and they form here one verb class with stems ending in *-e*. Verbs with an infinitive form on *-odh* are not attested in the data, which allows to incorporate these verbs in the *a*-stem verb group. The proposed classification is closer to Lagercrantz (1923) and Bergsland (1946) than later, more standardized descriptions that reflect a prescriptive language, e.g., Bergsland (1994), Magga and Magga (2012). In that respect, the earlier classifications were possibly more descriptive than later literature; see Ylikoski (2017: 322) for a similar stance on that matter.

Apart from the verbal system as a whole, several details such as allomorphy, lengthening of suffixes and phonotactic restrictions are discussed. Some topics, however, still require further investigation. Among these is the vowel lengthening in some verb forms (for instance in the third person singular present tense and in the dual forms of all verbs) and the vowel changes in trisyllabic verb stems; another is the morpho-phonological marking of the third person singular. Perhaps the most interesting topic is the role of vowel alternations in the language, which might be undergoing change.

Abbreviations

1, 2, 3—First, second, third person; ACC—Accusative case; AUX—Auxiliary; dem—Demonstrative; DU—Dual; GEN—Genitive case; INF—infinitive; ILL—Illative case; IRR—Irrrealis; LOC—Locative case; NEG—Negation; NOM—Nominative case; PL—Plural; PRS—Present; PST—Past; SG—Singular.

References

- Bergsland, Knut. 1946. *Røros-lappisk grammatikk: et forsøk på strukturell språkbeskrivelse*. Oslo: Aschehoug.
- Bergsland, Knut. 1982. *Sydsamisk grammatikk*. Oslo: Universitetsforlaget.
- Bergsland, Knut. 1994. *Sydsamisk grammatikk*. Karasjok: Davvi girji.
- Blokland, Rogier & Nobufumi Inaba. 2015. Negation in South Saami. In Matti Miestamo, Anne Tamm & Beáta Wagner-Nagy (eds), *Negation in Uralic languages* (Typological Studies in Language 108), 377–398. Amsterdam: John Benjamins.
- Corbett, Greville G. 2000. *Number*. Cambridge: Cambridge University Press.
- Feist, Tim. 2015. *A grammar of Skolt Saami* (Suomalais-Ugrilaisen Seuran Toimituksia 273). Helsinki: Suomalais-Ugrilainen Seura.
- Frändén, Marit, Björn Lundqvist & Karin Wilson. 2007. *Verbh: en sydsamisk verbhandbok*. Uppsala: Konsultförlaget/Uppsala Publishing House.
- Halász, Ignác. 1886. *Svéd-lapp nyelv. II. Jemtlandi lapp nyelvmutatványok* (Ugor füzetek. 8. szám). Budapest: Magyar Tudományos Akadémia.
- Hasselbrink, Gustav. 1944. *Vilhelminalapskans ljudlära: med särskild hänsyn till första stavelsens vokaler*. Uppsala: Almqvist & Wiksells.

- Hasselbrink, Gustav. 1981–1985. *Südlappisches Wörterbuch I-III* (Schriften des Instituts für Dialektforschung und Volkskunde in Uppsala. Ser. C4). Uppsala: Lundequistska bokhandeln/Dialekt-och folkminnesarkivet i Uppsala.
- Kowalik, Richard. 2023. *A grammar of spoken South Saami*. Phd dissertation. Stockholm: Stockholm university, Department of Linguistics.
- Lagercrantz, Eliel. 1923. *Sprachlehre des Südlappischen nach der Mundart von Weßen* (Kristiania Etnografiske Museum Bulletin 1). Kristiania: Kristiania Etnografiske Museum.
- Magga, Ole Henrik & Lajla Mattsson Magga. 2012. *Sørsamisk grammatikk*. [Kárásjohka]: Davvi girji.
- Sammallahti, Pekka. 1998. *The Saami languages: an introduction*. Kárásjohka: Davvi girji.
- Schmidt, Annette. 1983. *Young people's Dyirbal: an example of language death from Australia*. Cambridge: Cambridge University Press.
- Siegl, Florian. 2017. The essive in South Saami. In Casper de Groot (ed.), *Uralic essive and the expression of impermanent state* (Typological Studies in Language 119), 185–215. Amsterdam: John Benjamins.
- Sloetjes, Han & Peter Wittenburg. 2008. Annotation by category - ELAN. In *Proceedings of the Sixth International Conference on Language Resources and Evaluation (LREC'08)*, Marrakech, Morocco. European Language Resources Association (ELRA).
http://www.lrec-conf.org/proceedings/lrec2008/pdf/208_paper.pdf
- Ylikoski, Jussi. 2017. What made Proto-Germanic *jah ‘and’ an infinitive marker in westernmost Uralic? Observations from the Saami-Scandinavian border. *Norsk Lingvistisk Tidsskrift* 35(2): 291–327.
- Ylikoski, Jussi. 2018. The so-called relation forms of nouns in South Saami: A byproduct or remnant of Uralic *-mpV? *Finnisch-Ugrische Forschungen* 64. 6–71.

Inchoative verbs in Saami: Derivational types and their variation

Juha Kuokkala, University of Helsinki

1. Introduction

1.1. Saami verbal derivation and inchoatives

The Saami languages are known for their rich derivational morphology, especially when it comes to verbal derivation. A considerable share of the system of verbal derivation consists of aspectual derivatives, which refer to deverbal verbs changing the lexical aspect or Aktionsart of the base verb but not (considerably) affecting the actual meaning or valency of the verb. Aspectual derivatives can be classified on semantic grounds based on the modification they apply to the base verb; the following classes have been commonly used in Uralic studies (examples from North Saami, derivational morphemes marked in bold face):

- Frequentative: *láhp**adit*** ‘lose (many objects)’ ← *láhppit* ‘lose’, *báhč**alit*** ‘shoot (many times)’ ← *báhčit* ‘shoot’
- Momentative (punctual): *čolg**adit*** ‘spit once’ ← *čolgat* ‘spit’, *njuiket* ‘jump once’ ↔ *njuik**ut*** ‘jump many times’
- Subitive: *njuik**estit*** ‘jump quickly’ ← *njuiket* ‘jump once’
- Diminutive: *att**estit*** ‘give a little’ ← *addit* ‘give’
- Conative: *bokt**alit*** ‘try to wake’ ← *boktit* ‘wake’
- Inchoative: *buol’**lát*** ‘catch fire’ ← *buollit* ‘burn (intr.)’, *lohkag**oahtit*** ‘begin to read’ ← *lohkat* ‘read’

(See Sammallahti 1998: 91–93; Nielsen 1979 [1926]: 229–279, 283–286; as for South Saami, see Magga & Magga 2012: 131–166; Hasselbrink 1981–1985: 177–186; for an extensive study of Pite Saami verbal derivation, see Ruong 1943.)

The term **inchoative** is used here in accordance with its traditional use in Uralic studies, meaning a beginning aspect added to the meaning of a base verb, usually with a derivational suffix, and synonymous with the term **inceptive** in some traditions (also in Feist’s 2015 Skolt Saami grammar). This is in contrast to the use of “inchoative” in Indo-European studies, where it is a somewhat broader concept, also involving denominal state-change verbs such as *darken* or *liquefy*. In the present paper, the term **inchoative verb** is thus used synonymously with **inchoative derivative**, so that the concept is bound to a derivational relation to a base verb, not only to the semantics of ‘beginning’.

The Saami languages use a wide array of suffixes to derive inchoative verbs. The suffix most purely dedicated to the inchoative function is SaaN *-(š)goahti-* (with cognates in all Saami languages):

- *lohkat* ‘read’ → *lohkagoahtit* ‘begin to read’

But inchoative verbs can also be formed using suffixes with several aspectual functions (examples from North Saami):

- *-d-*: *šlahtádit* ‘begin to sleet’ ← *šlahttit* ‘sleet’
cf. *gahčádit* ‘fall repeatedly’ ← *gahččat* ‘fall, tumble down’
- *-al-*: *čuoigalit* ‘set off on skis’ ← *čuoigat* ‘ski’
cf. *báhčalit* ‘shoot (many times)’ ← *báhčit* ‘shoot’
or *boktalit* ‘try to wake’ ← *boktit* ‘wake’
- *-á(j)-*: *vuoddját* ‘begin to drive, drive off’ ← *vuodjit* ‘drive’
cf. *vadnát* ‘lie fully extended’ ← *vatnat* ‘be stretched out, become longer’

As can already be seen in the previous examples above, formally identical derivational suffixes can on different occasions carry (at least partly) different aspectual meanings. The aspectual meaning tends to be lexicalized, fixed with a certain derived lexeme, but may also vary contextually, and several of the possible meanings of a certain suffix may be present simultaneously. This contextual variation and (re)interpretation of the meaning is apparently a key factor in the diachronic change in the functions of the suffixes and, consequently, in the study of these developments.

1.2. Goals of the study

Lexical derivation has to some extent been described in most grammatical works on Saami languages, though the scope of such descriptions varies considerably. The goal of the current article is to compose an overall picture of one semantic group of verbal derivations, i.e., inchoatives, throughout the Saami language area, where the frequencies of different derivational suffixes are explicitly reported based on dictionary data. The data will be first analyzed per language (Section 2) and after that, each derivative type is treated briefly from a cross-linguistic perspective (Section 3). Based on these comparative results, an attempt will be made to outline the diachronic development of inchoative derivation in the Saami language group, presenting some general trends as well as some more detailed lines of development.

1.3. Data and method

The survey in the following section will be based on lexical data collected from the largest existing dictionaries for each Saami language, which mostly are descriptive dictionaries covering several dialects or dictionaries for standard language made from the data of such dialect dictionaries. For the lowest resourced languages, data has had to be collected from more varying sources. The main sources for each language are the following:

- South Saami: Bergsland & Magga 1993 (ca. 19,000 lexemes) – a general dictionary based on the comprehensive dialect dictionary of Hasselbrink (1981–1985) and some additional dialect records;
- Ume Saami: Schlachter 1958 (ca. 6,000 lexemes) – vocabulary recorded from a single informant of the Máláge/Malå dialect; with a few additions from Barruk (2018);
- Pite Saami: 1) The verbal derivative lists of Ruong (1943), combined with 2) Wilbur 2016 and the somewhat larger online word list (BB; ca. 6,000 lexemes) – a compact general dictionary based on a word list compiled by a group of native Pite Saami speakers – as well as 3) Halász 1896 (ca. 3,000 lexemes) – a dictionary based on the language of a single family;
- Lule Saami: O. Korhonen 2007 (ca. 19,000 lexemes) – a general dictionary based partly on central Lule Saami data from Grundström’s dialect dictionary (1946–1954) and partly the

author's own contemporary notes (a previous, more concise edition with an older orthography was O. Korhonen 1979);¹

- North Saami: Nielsen 1979 [1932–62] (ca. 25,000 lexemes) with supplementary data from Sammallahti 1989 – Nielsen's dictionary contains vocabulary from three areally distinct dialects (Buolbmát/Polmak, Kárášjohka/Karasjok and Guovdageaidnu/Kautokeino), Sammallahti's general dictionary is partly based on Nielsen's work and partly other, more contemporary sources;
- Inari Saami: E. Itkonen 1986–1991 (ca. 17,000 lexemes) – a comprehensive dialect dictionary;
- Skolt Saami: Sammallahti & Mosnikoff 1991 (ca. 11,000 lexemes) – a general-use dictionary based on Suõ'nn'jel–Če'vetjau'rr (Sevettijärvi) dialect but somewhat complemented with data from other dialects, using Itkonen 1958;²
- Kildin Saami: Kuruč 1985 (online version SRS-Kuruč; ca. 7,500 lexemes) and Antonova 2014 (online version SRS-Antonova; ca. 5,000 entries, excluding inflected forms and phrases given as headwords) – the largest general dictionaries of the literary language. The two dictionaries together contain probably about 9,000 different lexemes.³

Electronic versions of the dictionaries or their headword lists have been used, where possible (the headword and meaning data of the Ume and North Saami sources, for instance, are available in their entirety in the Álgut database). The scarce Pite Saami lexical resources are complemented with the abundant lists of verbal derivatives in Ruong's monograph (1943). Ter Saami lacks a lexical

¹ Another modern Lule Saami dictionary with by and large the same scope measured in headword count has been compiled by Anders Kintel (2012). Korhonen's dictionary has been selected as material for the current study because it appears to contain somewhat more derivatives, while Kintel's vocabulary is slightly more inclined towards (neologistic) compounds.

² After the data collection of the present study, a more comprehensive Finnish–Skolt Saami dictionary was published, whose source code is publicly available (Sámediggi 2020; Giellatekno 2020). That could be used as material in future studies, although it appears that there are not very many additions concerning derivatives in comparison to Sammallahti & Mosnikoff 1991.

³ Kuruč's dictionary is better suited for the current study than its followers (Kert 1986; Antonova 2014) not only due to its extent and quality but also because verbal derivatives are abundantly represented (2/3 of the lexemes in the dictionary are verbs!) and because derivatives of the same stem are conveniently arranged under a single dictionary entry. Antonova's material was used as complementary data, however, since it contains a number of verbs not present in Kuruč, most importantly *-(š)kue'dte*-inchoatives, which Kuruč leaves out altogether. – Due to spelling variants and identical words listed with two different wordclass labels (e.g., noun/adjective or adjective/adverb), the number of actual unique lexemes in Antonova's dictionary is considerably lower than the number of entries would indicate.

resource suitable for a comprehensive morphological survey, but Ter data is included when a reference is made to the current author's previous account on the variation of *-(š)*koatē*-inchoatives in Skolt, Kildin and Ter Saami, which was based on the extensive dialect dictionary of T. I. Itkonen (1958, = KKLS). Lexically poorly documented Akkala Saami is left out of consideration in this survey. The word forms have been converted into the current standard orthographies where the sources use an older spelling (Schlachter 1958; Nielsen 1979 [1932–62]; E. Itkonen 1986–1991), and a Latinized spelling variant is provided for the Kildin words beside the Cyrillic spelling (see details in Section 2.7).

The data was collected as follows: In the first phase, a programmatic search was used to pick out all verbs whose gloss (translation) contains a usual word meaning 'begin' or 'start' in the explanatory language, or a definition as an inchoative verb, such as the

viation "inch." This initial data was then manually checked, and only such verbs were kept that actually have at least a connotation of 'beginning' and that can be interpreted as deverbal derivatives (i.e., they have a verbal correlate that can be seen as their derivational base). In the second phase, supplementary harvesting was carried out by skimming through the words with a suffix previously known to form inchoative verbs in the language in question. This was necessary because, in many cases, the wording of the glosses conveys an inchoative meaning using other means than a "begin" verb, cf. for example SaaN (*girdit* 'fly' → *girdilit* glossed as 'fly away, fly up' or (*oaddit* 'sleep' → *oadđ'dát* 'fall asleep' in Nielsen 1979 [1932–62]. Finally, tables of cognate derivatives from different languages were compiled for each derivative type, and some checking was done to ensure that apparent gaps in the data were not caused by omissions in data gathering. The data tables are published as an online appendix at Zenodo.org.⁴

1.4. Representativity of the data

As the analysis relies on dictionary data, it should be kept in mind that the results reflect lexical frequencies rather than text (usage) frequencies. There may also be some factors stemming from the varying backgrounds of the dictionaries that affect the representation of different parts of the vocabulary. Possible sources of errors or inaccuracy include, for example, the limited genres of source texts (overrepresentation of folkloristic monologue narratives) or idiolectal language use – in the case of a small number of informants, personal preferences and habits may render the material less

⁴ <https://doi.org/10.5281/zenodo.5748180>

representative of the whole language community (cf. Schlachter 1958, which is based on the language of a single informant). The choices made by the dictionary compiler may also affect the result, since dictionaries can have different purposes. Descriptive dialect dictionaries usually aim to cover the vocabulary as exhaustively as the material permits, whereas general dictionaries made for practical use normally involve selection as to which words are essential for the dictionary user (in the current study, this concerns mostly the sources for Kildin, Skolt and Lule Saami). This is also one of the reasons I have used Schlachter 1958 for Ume Saami, and the new general dictionary (Barruk 2018) only as a complementary source: although more representative in terms of its material base, the latter has clearly left out many derivatives and compounds in favor of including more basic vocabulary items, which makes it unsuitable for studying the structure of the lexicon (see Kuokkala 2020). At the same time, it is possible that the abundance of verbal derivatives in Kuruč's Kildin Saami dictionary partly stems from systematic constructing based on the collaborators' native speaker intuition rather than actual language usage.

Although the number of occurrences in the data tells something about the *productivity* of a certain derivative type/suffix, it is not clear without a dedicated field study, to what degree a certain suffix is productive, i.e., can be used to form new lexemes, in the contemporary language. Therefore, I will avoid the term “productivity”, preferring instead to talk only about frequencies in the descriptive section of the paper, and will refer to possible long-term changes in productivity of various suffixes only in the comparative conclusions.

2. Inchoative derivatives by language

2.1. South Saami

Instead of North Saami, which often serves as a starting point for treatments of Saami languages in general, it might be illuminating to begin with South Saami and explore the inchoative systems of Saami languages in geographical succession. It is also known that South Saami has preserved many archaic linguistic features, and this may to some extent be true of the current subject as well.

The dictionary of Bergsland & Magga (1993) contains the following types of inchoative verbs. After the derivational suffix, the number of derivatives found as well as one or more example words are given. In the examples, the base verb (or a derivational correlate) of the derivative is given in brackets.

Table 1. South Saami inchoative verb types.

-gâtedh	54	e.g., <i>vååjnegaâtedh</i> ‘begin to get visible’ [<i>vååjnedh</i> ‘be visible’]
-(a)htjedh	18	e.g., <i>saejriedahtjedh</i> ‘begin to hurt’ [<i>saejriedidh</i> ‘hurt, ache’]
-jahtjedh / -jihjedh	42	e.g., <i>laavloejahhtjedh</i> ‘begin to sing’ [<i>laavlodh</i> ‘sing’], <i>obrijahtjedh</i> [= <i>obrijidh</i>] ‘begin to rain’ [<i>abrodh</i> ~ <i>åbrodh</i> ‘rain’]
-jidh	42	e.g., <i>bualajidh</i> ‘begin to burn (tr.)’ [<i>bueledh</i> ‘burn’], <i>nyölkijid</i> ‘begin to get hungry’ [<i>nealkodh</i> ‘be hungry, starve’]
-edh (: -a-)	4	e.g., <i>vuarhkedh</i> ‘begin to rub antlers’ [<i>vuerhkedh</i> ~ <i>vuetkedh</i> ‘rub antlers’]
-edh (*-î-)	3	e.g., <i>röökedh</i> ~ <i>reekedh</i> ‘start barking’ [<i>råakodh</i> ‘bark’]
-lidh	12	e.g., <i>tjoekelidh</i> ‘set off on skis’ [<i>tjoekjedh</i> ‘ski’]
-didh	10	e.g., <i>ravkedidh</i> ‘begin to dawn’ [<i>ravkedh</i> ‘dawn’]
-htidh	8	e.g., <i>miejkiehtidh</i> ‘begin to melt; melt quickly’ [<i>meajkedh</i> ‘melt, thaw’]
-gidh	1	<i>diektjegidh</i> ‘begin to thaw (of snow)’ [<i>deatjkedh</i> ‘thaw’]
-lgidh	2	e.g., <i>såakelgidh</i> ~ <i>såakalgidh</i> ‘begin to thicken (of milk)’ [<i>såakodh</i> ‘become thick, viscous’]
-skidh	1	<i>jealaskidh</i> ‘begin to revive’ [<i>jieledh</i> ‘live’]
-nidh	1	<i>haalanidh</i> ‘fly away’ [<i>haeliehtidh</i> ‘fly’]
-stidh	9	e.g., <i>biegkestidh</i> ‘begin to blow a little, for a while’ [<i>biegkedh</i> ‘blow (of wind)’]
-sidh	2	e.g., <i>goerkesidh</i> ‘reach an understanding’ [<i>guarkedh</i> ‘understand’]
-(l/r)ostedh	2	e.g., <i>gaarhtalostedh</i> ‘begin to be happy in a place’ [<i>gaerhtedh</i> ‘be fond of’]
-(d)ovvedh	5	e.g., <i>asvedovvedh</i> ‘be seriously frightened’ [<i>asvedidh</i> ‘be in horror’]

The most common and most neutral inchoative suffixes in South Saami are *-gåtedh* and *-jahhtjedh* ~ *-jihhtjedh* (used with parisyllabic stems), *-(a)htjedh* (with imparisyllabic stems) as well as *-jidh* (with parisyllabic stems). The last one is a quite regular continuation of the Proto-Saami **-j(e)-* suffix in that the vocalism of the derivative depends on the stem type (original stem vowel) of the base verb (cf. Bergsland 1994: 91–92; Magga & Magga 2012: 131–133):

bueledh^I → *bualajidh* (PSaa **puolē-* → **puolāje-*) ‘begin to burn, be kindled’
nealkodh → *nyölkijidh* (PSaa **nealkō-* → **nealkuje-*) ‘begin to be hungry’
doeltedh^{IV} → *doeltijidh* (PSaa **tuolte-* → **tuolteje-*) ‘begin to boil’

Occasionally, though, such derivatives can also be found where the suffix is simply attached to the base stem without applying the usual vowel alternations, e.g., *vaedtsiejidh* (normally *vaadtsajidh*) ‘walk away’ ← *vaedtsedh* ‘go, walk’ (Hasselbrink 1981–1985: 1390). Apparently such irregular forms have purposely been left out of Bergsland & Magga (1993).

The *-edh : -a-* type, which is historically a variant of the *-(a)jidh* type, contains a few verbs with inchoative semantics that do not have a morphologically clear base verb, but only complex correlates, such as *sjagkedh* ‘begin to dry’ ↔ *sjegkiehtidh* ‘dry off (quickly)’. The only example with a clear formal base verb is *vuarhkedh* ‘begin to rub antlers’ (← *vuerhkedh* ~ *vuetkedh* ‘rub antlers (of reindeer)’). To some extent, the same applies to verbs in *-didh*, cf. for example *skoejedidh* ‘begin to understand (of infants)’ (↔ *skuajasjidh* ‘understand (quickly)’).

The other suffixes listed above are employed less frequently in an inchoative function, and the derivatives tend to also bear other aspectual connotations, which are more typical of the said suffixes. Such suffixes include *-htidh* (subitive/inchoative), e.g., *miejkiehtidh* ‘begin to melt, melt quickly’ (← *meajkedh* ‘melt, thaw (intr.)’), *-stidh* (diminutive/inchoative), e.g., *biegkestidh* ‘begin to blow a little, for a while’ (← *biegkedh* ‘blow (of wind)’), *-lidh* (usually momentative/subitive/frequentative), e.g., *vaedtsielidh* ‘set off, go away’ (← *vaedtsedh* ‘go, walk’), *-lgidh* (usually passive), e.g., *spuallalgidh* ‘begin to get faded (of shoe-hay)’ (← *spualledh* ‘become faded’), and *-sidh* (usually automative), e.g., *valkesidh* ‘begin to head in one direction; gradually gather speed (of reindeer)’ (← *valkedh* ‘head off (from a gathering place; of a reindeer herd)’). In addition, *-ovvedh* (in southern dialects *-oevedh*; usually automative or translative) can in some cases be analyzed as inchoative, e.g., *nahkerdovvedh* ‘become sleepy’ (← *nahkerdidh* ~ *nihkierdidh* ‘be sleepy’, ↔ *nahkere* ‘drowsiness; dream’). Of these atypical suffixes, only *-stidh* is mentioned as inchoative in Magga & Magga (2012: 134). In his

account on verbal derivation, Hasselbrink (1981–1985: 177–186) mentions the inchoative use for other suffixes presented above except for *-htidh*, *-sidh*, *-stidh* and *-ovvedh*. The suffix *-(a)htjedh* (*-ttj'edh* in Hasselbrink's spelling) is also mentioned only as a diminutive suffix (ibid. 185).

Some other suffixes occurring with verbs denoting 'to begin something' seem to actually form denominal (translative) derivatives, and only very occasionally deverbal ones (mostly a single occurrence in the data). Such suffixes include **-gidh** (*dïektjegidh* 'begin to thaw (of snow)' ↔ *deaktjadidh* 'id.', *deatjkedh* 'thaw (of snow)'), **-nidh** (*haalanidh* 'fly away' ↔ *haeliehtidh* 'fly') and **-skidh** (*jealaskidh* 'begin to revive, to get well' ← *jieledh* 'live'). Here one can also mention the suffix cluster **-(l/r)ostedh** (< **-ęstuj-*), which forms translative or automative-passive verbs (cf. Magga & Magga 2012: 148, 157), but whose isolated derivatives may have inchoative-like meanings as well, such as *gaarhtalostedh* 'begin to be happy in a place' (cf. *gaerhtedh* 'be fond of, attached to').

2.2. Ume Saami

The following inchoative verb types can be found in Schlachter's dictionary of the Malå dialect (1958). In addition to the six *-gåhtiet* verbs in the actual dictionary part, seven more are found in the text samples of the same work. Other derivative types of the texts seem to be completely included in the dictionary part. The numbers in square brackets include additional items found in Barruk (2018).⁵

Table 2. Ume Saami inchoative verb types.

<i>-gåhtiet</i>	13 (6+7)	e.g., <i>guarragåhtiet</i> 'begin to follow' [<i>guarrat</i> 'climb; follow']
<i>-åhtjat</i>	14	e.g., <i>dåjjadåhtjat</i> 'begin to understand' [<i>dåjjadit</i> 'understand']
<i>-åjit</i>	11	e.g., <i>varddåjit</i> 'begin to bleed' [<i>varddiet</i> 'bleed']
<i>-ijit</i> (*-uj-)	2	e.g., <i>vyössijit</i> 'revive' [<i>viässuot</i> 'live']
<i>-ijit</i> (*-ej-)	5	e.g., <i>guvllijit</i> 'begin to hear' [<i>guvllat</i> 'hear']
<i>-at</i> (*-û-)	1 [2]	e.g., <i>gulddat</i> 'begin to snow' [<i>gualdduot</i> 'be driving snow']
<i>-lit</i>	5	e.g., <i>rahkkalit</i> inchoative of <i>rahkkat</i> 'be in the rut'

⁵ The word forms presented here have been adapted to the current orthography as used in Barruk (2018).

-rit	[1]	<i>sjállarit</i> ‘begin to twitter’ [<i>sjállat</i> ‘twitter; be glad’]
-dit	3	e.g., <i>tjühkkedit</i> ‘sit down’ [<i>tjåhkkáhit</i> ‘sit’]
-sit	[1]	<i>valggasit</i> ‘begin to head in one direction’ [cf. SaaS <i>valkedh</i> ‘head off (from a gathering place)’]
-stit	2	e.g., <i>läjdestit</i> inchoative of <i>läjððet</i> ‘lead’

A couple of derivatives glossed with a meaning ‘allmählich = gradually, little by little’ have also been counted as inchoatives here, e.g., *alddánáhtjat* ‘come gradually closer’ (← *alddánit* ‘come closer’). Those with no apparent aspect difference from the base verb (according to the gloss) have not been included, e.g., *varrasmáhtjat* ‘recover [sich erholen]’ (← *varrasmit* ‘get well [gesund werden]’) or *vuajjájit* ‘drive’ (← *vuöjjiet* id.) (cf. SaaN *vuodjit* ‘drive’, *vuoddját* ‘drive off’). Another example of a contracted *-û- : -uje- stem inchoative in addition to *gulddat* ‘start snowing’ (← *gualdduot* ‘be driving snow’) can be found in Barruk (2018): *tjyörrat* ‘start crying’ (← *tjiärruot* ‘cry’). A peculiar case is the inchoative of *tjuajggat* ‘ski’, which in Schlachter (1958 s.v. *tjuaigat*¹, *tjuaigat*²) has the infinitive and present 1SG forms identical to the base verb, but the present 3SG form (*tjuǞǞ-gǞǞ*) is indicative of an *-ǵj-stem. The derivatives in -lit preserve the second syllable vowel of the base verb, with the exception of *tjuödtjielit* ‘stand up’ (← *tjuadttjuot* ‘stand’).

2.3. Pite Saami

The combined data from Wilbur (2016; BB), Halász (1896) and the verbal derivative lists of Ruong (1943) contain inchoative verbs of the following types. The examples from Ruong have been converted into the current orthography, with extra-long geminate marks (‘) added to emphasize the changes in the stem.

Table 3. Pite Saami inchoative verb types.

-gáhtet	73	e.g., <i>låhkågáhtet</i> ‘begin to read’ [<i>låhkåt</i> ‘read’]
-áhtjat (/-atjgáhtet)	55	e.g., <i>vuasstaláhtjat</i> ‘begin to resist, oppose’ [<i>vuasstatlit</i> ‘resist, oppose’]
-á(ji)t	27	e.g., <i>bal’lát</i> ‘become scared’ [<i>ballat</i> ‘be scared’]
-ut	9	e.g., <i>tjer’rut</i> ‘begin to cry’ [<i>tjierrot</i> ‘cry’]
-álit	3	e.g., <i>ruag’gálit</i> ‘get hungry’ [<i>ruag’gat</i> ‘be hungry’]

<i>-(a)lit</i>	27	e.g., <i>vuorralit</i> ‘become crazy’ [<i>vuorrat</i> ‘be crazy’]
<i>-(l)dahttet</i>	11	e.g., <i>sugaldahttet</i> ‘quickly row off shore following smb.’ [<i>suhkat</i> ‘row’]
<i>-árit</i>	5	e.g., <i>stuvvsárit</i> ‘begin to play’ [<i>stuvvsit</i> ‘play (with one another)’]
<i>-dit</i>	7	e.g., <i>sjvahtjadit</i> ‘begin moving, move a little bit’ [<i>sjvahtjat</i> ‘move’]
<i>-ijdit</i>	5	e.g., <i>istijdit</i> ‘get enough time’ [<i>asstat</i> ‘have time’]
<i>-ánit</i>	2	e.g., <i>hállánit</i> ‘fly away’ [<i>háledit</i> ‘fly’]
<i>-(a)stit</i>	4	e.g., <i>vádtasstit</i> ‘(quickly rise and) get going’ [<i>vádtset</i> ‘go, walk’]

In the light of the recorded materials, *-gáhtet* is the most frequent inchoative suffix used with parasyllabic verb stems in Pite Saami. The same applies to *-áhhtjat* with imparisyllabic stems; according to Ruong (1943: 134), the suffix is “strongly productive” and his examples should serve only as a selection, and alongside *-áhhtjat*, the combined suffix *-atjgáhtet* is used just as often.⁶

Contracted verbs in *-á(ji)t* and *-ut* form inchoatives from base verbs with *a/e-* and *o-*stems, respectively (*buahttsát* ‘become sick’ ← *buohtsat* ‘be sick’; *dur’rut* ‘start to argue/fight’ ← *dårrot* ‘fight’). The *-át* verbs appear to be exclusively inchoative, whereas the *-ut* type also has a wide use in passive-like derivatives (cf. Ruong 1943: 102–103, 113–117).⁷

In contrast to North Saami, for instance, the inchoatives in *-(a)lit* do not seem to be especially restricted in terms of the semantics of the base verbs (e.g., *stiehpalit* ‘start raining (locally)’ ← *stiehpat* ‘rain’; *luodjalit* ‘begin to complain’ ← *luodjot* ‘complain’). According to Ruong (1943: 230), they have a more momentative or subitive connotation than the inchoatives in *-gáhtet*. The base verb is usually a bisyllabic *a/å-*stem, only occasionally an *e-* or *o-*

⁶ In another section, Ruong (1943: 219) also states that “the suffix *-gøhtet* can always, without a change in the meaning, be replaced with the double suffix *-ãžgøhtet*” (“Das Suffix *-gøhtet* lässt sich immer, ohne Veränderung der Bedeutung, durch das Doppelsuffix *-ãžgøhtet* ersetzen”). This must have been somehow misworded since there is no indication of any actual use of the suffix *-atjgáhtet* with parasyllabic stems in Ruong’s own or other Pite Saami language materials. It is not even clear how the suffix should be attached to such stems.

⁷ In the printed dictionary of Wilbur (2016), inchoative verbs in *-á-* and *-u-* seem to have been largely left out. Interestingly, though, one can note that e.g., *ballagáhtet* ‘begin to become afraid’ is included, while *ballát* ‘become scared’ is not. Actually, looking at this verb, it appears that in the dictionary editing phase the inchoative aspect has been erroneously transferred to the base verb *ballat* (‘become scared’ in Wilbur 2016), as in other sources its meaning is the stative ‘be scared’ (BB; Ruong 1943: 102; Halász 1896: 91).

stem, in which case the stem vowel is replaced by *-a-*. In a couple of cases, an inchoative derivative has been formed with the suffix *-álit* (e.g., *vávvalit* ‘begin talking dirty’ ↔ *vávaldit* ‘talk dirty’).

Ruong (1943: 147) classifies the derivatives in *-ldahttet* as “mutativ-inchoativ-subitiv”. Since this suffix cluster also adds a causative function to the meaning of the base verb (e.g., *duoldaldahttet* ‘quickly bring to a boil’ ← *dualldat* ‘boil (intr.)’), it cannot be considered purely inchoative, perhaps with the exception of the single derivative *mujjtáldahttet* ‘suddenly recollect’ ← *mujjtet* ‘remember’ (cf. North Saami *-ldahttit*).

The suffix *-ijdit*, when used to derive verbs denoting a change of state, is, according to Ruong (1943: 227–228), primarily denominal and expresses a gradual change. Still, some of these derivatives are better analyzed as deverbal inchoatives even if they do have a nominal correlate as well, for example *mujtijdit* ‘recall, start remembering better’ (← *mujjtet* ‘remember’ ↔ *mujjto* ‘memory’). Interestingly, Wilbur’s material (BB) also contains a few clearly inchoative derivatives with the suffix *-dit* (e.g., *skellbidit* ‘begin to shake’ ← *skellbit* ‘shake’) – these seem to be generally unknown to Ruong (1943: 188–208), although the inchoative posture verbs *tjuhkidit* ‘sit down’ (← *tjáhkkahit* ‘sit’) and *vel’lidit* ‘lie down’ (← *väl’lahit* ‘lie’) are presented in comparison to *-hit* verbs (ibid. 138–139).

The rarer inchoative types *-árit* and *-ánit* mainly consist of correlative derivatives, such as *bánntsárit* ‘begin to act fiercely, to struggle’ (↔ *bánntsurit* ‘struggle’) and *lámngánit* ‘start chattering’ (↔ *lámngátit* ‘chatter, ramble on’). A few examples of inchoatives in *-(a)stit* can also be found.

2.3. Lule Saami

Inchoative verb types found in the dictionary of O. Korhonen (2007) can be summarized as follows:

Table 4. Lule Saami inchoative verb types.

<i>-goahtet</i>	39	e.g., <i>lávlllogoahtet</i> ‘start singing’ [<i>lávllot</i> ‘sing’]
<i>-sjoahtet</i>	2	e.g., <i>rássjodasjoahtet</i> ‘start raining’ [<i>rássjodit</i> ‘rain’]
<i>-(a)htjat</i>	21	e.g., <i>liegganahtjat</i> ‘begin to get warm(er)’ [<i>liegganit</i> ‘get warm(er)’]
<i>-át</i>	45	e.g., <i>lávllát</i> ‘start singing’ [<i>lávllot</i> ‘sing’]
<i>-ut</i>	8	e.g., <i>guohttut</i> ‘begin to graze’ [<i>guohttot</i> ‘graze’]
<i>-lit</i>	20	e.g., <i>lahpalit</i> ‘begin to move’ [<i>lahpat</i> ‘move’]

-(<i>á</i>)rit	10	e.g., <i>giehtttárit</i> ‘begin to tell’ [<i>giehttot</i> ‘tell’, <i>giehtto</i> ‘story; occasion’]
-dit	6	e.g., <i>vádtásádit</i> ‘start off (by foot)’ [<i>vádtset</i> ‘go, walk’, <i>vádtását</i> ‘start off’]
-ijdit	14	e.g., <i>gálgijdit</i> ‘begin to flow’ [<i>gálggát</i> ‘flow’]
-git	2	e.g., <i>muovagit</i> ‘begin to smell sour/musty’ [<i>muovvat</i> ‘turn a bit sour’, <i>muovas</i> ‘sour, musty (of a smell)’]
-nit	2	e.g., <i>hál’lánit</i> ‘begin to slant’ [<i>hállit</i> ‘slant, tilt’]
-stit	1	<i>biekkastit</i> ‘begin to blow; rise (of wind)’ [<i>bieggat</i> ‘blow (of wind)’]

According to this data, *-goahtet* and *-át* are about equally frequent inchoative suffixes used with parisyllabic verb stems. There do not seem to be any semantic or morphosyntactic restrictions or tendencies as to which base verbs can take these suffixes. The suffix *-át* attaches mostly to *e-* or *a-*stems, replacing the stem vowel (e.g., *muohtttát* ‘begin to snow’ ← *muohttet* ‘to snow’, *gul’lát* ‘begin to hear, get to hear’ ← *gullat* ‘hear’), and sometimes also to *o-*stems (*lávllát* ‘begin to sing’ ← *lávllot* ‘sing’). A corresponding neutral inchoative suffix used with imparisyllabic base stems is *-(a)htjat* (e.g., *báhtarahjtat* ‘begin to escape’ ← *báhtarit* ‘escape’); a rarer type is *-(a)sjgoahtet* (*dálkkuhasjgoahtet* ‘begin to be bad weather’ ← *dálkkuhit* ‘be bad weather’), also attested as *-(i)sjgoahtet* (Kuokkala 2019: 161). The suffix *-ut* occurs with *o-*stem base verbs. One more type specialized for inchoative verbs is *-(á)rit*. The other suffixes listed are rarer in inchoative function and usually have other aspectual meanings.

In Kintel’s description of Lule Saami verbal derivation (Kintel 1991: 30–49), the general picture of inchoative-building suffixes is more or less the same. Notably, though, Kintel does not include the suffix *-lit*, which often occurs in inchoative verbs denoting motion, and in some cases has subitive or frequentative connotations. On the other hand, Kintel mentions the suffixes *-nit* and *-stit*, which occur only once in our inchoative data, but does not mention *-dit* or *-git*. Kintel presents one inchoative suffix not found in Korhonen’s dictionary, namely *-ahtjtát* (*basádahtjtát* ‘begin to wash oneself’ ← *basádit* ‘wash oneself’; Kintel 1991: 49), which appears to be a combination of two frequent inchoative suffixes, *-(a)htjat* and *-át*.

2.4. North Saami

The following types of inchoative verbs are found in Nielsen's dictionary (1979 [1932–62]). The figures include only a few additions from Sammallahti's (1989) data in the most frequent types. Meaning information from both dictionaries has been used.

Table 5. North Saami inchoative verb types.

-goahtit	4	e.g., <i>jápmigoahtit</i> 'be nearly dying' [<i>jápmi</i> 'die']
-šgoahtit	4	e.g., <i>jurddašišgoahtit</i> 'begin to think, bethink oneself that' [<i>jurddašit</i>]
-át	39	e.g., <i>muohtát</i> 'begin to snow' [<i>muohtti</i> 'snow']
-ot	2	e.g., <i>čir'rot</i> 'begin to cry' [<i>čierut</i> 'cry']
-et	4	e.g., <i>duldet</i> 'begin to boil, boil up' [<i>duoldat</i> 'boil']
-lit	40	e.g., <i>gizzulit</i> 'start, set off, climbing' [<i>gizzut</i> 'climb']
(-ldahttit)	12	e.g., <i>vuojaldahttit</i> 'swim off after (to overtake); swim off with (taking) smb./smth.' [<i>vuodjat</i> 'swim']
(-lahttit)	5	e.g., <i>vuodjalahttit</i> '= vuojaldahttit; cause to swim off' [<i>vuodjat</i>]
-(á)dit	6	e.g., <i>šlahtádit</i> 'set in to sleet, begin to sleet' [<i>šlahtti</i> 'sleet']
-iidit	32	e.g., <i>dovddiidit</i> 'begin to know, feel, etc., (really) get to know' [<i>dovdat</i>]
-ihit	4	e.g., <i>cielihit</i> 'suddenly begin to bark' [<i>ciellat</i> 'bark']
-git	15	e.g., <i>ceakkagit</i> 'begin to stand' [<i>ceaggát</i> 'stand upright', etc.]
-áskit	15–23	e.g., <i>jodáskit</i> 'start moving, migrating' [<i>johtit</i>]
-sit	7	e.g., <i>modjosit</i> 'begin to smile' [<i>moddját</i> 'be smiling', <i>modji</i> 'smile']
-(a)nit	4	e.g., <i>cahkkanit</i> 'kindle, begin to burn' [<i>cahkat</i> 'smoulder']
-ánit	24	e.g., <i>vielgánit</i> 'begin to move, set off (of smth. white)' [<i>vielgut</i> 'appear white (while moving)', <i>vielgat</i> 'white']
-stit	13	e.g., <i>jielastit</i> 'begin to blow gently' [<i>jiellat</i> ; <i>jiella</i> 'gentle breeze']

<i>-(á)stallat</i>	5	e.g., <i>vuoddjállat</i> ‘get ready to drive; drive off, one after another’ [<i>vuodjit</i>]
<i>-(i)duvvat</i>	2+?	e.g., <i>ásaiduvvat</i> ‘begin to feel at home’ [<i>ássat</i> ‘dwell; settle (down)’]
<i>-stuvvat</i>	35	e.g., <i>njamastuvvat</i> ‘(begin to) want to suck’ [<i>njammat</i> ‘suck’]

The North Saami data yields as many as twenty distinct derivative types with some sort of deverbal inchoative functions, though it must be noted that a number of these represent historically identical suffixes, and not all can be seen as strictly inchoative suffixes according to our definition in the introductory section. Specifically, the derivatives in *-ldahttit* and *-lahttit* are not simple inchoatives with respect to the base verb, but rather transitive modifications that have a causative function ‘to make someone start doing something’ or, especially in the case of the *-ldahttit* suffix, ‘to set off carrying something/someone’ or ‘to set off to overtake someone/something’. Most of these have a *-lit* inchoative as their possible immediate base verb (e.g., *vuojaldahttit* ↔ *vuodjalit* ‘swim off’), so they could also be analyzed as transitive derivatives in *-dahttit/-httit*. The derivatives in *-(á)stallat*, which tend to have a meaning ‘to prepare to do something’, are even more clearly frequentative re-derivations of *-át*-inchoatives (cf. *vuoddjállat* ↔ *vuoddját* ‘drive off’) rather than an independent inchoative type.

Some of the derivative types listed are not unambiguously deverbal, as their members regularly belong to a word family containing both a simple verb and a simple noun analyzable as a derivational base, such as *-áskit*: *soadáskit* ‘go to war, get into a fight’ ↔ *soahti* ‘war, struggle’ ↔ *soahtat* ‘be at war, fight’. Further such types are *-git*, *-nit* and *-sit*, which connect to word families usually consisting of verbs, nouns, adjectives and adverbs denoting some kind of state, posture, etc., where it is often difficult to decide which of the correlates would be the most primary from a derivational standpoint. For example, the verbs *ceakkagit* ‘begin to stand’ and *ceakkanit* ‘rise upright; begin to bristle’ can be correlated with the adjective *ceakko(-)* ‘upright; steep’, the adverb *ceaggut* ‘in an upright position; steep’, the noun *ceagga* ‘obstinacy’ as well as the verbs *ceaggát* ‘stand upright’, *ceaggut* ‘raise up, set up (several things)’, etc. The *-ánit* type especially contains verbs meaning ‘to start off in a certain state, appearing such-and-such’, which connect to families with several wordclasses, but can on formal grounds most closely be associated with the corresponding verbs in *-át*, such as *vielgánit* ‘set off (of something white)’ ↔ *vielgát* ‘look white or light’ (↔ *vielgat* ‘white’) or *barfánit* ‘set off

(being furry, bushy)' ↔ *barfát* 'lie/stand furry, bushy somewhere' (↔ *barffas* 'furry, bushy' ↔ *barfi* 'a furry, bushy one').

Several of the listed derivative types have primarily momentative and/or subitive semantics, which is more or less present also in the derivatives with an inchoative meaning. These types include **-(i)hit** (e.g., *golggihit* 'suddenly run (out) or float away' ← *golgat* 'run (of liquid), float (with the current), leak, etc.'), **-stít** (e.g., *oakkøstít* 'fish/troll a little; begin to fish/troll' ← *oaggut* 'fish with a rod, troll'), **-ádit** (e.g., *biškkádit* 'suddenly start screaming' ← *biškut* 'scream, crow, shriek') and **-lit** (e.g., *suhkalit* 'row a little; row away; go for a short row' ← *suhkat* 'row'); the last one is used especially with motion verbs. The verbs in **-adit** are rather frequentative-conative than inchoative in terms of their semantics (cf. *dámádit* 'be about to, try to tame, train; keep taming, training' ← *dápmat* 'tame, train'), and have thus been left out of the listing.

The verbs in **-uvvat** constitute several subtypes, of which almost all consist of denominal derivatives (cf. e.g., *dálūiduvvat* 'get a farm or house; begin to stay at home' ← *dállu* 'farm, house'; *čakčaluvvat* 'begin to be autumnal; be surprised by the autumn' ← *čakča* 'autumn'). There are, nevertheless, a couple of deverbal inchoatives in **-(i)duvvat** (e.g., *ásaiduvvat* 'begin to feel at home' ← *ássat* 'dwell; settle (down)') and the type **-stuvvat**, which operates widely on verb bases (e.g., *barggastuvvat* 'desire, want to work; begin working'). These derivatives often have the meanings of wanting or being close to doing/happening in addition to or rather than inchoativity. The corresponding suffix to **-uvvat** attaching to bisyllabic roots, **-ut**, seems to form only denominal translatives (e.g., *seadđut* 'begin to get a little dark, get dim' ← *seadđi* 'dimness, obscurity' ↔ *seadās* 'dim, unclear'). Another type that has been left out of the table is **-(l/s)mit**, which often does have verbal correlates too, but whose derivatives are rather to be analyzed as denominal translatives, as the intermediate consonant also indicates, cf. for instance *jođálmit* 'become quicker' ← *johtil* 'quick, rapid, which runs easily' ↔ *johtit* 'move, go, travel'.

In addition to **-(š)goahtit** (see below), the most clearly deverbal inchoative suffix in North Saami is **-át**, which in addition to the historically connected **-i**-verbs (cf. *vardát* 'begin to bleed' ← *vardit* 'bleed') currently attaches to other bisyllabic stem types too (*gul'lát* 'begin to hear' ← *gullat* 'hear', *goal'lát* ← *goallut* 'be/feel cold'). Other descendants of the historical ***-Vj-** suffix show inchoative semantics only sporadically: for instance **-ot** in the verb *čir'rot* 'begin to cry' ← *čierrut* 'cry' and **-et** in *lávlet* '(suddenly) begin singing, sing a little' (← *lávlut* 'sing'); the derivatives in **-et** tend to have momentative or subitive connotations as well. The suffix **-iidit** is somewhat ambivalent in that a number of derivatives (15) can or must be analyzed as

simple verb-based derivatives (e.g., *golggiidit* ‘begin to run (of liquid), to float’ ← *golgat* ‘run (of liquid), float (with the current), leak, etc.’), whereas others (19) have only nominal or complex verbal correlates (e.g., *dilddiidit* ‘begin to quake, tremble’ ← *dilddas* ‘quaking, trembling ↔ *dilddistit* ‘quake, shake, tremble’).

The appearance of **-(š)goahtit** verbs in Nielsen’s dictionary is evidently not representative of their actual use: only a few derivatives with a more lexicalized meaning are included as headwords, and otherwise only the suffix *-goahtit* (s.v. *-goattet*) is given to explain the highly productive derivative type (cf. Nielsen 1979 [1926]: 247–248). According to Nielsen (ibid.) and Nickel (1994: 262–263), the inchoative suffix *-(š)goahti-* can attach to (nearly) all verbs regardless of semantics or valency, and the allomorph distribution is based on the syllable count of the base stem: the variant *-goahti-* attaches to parisyllabic and contracted stems, while *-šgoahti-* is used with imparisyllabic stems (in eastern dialects, also with contracted stems). For example:

<i>lohka-t</i> → ‘read’	<i>lohka-goahti-t</i> ‘begin to read’	(bisyllabic)
<i>liiko-t</i> → ‘like’	<i>liiko-(š)goahti-t</i> ‘begin to like’	(contracted)
<i>muitali-t</i> → ‘tell’	<i>muitali-šgoahti-t</i> ‘begin to tell’	(trisyllabic)

As we have seen in the previous sections, a similar distribution is typical of other western Saami languages as well: the cognates of the *-goahti-* suffix are used only with parisyllabic (= bisyllabic and quadrisyllabic) base verbs, and imparisyllabic verbs require a different suffix (e.g., *-ahtja-*, *-áhtja-*) or an “extended” variant of *-goahti-* resembling North Saami *-(i)šgoahti-* (Pite *-atjgáhte-*, Lule *-asjgoahte-* ~ *-isjgoahte-*). I have previously argued that these “extended” variants of *-goahti-* historically represent a combined suffix containing the *-ahtja-* suffix as their initial part (as already suggested by Ruong 1943; see Kuokkala 2019 and the concluding section of this article).

2.5. Inari Saami

The inchoative verbs in the Inari Saami dictionary of E. Itkonen (1986–1991) fall into the following derivative types:

Table 6. Inari Saami inchoative verb types.

-gyettiđ / -guáttiđ	42	e.g., <i>tuáru gyettiđ</i> ‘begin to fight’ [<i>tuárruđ</i> ‘fight’]
-škyettiđ / -škuáttiđ	212	bisyll. 98, e.g., <i>vuojáškyettiđ</i> ‘begin to swim’ [<i>vuojád</i> ‘swim’] trisyll. 52, e.g., <i>määlistškyettiđ</i> ‘begin to cook’ [<i>määlistiđ</i> ‘cook’] quadrisyll. 27, e.g., <i>njurgiištâlškyettiđ</i> ‘begin to whistle’ [<i>njurgiištállâđ</i> ‘whistle’] contracted (-j-) 22, e.g., <i>kulluuškuáttiđ</i> ‘begin to be audible’ [<i>kulluđ</i> ‘be audible’] contracted (-v-) 12, e.g., <i>šišneluuškyettiđ</i> ‘begin to get hairless’ [<i>šišneluđ</i> ‘become hairless’]
-áđ (-ađ)	13	e.g., <i>páštađ</i> ‘begin to shine’ [<i>páštiđ</i> ‘shine (of the sun, moon)’]
-iđ (: -ej-)	5	e.g., <i>kulgiđ</i> ‘begin to flow’ [<i>kulgâđ</i> ‘flow’]
-liđ	35	e.g., <i>vuoijâliđ</i> ‘go swimming, start swimming’ [<i>vuoijâđ</i> ‘swim’]
-(â)âhčâđ	2	e.g., <i>rohâlahčâđ</i> ‘come into the rut’ [<i>rohâliđ</i> ‘be in the rut’]
-âdiđ	4	e.g., <i>nohâdiđ</i> ‘fall into sleep’ [<i>nukkâđ</i> ‘doze’]
-ijdiđ	5	e.g., <i>ruoškiđiđ</i> ‘begin to crackle’ [<i>ryeškiđ/ruoškâđ</i> ‘crackle’]
-(â)gidiđ	18	e.g., <i>njuigâgidiđ</i> ‘go into a sulk’ [<i>njuigâđ</i> ‘sulk’]
-âskiđ	8	e.g., <i>tuárâskiđ</i> ‘begin to fight’ [<i>tuárruđ</i> ‘fight (v.)’, <i>tuáru</i> ‘fight (n.)’]
-niđ	3	e.g., <i>čokâniđ</i> ‘sit down’ [<i>čokkâđ</i> ‘sit’]
-stiđ	4	e.g., <i>ruotâstiđ</i> ‘break into a gallop’ [<i>ruottâđ</i> ‘run, gallop’]
-(ii)stállâđ	3	e.g., <i>tuoldiistállâđ</i> ‘begin to boil any moment’ [<i>tuoldâđ</i> ‘boil, bubble’]
-studđ	4	e.g., <i>kyedistudđ</i> ‘begin to calve’ [<i>kyeddiđ</i> ‘calve’]

According to this data, by far the most frequent Inari Saami inchoative suffix is *-škyettiđ* (with the phonological variant *-škuáttiđ*)⁸, which attaches to virtually all kinds of verbal stems, even the monosyllabic *leđe* ‘be’ → *leškyettiđ* ‘begin to be’, which is unique to Inari Saami and some of the easternmost subdialects of North Saami (Nielsen 1979 [1932–62] s.v. *læšgoattet* records the derivative from Porsanger). The shorter suffix *-gyettiđ* (~ *-guáttiđ*) appears in a couple of pure inchoative derivatives, such as *tuárugyettiđ* ‘begin to fight’ (← *tuárruđ* ‘fight’), but it is mainly used in expressive verbs, which usually do not have a clear base verb, and have subitive or momentative as often as inchoative functions (e.g., *kommágyettiđ* ‘to boom (once)’, cf. *komijdiđ* ‘to boom, rumble’; see Kuokkala 2019: 160–161).

The inchoatives in *-áđl/-ađ* (contracted *-áj*-stems) comprise a dozen derivatives with more or less lexicalized meanings and mostly corresponding derivatives in other Saami languages, e.g., *muštáđ* ‘recall, start to remember’ (← *mušteđ* ‘remember’), *palláđ* ‘be frightened, get a fright’ (← *poolláđ* ‘be afraid’). The few inchoatives in *-iđ* (contracted *-ej*-stems) make a similar impression, e.g., *tuoldiđ* ‘begin to boil’ ← *tuoldáđ* ‘boil, bubble’ (cf. SaaS *doeltijidh* id.).

A couple of derivatives reminiscent of the western Saami inchoative type *-(a/á)htjat* can be found with the suffix *-(ál)áhčáđ*; these, however, correlate to *-áliđ* verbs, which already have more or less inchoative meaning (*rohâláhčáđ* ‘begin to come into the rut’ ← *rohâliđ* ‘come into the rut’). Usually the *-(ál)áhčáđ* derivatives are frequentative(-continuative) in meaning, although they may also have connotations of intention or being about to do something, as in *vuájáláhčáđ* ~ *vuájáluhčáđ* ‘be about to drive off, intend to drive off right away’.

The suffix *-liđ* is quite frequent in verbs denoting the start of a motion, e.g., *suhâliđ* ‘start rowing’ ← *suuhâđ* ‘row’ (cf. North Saami *suhkalit* id. and others), but there are also inchoatives with other kinds of meanings, e.g., *párguliđ* ‘start bawling’ (← *párguđ* ‘shout, bawl’) and *kuđđâliđ* ‘begin to spawn’ (← *kuuđđâđ* ‘spawn’). Inchoatives in *-áđiđ* are rare, and despite the lack of an obvious base verb, I have included here the verb *suláđiđ* ‘start to row, row off’, which has a cognate in Skolt (*solleed*) and Kildin Saami (*sullde*, KKLS *sulŋde*^δ, id.) and might be an irregularly shortened derivative of the previously mentioned *suhâliđ* or the base verb *suuhâđ* (unless it is

⁸ The variant *-(š)kuáttiđ* is the regular cognate of North Saami *-(š)goahtit* (< Proto-Saami **-koatē-*); in the more frequent variant *-(š)kyettiđ*, the diphthong has gone over to the historical **uo* series.

etymologically connected with SaaI *sullâ*, N *sulla* ‘resemblance; direction’ and only secondarily associated with the verb ‘to row’).

The occasional inchoatives in **-ijdiđ** do not always have a formally clear base verb (e.g., *tuvhijdiđ* ‘begin to ache’ ← *tuvhestiđ* ‘ache (frozen hand or foot while warming again)’). The **-(â)giđ** derivatives interpreted here as inchoative most often denote getting into a position or state, while a verbal correlate in **-áđ** means being in the said position (e.g., *kuvrâgiđ* ‘stoop down’ ← *kuvráđ* ‘sit hunched up, squat’). The derivative type **-âskiđ** contains only a few undoubtedly deverbal inchoatives, e.g., *tájâskiđ* ‘go mad, become insane’ ← *táâijud* ‘ramble, rave, fool around’, while denominal translatives are more typical of it (e.g., *kodâskiđ* ‘go wild (in the forest)’ ← *kodde* ‘wild reindeer’).

The derivatives in **-stiđ** are usually diminutive and have inchoative meanings just occasionally, e.g., *kulnâstiđ* ‘wither a little; begin to wither’ (← *kulnâđ* ‘wilt, wither’). Some inchoative-like meanings occur also with **-(ii)stállâđ** derivatives, which otherwise tend to be frequentative or continuative, e.g., *tuoldiistállâđ* ‘begin to boil any moment’, which come semantically close to the conative occurrences of the same type, e.g., *hettiistállâđ* ‘(try to) prevent’ (← *hettiđ* ‘prevent, hold back’). Finally there are inchoative derivatives in **-stud**, such as *čičarustud* ‘feel like crying; begin to cry’ (← *čičärrud* ‘cry’), which also directly connect to a larger group of derivatives bearing a meaning of intention, wanting, or feeling like doing something, e.g., *njomâstud* ‘want to suck’ (← *njommâđ* ‘suck’).

Some inchoative verb types allow further derivations to be based on them; these include causatives based on **-lit** and **-stit** inchoatives, e.g., *ruotâstiđ* ‘break into a gallop’ → *ruotâstittiđ* ‘spur (a reindeer) into a gallop’ and *vuojjâliđ* ‘swim off’ → *vuojjâlittiđ* ‘swim off after someone’ (cf. the North Saami *vuodjalahttit*, *vuojaldahttit* types above). Some causative re-derivations of inchoative **-gyettiđ/-guättiđ** verbs can be found, e.g., *tuárugyettiđ* ‘begin to fight’ → *tuárugyedettiđ* ‘attack, pick a fight with someone’, but more often these are based on non-inchoative **-gyettiđ/-guättiđ** verbs, such as *skillâgyettiđ* ‘clink (once, intr.)’ → *skillâgyedettiđ* ‘clink, ping (tr.)’; also diminutive re-derivations can be found, e.g., *šuddâgyettiđ* ‘hiss (once/suddenly)’ → *šuddâgyedestiđ* ‘hiss a little’.

2.6. Skolt Saami

The following types of inchoative derivatives are found in the Skolt Saami dictionary of Sammallahti & Mosnikoff (1991).

Table 7. Skolt Saami inchoative verb types.

-gue'tted / -vue'tted	1	<i>rämmgue'tted</i> ~ <i>rämmvue'tted</i> ‘be delighted; rejoice’
-škue'tted	125	bisyll. 77, e.g., <i>looggškue'tted</i> ~ <i>logškue'tted</i> ‘begin to read’ [<i>lookkâd</i> ‘read’] trisyll. 33, e.g., <i>kaaupšeškue'tted</i> ‘begin to sell’ [<i>kaaupšed</i> ‘sell’] quadrisyll. 10, e.g., <i>vaalmštõlškue'tted</i> ‘begin to prepare’ [<i>vaalmštõllâd</i> ‘prepare’] contracted (-ee-) 5, e.g., <i>tuēj'jeeškue'tted</i> ‘begin to work’ [<i>tuēj'jeed</i> ‘work’]
-jed (*-āj-)	8	e.g., <i>vuäd'đjed</i> ‘go to sleep; fall asleep’ [<i>vue'dded</i> ‘sleep’]
-jed (*-ij-)	2	e.g., <i>šu'vved</i> ‘begin to rustle, hum’ [<i>šuuuvâd</i> ‘rustle, hum’]
-led	10	e.g., <i>ķe'rdled</i> ‘fly off’ [<i>ķe'rded</i> ‘fly’]
-lõõččâd	2	e.g., <i>vä'zzlõõččâd</i> ‘go walking’ [<i>vä'zzed</i> ‘walk’]
-eed	4(?)	e.g., <i>nååkkeed</i> ‘fall asleep’ [<i>nokkâd</i> ‘doze’]
-sted	3	e.g., <i>u(u)rčsted</i> ~ <i>u(u)'rčsted</i> ‘start running’ [<i>urččâd</i> ~ <i>u'rččâd</i> ‘run’]
-(st)õõvvâd	2/3	e.g., <i>poorstõõvvâd</i> ‘begin to feel like eating’ [<i>poorrâd</i> ‘eat’, <i>poorsted</i> ‘eat (hastily)’]

Judging by this data alone, the representative of the simple **-koatē*- inchoative suffix (-*gue'tted* ~ -*vue'tted*) would seem to be entirely fallen out of deverbal use, occurring only in one rather denominal verb *rämmgue'tted* ‘be delighted; rejoice’ (cf. *rämm* ‘joy’). In KKLS, however, a small number of such deverbal inchoatives can be found, for example *kaalgue'tted* ‘begin to follow (of an animal)’ ← *kaallâd* ‘follow, allow to be led’ (KKLS: 84; KoSaS only has *ka(a)lškue'tted* ‘id.’). The suffix variant -*škue'tted*, again, appears to be highly frequent in building inchoatives regardless of the stem type of the base verb. The use of the **-(š)koatē* suffixes in Skolt Saami and more eastern Saami varieties will be shortly revisited from a comparative perspective in Section 3.2.

The inchoative verbs in *-jed* fall into two groups: *e*-stem base verbs form derivatives with an unpalatalized root and stem vowel *u* (e.g., *kââ'cced* 'be awake' → *kâccjed* 'wake (intr.)'), while *â*-stem base verbs form derivatives with a palatalized root and stem vowel *e* (e.g., *tuõlddâd* 'boil (intr.)' → *tuõ'lddjed* 'boil up'); these correspond to earlier **-âj-* and **-ij-* (< **-ej-*) stems, respectively.⁹ A stem-wise exceptionally formed case is (*puõccâd* 'be ill' →) *puâc* 'cjed' 'get ill' (= North Saami *buohccát*).

Skolt Saami inchoatives in *-led* are formed exclusively of verbs denoting motion, and they contain the unaltered stem of the base verb similarly to *-led* subitives, some even have a subitive connotation (e.g., *suukkled* 'begin to row (quickly)' ← *suukkâd* 'row').¹⁰ Similarly to Inari Saami *-(â)lâhčâđ* verbs, Skolt Saami derivatives in *-lõõččâd* are mainly diminutive-frequentative in semantics, and the few inchoative ones can be seen as extended versions of corresponding inchoatives in *-led*: (*vue'jjed* ~ *vuejjad* 'drive' →) *vuej'jled*, *vuej'jlõõččâd* 'drive off'; (*vâ'zzed* 'walk' →) *vâ'zzled*, *vâ'zzlõõččâd* 'go walking'; according to KKLS (p. 706), the meaning of the last one is actually 'be about to go visiting' and thus expresses rather intention than the actual start of action.

The few inchoative-like verbs in *-eed* represent heterogenous derivational types – these include two distinct derivatives of *nokkâd* 'doze': *nââkkeed* 'fall asleep' and *no'kķeed* 'id.'. The unclear case *solleed* ~ *sõlleed* 'begin to row' has a cognate in Kildin and Inari Saami; see discussion of Saal *sulâdiđ* above. A couple of derivatives with the momentative/subitive suffix *-sted* show inchoative semantics as well, e.g., *vuâgsted* 'set off at a slow trot; trot for a while' (← *vuâkkad* 'jog, trot slowly'). There is also one example with the suffix cluster *-sâstted*: *tiârsâ'stted* 'gallop a little, start galloping' (← *tiârrâd* 'gallop, run'). These derivatives are mainly formed of verbs denoting movement by foot. Another rare inchoative type based on the *-st-* suffix is *-stõõvvat*, corresponding to the North Saami *-stuvvat* type: *vuâ'đstõõvvâd* 'begin to feel sleepy' (← *vue'đded* 'sleep'). Also the variant *-šõõvvâd* occurs in one verb, *põškkšõõvvâd* 'begin to feel the need to shit' (← *pâ'šķked* 'shit').

⁹ The unpalatalized *-jed* verb class is a merger of the Proto-Saami **-âj-* and **-uj-* verb classes, where the present 3sg stem vowel *a* is a regular reflex of both original classes (*kâccjed* : *kâccai* 'wake' < **kocâje-*; *kuâđ'đjed* : *kuâđ'đai* 'be left' < **kuođuje-*), whereas the imperative 2sg and connegative forms have generalized the stem vowel **û* of the **-uj-* class (*koccu* '(doesn't) wake' as *kuõđ'đu* 'is (not) left' < **kuođûk*) (cf. Sammallahti & Mosnikoff 1991: 173–175, 201).

¹⁰ Compare for instance the subitive of *mu'rded* 'fold': *mu'rdled* 'fold (quickly)' vs. the unpalatalized stem in the frequentative *murdled* 'keep folding'.

2.7. Kildin Saami

The inchoative verb types found in the combined material of Kuruč (1985 / SRS-Kuruč) and Antonova (2014 / SRS-Antonova) are summarized in the table below. To facilitate the presentation and comparison with data from other Saami languages, a Latin transliteration is used in the table.¹¹ In the text, the Cyrillic orthography forms are also given for reference.

Table 8. Kildin Saami inchoative verb types.

-gue'dte / -kue'dte	33 (-36)	e.g., <i>laavvlgue'dte</i> 'begin to sing' [<i>laavvle</i> 'sing']
-škue'dte	42	bisyll. 20, e.g., <i>tiideškue'dte</i> 'begin to know; recognize' [<i>tii'dte</i> 'know'] trisyll. 17, e.g., <i>vuellneškue'dte</i> 'begin to go down' [<i>vuellne</i> 'go down'] quadrisyll. 5, e.g., <i>lunntnenteškue'dte</i> 'begin to burst' [<i>lunntnente</i> 'burst']
-je (*-āj-)	9	e.g., <i>vuadtje</i> 'go to sleep; fall asleep' [<i>vue'dte</i> 'sleep']
-je (*-ij-)	7	e.g., <i>tuu'lltje</i> 'start to boil; boil up' [<i>tuullte</i> 'boil']
-je (*-uj-)	9	e.g., <i>oorrje</i> 'sit down; calm down' [<i>årre</i> 'sit; stay']
-je (anal.)	21	e.g., <i>laaffje</i> 'begin to rot' [<i>laaffe</i> 'rot']
-le	18	e.g., <i>ki'rrtle</i> 'fly off' [<i>ki'rrte</i> 'fly']
-la'hče	2	e.g., <i>šeenyja'hče</i> 'begin to storm' [<i>šeenyne</i> 'storm, be stormy weather']
-la'hte	1	<i>poorrkla'hte</i> 'start snowing heavily' [<i>poorrike</i> 'snow heavily', <i>poorrk</i> 'blizzard']

¹¹ The transliteration used here approaches a transcription in that some phonemic ambiguities of the orthography have been manually resolved (⟨ä⟩ > (j)ea / jaa; ⟨oa⟩/⟨ōā⟩ > ā / āā) and palatalization is indicated with a palatalization mark before the whole consonant center. It closely resembles the practice of Rießler (2022) except for the use of *i* instead of *ö* for the close central vowel ⟨ɨ⟩ and *gk*, *bp*, *dt* instead of *gg*, *dd*, *bb*. Note that *e*, *ee* ⟨ɛ, ɛ̃⟩ are primarily central vowels [ɛ ɛ̃] (cf. E. Itkonen 1971: 87–89).

-te (*-ijt-)	1	<i>ko'lte</i> ‘begin to flow, run’ [<i>koollke</i> ‘flow, run’]
-ne (*-ntē-)	4	e.g., <i>azne</i> ‘start to rise (water, at high tide, quickly)’ [<i>a'ze</i> ‘rise (of water)’]
-se (*-stē-)	1	<i>tii'dse</i> ‘find out’ [<i>tii'dte</i> ‘know’]
-selle	2	e.g., <i>piiŋŋkselle</i> ‘begin to blow (of wind)’ [<i>piiŋŋke</i> ‘blow (of wind)’]
-(l)uvve	24	e.g., <i>vuajjpluvve</i> ‘begin to wrestle’ [<i>vuajjpe</i> ‘wrestle’]

The *-(š)kue'dte* inchoatives, probably deemed to be a productive morphological category, seem to have been completely left out of Kuruč's material. Antonova's material, by contrast, generally includes quite few verbal derivatives, but *-(š)kue'dte* inchoatives are represented in considerable numbers, possibly as an attempt to compensate for their absence in Kuruč's dictionary. Due to this imbalance in the original data and the selections made in the editing processes, the frequencies of different derivative types in the combined data can only be taken as indicative.

Inchoatives in *-gue'dte/-kue'dte* appear to be rather frequent with (originally) bisyllabic base stems, e.g., *laavvlgue'dte* «лāввлгүэдтэ» ‘begin to sing’ (← *laavvle* «лāввлэ» ‘sing’). A few such derivatives do not have a formal bisyllabic base verb but correlate with noun stems and their verbal derivatives, such as *čuijjgue'dte* «чүййгүэдтэ» ‘start ringing’ (↔ *čuijj* «чүйй» ‘chime, ringing’, *čuijne* «чүйнэ» ‘ring; start ringing’). The suffix *-škue'dte* is used to derive inchoatives irrespective of the syllable count of the base verb; it may appear alongside a synonymous *-gue'dte/-kue'dte* inchoative, e.g., *luujhkue'dte* «лүйххкүэдтэ» ~ *luujhkeškue'dte* «лүйххкэшкүэдтэ» ‘start crying’ (← *luujjhke* «лүййжкэ» ‘cry’).

The Kildin Saami derivatives with the suffix *-je* can, in spite of the reductions and analogies in second syllable vocalism, be divided into distinct stem vowel groups (*-āj-, *-ij-, *-uj-) by examining and comparing the vowel and consonant center qualities to those of the base verb (and to cognates in other Saami languages). All the inchoatives recognizable as *-āj-stems are derived from *ē-verbs and largely coincide with the Skolt inchoatives of the same type, e.g., *puell,je* «пуэллѐ» ‘kindle; go up in flames’ (← *puu'lle* «пүүлѐ» ‘burn (intr.)’). Most *-ij-inchoatives derive from *e-stems (e.g., *poo'rrkje* «пōрркѐ» ‘start snowing heavily’ ← *poorrke* «пōрркэ» ‘snow heavily’) and *-uj-inchoatives from *ō-stems (e.g., *oollvje* «ōллвѐ» ‘let out a

howl' ← *állve* <оаллвэ> 'howl'), although there are single exceptions with *-ij-inchoatives from an *ō-stem (*koo'bbl,je* <кōббълье> 'come to the surface' ← *kåbble* <коабблэ> 'float') or *-uj-inchoatives from an *ē-stem (*leeffkje* <лѣффкъе> 'start coughing' ← *la'ffke* <лаффъкэ> 'cough'). In addition to these groups, which reflect regular Proto-Saami derivational processes, there is a considerable number of inchoative verbs where the suffix *-je* seems to be directly attached to the unchanged Kildin base verb stem, resulting in derivatives that from a historical perspective are irregular, even irreconstructible, such as *luujjhke* <лӯјјкъе> 'start crying' (← *luujhke* <лӯјјкэ> 'cry') and *laaffje* <лāффъе> 'begin to rot' (← *laaffe* <лāффэ> 'rot'). In some cases, both a historically regular derivative (*állve* <оаллвэ> 'howl' → *oollvje* <ōллвѣе> 'let out a howl') and an analogical one (*állvje* <оаллвѣе> 'start howling') are present; the former seems to carry more momentative connotations whereas the latter are more purely inchoative in meaning.

Inchoatives in *-le* are mostly used with motion verbs and show no sound alternation with respect to the base verb, e.g., *a'zle* <адзълэ> 'begin to rise (of water)' (← *a'ze* <адзе> 'rise (of water)'). Thus, they are phonologically similar to the subitive-diminutive derivatives in *-le* < *-lē- (*a'nntle* <аннѣтлэ> 'give (quickly)' ← *a'nnte* <аннѣтэ> 'give') and differ from the frequentatives in *-le* < *-elē- (*eenntle* <ѣннтлэ> 'give (continuously/sometimes)'). The connection is also obvious with derivatives having both subitive/diminutive and inchoative uses, such as *šeenntle* <шѣннтлэ> 'grow up (quickly); grow a bit; start to grow' (← *šeennte* <шѣннтэ> 'grow, etc.').

The data contains one Kildin inchoative verb in *-te*, which corresponds to North Saami *-iidit*: *ko'lkte* <кольктэ> 'begin to flow, run' (← *koollke* <кōллкэ> 'flow, run'), cf. SaaN *golggiidit* 'id'. The single example of the complex suffix *-la'hte* in *poorrkla'hte* <пōррклахтэ> 'start snowing heavily' (↔ *poorrke* <пōрркэ> 'snow heavily') may instead be based on the noun *poorrk* 'blizzard', as the suffix is mainly attested in denominal use. One or two verbs with the suffix *-la'hče* can also be analyzed with reservations as inchoatives, e.g., *šeeŋla'hče* <шѣн҅улахче> 'begin to storm' (↔ *šeeŋne* <шѣн҅нэ> 'storm, be stormy weather', *šeeŋŋ* <шѣн҅н҅> 'storm, bad weather').

There are a few inchoative verbs with the suffix *-ne*, for example *azne* <азнэ> 'start to rise (of water, at high tide, quickly)' (← *a'ze* <адзе> 'rise, of water'). The unpalatalized weak-grade consonant centre shows that the suffix goes back to the Proto-Saami frequentative suffix **-nte-* and corresponds to *-(á)d-* in the more western languages. A single case of *-se* (**-stē-*) can also be found: *tii'dse* <тйдѣсэ> 'find out' (← *tii'dte* <тйдтэ> 'know'). The complex suffix *-selle* in *piinŋkselle* <пйн҅уксэллэ> 'begin to blow (of wind)' and *jeallselle* <яллсэллэ> 'be resurrected' (← *jie'lle* <ѣлле> 'live') seems to reflect

*-sele-, provided that these derivatives are really deverbal and not based on the nouns *piinŋk* <пйиңк> ‘wind’ and *jeall* <ялл> ‘life’.

The suffix *-uvve* appears as the variants *-luvve* (most frequent), *-juvve*, *-xuvve* and *-šuvve*. For many of the individual derivatives, there exists a corresponding *-le/-je/-xe/-še* derivative, which at least formally can be seen as the immediate base of the *-uvve* verb; compare for instance *saarnluuvve* <сāррнлуввэ> ‘begin to talk’ ← *saarnle* <сāррнлэ> ‘talk (for some time)’, *saarnne* <сāррнэ> ‘talk’ and *illšuvve* <ыллшуввэ> ‘begin to sting, burn’ ← *illše* <ыллшэ> ‘sting, burn (of a wound)’.

Table 9. Frequencies of inchoative verb types in Saami languages (based on selected dictionaries).

	nV	S	U	P	L	N	I	Sk	K
<i>*-koatē-</i>	-	54	13	73	39	(4)	42	1	33
<i>*-čkoatē-</i>	-	-	-	(10)	2	(4)	212	125	42
<i>*-ekčē- (-ākčē-)</i>	-	18	14	55	21	-	-	-	-
<i>*-jekčē-</i>	42	42	-	-	-	-	-	-	-
<i>*-lekčē-</i>	6	-	-	-	-	-	2	2	2
<i>*-j-</i>	151	49	11	36	53	45	18	10	47
<i>*-āje-</i>	80	20	11	27	45	39	13	8	10
<i>*-ejē- > *-ijē-</i>	28	19	-	-	-	4	5	2	7
<i>*-ujē-</i>	22	10	-	9	8	2	-	-	9
Analog. <i>-j-</i>	21	-	-	-	-	-	-	-	21
<i>*-l-</i>	119	12	7	30	20	41	33	11	18
<i>*-ēlē-</i>	13	5	3	2	2	5	4	3	7
<i>*-ōlē-</i>	32	-	1	-	2	16	16	4	5
<i>*-elV-</i>	68	5	3	25	15	20	13	4	6
<i>*-ālV- / *-ilV-</i>	5/1	1/1	-	3	1	-	-	-	-
<i>*-ārV-</i>	13	-	-	5	10	-	-	-	-
<i>*-ne-</i>	34	1	-	2	2	28	3	-	1
<i>*-enē-</i>	4	-	-	-	-	4	1	-	-
<i>*-āne-</i>	30	1	-	2	2	24	2	-	1
<i>*-nte-</i>	31	9	3	7	6	5	4	4	5
<i>*-ijtV-</i>	41	-	-	5	14	31	5	-	1
<i>*-(ē)htē-</i>	11	8	-	-	-	4	-	-	-
<i>*-enkv-</i>	35	1	-	-	2	15	18	-	-
<i>-lg-</i>	2	2	-	-	-	-	-	-	-
<i>*-āske-</i>	27	1	-	-	-	23	8	-	-

*-(e)se-	9	2	1	-	-	7	-	-	-
*-stē-	30	9	2	4	2	14	4	3	1
*-stuve-	41	-	-	-	-	35	6	2	-
*-(j/l)tuve-	3	5	-	-	-	3	-	-	-
*-(l)uve-	24	-	-	-	-	-	-	-	24

3. Derivational types across Saami languages

3.1. General figures

In this section, the use of each derivational type or suffix found in inchoative functions in Saami languages is treated from the perspective of the whole language area; areal differences are analyzed and groups of most widespread individual derivatives are pinpointed as possible indications of the early existence of the derivational type in question. Table 9 above summarizes the frequencies of each inchoative verb type in each language as found in the data sources of the current article. It must once again be stressed that due to the variable nature and extent of the sources, the figures for each language are not directly comparable and should rather be taken as indicative of tendencies. The most clearly underrepresented cases, where a highly productive derivational type is represented by just a few examples, are marked in parentheses. Some of the types (*-j-, *-l-, *-n-) are further split into sub-types based on vocalism. The column **nV** indicates the number of individual base verb + suffix variant combinations (an etymologically identical base + suffix in several languages counts as one).

3.2. *-koatē-, *-čkoatē-

As inchoative formation with the suffixes continuing Proto-Saami *-koatē- (or its extended version *-čkoatē-) is apparently a highly productive, regular morphological process in all Saami languages, there is little point in listing and matching single derivatives across languages. The number of recorded derivatives reflects mainly the general extent of the data available for each language. The proportions in the use of the suffix variants are, however, a matter of interest. The general picture is that *-koatē- (North Saami -goahtit, etc.) is used with parisyllabic base verbs and the extended *-(č/š)koatē- with imparisyllabic base verbs in the northwestern languages (from Pite Saami to North Saami); South and Ume Saami lack the extended suffix and use only the suffix *-ekčē- (-ahtjedh) with imparisyllabic verbs, which is also common

alongside *-(č/š)koatē- in the other western varieties up to Lule Saami and appears to be very productive as well. In the eastern languages, by contrast, the extended suffix *-škoatē- has taken the place of a general-use inchoative suffix irrespective of the syllable count of the base verb. In Inari, Skolt and Kildin Saami, the simple *-koatē- is used only in a restricted set of derivatives often having some kind of a lexicalized meaning, in Inari Saami also with subitive/momentative verbs of onomatopoeic stems. In Ter Saami, however, the system resembles that of North Saami, and the extended *-škoatē- suffix is found with parisyllabic base stems only occasionally. (For a more detailed account, see Kuokkala 2019: 155–163.)

3.3. *-ekčē- (> *-ehčē-)

In western Saami languages, from South Saami up to Lule Saami, a very common inchoative suffix used with imparisyllabic base verbs is *-ehčē- (S *-ahtjedh*, L *-ahtjat*, in Ume and Pite Saami a secondarily lengthened *-áhhtjat*). South Saami also has a suffix variant *-jahtjedh* ~ *-jihhtjedh* attaching to parisyllabic stems, which probably contains another common inchoative suffix *-j-* in the beginning. The suffix is without a doubt historically the same as the frequentative suffix *-ekčē- : *-ehčē-, which in the modern Saami languages occurs usually only with parisyllabic base verbs (e.g., SaaL *váttsatjit*, N *váččašit* ‘go slowly or to and fro’ ← N *vázzit* ‘walk’; cf. M. Korhonen 1981: 337; Sammallahti 1998: 92). Only in eastern Saami varieties does it have a successor in an imparisyllabic context, namely in the complex frequentative(-inchoative) suffix SaaL *-lâhčâ-* ~ *-luhčâ-* ~ Sk *-lõõččâ-* ~ K *-la'hče*. (About the relationship of *-ekčē- with the extended *-(č/š)koatē- suffix, see Section 4 and Kuokkala 2019).

3.4. *-j- (*-āje-, *-eje-, *-uje-)

The *-j- suffix family can be said to be the second most popular inchoative suffix family throughout the Saami languages. The current data contains about 280 individual derivatives of 135 base verbs representing all Saami languages. As regards derivational morphophonology, the expected pattern between the base and the derivative stem vowel, reconstruable to Proto-Saami as *-ē- → *-āje- (~ *-â-), *-e- → *-eje- (~ *-î-), *-ō- → *-uje- (~ *-û-), is quite regularly preserved in South and Ume Saami (cf. Sections 2.1, 2.2). Looking at individual verbs, the same pattern occurs in all widespread *-j-inchoatives. The following derivatives are known in virtually all Saami languages (example forms from North Saami, language distributions in brackets):

<i>eallit</i> ‘live’	→ <i>eal’lát</i> ‘revive’	[S – P L N I Sk K T] ¹²
<i>johtit</i> ‘move, travel’	→ <i>johttát</i> ‘start moving’	[S U P L N I Sk K T]
<i>gohcit</i> ‘be awake’	→ <i>gohccát</i> ‘wake’	[S U P L N I Sk K T]
<i>muitit</i> ‘remember’	→ <i>muitát</i> ‘come to remember’	[S U P L N I Sk K –]
<i>oaddit</i> ‘sleep’	→ <i>oad’đát</i> ‘fall asleep’	[S U P L N I Sk K T]
<i>buollit</i> ‘burn (intr.)’	→ <i>buol’lát</i> ‘catch fire’	[S U P L N I Sk K –]

The following derivatives also have a distribution from South/Ume Saami to eastern Saami:

<i>muohttit</i> ‘snow’	→ <i>muohttát</i> ‘begin to snow’	[– U P L N I – – –]
<i>vardit</i> ‘bleed’	→ <i>vardát</i> ‘begin to bleed’	[S U P L N I – – –]
<i>duoldat</i> ‘boil’	→ <i>duldet</i> ‘begin to boil’	[S – – – N I Sk K T]
<i>orrut</i> ‘be still, stay, dwell’	→ <i>or’rot</i> ‘settle down’	[S – P L N – – K –]

There are also similarly regular **-j*-derivatives with wide distributions from North Saami westwards (*vázzát* ‘start walking’, *vuoddját* ‘drive off’, *vuordát* ‘begin to expect’, *čir’rot* ‘begin to cry’) or eastwards (*ahccát* ‘begin to rise (of water)’, *arvát* ‘begin to rain’, *máhttát* ‘begin to know; learn’, *bigget* ‘begin to blow (of wind)’, *borget* ‘start driving snow’). A number of **-ō-* → **-uje-* inchoatives are attested in the western group of South, Pite and/or Lule Saami (e.g., SaaL *doarrot* ‘fight, quarrel’ → *doar’rut*, P *dur’rut*, S *dåârrijidh* ‘begin to fight’), six of which can be found in more than one language.

The system of **-j*-inchoatives has undergone wide-scale analogical developments in two areas. First, the central languages North, Lule and Pite Saami have generalized the *-á(j)*-suffix so that it can also be attached to stems other than **ē*-stems, such as SaaN *gullat* ‘hear’ → *gul’lát* ‘begin to hear’ or *doarrut* ‘fight’ → *doar’rát* ‘begin to fight’. Many of the analogical derivatives are common to two or three of these languages (P–L–N: *čuoigát*, *gul’lát*, *bal’lát*, *biv’vát*, *buohccát*, *skihppát*, *dovdát*; L–N: *rahkkát*, *suhkkát*, *doar’rát*; P–L: *tsahkkát*, *duolldát*, *sjnjuv’vát*).¹³ Single verbs are also attested in Inari (*pallad’* ‘get a fright’) and Skolt Saami (*puâc’cjed* ‘fall ill’).

Second, Kildin Saami has generalized the suffix *-je* so that a large number of inchoatives have been formed by simply attaching the suffix to the

¹² The distribution figures have been supplemented with Ter Saami data from KKLS. See language abbreviations at the end of the article.

¹³ Meanings: P–L–N: *čuoigát* ‘set off on skis’, *gul’lát* ‘begin to hear’, *bal’lát* ‘become scared, get a fright’, *biv’vát* ‘warm up (after being cold)’, *buohccát* ‘become ill’, *skihppát* ‘become unwell, fall ill’, *dovdát* ‘recognize’; L–N: *rahkkát* ‘begin to rut’, *suhkkát* ‘begin to row’, *doar’rát* ‘begin to fight’; P–L: *tsahkkát* ‘begin to glow; ignite’, *duolldát* ‘start boiling’, *sjnjuv’vát* ‘flare up’.

contemporary Kildin base verb stem, without applying the initial-syllable vowel changes that the suffix change would entail from a historical point of view (21 recognizable cases in our data). Notably, though, Kildin Saami also has a group of its own historically regular **-ō-* → **-ujē-* inchoatives (e.g., *oollvje* <ОЛЛВЪЕ> ‘give a howl’ ← *ǎllve* <ОАЛЛВЭ> = SaaN *holvut* ‘howl’), which might at least partly be remains of an older layer of derivatives that have disappeared from the neighboring languages, considering that Kildin still has a number of regular **-ē-* → **-ājē-* and **-e-* → **-ejē-* inchoatives in common with Skolt and Inari (as well as North) Saami but not found elsewhere.

As for the origins of the inchoative type, one should note that verbs with a suffixal **-j-* can have several different functions in the Saami languages. Denominal derivatives (with the generalized suffix variant **-ejē-* ~ **-î-*) can, for instance, mean using or producing the referent of the base word, such as SaaN *murjet*, S *möörjedh* ‘pick berries’ ← SaaN *muorji*, S *muerjje* ‘berry’ (cf. M. Korhonen 1981: 331–333). Deverbal **-ujē-* ~ **-û-* verbs are mostly automative-passive verbs (e.g., SaaN *gullot* ‘be audible’ ← *gullat* ‘hear’), where the suffix **-ujē-* < **-ōj(ē)-* is apparently identical to Finnic *-u/-y-* (Finnish *kuulua* ‘be audible’ ← *kuule-* ‘hear’) and goes back to **-ow(i)-* < stem vowel + **-w-* (Sammallahti 1998: 84–85). Deverbal **-ejē-* ~ **-î-* verbs have a momentative meaning, such as SaaN *njuiket* ‘jump (once)’ (as opposed to frequentative/continuative *njuikut* ‘jump (several times)’). It is thus possible that the inchoative function is ultimately based on (momentative?) aspectual use, as in the case of several other inchoative types. A possibly related derivational suffix going back to **-j-* is known in many Uralic branches with a slightly different aspectual meaning, including the frequentative-continuative suffixes in Finnic (Finnish *hyppiä* ‘jump (continuously)’ vs. *hypätä* ‘jump (once)’), Mordvin (Erzya *tośtajems* ‘push (repeatedly), jostle’ vs. *tośiadems* ‘push, shove’) and Permian (Udmurt *vošjany* ‘keep changing’ vs. *voštyny* ‘change’) (cf. Bartens 1999: 159; 2000: 281).

3.5. **-lē-*

The inchoative derivatives with an *-l-* suffix are formally identical to the subitive-diminutive derivatives with the **-lē-* suffix; the stem shows strong-grade consonantism and no vowel changes with respect to the base verb (cf. M. Korhonen 1981: 333–334). Only a small number of derivatives in western Saami languages show signs of a generalized *-alit* or *-álit* inchoative suffix (e.g., SaaP *luodjalit* ← *luodjot* ‘complain’). The following *-l-*inchoatives occur in at least four Saami languages:

<i>čuoigat</i> ‘ski’	→ <i>čuoigalit</i> ‘set off on skis’	[S – – – N I Sk K T]
<i>čuovvut</i> ‘follow’	→ <i>čuovvulit</i> ‘start following’	[– U – L N I Sk – –]
<i>girdit</i> ‘fly’	→ <i>girdilit</i> ‘fly away’	[– – – – N I Sk K –]
<i>oahkut</i> ‘stroll’	→ <i>oahkulit</i> ‘start walking’	[– – – – N – Sk K T]
<i>suhkat</i> ‘row’	→ <i>suhkalit</i> ‘row away’	[S – – – N I Sk K T]
<i>vázzit</i> ‘go on foot’	→ <i>vázzilit</i> ‘set off’	[S – – – N I Sk K T]
<i>vuodjit</i> ‘drive’	→ <i>vuodjēlit</i> ‘drive off’	[S – P L N I Sk K T]
<i>vuodjat</i> ‘swim’	→ <i>vuodjalit</i> ‘swim off’	[– – – – N I Sk K T]
<i>čuožžut</i> ‘stand’	→ <i>čuožžilit</i> ‘rise, get up’	[S U P L N – – – –]
<i>čuožžut</i> ‘stand’	→ <i>I čuázžulid</i> ‘stand up’	[– – – – – I Sk K T]

All the *-l*-inchoatives with a wide distribution (except for *čuožžilit*, about which more below) involve verbs denoting motion, usually active motion of a human being. The same applies to the derivatives found only in North, Inari or Skolt Saami. In Kildin Saami, *-l*-inchoatives may also take verbs of other kinds of semantics, for instance *šeennte* ⟨ШЭННТЭ⟩ ‘grow’ → *šeenntle* ⟨ШЭННТЛЭ⟩ ‘begin to grow’. In Pite and Lule Saami, the derivational type seems instead to favor verbs of somewhat affective semantics, especially those whose meaning involves a prominent sound or being insane (e.g., SaaP *jabmalit*, *klibmalit* ‘start making noise’; P *dárralit*, *dájjsalit*, L *dájssalit* ‘become hysterical, crazy’).

The verb *čuožžilit* ‘rise, get up’ is also exceptional as concerns its phonological relationship to the base verb, *čuožžut* ‘stand’. It is possible that this is a very old coinage based on the original illabial-vowel verb stem **čaŋća-* (cf. South Estonian *saisa-* ‘stand’) ? > Proto-Saami **čuoŋčē-*. However, the vocalism of the eastern Saami derivatives SaaI *čuázžulid*, Sk *čuęžžled* [tšųęđđíšled], K *čuennčle* ⟨ЧУЭННЧЛЭ⟩, T *čjeŋžaled* ‘stand up’ (KKLS: 684:13) points to the proto-form **čuoŋčōlē-* with a regular correspondence to the contemporary base verb.

3.6. **-ārV-*

Inchoative verbs in *-árit* (~ *-árít*) are found only in Pite and Lule Saami, including a couple of common verbs (SaaP *stuvvsárit*, L *stuvssárit* ‘begin to play’; SaaP *gájbmárit*, L *gájmmárit*¹⁴ ‘become unstable, insane’). Their logical base verbs represent varying stem types, including the rare *-ur*-verbs such as Pite *gājbmurit* ‘be in an agitated state’. Since *-r-* is not otherwise known as a verb-forming suffix in Saami languages, this inchoative type must

¹⁴ The Lule Saami verb is found only in the dictionaries of Kintell (2012) and Grundström (1946–1954 s.v. *kai’mārit*).

have come about through analogy with verbs from other backgrounds, compare for instance SaaL *duhpárit* ‘clog up in wet snow’ (← *duber* ‘wet snow’) or *báhtarit* ‘flee’ (< Proto-Uralic **pätäri-*; see Aikio 2013: 162).

3.7. **-ne-*

The verb-forming suffix **-ne-* is usually found in transitive or reflexive/passive verbs and is presumed to ultimately go back to Proto-Uralic **-mi-*, cf. SaaN *buorránit* ~ Finnish *parane-* ‘get better’ (M. Korhonen 1981: 335–336). In inchoative function, the suffix mostly appears sporadically in different Saami languages. Only in North Saami has it been more productive in the suffix *-ánit*, corresponding to verbs in *-át* which denote the state of having a definite appearance, such as *barfánit* ‘set off (being furry, bushy)’ ↔ *barfát* ‘lie/stand furry, bushy somewhere’ (↔ *barffas* ‘furry, bushy’ ↔ *barfi* ‘a furry, bushy one’). Two derivatives appear in more than one language, namely (S, P,) L *hállánit* ‘fly off’ (← *háledit* ‘fly’) and N *cahkkánit*, I *cokkâniđ* ‘begin to burn’ (← N *cahkat* ‘smoulder’). Kildin Saami *cuagkne* (цyагкнэ) ‘begin to fall (of water level)’ (← *cuagke* цyагкэ ‘fall (of water level)’) belongs to this type as well, judging by its strong-grade stem.

3.8. **-nte-*

The inchoative verbs with the suffix *-d-* up to Inari Saami, as well as the corresponding *-eed-*verbs in Skolt Saami and *-ne-*inchoatives in Kildin Saami, seem to generally go back to the frequentative suffix **-nte-* (cf. M. Korhonen 1981: 330–331). There are few derivatives of this type common to several languages; the most widespread ones are (P, U,) L *tjuottjadit* ‘stand up’ (← *tjuodtjot* ‘stand’), which has an analogical stem vowel, and the two posture verbs SaaN *čohkkedit* ‘sit down’ (← *čohkkát* ‘sit’) and *vel’ledit* ‘lie down’ (← *veal’lát* ‘lie’; both S, U, P, L, N), which are formed with a contracted **î-*stem. Other **î-*stem derivatives include Skolt *lie’n’njeed* ‘lie down’ (← *lien’njad* ‘lie, rest’) and *no’kķeed* ‘fall asleep’ (← *nokkâd* ‘doze’). Besides these, strong-grade formations are found in Lule *guoddádít* ‘(begin to) support (of snow crust)’ (← *guoddet* ‘carry’) and *váđtsádít* ‘set off (of many)’ (← *váđtset* ‘walk’), which are probably based on *-á(j)-*inchoatives (cf. *váđtsát* ‘set off walking’). Usually, though, *-d-*type inchoatives show a weak-grade stem as expected before the **-nte-* suffix, both when the stem vowel is analogical (e.g., SaaL *čiärádiđ* ‘start crying (suddenly)’ ← *čiärruđ* ‘cry, weep’) or regular with respect to the base verb (SaaN *šlahtádít* ‘begin to sleet’ ← *šlahttiit* ‘sleet’; SaaK *azne* [ʔɛcāntɛ-] ‘begin to rise (of water)’ ← *a’ze* [ʔɛcē-] ‘rise (of water)’).

3.9. *-ijtV-

The derivational suffix *-ijt- appears in inchoative verbs in Pite, Lule, North and Inari Saami, and individual verbs are found only in one or two languages each. An exception is the verb SaaN *golggiidit* ‘begin to flow, run’ (← *golgat* ‘flow, run’), which has attestations in all four of these languages and also in Kildin Saami (*ko'lkte* «кольктэ» id.). Besides *ē-stems, base verbs include *ē- and *ō-stems, but especially in North Saami, many of these inchoatives lack a formal base verb and instead correlate with other complex verbs and possibly also nouns. As regards semantics, many of these verbs denote the beginning of some kind of physical or mental state.

3.10. *-(ē)htē-

Inchoative verbs with suffixal -ht- are found in South Saami, with a couple of verbs also occurring in North Saami; there is one possibly common derivative: SaaS *gelkiehtidh* ‘begin to move off in a long line’, N *golggihit* ‘suddenly run (out) or float away’ (← S *galkedh*, N *golgat* ‘run, float’). The suffix brings with it a preceding vowel *ē just like the identical subitive-momentative suffix *-ēhtē- (cf. SaaS *tjelkiehtidh*, N *čalggihit* ‘suddenly go through (ice, etc.)’ ← S *tjalkedh*, N *čalгат* ‘go through, sink’). Many of the -ht-inchoatives also have momentative connotations.

3.11. *-(ē)ŋkV-

North and Inari Saami have a distinguishable type of inchoative verbs that occur with the suffix -(a)g- (< *-(ē)ŋk-) and mostly denote going into a (physical) state or posture. Most have a (semantic) base verb in -á(j)-, but they may also have nominal and adverbial correlates, for instance SaaN *hirragit* ‘begin to stick out or sprawl’ ↔ *hir'rát* ‘stand or lie sticking out or sprawling’ ↔ *hirras* ‘(adj.) sticking out; sprawling’ ↔ *hirril* ‘(adv.) id.’. Notably, these 15 North and 18 Inari Saami derivatives include no common verbs. A couple of stray -g-inchoatives can be found in South and Lule Saami, too. In general, the -g- verbal derivative type appears to be originally denominal and possibly goes back to the Pre-Saami translative suffixes *-m(i)- + *-ka- > Proto-Saami *-ŋkē-.

3.12. *-āskē-

Inchoatives with the suffix -āsk- also occur almost exclusively in North and Inari Saami. Some of the derivatives are common to the two languages: SaaN *riiddáskit*, I *riidáskiđ* ‘come into conflict’ (← N *riidalit* ‘quarrel (v.)’, *riidu*

‘quarrel (n.)’); SaaN *doaráskit*, I *tuáráskiđ* ‘begin to fight’ (← N *doarrut* ‘fight’); SaaN *fákkáskit*, I *fágáskiđ* ‘begin to wrestle’ (← N *fággádit* ‘wrestle’, *fággi* ‘wrestling’). One does, however, occur in South Saami as well: SaaS *jealaskidh* ‘begin to recover one’s health’, N *ealáskit*, I *iäláskiđ* ‘come back to life, regain consciousness’ (← N *eallit* ‘live’). Otherwise, the suffix **(ā)skē-* is used to form denominal transitive verbs (M. Korhonen 1981: 336).

3.13. **(e)se-*

Inchoatives in *-sit* in North Saami are similar to those in *-iidit* and *-git* in that they correlate with word families denoting states and postures and containing members belonging to several word classes. A few *-s-*inchoatives are also found in South and Ume Saami, but they represent verbs of more general semantics (SaaS *goerkesidh* ‘reach an understanding’; SaaS *valkesidh*, U *valgg(a)sit* ‘begin to go in one direction’). The suffix **-se-* is normally used to form reflexive-automotive verbs (M. Korhonen 1981: 336; Magga & Magga 2012: 147).

3.14. **-stē-*

A few inchoatives with the suffix **-stē-* are found in almost all Saami languages, most in South (9) and North Saami (14). Two derivatives are more widely spread: (S, U, P, L,) N *biekkastit* ‘begin to blow (of wind)’ (← *bieggat* ‘blow (of wind)’) and (S, U, P, L,) N, (I) *ruohtastit* ‘begin to gallop’ (← *ruohttat* ‘gallop’). Besides inchoativity, the derivatives often have momentative, subitive or diminutive semantics as well, which are actually the main functions of the deverbal **-stē-* suffix (cf. M. Korhonen 1981: 337).

3.15. **-stuve-*, **-uve-* (< **-uove-*)

North, Inari and Skolt Saami have a group of verbs with the suffix SaaN *-stuvva-*, I *-stu(v)-*, which bear inchoative connotations. Inchoativity can be seen as a further development of the usual semantics of the verbs with these suffixes, which include feeling the need to, wanting, intending or being close to doing something; consider, for instance, the meanings of SaaI *čiärustuđ* ‘feel like crying; begin to cry’ (← *čiärruđ* ‘cry’). South, North and Kildin Saami also have a number of other inchoative verbs with the **-uve-* ~ **-uove-* element, which usually has passive, automotive and transitive functions.¹⁵

¹⁵ The reconstruction **-uove-* is mainly based on the southern South Saami variant *-oeve-* and appears to reflect a more archaic form from which the variant **-uve-* has been simplified.

4. Conclusions and discussion

The SaaN *-goahitit* (**-koatē-*) suffix is the most widely used means of inchoative derivation throughout the Saami language area. In Inari–Skolt–Kildin Saami, the extended variant **-č/škoatē-* has taken the place of a general-purpose inchoative suffix. Other widely used suffixes forming purely inchoative derivatives are **-j-*, represented by so-called contraction verbs in most Saami languages and generalized as the *-á(j)-* derivational type in North, Lule and Pite Saami, as well as *-ahtja/-áhija-* (**-ekčē-*), which is used with imparisyllabic (trisyllabic) base verbs from South Saami up to Lule Saami. The rest of the suffixes tend to have other simultaneous aspectual implications besides inchoativity or special inclinations concerning the base verb semantics, which may be areally more or less restricted. Despite its divergent developments, at least the **-lē-* inchoative type (SaaN *vázzilit*) must also already have been established in Proto-Saami, judging by the number of widely known common derivatives (Section 3.5). Actually, the distributional figures show that the **-lē-* inchoatives of several motion verbs (N *vázzilit*, *čuoigalit*, *suhkalit*) must have been secondarily replaced with **-ā-* inchoatives (SaaN *vázzát*, *čuoigát*, *suhkkát*) in the northwestern Saami languages Pite–Lule(–North). Within North Saami, too, the *-á-* variants are generally known only in the western dialects (Guovdageaidnu; according to Nielsen 1979 [1932–62]).

In several cases, it can be observed that the inchoative function has been reinforced by adding two consecutive suffixes used as inchoatives, such as in SaaL *roohâđ* ‘be in the rut’ → *rohâliđ* ‘come into the rut’ → *rohâlâhčâđ* ‘begin to come into the rut’ (*-l-* and *-âhčâ-* also often act as frequentatives). This has led to new productive derivational types, which have been favored by the common Saami inclination towards wordforms consisting of bisyllabic feet, especially in verbal morphology (cf. M. Korhonen 1981: 259–260). Consider the combination of the suffixes *-j-* + *-ahtje-* > SaaS *-jahtje-* when used with a bisyllabic stem: now, for example, the frequent PRS.3SG form *doel.ti|jah.tja* ‘begins to boil’ is prosodically more favorable than that of the plain *-j-* inchoative, *doel.ti.je*. The same factor must also have in part contributed to the amalgamation of the suffixes **-(e)kčē-* + **-koatē-* > **-(e)hčēkoatē-*: PRS.3SG **pāh.ken|te.k.čē|koa.tā* ‘begins to warm’ is again more favorable than **pāh.ken.te|koa.tā* (the **-koatē-* suffix always carries secondary word stress and thus starts a new foot).¹⁶ In this case, though, the prosodic advantage was

Sammallahti (1998: 84–85, 93) considers the suffix **-uovē-* a reflex of the Finno-Saamic passive-reflexive **-ov(i)-* < Proto-Uralic (stem vowel +) **-w-*.

¹⁶ Erroneous reconstructions and their meanings in Kuokkala (2019: 175) should be replaced with these.

later lost when the complex suffix was reduced to **(ē)čkoatē-* (SaaP *-atjgâhtet*) > **-škoatē-* (SaaN *-šgoahtit*, etc.), and the way was open for generalization, leading the new suffix variant to largely replace the old **-koatē-* in all syllable contexts in eastern Saami.

Many of the Saami inchoative suffixes are formally identical to derivational suffixes carrying other aspectual functions, which are partly also present in the corresponding derivatives classified here as inchoative. These include the momentative/subitive/diminutive suffixes **-lē-*, **-ēhtē-*, **-stē-*, frequentative **-nte-*, **-ēhčē-*, reflexive/passive/automative **-ne-*, **-se-* and translative **-ne-*, **-(ē)ŋk-*, **-(ā)ske-*. As the other functions are lexically and geographically more widespread, the obvious conclusion is that the inchoative function is secondary and developed from the other meanings. In one case, however, it appears that an originally inchoative suffix has taken over a momentative function, namely the Inari Saami suffix *-gyettiđ* (see Section 2.5).

The change in function is easy to understand in the case of subitive-momentative derivatives, as especially with onomatopoeic and motion verbs, subitive or momentative aspect, i.e., sudden or brief occurrence, pragmatically also means the beginning of a sound or a movement and can thus develop into a lexicalised inchoative meaning. Frequentative aspect, on the other hand, has a tendency to secondarily denote attempt or intention, which again have a direct implicational link to a briefly beginning action. Connecting links of these semantic chains can be illustrated with contemporary Saami vocabulary; consider for instance SaaN *cielihit* = *cieládit* ‘let out a bark; suddenly begin to bark’ (momentative/inchoative); *viehkaliit* ‘run quickly (to somewhere); run away’ (subitive/inchoative); *beasadiit* ‘get loose (several times); try to escape’ (frequentative/conative); *čuogastuvvat* ‘want to / begin to gobble, lap’ (intentional/inchoative). Finally, the connection of passive/automative and translative functions with inchoativity has to do with the notion of change of state, expressed either verbally or nominally, which, when focusing on the result, means simply the beginning of a new state. To mention a parallel case, the Karelian reflexive/passive/translative verb suffixes *-utu-/yty-* and *-vu-/vy-* have also acquired an inchoative use (e.g., *elevyö* ‘revive; begin to live’ ← *eleä* ‘live’). The expansion of an originally denominal derivative type, such as translatives, into deverbal use may have happened with the support of such derivatives that have an associatable correlate in both word classes, consider for instance SaaN *ražáskit* ‘begin working hard’ ← *rahča* ‘hard work’, *rahčat* ‘work hard, struggle’.

Thus, nearly all the means of inchoative derivation in Saami languages seem to stem from suffixes with other aspectual functions, and although a few of the derivative types do have a wide distribution and widely known

individual derivatives in the language area, we can hardly assume more than 3–4 of the 14 suffixes listed in Section 3 to have constituted even a liminally established inchoative derivational type in Proto-Saami. One or two old inchoatively used derivatives may have later provoked the expansion of an inchoative type, as in the case of **-stē-* and SaaN *biekkastit* ‘begin to blow (of wind)’, *ruohtastit* ‘begin to gallop’; otherwise, we must largely assume parallel developments. The **-koatē-* inchoative suffix has in the past been etymologically connected to Veps *-gande-*, *-škande-* and Mordvin *-kadā-*, *-gadā-* inchoatives, but the comparisons are problematic from the perspective of historical phonology, and in view of the fact that the Saami suffix acts prosodically like a compound part, it seems probable that it developed during the Proto-Saami period from an independent lexeme, possibly the verb **poatē-* ‘come’ (see Kuokkala 2019). The Mordvin momentative/inchoative suffix *-ž(ə)-*, which has been compared with SaaS–L *-ahtja-/ -āhtja-* (Hallap 2000 [1955]: 138), must also be separated on phonological grounds, as the Inari Saami cognate *-(l)āhčā-* points to original **-kč-* and a connection with the Finnic frequentative *-ks(i)-* (e.g., Finnish *pureksia* ‘chew’ ← *pure-* ‘bite’). To conclude, all the derivational means of expressing inchoativity in Saami languages appear to have developed during the separate development of the Saami branch of the Uralic language family, most of them even after the breakup of the common Proto-Saami language.

Abbreviations

adj. —adjective; adv. —adverb; intr. —intransitive (verb); n. —noun; (Saa)I—Inari Saami (Saa); K—Kildin Saami; (Saa)L—Lule Saami; (Saa)N—North Saami; (Saa)P—Pite Saami; (Saa)S—South Saami; (Saa)Sk—Skolt Saami; (Saa)U—Ume Saami; (Saa)T—Ter Saami; tr. —transitive (verb); v. —verb.

References

- Aikio 2013 = Luobbal Sámmol Sámmol Ánte (Ante Aikio). 2013. Studies in Uralic etymology I: Saami etymologies. *Linguistica Uralica* 49(3). 161–174.
- Álgu = Álgu. Sámegielaid etymologáš diehtovuoddu / The etymological database of the Saami languages (online database). Helsinki: Institute for the Languages of Finland. (<http://kaino.kotus.fi/algu/>)
- Antonova, Aleksandra A. 2014. *Saamsko-russkij slovar’*. Murmansk: Lema.
- Barruk, Henrik. 2018. *Báhkuogirjje: Ubmejesámien-dáruon, Dáruon-ubmejesámien / Ordbok: Umesamisk-svensk, Svensk-umesamisk*. Umeå.
- Bartens, Raija. 1999. *Mordvalaiskielten rakenne ja kehitys* (Mémoires de la Société Finno-Ougrienne 232). Helsinki: Société Finno-Ougrienne.
- Bartens, Raija. 2000. *Permiläiskielten rakenne ja kehitys* (Mémoires de la Société Finno-Ougrienne 238). Helsinki: Société Finno-Ougrienne.
- BB = [Wilbur, Joshua (ed.):] Bidumsáme Báhkogirje / Pitesamisk ordlista / Pite Saami Word List (online database). (<http://saami.uni-freiburg.de/psdp/pite-lex/>) (version from 2020-04-24).

- Bergsland, Knut. 1994. *Sydsamisk grammatikk*. 2. utgave. Karasjok: Davvi Girji.
- Bergsland, Knut & Lajla Mattsson Magga. 1993. *Aarjelsaemien-daaroen baakoegærja / Sydsamisk-norsk ordbok*. [Lakselv:] Idut.
- Feist, Timothy. 2015. *A grammar of Skolt Saami* (Mémoires de la Société Finno-Ougrienne 273). Helsinki: Société Finno-Ougrienne.
- Giellatekno. 2020. Neahttagisánit: Source material: fin ↔ sms. (<https://saan.oahpa.no/about/>)
- Grundström, Harald. 1946–1954. *Lulelappsk ordbok / Lulelappisches Wörterbuch 1–4* (På grundval av K. B. Wiklunds, Björn Collinders och egna uppteckningar utarbetad av Harald Grundström (Skrifter utgivna genom Landsmåls- och Folkminnesarkivet i Uppsala, Ser. C:1). Uppsala: A.-B. Lundequistska Bokhandeln.
- Halász, Ignác. 1896. *Pite lappmarki szótár és nyelvtan* (Svéd-lapp nyelv VI). Budapest: Magyar Tudományos Akadémia.
- Hallap, Valmen. 2000 [1955]. *Verbaaltuletusufiksidi mordva keeltes* (Eesti Keele Instituudi toimetised 5). Tallinn: Eesti Keele Instituut.
- Hasselbrink, Gustav. 1981–1985. *Südlappisches Wörterbuch 1–3* (Skrifter utgivna genom Dialekt- och Folkminnesarkivet i Uppsala, Ser. C:4.). Uppsala: A.-B. Lundequistska Bokhandeln.
- Itkonen, Erkki. 1971. Ehdotus kildininlapin šongujn murteen fonemaattiseksi transkriptioksi. In Erkki Itkonen et al., *Lapin murteiden fonologiaa* (Castrenianumin toimitteita 1), 87–110. Helsinki.
- Itkonen, Erkki. 1986–1991. *Inarilappisches Wörterbuch 1–4* (Lexica Societatis Fenno-Ugricae 20). Helsinki: Société Finno-Ougrienne.
- Itkonen, T.I. 1958. *Koltan- ja kuolanlapin sanakirja / Wörterbuch des Kolta- und Kolalappischen I–II* (Lexica Societatis Fenno-Ugricae 15). Helsinki: Société Finno-Ougrienne.
- Kert, Georgij M. 1986. *Slovar' saamsko-russkij i russko-saamskij*. Leningrad: Prosveščenie.
- Kintel Anders. 1991. *Syntaks og ordavledning i lulesamisk*. [Kautokeino:] Samisk utdanningsråd.
- Kintel, Anders. 2012. *Julevsáme-dárro báhkogirjje*. (<http://gtweb.uit.no/webdict/ak/smj2nob/>)
- KKLS = See Itkonen, T.I. 1958.
- Korhonen, Mikko. 1981. *Johdatus lapin kielen historiaan* (Suomalaisen Kirjallisuuden Seuran Toimituksia 370). Helsinki: Suomalaisen Kirjallisuuden Seura.
- Korhonen, Olavi. 1979. *Báhkogir'je julevusámes dárroi, dáros julevusábmái / Lulesamisk svensk, svensk lulesamisk ordbok*. Uppsala: Almqvist & Wiksell.
- Korhonen, Olavi. 2007. *Báhkogirjje: Julevusámes dárroi, dáros julevusábmái / Lulesamisk svensk, svensk lulesamisk ordbok*. Jokkmokk: Sámiij áhpadusguovdásj.
- Kuokkala, Juha. 2019. Saami -(š)goahti inchoatives: their variation, history, and suggested cognates in Veps and Mordvin. *Journal de la Société Finno-Ougrienne* 97. 153–181.
- Kuokkala, Juha. 2020. A new general dictionary of Ume Saami (Review of Barruk 2018). *Finnisch-Ugrische Forschungen* 65. 195–201.
- Kuruč, Rimma D. (ed.). 1985. *Saamsko-russkij slovar' / Sām-rušš soagknehk'*. Moskva: Russkij jazyk.
- Magga, Ole Henrik & Lajla Mattsson Magga. 2012. *Sørsamisk grammatikk*. (Karášjohka:) Davvi Girji.
- Nickel, Klaus Peter. 1994. *Samisk grammatikk*. 2. utgave. (Karasjok:) Davvi Girji.

- Nielsen, Konrad. 1979 [1926]. *Lærebok i lappisk I: Grammatikk*. Oslo: Universitetsforlaget.
- Nielsen, Konrad. 1979 [1932–62]. *Lappisk (samisk) ordbok / Lapp dictionary* 1–5. Based on the Dialects of Polmak, Karasjok & Kautokeino. 2. opplag. (Instituttet for sammenlignende kulturforskning, Serie B: 17). Oslo.
- Rießler, Michael. 2022. Kildin Saami. In Marianne Bákro-Nagy, Johanna Laakso & Elena Skribnik (eds), *The Oxford guide to the Uralic languages*, 219–239. Oxford: Oxford University Press.
- Ruong, Israel. 1943. *Lappische Verbalableitung dargestellt auf Grundlage des Pitelappischen*. Uppsala.
- Sámediggi. 2020. *Koltansaamen ja pohjoissaamen digitaaliset sanakirjat valmistuneet*. (<https://www.samediggi.fi/2020/08/25/koltansaamen-ja-pohjoissaamen-digitaaliset-sanakirjat-valmistuneet/>) (2020-08-25).
- Sammallahti, Pekka. 1989. *Sámi-suoma sátnegirji / Saamelais-suomalainen sanakirja*. Ohcejohka: Jorgaleaddji.
- Sammallahti, Pekka. 1998. *The Saami languages: An introduction*. Kárášjohka: Davvi Girji.
- Sammallahti, Pekka & Jouni Mosnikoff. 1991. *Suomi-koltansaame sanakirja / Lää'dd-sää'm sää'nnke'rrj*. Ohcejohka: Girjegiisá.
- Schlachter, Wolfgang. 1958. *Wörterbuch des Waldlappendialekts von Malå und Texte zur Ethnographie* (Lexica Societatis Fenno-Ugricae 14). Helsinki: Société Finno-Ougrienne.
- SRS-Antonova = Saamsko-russkij slovar'. Avtor A.A. Antonova. DSL-format. (<http://slovari.saami.su/prilozheniya/skachat/file/17-dsl-format-saamsko-russkij-slovar-avtor-a-a-antonova.html>) (accessed 2021-01-30).
- SRS-Kuruč = Saamsko-russkij slovar' (kil'dinskij dialekt). Pod redakciej Kuruč R.D. (<http://slovari.saami.su/slovari/saamsko-russkij-slovar-kuruch/Saami-Russian-dictionary-Kuruch-1/>) (accessed 2021-01-17).
- Wilbur, Joshua (ed.). 2016. *Pitesamisk ordbok samt stavningsregler* (Samica 2). Freiburg: University of Freiburg, Department of Scandinavian Studies.

Morphological complexity in Saami: A historical-typological view

Johanna Nichols, University of California, Berkeley

1. Introduction

Saami inflection is very complex, but not evenly so in all parts of the morphology and not equally so in all the languages.¹ Complexity can be measured in various ways, of which the two most commonly used in linguistics will be used here (for kinds of complexity see Dahl 2004; Miestamo 2008; Kusters 2008; Sinnemäki 2011; Anderson 2015; Audring 2017; and many others). One is *enumerative complexity* (also known as *inventory complexity*, *taxonomic complexity*, and other terms): the count of items in a system, such as the number of phonemes in the inventory, the number of cases in noun declension, the number of tenses in verb inflection, and so on. Actually counting all of these things rapidly becomes cumbersome, and leads to comparing unlikes: a language will be considered morphologically complex if it marks local relations with a set of local cases such as inessive, illative, and elative, but not if it marks the same relations with postpositions; the answer is not adding a count of postpositions, since adpositions are usually an open class; and so on. Therefore, the count of enumerative complexity used here limits the survey to just three basic arenas: the phoneme inventory, the inflectional categories of the verb (as laid out in Bickel & Nichols 2013), and a similar inventory for nouns (Sokur & Nichols 2019) (for the whole survey see also Nichols 2019; Nichols & Bentz 2019).

The other measure is what has usually been known as *informational complexity* or *Kolmogorov complexity*, recently *canonical complexity* (Nichols 2020): the amount of information required to describe a system. The amount of information, in the form of prose discussion and listing and

¹ Research and initial drafting done while Helsinki University Humanities Visiting Professor, 2017–2019. Supported in part by a grant from the Russian Academic Excellence Project 5-100 to the Higher School of Economics, Moscow. I thank the SAALS 4 participants for comments and discussion.

annotation of different paradigms, needed to describe the Russian case system, with its four declension classes depending on the gender and arbitrary classifications and with its several instances of case syncretism (e.g., *-u* marks accusative, dative, genitive, or locative singular depending on the declension class; and the markers of accusative are *-u*, *-a/-Ø*, and *-o* in different paradigms), is much greater than what is required to lay out the set of case endings for a Saami language. This kind of complexity amounts to a measure of the non-transparency in a system, and non-transparency is what is invoked in most work seeking correlations between language complexity and sociolinguistic factors (e.g., Dahl 2004; Trudgill 2011). For this reason, though labor-intensive to survey, it is probably more valuable in the long run.

Most measures of informational complexity are computationally based (e.g., Juola 1998; Bentz 2018; Ehret & Szmrecsanyi 2016) and require large corpora. I have proposed a grammar-survey measure (Nichols 2020) making use of the fact that the needed information is minimal in a system that follows biuniqueness, the one-form-one-function ideal of structural analysis and canonicity theory (Corbett 2007, 2015, and several others), and increases as the imbalance between forms and functions increases. Thus, the Russian example just cited, with three forms for the one function of marking the accusative case, is more complex than a comparable segment of Saami case inflection: Aikio and Ylikoski (2010: 39) describe two sets of the case endings for respectively consonant-final and vowel-final stems of North Saami, but these are phonologically predictable variants and not separate declension classes. (For a comparison of Russian to Mongolian, which like Saami has a single noun declension class with phonologically predictable variants, see Nichols 2020.) Since departures from one-form-one-function are non-canonical in the approach of Corbett and others, this complexity measure can be called *non-canonicity-based complexity*, or *canonical complexity* for short.

Keeping a survey of canonical complexity manageable requires limiting the morphology surveyed to a sample of the grammar. I survey noun inflection marking the core relations of subject, direct object, indirect object, and possessor, for singular and plural numbers only; personal pronoun inflection for the same roles, again for singular and plural only; and verb inflection for argument indexation (three persons, singular and plural only) and TAM in the simplest present tense and the simplest past tense (the perfective-like one if the language has more than one). For all I consider default valence or argument marking only: subjects in the strict traditional morphosyntactic sense, which in Uralic languages means nominative case and verb agreement; direct objects in the strict sense, i.e., accusative in many languages, and not

the various oblique object cases found, e.g., with verbs of emotion or perception in many languages; and so on. Among these parts of the grammar the survey counts the number of non-biunique patterns such as the number of declension classes (different sets of endings), the number of conjugation classes (likewise), the number of different stems per lexeme, the number of different stem classes in the language, the numbers of all of those that are unpredictable, and numbers of such things as syncretisms, suppletion, allomorphy, portmanteaus, and zero markers other than nominative singular.

2. Complexity levels in Saami languages

2.1. Inventory complexity

The measure of inventory complexity used here (and also used in Nichols 2019; the morphological part is used in Nichols & Bentz 2019) counts these things: in the phoneme inventory, the number of contrastive consonant types and vowel qualities; presence or absence of tones and phonation types; and syllable complexity measured as the maximum number of consonants per syllable; verb inflectional synthesis (number of categories per word) following Bickel & Nichols (2013); in noun inflection, the presence vs. absence of gender, numeral classifiers, number, and possessive classes; in syntax, the number of basic alignments and basic word orders (further measures for syntax, and measures of derivational morphology, are under development). The results for the Saami languages surveyed are shown in Table 1 below.²

As these results show, Saami languages vary somewhat in their complexity, and on the whole they are similar to the worldwide average and somewhat lower for morphological complexity. Their morphological inventory complexity includes, for nouns and pronouns, only case and number; also possessor person-number for nouns; and for verbs, one argument, TAM, and voice. Possessive person-number inflection is counted as absent in most Saami languages; where remnants of possessive inflection are preserved they appear to be fossils frozen for use only on certain kin terms, and not a regular inflectional category.³

² The full database for this paper will be published in an open repository, by approximately the end of 2025.

³ Grammars give the impression that it is regular on kin terms and some body parts for Kola Saami languages – which would then have inflectional possessive morphology and also an alienability opposition, following the cross-linguistically common pattern of head marking on kin terms and/or body parts and dependent marking (genitive case of possessor) for others. However, I did not count this possessive marking as inflection. Janda and Antonsen 2016 show that loss of possessive inflection is well underway in North Saami but do not claim that it is

Table 1. Enumerative complexity in Saami languages, separately tabulated for all metrics and only morphological ones. The lower portion shows mean enumerative complexity for those Saami languages, for other Uralic languages, and for all languages in my database (N = 245).

<i>Language</i>	<i>All</i>	<i>Morphology</i>
South	18	7
Pite	20	8
Lule	18	7
North	18	8
Skolt	20.5	9
Kildin	21.5	9
Saami mean	19.3	8.0
Other Uralic mean	17.7	8.7
World mean	19.1	9.8

2.2. Canonical complexity

Canonical complexity in inflectional morphology, for the three parts of speech surveyed here, essentially boils down to two targets. One is the number of different inflectional classes (i.e., sets of endings) for singular inflection of nouns, singular and plural inflection of personal pronouns, and person-number inflection in the two basic verb tenses surveyed here; and the number of these that are unpredictable, i.e., lexically specified and arbitrary. For each of the three parts of speech there is generally just one inflectional class in the Saami languages, so there is no unpredictability. (Note that phonologically predictable allomorphy is not counted here, as it is automatic.) The other survey target is the number of different stem forms and their predictability, and here Saami languages have more complexity, often a great deal. For all but South Saami, there are two or three dimensions of stem complexity in inflection: patterns of consonant gradation, patterns of vowel alternation (or ablaut), and, for the Kola languages, alternation of plain and palatalized consonants. Each set of combined consonant and vowel changes defines a stem paradigm or class; for example, in Kildin Saami, basic strong grade shifts to weak in the genitive and accusative singular, and this occurs both with and

absent. Aikio & Ylikoski (2010): 70ff. describe it as present. Kert (1971) gives full paradigms for Kildin Saami without comment, but Rießler (2022: 226) describes possessives in today’s speech, about two generations later, as lost except for remnants. Wilbur (2014: 108) describes them as “nearly fallen out of use” for Pite. Without frequency counts for the other languages I could only treat this as a yes/no variable and, assuming that the situation in these two smaller, more endangered ones is typical, I coded all the varieties as lacking possessive marking.

without vowel alternations, a total of two paradigm types with gradation and two without. (I provisionally lump vowel alternations into a single broad category coded as present vs. absent, which is probably an oversimplification.) (For more discussion and examples see §2.3.) Some grammars lay out the full inventory of paradigm types (e.g., Kert 1971 for Kildin Saami; Morottaja & Olthuis 2020 for Inari; Spiik 1989 for Lule; Wilbur 2015 for Pite). For Skolt Saami, Feist (2015) presents most of the inventory and notes that it is probably not complete. For Akkala, Zajkov (1987) presents the attested alternation patterns but not the possible paradigm types; the impression gained is that there is only one gradation paradigm for nouns and all nouns have gradation, but this is not stated. Tereškin (2002) describes a single gradation pattern and two palatalization patterns (which appear to be in complementary distribution) for nouns, giving the impression that this is the full inventory.

Many grammars list separately the alternations found in vowel-final and consonant-final stems, and/or stems with even vs. odd numbers of syllables. In such cases I took the total number of different patterns found among the stem shapes. (Usually, one stem shape exhibits all types and others have only a subset of those.)

In such counts it is important to distinguish between the *patterns*, on the one hand, and on the other the *factors delimiting their distribution*, or *observations that can be made on their distribution* (henceforth *distributional observations*). The patterns are the alternation types (the gradation and ablaut patterns) as they occur in different cases or person-number forms; these are what is counted here. The distributional observations are the number of syllables and whether the stem ends in a consonant or a vowel. These are not alternation patterns and not predictive or causal factors but grammatical observations about which patterns occur where. Of course, the distributional factors are descended from contextual properties that historically conditioned the alternation in the first place, i.e., they *were* conditioning factors. They may be important for language teaching and pedagogical presentations, and covering them is essential if a grammar is to be useful to historical linguistics, but they are not causal now and are not what is measured in this study.⁴

In most modern Saami languages, gradation and vowel patterns are not predictable (e.g., from the phonological environment or grammatical categories). They can be described as morphologized, or as lexically or morphologically specified. Exceptions are South Saami, which has no gradation, and North Saami, where there may be some phonological

⁴ They could be used in a differently designed survey of complexity, e.g., counting the number of different constraints on the distribution of gradation patterns in different languages. By that token Akkala Saami would be maximally simple: at least for verbs, syllable count has no role in the distribution of alternation types (Zajkov 1987: 144), while in the others it does.

predictability of gradation. (The description of North Saami in Aikio & Ylikoski (2014: §4.2) gives the impression that vowel-final and consonant-final stems are predictable from syllable count and may predict some alternation patterns, but these factors are probably best viewed as distributional observations and not as conditioning. Nickel & Sammallahhti (2011: 70–71) use two gradation classes and two vowel alternation classes.)

Consider now the number of stem classes in Saamic vs. Finnish noun inflection. In Saamic they range from one in South Saami (where consonant gradation is absent) to many in Skolt or Pite. In Finnish, noun stems vary due to consonant gradation, but in Finnish the gradation is almost always predictable from the syllable structure: weak grade before a closed syllable, strong before an open syllable.⁵ Vowel harmony in Finnish is phonologically predictable, while the vowel alternations of Saamic are not.

Thus, while Finnish and most Saami languages both have consonant gradation and vowel harmony or alternation, producing allomorphy in both languages, most Saami languages have high canonical complexity because the variation is not predictable, while Finnish canonical complexity is low. Table 2 compares Saami to Finnish, Russian, and English, and Table 3 gives the numbers of stem classes for nouns in the Saami languages. Table 4 gives canonical complexity levels for all the Saami languages surveyed, and Table 5 gives mean complexity levels for Saami, the rest of Uralic, and all the languages in my database.

Table 2. Complexity of noun inflection in Saami and selected other languages.

	<i>Saami</i>	<i>Finnish</i>	<i>Russian</i>	<i>English</i>
Ending classes	1	1	4 *	2 ***
Unpredictable	–	–	4	2
Stem classes	1–11	1	1 **	2 ***
Unpredictable	All	–	–	2

* The ending sets now increasingly called “masculine”, “neuter”, and “feminine” plus the traditional “third declension” (now fourth; words like *set’* ‘net’, *put’* ‘path’).

** Stress alternations are not counted.

⁵ Predictable, that is, provided the transcription is sufficiently abstract to account for stem shapes such as *tiede*, oblique *tietee-* ‘science’, where the stem-final syllable is open in both at the surface level but not underlyingly: {tie.teC} *tiede* vs. {tie.te.Ce-} *tietee-*, where C is an abstract consonant that has abstract zero as its weak grade but phonological zero, abstract non-zero unspecified consonant as its strong grade.

*** Nouns like *cat, dog, house* vs. ones like *tooth, mouse*. The second set have a zero plural suffix and a vowel change (*teeth* vs. *cats*).

Table 3. Noun stem classes in Saami languages. For what are sometimes called class suffixes or class markers, see §2.3.

<i>Language</i>	<i>Classes</i>	<i>Source</i>	<i>Comment</i>
Ter	1	Tereškin 2002	
Kildin	11	Kert 1971: 143	
Skolt	10	Feist 2015: 158	6 large classes
Akkala	2 or 3	Zajkov 1987	1 gradation, 2 vowel (+ non-gradation?)
North	4	Aikio & Ylikoski 2010: §4.2	(see text)
Lule	6 or 8	Spiik 1989	
Pite	4 or 5	Wilbur 2014: 98ff.	
South	1	Bergsland 1994	

Table 4. Canonical complexity in Saami languages. All = all measures. Strict = removing any (such as the categories involved) that might be considered inventory complexity.

<i>Language</i>	<i>All</i>	<i>Strict</i>
South	24	15
Pite	62	45
Lule	57	37
North	45	30
Inari	54	37
Skolt	57	41
Kildin	46	33
Mean	49.3	34.0

Table 5. Mean canonical complexity for Saami and other languages. World = all 116 languages surveyed so far. (New coding principles have been implemented recently, but most of the non-Uralic languages still follow the older principles, so the World total shown here is approximate.)

	<i>All</i>	<i>Strict</i>
Saami	49.3	34.0
Other Uralic	25.9	17.2
World	32.4	26.6

2.3. Problems of grammatical analysis

Only seven Saami languages could be fully surveyed for Tables 3–5. Akkala and Ter were surveyed, but available grammars do not give all the information needed. For most of those that are included, uncertainties remain. Many grammars can be described as comparative-historical in their primary orientation: the presentation of inflectional paradigms focuses on syllable count and stem-final vowel vs. consonant, the factors that historically triggered consonant gradation and some vowel alternations and are necessary for reconstructing the historical evolution of Saami. These are essential in any Saami description, and are required for comparative-historical work, but they are not the full picture. Synchronically, these stem shape properties are the distributional observations on stem classes. Not all grammars enumerate and exemplify stem classes in the strictly synchronic general-linguistic sense used here, and none make clear whether they are contrastive or in complementary distribution.

A case in point is what are often called “class suffixes” or “classification” or “class markers” in Saami noun inflection. These are a set of vowels that appear in vowel-final stems. Which stem takes which vowel is unpredictable, but each stem has only one vowel (which undergoes automatic phonological variation in some contexts in some languages). They are generally found throughout the inflectional paradigm. They are found in at least some derived forms (coverage of derivation in most grammars is not full, so this is not certain). Many are consistent in Saami cognate sets, and often appear to be inherited from Proto-Uralic.⁶ The grammars I have read show that they are identifiable but not that they are morphologically segmentable.⁷ I provisionally analyze them as stem-final or root-final elements, not separate morphemes, so they do not contribute to inflectional complexity.

Another is the number of manners of obstruent articulation. Most grammars list only two, voiceless and voiced, implicitly treating long or geminate consonants as two-consonant sequences. However, long consonants figure in consonant gradation as strong counterparts to singleton weak grades (e.g., Lule 1sg *guláv*, 3sg *gullá* ‘hear’; Spiik 1989: 27, 31), and while they can arguably be syllabified as C.C in intervocalic position (e.g., *gul.lá*), a syllable

⁶ Proto-Uralic lexical roots are usually CV(C)CV in shape (Aikio 2021: §1.3.4). An example of a *CVCV word with such a vowel is PU **kala* ‘fish’, Saami **kōlē*, e.g., S Saami *guōlie*, N Saami *guolle*, Kildin *kū’ll* (Lehtiranta 2001: set 511).

⁷ For instance, in Pite Saami (Wilbur 2014: 98–108) there are three noun classes: one with a final vowel, one with final -Vj, and one consonant-final. The final vowels are visible in the nominative singular, distinguishable from each other there and in their morphophonemic interaction with inflectional suffixes, and distinguishable from -Vj or lack of a final vowel. Inflection and derivation use the stem with its class suffix as base, so the class suffix behaves as though it were an unsegmentable root-final element rather than a separate morpheme.

boundary internal to the geminate is impossible when they occur word-finally (e.g., Kildin Nsg *vuedd* ‘ground’, Rießler 2022: 225, Table 13.3) and in some internal clusters (Akkala loc.sg. *puaccakas* ‘part of reindeer harness’, surely {*puacc.kas*}) (Zajkov 1987: 69).

These three points – the full list of gradation and vowel-alternation classes, whether class suffixes are demonstrably morphemes rather than stem-final vowels, and the phonological status of long consonants – are usually not all covered fully in Saami grammars and for the present survey had to be extracted or inferred from partial or scattered description and inspecting what examples are given. To construct a single consistent framework I use not holistic classes but a componential analysis based on which gradation patterns, which vowel alternation patterns, and which (if any) other factors do and do not cooccur. The following gives, as a sample, the decisions behind the numbers of stem classes for noun inflection in three of the languages surveyed.

For Kildin Saami, Kert (1971) lists 9 basic declension classes, plus two more for longer words. The classes are based on consonant gradation, consonant palatalization, and vowel alternation. For the nine basic classes there is one pattern of consonant gradation in the singular paradigm: strong in the nominative, essive, and dative, weak elsewhere.⁸ (*Weak* and *strong* are standard terms based on alternations and not on phonetics. The strong grade, found inter alia in the nominative singular of nouns, can be taken as base or underlying; it undergoes lenition-like changes in the weak grade contexts. The output is predictable given the underlying consonant. Any given phoneme may be either weak or strong depending on the underlying consonant: for example, single voiceless stops are strong grade that becomes voiced and/or continuant in the weak grade, and they are the weak grade of long or geminate stops.) Vowel-final stems and stems with three or more syllables lack gradation. Since I am treating syllable count and type of final segment as distributional observations, I lump all stems together and analyze Kildin as having two gradation classes: gradation and no gradation. Which words have gradation is an arbitrary lexical property, and in words that have it the distribution of gradation among the cases is absolutely consistent and morphomic rather than predictable.

Orthogonally to consonant gradation, stems either do or do not have vowel alternation, for which there is again one pattern: a change of vowel quality in the dative singular. Which vowel appears in the dative is predictable from the vowel of the other stems, but whether that vowel change occurs in the dative is not predictable.

⁸ Here and below, I use the term dative for the case variously called dative, dative-illative, and illative in different grammars.

The canonical complexity count for noun stem classes has three entries: maximum number of stems per lexeme, number of stem classes per language, and number of those stem classes that are unpredictable. The inflectional paradigms surveyed for nouns are the singular of the cases for A, S, O, G, and T arguments, as well as possessor: nominative, accusative, genitive, dative. For Kildin Saami, in those four cases there are up to three possible stem forms, schematically shown in Table 6:

Table 6. Kildin Saami noun stem gradation and ablaut, and two examples (Kert 1971: 141–142). /*ɲ*, /*t̪*/ = palatalized consonant.

	Consonant gradation	Vowel ablaut	‘fish’	‘tree’
NOM	Strong	basic	<i>kuɲɲ</i>	<i>murr</i>
ACC	Weak	basic	<i>kuɲ</i>	<i>mur</i>
GEN	Weak	basic	<i>kuɲ</i>	<i>mur</i>
DAT	Strong	changed	<i>kuell-a</i>	<i>muɲɲ-e</i>

There are four stem classes in the language: ± gradation and ± ablaut. Lack of gradation occurs only with consonant-final stems and trisyllabic or longer ones, but these are contextual and not predictive factors. Therefore, all four stem classes are unpredictable. Thus, noun stems contribute 11 complexity points to the Kildin total: 3 for stems per lexeme and 4 for both classes per language and unpredictability of those classes.

Lule Saami has different patterns and interactions:

Gradation, strong > weak	(1 pattern)
Gradation, weak > strong; ± ablaut	(2 patterns)
No gradation	(1 pattern)
Total	4 stem classes, all unpredictable

Skolt Saami has more patterns and factors: two gradation patterns (nominative strong, genitive-accusative weak, as in Kildin; nominative weak, others strong), of which the former has ± stem-final vowel loss and the latter has ± palatalization in the nominative singular. Disyllables contrast in whether it is the medial or the final consonant that undergoes gradation and whether there is gradation or not. Feist sets up 7 stem classes (plus others specific to particular derivational suffixes, not counted here). I count the following:

Strong > weak ± stem-final vowel loss,	
± C2 gradation, ± Dat.sg. strong+	4 patterns
Weak > strong ± palatalization in Nom.sg.	2 patterns

No gradation
Total

1 pattern
7 classes

Not all are found in words with all syllable counts, but that is what I call a contextual factor. C2 gradation has to do with the fact that in one of the classes of Feist (2015) it is the final rather than the second consonant or sequence that undergoes gradation. The dative (Feist: illative) has a distinct super-strong grade (strong+) in some nouns.

3. Correlations of complexity with non-linguistic factors

I looked for correlations of complexity with three non-linguistic factors that are easily measured or looked up and which figure in many discussions of linguistic complexity as sociolinguistic factors fostering increased complexity.

(1) Latitude, which also has implications for climate (colder and drier to the north) and ecology (forest in the south, tundra in the north), which determine resource richness and directly impact population size, carrying capacity of the land (range size per capita needed for survival), and prosperity in general. Other things being equal, richer resources and higher prosperity attract others to identifying with the group and adopting its language. The firmly reconstructed prehistory of Saami involves its appearance in Scandinavia beginning perhaps 1500 years ago, followed by a Saami spread to cover all the known historical territory, and absorption of the previous inhabitants of that area into the Saami ethnic and speech communities (Aikio 2012, 2022). More recently, the introduction of reindeer herding made possible large herds and production of surplus on tundra land, the northernmost ecology, so the high latitudes became more prosperous.

(2) Population. Various figures on population size are available, of which I mainly used Krauss 1997. I used the figures not for number of speakers of the language but for overall population size (speakers and non-speakers), since the overall size is a better reflection of the group's size before the speedup of language shift and extinctions in the last several decades. I measured the size not as total or rounded total figures but as the magnitude of the population: hundreds, thousands, tens of thousands. Krauss's figures generally reflect the situation in the mid to late 20th century.

(3) Range. I estimated range sizes by inspecting maps (Rantanen et al. 2022), taking a range the size of Ume or Pite as the basic unit, and estimating by eye the sizes of the other ranges as fractions or multiples of those. This measure is obviously Saami-specific and requires good language and dialect

maps, and it is very approximate, but it is quite adequate as a quick-and-dirty measure.

Table 7. Complexity means and correlations with non-linguistic factors. Canonical complexity for the world sample is based on an earlier coding system and are not exact. Range size estimates are available only for Saami. Population figures for the world sample are not available (see text). Bold = statistically significant ($p \leq .05$); italics = fairly close to significant.

	Enumerative:		Canonical:		
	All	Morph. All	Strict		
Mean complexity:					
Saami mean	19.3	8.0	49.3	34.0	
Other Uralic mean	17.7	8.7	25.9	17.2	
World mean	19.1	9.9	32.4	26.6	
Correlation coefficients:					Expected:
Saami (N=6 (enumerative) or 7 (canonical); 0.73 or 0.67 required for $p = 0.05$).					
Latitude	0.216	0.520	0.702	0.664	+
Population	-0.622	-0.267	-0.111	-0.155	-
Range	-0.485	-0.078	-0.542	-0.552	-
Other Uralic (N=9 (enumerative) or 12 (canonical); 0.58 or 0.50 required)					
Latitude	0.346	0.627	0.394	0.353	+
Population	-0.186	-0.580	-0.346	-0.270	-
World (N = 245 (enumerative) or 117 (canonical); 0.10 or 0.12 required)					
Latitude	0.153	-0.068	0.197	0.171	+

Correlations of complexity levels with these factors, for Saami and other languages, are shown in Table 7 above. Population figures for the entire database have not been compiled yet.⁹ A quick and easy calculation of relative

⁹ Nichols (2009) found a highly significant negative correlation worldwide between population size and an earlier and less complete version of enumerative complexity, using a smaller language sample, but showed that it was an artifact of population size differences between very

range sizes like that used here for Saami is problematic for the rest of Uralic and not feasible for the whole world, so correlations with range are not calculated for the other languages. There are two relevant measures of adequacy: statistical significance, hard to reach with a small sample but nonetheless reached for canonical complexity; and whether correlation coefficients are positive or negative as expected. Expected correlations are as shown in the table.

Latitude: A positive correlation is expected (higher latitude = higher complexity), since higher latitude generally means greater distance from centers of spread and economic and cultural centers. (Bentz 2016 finds a worldwide correlation between low latitude and low complexity.) Positive correlations with latitude were found within Saami, significant for canonical complexity, appreciable for morphological enumerative complexity, but very weak for overall enumerative complexity. The largest difference between overall and morphological enumerative complexity is that measures of phonological complexity were included in the overall figure, and this suggests that phonological complexity does not correlate positively with latitude. I did a separate count of phonological complexity, using the more elaborated measure of Hartmann & Nichols in press, and here the correlation was negative (-0.381) but weak, confirming that there is no correlation between latitude and phonological complexity. The correlation for only morphological enumerative complexity is more similar to those for canonical complexity (unsurprisingly, as the measures of canonical complexity include only morphology).

Population: A negative correlation is expected (smaller population = higher complexity). The linguistic literature finds mixed results on whether such a correlation exists worldwide (pro: e.g., Hay & Bauer 2007 for phonological complexity; con: Nichols 2009 for overall enumerative complexity [see again footnote 8]; Donohue & Nichols 2011 for phoneme inventory size), but it continues to be a good hypothesis, since small populations are often sociolinguistically isolated and have not recently absorbed appreciable numbers of adult L2 learners whose presence can decomplexify a language (for this effect see below). Mostly negative correlations were found as expected (close to significant for overall enumerative complexity; others negligible).

large areas: mean population sizes are much larger in Eurasia and Africa than in the Pacific or the Americas, and mean enumerative complexity levels are higher in the Pacific and (especially) the Americas. That imbalance accounted for the negative correlation; there was no correlation within any of the three large areas. All of those findings would undoubtedly be true with the current coding and sample as well.

Range size: Negative correlations are expected, for the same reasons as for correlation with population magnitude. Negative correlations were found, most of them appreciable.

These factors are of interest not in themselves but because of the sociolinguistic and economic factors that are likely to have caused them. Those include such things as known language expansions and whether they primarily involved language shift (and consequent absorption of adult L2 learners); known relative economic strength of different groups; number and nature of contact situations; and others. No attempt has been made to measure such factors here, since collecting the data would be a major research project in itself and would require expertise in Saami ethnography and history. A known relevant episode is the relatively recent expansion of North Saami with the introduction of reindeer herding involving language shift (of other Saami populations to North Saami) and territorial expansion to utilize the tundra. This would have meant absorption of adult L2 learners, which can decomplexify a language as noted above, and in fact North Saami is low in both enumerative and canonical complexity compared to the other Saami varieties. South Saami is even less complex in all measures, possibly because of its historical sociolinguistic status but most likely because it diverged first from the rest of Saami (Aikio 2012), before consonant gradation arose.

Principal findings for the three non-linguistic factors are these. With one exception, correlations were positive or negative as expected. The correlations of complexity with latitude and population magnitude found for Uralic are somewhat different for the rest of Uralic, but appreciable, and significant for morphological enumerative complexity. For the whole world, correlations with latitude are highly significant except for morphological enumerative complexity (which is furthermore the only case of sign reversal in these correlations).

4. Discussion

The positive correlation of complexity with latitude found for Saami finds support in the rest of Uralic and strong support worldwide. The negative correlation with population magnitude, for enumerative complexity only, is partly supported in the rest of Uralic. All in all, the other findings give plausibility, and sometimes support, to what was found for Saami and make it worthwhile to continue interpreting them. It is reasonably safe to generalize that grammatical complexity in Saami increases at higher latitudes, with smaller populations, and with smaller range sizes. All three are likely correlates of a sociolinguistic situation similar to what Trudgill (2011) terms

sociolinguistic isolation: no recent history of absorbing speakers of other languages. Absorbing adult L2 speakers generally decomplexifies a language, but otherwise complexity gradually builds up in intergenerational transmission. Languages which do not absorb adult learners are under no pressure to decomplexify, and can theoretically complexify without limit (though a situation with zero shift to the language and nonzero shift from it means that extinction is inevitable, putting a practical limit to the complexity). Note that sociolinguistic isolation does not mean isolation in the sense of minimal contacts; sociolinguistically isolated languages are often in multiple close contacts but not as dominant or expanding languages. (This is true of the eastern Caucasus, for instance.) They seem to often be in locations that are peripheral to centers of spread, commerce, and population growth. This general picture describes Saami languages, spoken by fairly mobile populations that were in contact with other languages, Saami and non-Saami, but in the ultimate peripheral location at the northern edge of the habitable world.

Ordinarily, findings can be firmed up by expanding the sample, but in the case of Saami the remaining languages – Ume, Akkala, and Ter – have only partial descriptions and could not be surveyed fully enough to include them here. It is easy to see that Akkala and Ter, if included, would upset some of the patterns established above. Though they are far northern languages they have very small range sizes for their latitude, and their gradation systems and phonological inventories appear to be relatively simple compared to those of adjacent Skolt and Kildin (and their phoneme inventories no more complex). Their populations, on the other hand, are very small as predicted for their latitude. It is possible that some of the non-complexity is due to incomplete enumeration of gradation classes in the existing descriptions. Simplification can also be a language death symptom, and that seems more likely in the case of Akkala and Ter: both seem likely to have been losing speakers to their neighbors, and both had their economic institutions and community centers destroyed during collectivization and much of their population scattered (in the case of the hunter-fisher Akkala, scattered to towns or herding collectives, putting them in an unfamiliar social and economic situation); the extent of these losses appears to have been greater than for the Kildin and Skolt. Perhaps comparative documentation of factors like these could be assembled for more Saami groups to improve the qualitative nature of the correlations.

Extensive ethnographic, demographic, historical, and archaeological work has been done on such matters as Saami migration routes, whether they were for transhumant herding, hunting and trapping, or fishing, and their seasonal

schedules, geography, and distance covered. Such things could be brought together to give a much richer and more precise description of range size.

5. Conclusions

The unified descriptive framework for assessing Saami canonical complexity includes these points:

- Inflectional paradigms consist of stem, ending, and various distributional observations that delimit distributions by stem shape, syllable count, etc. but not do not predict either stem alternations or form of ending. For the most part, the stem alternations of consonant gradation and vowel ablaut are lexically specified and not predictable.

- Stem classes are defined by consonant gradation and vowel ablaut. The gradation and ablaut classes are defined abstractly: consonant grades are strong vs. weak; for lack of any standard terminology for vowel grades I have called them base (found in the nominative singular) vs. changed. A pattern is a set of consonant or ablaut grades stating which grade appears in which paradigmatic position (which case form, etc.).

- Distributional observations include the syllable count, whether the stem is consonant-final or vowel-final, and whether there is a stem extension in some of the forms.

- Ending classes are defined by same vs. different sets of endings. In Saami languages there is only one ending class for nouns and usually only one for verbs. Allomorphy in endings is usually phonologically conditioned, i.e., predictable.

- Canonical complexity is defined as the number of stem classes, the number of ending classes, and the number of different patterns of syncretism, allomorphy, zero endings (other than nominative), etc.

The main findings on the distribution and correlations of complexity in Saami are these:

- Enumerative complexity in Saami is slightly higher than in other Uralic, but unremarkable overall.

- Canonical complexity is high overall in Saami. It is due to morphologization of what were originally allophonic or subphonemic alternations. It is primarily a matter of stem alternations; endings are simple and regular.

- It is easier to identify the lowest-complexity Saami languages than the highest-complexity ones: South and North Saami have markedly lower complexity. For North Saami this probably reflects expansion to absorb other Saami speakers and, for recent years, standardization and use in media. For

South Saami sociolinguistic factors must be relevant, such as serving as contact frontier with North Germanic from early times on; but the early separation of South Saami from the rest must have meant that it simply did not undergo the post-Proto-Saami sound changes that eventually resulted in gradation and ablaut alternations.

- Complexity correlates positively with more northern latitude and smaller range and/or population. These three factors are interconnected and point to a sociolinguistic situation involving minimal pressure to decomplexify and transmission mostly by childhood L1 acquisition and infrequently by adult L2 acquisition, a situation associated with small and peripheral languages.

A final observation is that the analysis of linguistic complexity requires good grammatical descriptions, good understanding of the terminology and conventions of the descriptive tradition for those grammars, and knowledge of the non-linguistic factors that may explain the distribution of complexity levels across languages. Good descriptions in traditional and structural terms are available for many Saami languages, but there is still a need for comprehensive, typologically oriented reference grammars of as many of the languages as possible. I can vouch for the soundness and tested usefulness of the definitions and coding system used here, but good hypotheses and good explanations will require expertise in Saami linguistics, sociolinguistics, history, ethnography, and probably other fields, and for best results they should be initiated by Saami specialists themselves.

References

- Aikio, Ante = Luobbal Sámmol Sámmol Ánte. 2012. An essay on Saami ethnolinguistic prehistory. In Riho Grünthal & Petri Kallio (eds), *A Linguistic Map of Prehistoric Northern Europe* (Suomalais-Ugrilaisen Seuran Toimituksia 266), 63–117. Helsinki: Société Finno-Ougrienne.
- Aikio, Ante = Luobbal Sámmol Sámmol Ánte. 2022. Proto-Uralic. In Marianne Bákro-Nagy, Johanna Laakso & Elena Skribnik (eds), *The Oxford guide to the Uralic languages*, 3–27. Oxford: Oxford University Press.
- Aikio, Ante & Jussi Ylikoski. 2010. *The structure of North Saami*. Course handout.
- Anderson, Stephen R. 2015. Dimensions of morphological complexity. In Matthew Baerman, Dunstan Brown & Greville G. Corbett (eds), *Understanding and measuring morphological complexity*, 11–26. Oxford: Oxford University Press.
- Audring, Jenny. 2017. Calibrating complexity: How complex is a gender system? *Language Sciences* 60. 53–68.
- Bentz, Christian. 2018. *Adaptive languages: An information-theoretic account of linguistic diversity* (Trends in Linguistics. Studies and Monographs. Volume 316). Berlin: de Gruyter.

- Bentz, Christian. 2016. The low-complexity belt: Evidence for large-scale language contact in human prehistory? *The Evolution of Language: Proceedings of the 11th International Conference (EVLANG 11)*.
https://evolang.org/neworleans/pdf/EVLANG_11_paper_93.pdf
- Bergsland, Knut. 1994. *Sydsamisk grammatikk*. Karasjok: Davvi Girji.
- Bickel, Balthasar & Johanna Nichols. 2013. Inflectional synthesis of the verb. In Matthew S. Dryer & Martin Haspelmath (eds) *WALS Online* (v2020.4) [Data set]. Zenodo. <https://doi.org/10.5281/zenodo.13950591>
- Corbett, Greville. 2007. Canonical typology, suppletion, and possible words. *Language* 83(1). 8–42.
- Corbett, Greville. 2015. Morphosyntactic complexity: A typology of lexical splits. *Language* 91(1). 145–193.
- Dahl, Östen. 2004. *The growth and maintenance of linguistic complexity* (Studies in Linguistic Companion Series 71). Amsterdam: Benjamins.
- Donohue, Mark & Johanna Nichols. 2011. Does phoneme inventory size correlate with population size? *Linguistic Typology* 15(2). 161–170.
- Ehret, Katharina & Benedikt Szmrecsanyi. 2016. An information-theoretic approach to assess linguistic complexity. In Raffaella Baechler & Guido Seiler (eds), *Complexity, Isolation, and Variation* (linguae & litterae, volume 57), 71–94. Berlin: de Gruyter Mouton.
- Feist, Timothy. 2015. *A grammar of Skolt Saami* (Suomalais-Ugrilaisen Seuran Toimituksia 273). Helsinki: Suomalais-Ugrilainen Seura.
- Hartmann, Frederik & Johanna Nichols. (In press.) Geospatial effects on phonological complexity in the world's languages. *Linguistic Typology*.
- Hay, Jennifer & Laurie Bauer. 2007. Phoneme inventory size and population size. *Language* 83(2). 388–400.
- Janda, Laura A. & Lene Antonsen. 2016. The ongoing eclipse of possessive suffixes in North Saami: A case study in reduction of morphological complexity. *Diachronica* 33(3). 330–366.
- Juola, Patrick. 1998. Measuring linguistic complexity: The morphological tier. *Journal of Quantitative Linguistics* 5. 206–213.
- Kert, Georgij M. 1971. *Saamskij jazyk (kil'dinskij dialekt)*. *Fonetika, morfologija, sintaksis*. Leningrad: Nauka.
- Krauss, Michael E. 1997. The indigenous languages of the North: A report on their present state. In Hiroshi Shoji & J. Janhunen (eds), *Northern minority languages: Problems of survival*, 1–34. Osaka and Fairbanks: National Museum of Ethnology and Alaska Native Language Center, University of Alaska Fairbanks.
- Kusters, Wouter. 2008. Complexity in linguistic theory, language learning, and language change. In Matti Miestamo, Kaius Sinnemäki & Fred Karlsson (eds), *Language complexity: Typology, contact, change* (Studies in Language Companion Series, 94), 3–22. Amsterdam: Benjamins.
- Lehtiranta, Juhani. 2001. *Yhteissaamelainen sanasto* (Suomalais-Ugrilaisen Seuran Toimituksia 200). Helsinki: Suomalais-Ugrilainen Seura.
- Miestamo, Matti. 2008. Grammatical complexity in cross-linguistic perspective. In Matti Miestamo, Kaius Sinnemäki & Fred Karlsson (eds), *Language complexity: Typology, contact, change* (Studies in Language Companion Series, 94), 23–41. Amsterdam: Benjamins.

- Morottaja, Matti & Marja-Liisa Olthuis. 2020. *Inarinsaamen koulukielioppi*. Aanaar/Inari: Säämi kulttuurkuávdáš Sajos.
- Nichols, Johanna. 2020. Canonical complexity. In Peter Arkadiev & Francesco Gardani (eds), *The complexities of morphology*, 163–192. Oxford: Oxford University Press.
- Nichols, Johanna. 2019. Why is gender so complex? New typological considerations. In Francesca Di Garbo, Brunol Olsson & Bernhard Wälchli (eds), *Grammatical gender and linguistic complexity. Volume 1: General issues and specific studies*, 63–92. Berlin: Language Sciences Press.
- Nichols, Johanna. 2009. Linguistic complexity: A comprehensive definition and survey. In Geoffrey Sampson, David Gil & Peter Trudgill (eds), *Language Complexity as an Evolving Variable*, 110–125. Oxford: Oxford University Press.
- Nichols, Johanna & Christian Bentz. 2019. Morphological complexity of languages reflects the settlement history of the Americas. In Katerina Harvati, Gerhard Jäger & Hugo Reyes-Centano (eds), *New Perspectives on the Peopling of the Americas*, 13–26. Tübingen: Kerns.
- Nickel, Klaus Peter & Pekka Sammallahti. 2011. *Nordsamisk grammatikk*. Davvi Girji.
- Rantanen, Timo, Harri Tolvanen, Meeli Roose, Jussi Ylikoski & Outi Vesakoski. 2022. Best practices for spatial language data harmonization, sharing and map creation: A case study of Uralic. *PLoS ONE* 17(6). e0269648. <https://doi.org/10.1371/journal.pone.0269648>
- Rießler, Michael. 2022. Kildin Saami. In Marianne Bákro-Nagy, Johanna Laakso & Elena Skribnik (eds), *The Oxford guide to the Uralic languages*, 219–239. Oxford: Oxford University Press.
- Ross, Malcolm D. 1996. Contact-induced change and the comparative method: Cases from Papua New Guinea. In Mark Durie & Malcolm D. Ross (eds), *The comparative method reviewed*, 180–217. New York: Oxford University Press.
- Sinnemäki, Kaius. 2011. *Language universals and linguistic complexity: three case studies in core argument marking*. PhD dissertation, Helsinki: Department of Modern Languages.
- Sokur, Elena & Johanna Nichols. 2019. *Noun vs. verb inflectional synthesis: A complexity tradeoff?* Presented at SLE 2019 meeting, Leipzig.
- Spiik, Nils Eric. 1989. *Lulesamisk grammatik*. Luleå: Sameskolsyresen.
- Tereškin, Sergej N. 2002. *Jokan'gskij dialekt saamskogo jazyka*. PhD dissertation. Sankt-Peterburg: Rossijskij gos. pedagogičeskij universitet im. A. I. Gercena.
- Thurston, William R. 1989. How exoteric languages build a lexicon: Esoterogeny in West New Britain. In Ray Harlow & Robin Hooper (eds), *VICAL I – Oceanic languages: Papers from the Fifth International Conference on Austronesian Linguistics*, 555–579. Auckland: Linguistic Society of New Zealand.
- Trudgill, Peter. 2011. *Sociolinguistic typology: Social determinants of linguistic structure and complexity*. Oxford: Oxford University Press.
- Wilbur, Joshua. 2014. *A grammar of Pite Saami* (Studies in Diversity Linguistics, No 5). Berlin: Language Science Press.
- Zajkov, Pekka M. 1987. *Babinskij dialekt saamskogo jazyka (fonologo-morfologičeskoe issledovanie)*. Petrozavodsk: Karelija.

Appendix. All entries for complexity measures and non-linguistic factors. Extracts appear in tables in the text. Canonical complexity values for World (*italicized*) are based on an older coding system and are not exact. Complexity entries are totals; full breakdowns will be in the database to be published c. 2025.

<i>Languages</i>	<i>Complexity</i>				<i>Non-linguistic factors</i>			
	Enumerative		Canonical		Lat.	Long.	Population	Range
	All	Morph.	All	Strict			magnitude	
South Saami	18	7	24	15	62.4	11.2	4	3
Ume Pite	20	8	62	45	66.0	17.5	4	1
Lule	18	7	57	37	68.0	16.0	4.5	1.5
North Saami	18	8	45	30	69.0	24.0	5	5
Inari			54	37	68.5	27.0	3	0.5
Skolt	20.5	9	57	41	69.4	29.0	4	2
Akkala					67.9	32.4	3	0.5
Kildin	21.5	9	46	33	67.0	37.0	4	2
Ter					67.3	38.5	3	0.8
Finnish	16	7	25	15	62.0	25.0	7	
Estonian	20.5	9			59.0	26.0	7	
Erzya	17	8	24	15	53.0	45.5	5	
Mari	15	6	22	16	56.6	48.0	6	
Udmurt	14.5	6			57.5	52.5	6	
Komi	16	7			62.0	50.0	6	

Hungarian	19	8	30	22	47.0	20.0	7
Mansi	18	11	22	16	62.0	62.0	5
E Khanty	17	9	26	14	64.0	78.0	4
N Khanty			20	11	66.3	66.4	5
Selkup	17	9			65.0	82.0	4
T Nenets	20	11	31	21	69.0	72.0	5
Nganasan	22	13	33	25	71.0	93.0	4

Means:

Saami	19.3	8.0	49.3	34.0
Non-Saami Uralic	17.7	8.7	25.9	17.2
All Uralic	18.2	8.4	36.1	24.6
World	19.1	9.9	32.4	26.6

Notes:

Pite: The very high entries for canonical complexity are contributed mostly by the verb, which is much more complex than the noun (in the rest of Saami they are similar in complexity or nouns are somewhat more complex).

Kildin: Population figure based on Kert 1971 and Rießler 2022.

Other figures for population magnitude based chiefly on population figures in Krauss 1997.

Magn.: Order of magnitude of the population (e.g., 4 = numbering in thousands, 5 = ten thousands, etc.).

Bivalent verb classes in Skolt Saami: A pilot study

Ksenia Shagal, LMU Munich & Alena Blinova, independent researcher

1. Introduction

In this paper, we investigate valency classes of bivalent verbs in Skolt Saami, that is, classes of two-argument verbs with identical coding frames for arguments. The study focuses on two-place predicates because they appear most suitable for cross-linguistic comparison and allow the detection of relevant typological tendencies. As explained in Say (2018b: 8–9), high-frequency three-argument verbs are often few in number, and their heterogeneity may hamper cross-linguistic generalizations, while one-argument verbs rarely fall in more than two or three classes (and in some languages even constitute a single monovalent class). In contrast, two-argument verbs are commonly numerous in a language and frequent in speech, and at the same time they are cross-linguistically prone to show deviant valency behaviour (see Bickel et al. 2014). The study, thus, considers both prototypically transitive predicates, such as ‘break’ or ‘kill’ (*Peter broke a stick, Peter killed Mary*), and predicates that at least in some languages can be expected to be intransitive, such as ‘listen’ (*Peter is listening to the radio*), ‘be afraid’ (*Peter is afraid of the dog*), ‘get stuck’ (*The glass got stuck to the table*), or ‘be different’ (*My shirt is different from yours*).

The research questions and the methodology are largely defined by the project coordinated at the Laboratory for Typological Study of Languages at the Institute for Linguistic Studies (Russian Academy of Sciences) in Saint Petersburg and reported in the volume edited by Say (2018a). This larger project aimed, firstly, at developing methods to measure the degrees of (dis)similarity in the organization of valency classes across languages. The aspects considered in the study ranged from more holistic properties, such as the lexical extent of transitive predicates, to more complicated ones, such as the lexical distribution of verbs among valency classes as such (see overview of the goals in Say 2014: 118–119). Secondly, the proposed techniques were applied to classes of bivalent verbs in a sizeable sample of languages in order to reveal areal and genetic patterns (for a survey of European languages, see Say 2014). By the time the volume was published, data were collected for 78

languages, 67 of which represented Northern Eurasia. For each language, an expert filled in a questionnaire containing contextualized uses of 130 bivalent predicates, and annotated the sentences for morphosyntactic devices used for coding their two arguments. All the data collected within the project are published online in Say (2020).

Our article has two major goals. First, it contributes to the aforementioned project by providing data on Skolt Saami, a language belonging to the East Saami group within the Uralic language family. Primarily, we describe valency classes of the Skolt Saami bivalent predicates. We also place Skolt Saami in a broad typological context and comment on how the new data relate to the typological generalizations formulated within the larger project. Second, we discuss the Skolt Saami data from a contact and areal perspective. As argued in Say (2018c: 575–602), contact influence is visible in the behaviour of bivalent verbs on several levels. In particular, if we compare the share of transitive verbs (transitivity ratio) across languages, the results converge to large-scale areal patterns. Among the languages of Northern Eurasia, languages spoken in Western Europe, Siberia and the Far East have the highest ratios of transitives irrespective of the family they belong to, while the lowest ratios are observed in genealogically diverse languages spoken in Eastern Europe and in the Caucasus (Say 2018c: 576–579). At the same time, the way predicates are arranged into individual valency classes also reflects contact histories of these languages, but on a more local level. For instance, Basque patterns together with the surrounding Romance languages, Irish is similar to English and other Germanic languages, and Tungusic languages are much closer to Chukchi and Forest Enets than to their fellow members of the hypothetical proposed Altaic family, Turkic and Mongolic (Say 2018c: 590–597).

Contact and areal effects are particularly salient in the Uralic language family. As shown by Shagal & Say (2018), Uralic languages follow certain areal patterns with respect to both the share of transitive predicates and individual valency classes. For example, Forest Enets is in many ways similar to other languages of Siberia (see also Ovsjannikova 2018), Hungarian shares some features with Standard Average European languages (see Haspelmath 2001 on this notion), while the Permian, Mordvinic and Mari languages have a great deal in common with Russian, with which they have been in long-term contact. Skolt Saami can also be expected to show some relevant phenomena due to its contact history. For several centuries, it has been in close contact with Russian, and its earlier important contacts include Karelian and North Germanic (for an overview of East Saami contacts, see Rießler 2009; Blokland & Rießler 2011). Currently, Skolt Saami is spoken by approximately 300 people in the northeast of Finland and by some 30 people in Russia, on the Kola Peninsula (Feist 2015: 20–22). In this study, we only worked with speakers from Sevettijärvi (Municipality of Inari, Finland), and for this variety a recent major contact language has been Finnish, after most Skolt Saami speakers resettled in Finland during World War II. Other varieties of Skolt

Saami, which might reflect more of the Russian or possibly Norwegian influence, are outside the scope of this study.

The article is organized as follows. In Section 2, we describe the procedure of gathering and annotating the Skolt Saami data. Section 3 provides an overview of the classes of bivalent predicates in Skolt Saami based on their morphosyntactic and semantic properties. Section 4 puts Skolt Saami into the broader picture of Uralic languages outlining common patterns and dissimilarities in the organization of bivalent verb classes. Finally, Section 5 presents conclusions drawn from the research, and considerations for further studies.

2. Data collection and annotation

The data for this study were collected using a questionnaire that contained 130 bivalent predicates (a detailed description of the questionnaire is available in Say et al. 2018; the full Skolt Saami data are published in Shagal & Blinova 2020). In accordance with the research program of the larger typological study, the authors of the questionnaire aimed to represent as widely as possible various meanings that are likely to be expressed by intransitive predicates. However, several highly transitive meanings were also included, so that they could serve as a predictable background, or a standard of comparison.

Most sentences in the questionnaire are accompanied by specific context information which itself does not need to be translated but allows for a more accurate comparison between the Skolt Saami data collected for this project and the data from other languages. It also reduces possible variation in argument realization. For example, the verb *pääččad* ‘shoot’ in Skolt Saami can behave both transitively and intransitively, as shown in (1) and (2) below (the same type of variation is also present in the English translations). The project, however, only considers the example in (1) expressing the unsuccessful attempt of shooting for this predicate, which is specified by the context given in square brackets.¹

(1)	<i>Peâtt</i> Peter.SG.NOM ‘Peter shot at the bird. [He missed.]’	<i>pääčča-i</i> shoot-PST.3SG	<i>lââ’d</i> bird.SG.GEN	<i>vuâstta.</i> towards
(2)	<i>Käärnõs</i> raven.SG.NOM muu ”. 1SG.ACC ‘The raven says, “Don’t shoot me!”’ (Moshnikoff 1992: 72)	<i>ceäkk:</i> say.PRS.3SG	<i>“Jeä’l</i> NEG.IMP.2SG	<i>pääjj</i> shoot.CNG

¹ Unless indicated otherwise, all examples in the article come from our own field notes.

The data presented in this article were obtained primarily from an 80-year-old speaker from Sevettijärvi with kind assistance of Tim Feist. The speaker worked mainly with the English version of the questionnaire. In some cases, the speaker offered more than one possible translation, which could involve variation either in the choice of predicate, as in (3)–(4), or in the choice of argument encoding, as in (5)–(6).

(3)	<i>Peâtt</i> Peter.SG.NOM ‘[Peter had an argument with Maria, Maria went away, Peter went after her.] Peter caught up with Maria.’	<i>tää'vt-i</i> grab-PST.3SG	<i>Mää'ryj.</i> Maria.SG.ACC
(4)	<i>Peâtt</i> Peter.SG.NOM ‘Peter caught up with Maria.’	<i>õõst-i</i> reach-PST.3SG	<i>Mää'ryj.</i> Maria.SG.ACC
(5)	<i>Teârmmaztoll</i> lightning.SG.NOM ‘Lightning hit the house.’	<i>kõõskst-i</i> flash-PST.3SG	<i>põ'rtt-e.</i> house-SG.ILL
(6)	<i>Teârmmaztoll</i> lightning.SG.NOM <i>ool.</i> onto ‘Lightning hit the house.’	<i>kõõskst-i</i> flash-PST.3SG	<i>põõrt</i> house.SG.GEN

In order to make our data suitable for cross-linguistic comparison (especially quantitative), it was necessary to choose only one alternative in each case. The main criteria were naturalness and preciseness of translational equivalence to the stimulus, although in some cases these did not suffice for an undoubted decision, so the choice was still somewhat arbitrary.²

After the data were collected, the translations were annotated for the morphosyntactic devices used for argument encoding. Since the study deals with bivalent predicates, all the sentences in the questionnaire, naturally, are expected to have at least two nominal dependents. For each predicate, one of the dependents was considered the first argument, and the other the second. The decision on which one is which was made based on their semantic properties, so that the first argument would be close to A, or Proto-Agent, and the second would be close to P, or Proto-Patient, in the sense of Dowty (1991). For instance, in many cases, the first argument was the only animate

² In principle, language-internal variation in argument encoding could itself become a valid research question (see, for example, a detailed account of variation in Forest Enets in Ovsjannikova 2018). However, since we were only working with one speaker, any data would hardly be representative of the whole community, and therefore we do not discuss this issue in this article.

participant (e.g., Peter in ‘Peter wrote the letter’), because it is the only participant with volitional involvement in the event.

For Skolt Saami, typical argument encoding devices are case or postpositional marking on nominals (dependent-marking) and verbal agreement with the nominative participant (head-marking). For example, in (7), the first argument (Peter) is encoded by the nominative case, and the second argument (the clouds) is encoded by the postposition *árra* ‘towards’, which, like all postpositions in Skolt Saami, requires genitive marking on the noun. The predicate agrees with the nominative argument in person and number.

- (7) *Peátt* *lij* *ķiičče-men*
 Peter.SG.NOM be.PRS.3SG watch-PROG.PTCP
põõlv-i *árra.*
 cloud-PL.GEN **towards**
 ‘Peter is looking at the clouds [trying to determine if it will rain].’

In this article, valency frames are labelled in terms of dependent-marking devices only. Thus, the pattern in (7) is characterized as ‘first argument is in the nominative case, second argument takes the postposition *árra* with the genitive’, or, for convenience, simply NOM_GEN *árra*.³

Out of 130 sentences in the questionnaire, we were able to establish translational equivalents for 120 of them. Some gaps were due to the fact that the speaker could not think of any natural way to express the intended predicate meaning in Skolt Saami, or was not completely happy with any of the options. This was the case with the meanings ‘depend’, ‘lack, be short (of smth)’, ‘punish’, ‘be squeamish’, ‘enjoy’, ‘upset’, and ‘have a grudge’.⁴ In the rest of the cases, one of the predefined arguments could not be expressed in the same clause, as in example (8)⁵, or was not a clause-level constituent, but occurred as an adnominal dependent, as in (9), or as a part of a compound, as in (10).

³ In all our Skolt Saami data, the first argument always comes first in a sentence, so the label also reflects the neutral word order in the respective valency frame.

⁴ For stimulus sentences containing some of these verbs, we received suggestions, but none of them could be regarded as separate predicates. For example, the translation offered for the sentence ‘Memory depends on age’ was exactly the same as for ‘Age influences memory’, and it contained the verb *vaikkted* meaning ‘influence’, which required a different assignment of the first and the second argument status; see example (19) in Section 3.5. The stimulus ‘Peter is squeamish about dirty dishes’ was translated using the negated verb *tu’kķeed* ‘like’ (lit. ‘Peter does not like dirty dishes’), which is clearly not an elementary expression that we were looking for.

⁵ Notice that English also lacks a monoclausal way of expressing the required meaning. Nevertheless, many languages do have it; consider *püüdma* in Estonian, *kundšems* in Erzya, *kutnj* in Zyrian Komi and Permyak Komi, *fangen* in German, or *lovit’* in Russian.

- (8) *Peâtt* *strääžž* *[što* *tää'vte'-čč-i*
 Peter.SG.NOM try.PRS.3SG COMP grab-COND-3SG
kaazz].
cat.SG.ACC
 ‘[The cat sprang out of the house.] Peter is trying to catch the cat.’
- (9) *Peâtt* *lij* *Määr’j*
 Peter.SG.NOM be.PRS.3SG **Maria.SG.GEN**
taaurõš.
 friend.SG.NOM
 ‘Peter is Maria’s friend.’
- (10) *Peâtt-ast* *lij* *vuei’vv-kõpp*.
 Peter-SG.LOC be.PRS.3SG **head-disease.SG.NOM**
 ‘Peter has a headache.’

The 120 predicates for which we obtained translations of the stimulus sentences were grouped into non-overlapping classes. Two predicates were seen as belonging to the same valency class if and only if their first and second arguments were coded by the same means correspondingly (different postpositions being counted as different encoding devices). The next section presents all the valency classes that occurred in the Skolt Saami data.

3. Bivalent verb classes

3.1. Overview

With respect to argument encoding, bivalent predicates in Skolt Saami fall into 17 classes (the whole system with examples of predicates belonging to individual classes is summarized in Table 1 in Section 3.19). One of them is the class of transitive verbs (defined in this project as the class containing prototypical transitive verbs such as ‘break’ or ‘kill’; see Say 2018a: 13; Haspelmath 2015: 137–138 for the justification of this approach), and the rest comprise intransitive verbs. Intransitive verbs differ with respect to two major criteria that will be discussed in what follows, namely the formal means of encoding their arguments and the locus of non-transitivity.

The morphosyntactic devices involved in encoding arguments of bivalent predicates in Skolt Saami are cases, postpositions, and conjunctions. Out of the nine cases present in the language, seven are used in our data as means for argument encoding. All of them (that is, nominative, accusative, genitive, comitative, locative, illative, and essive) can encode the second participant of the situation, while only three (nominative, locative, and accusative) can be used for encoding the first participant. The only two cases that are not involved in coding arguments of bivalent verbs are the abessive and the partitive. The

abessive expresses the absence of an instrument or the absence of a person or object accompanying another (e.g., ‘the boat won’t move without an oar’). The partitive mostly occurs after numerals greater than six and after quantifiers, such as *muä'dd* ‘several’ and *mängg* ‘many’, although it is occasionally used in some other minor constructions; see Feist (2015: 254–255). Both cases are rare in Skolt Saami in general (Feist 2015: 252, 254–255). Classes of predicates that exclusively use case marking for argument encoding are discussed in Sections 3.2–3.5 and 3.8–3.10. Postpositions are only used to express the second participant, and they govern the nominal in the genitive case. Out of eight postpositions involved in argument encoding, six have spatial meaning (*ool* ‘on(to)’, *tuákka* ‘behind’, *râast* ‘across’, *vuälla* ‘under’, *vuâstta* ‘towards’, and *ârra* ‘towards’), one, *vuákka* ‘like’, expresses resemblance, and one, *diõtt* ‘because of’, mainly expresses reason. Nevertheless, postpositions with a primary spatial meaning can also be used with verbs of emotion, e.g., *suttõõttâd (ool)* ‘get upset (with)’ or *lee'd suttâm (ool)* ‘be angry (with)’. Classes of predicates that use postpositions for argument encoding are discussed in Sections 3.6–3.7 and 3.11–3.16. The two predicates expressing their second participants using conjunctions *ko* ‘than’ and *mâte* ‘like’ are presented in Sections 3.17 and 3.18. It should be noted that the type of formal means used for argument expression (case vs. postposition vs. conjunction) does not itself have any significant impact on the semantics of the respective predicates, so the classification outlined above is purely formal.

The locus of non-transitivity of a specific predicate is the participant or participants of this predicate encoded by means not used in the transitive construction (see Say 2018b: 59 on this notion). For Skolt Saami, this means encoding that is neither nominative for the first participant nor accusative for the second participant, since the transitive valency frame is NOM_ACC, as shown in Section 3.2. Cross-linguistically, most intransitive verbs have valency frames with P locus, that is, the encoding of their second participant deviates from the transitive frame. In Say’s (2018c: 565–566) data, this locus is characteristic of 85.2% of all intransitive predicates. In Skolt Saami, it is also by far the most common, being attested in 48 cases, or 90.6% of the 53 intransitive verbs. Valency frames with A locus, where it is the first participant that is encoded “intransitively”, constitute 12.7% in Say’s data, and 7.5% in our data (four predicates, all belonging to one class discussed in Section 3.8). In conformity with the Relevance Principle introduced by Malchukov (2006: 335), all the predicates with A locus of non-transitivity in Skolt Saami have first arguments that deviate from a prototypical A participant in that they are non-volitional, e.g., predicates meaning ‘have (a disease)’, or ‘have enough (of smth)’. Finally, 2.1% of the intransitive predicates investigated in the cross-linguistic study had A&P locus, that is, neither of their arguments was expressed in the same way as arguments of transitive verbs in respective languages. In Skolt Saami, there is only one predicate with A&P locus (1.9% of all intransitive predicates), *kääčceed* ‘call, refer (to smth)’; see Section 3.9.

According to preliminary typological observations discussed in Say (2018c: 584), predicates with A&P locus tend to be stative, and the Skolt Saami data seem to be in line with this tendency.

In the sections to follow, we present all the 17 valency classes of Skolt Saami two-place predicates in descending order of the number of members they comprise. For each class, we describe the valency frame and provide at least one sentential example. In addition, we point out certain semantic tendencies that can be traced in some of the classes. The table featuring all the classes with examples of the predicates they include is given in Section 3.19.

3.2. Transitive predicates: NOM_ACC

In a transitive construction in Skolt Saami, the subject receives nominative encoding, and the direct object is marked with the accusative case. Similarly to most other Saami languages (and unlike South Saami, as discussed in Kroik 2016), Skolt Saami does not show any differential object marking, so the valency frame is always NOM_ACC, as illustrated in (11).

- (11) *Peâtt* *kooll* *musiikk*.
Peter.SG.NOM hear.PRS.3SG music.SG.ACC
‘Peter hears the music.’

Transitive verbs are by far the largest class of bivalent predicates in Skolt Saami. Out of 120 predicates, 67, or 55.8%, are transitive (the share of transitive verbs in Skolt Saami is discussed from a typological and Uralic perspective in Section 4). As is typical of many other languages (see Say 2014: 120–121), the class of transitive verbs in Skolt Saami cannot be characterized in semantic terms in its entirety. It encompasses both predicates that are semantically highly transitive (e.g., *poorrâd* ‘eat’ in ‘Peter ate an apple’, or *poonnjâd* ‘bend’ in ‘Peter bent the branch’) and predicates that lack most of prototypical transitive properties, such as actionality, telicity, volitionality, or affectedness of the P participant (e.g., *pirrôôttâd* ‘surround’ in ‘walls surround the city’ or *kađšted* ‘envy’ in ‘Peter envies Maria’). As for the semantic range of second participants, apart from Patients, other semantic roles that commonly feature in a transitive construction are Theme (‘stone’ in ‘Peter threw a stone’ with the verb *kôôškeed* ‘throw’) and Stimulus (‘Maria’ in ‘Peter misses Maria’ with the verb *â’k’kôôššâd* ‘miss’).

3.3. NOM_COM

Among the intransitive predicates, the largest class consists of 11 predicates whose first argument is in the nominative, and the second in the comitative, as illustrated in (12).

- (12) *Peâtt* *teeivõõđ-i*
 Peter.SG.NOM encounter-PST.3SG
Määrja-in.
 Maria.dim-SG.COM
 ‘Peter encountered Maria [accidentally, on the street].’

The class shows a certain consistency in terms of semantics. Seven predicates, namely *teeivõõttâd* ‘encounter’, *tuärrad* ‘fight’, *tobdstõõttâd* ‘get to know’, *mainsted* ‘speak’, *lee'd õõutmiõllsaž* ‘agree’, *reeidcõõttâd* ‘have a quarrel’ and *rä'kštõõvvâd* ‘fall in love’, have two animate participants. All of these meanings except for ‘fall in love’ belong to a typologically stable cluster of symmetric predicates. In other words, the typological study showed that they are commonly expressed by a set of verbs with the same valency frame (Say 2018c: 569, 574). Another predicate belonging to the same cluster is *seäkkned* ‘get mixed’, which also denotes a symmetric event, but involving two inanimate participants; see example (13). One more predicate of this type is *suâppâd* ‘match’, as shown in (14).

- (13) *Miõtt* *seäkkn-i* *mie'lķ-in.*
 honey.SG.NOM get.mixed-PST.3SG milk-SG.COM
 ‘[Some honey was put in a jar of milk and stirred.] The honey got mixed with the milk.’

- (14) *Tät* *puägganj* *suâpp*
 PROX.SG.NOM belt.SG.NOM match.PRS.3SG
puârast muu *kååut-in.*
 well 1SG.GEN skirt-SG.COM
 ‘[A girl tries a belt in the shop and says, ‘Great,] this belt matches my dress.’

The remaining two predicates in this class are *čuõppõõttâd* ‘cut oneself’ and *siõrrâd* ‘play’. In both cases, the second participant can be seen as an instrument, which is normally encoded by the comitative in Skolt Saami (Feist 2015: 250), as shown in (15).

- (15) *Peâtt* *seârr* *kitara-in.*
 Peter.SG.NOM play.PRS.3SG guitar-SG.COM
 ‘[What is Peter doing?] Peter is playing the guitar.’

The co-occurrence of these two types of verbs in one valency class is, again, in line with typological research which shows that it is cross-linguistically common for arguments denoting an instrument to receive the same type of morphosyntactic encoding as arguments denoting a companion, or a participant in a symmetric situation (Narrog & Ito 2007; Say 2018c: 569). It should be noted though that other predicates with instrument/means/cause as their second participant (yet another cluster discerned in Say 2018c: 569) do

not pattern together in Skolt Saami, but are distributed across different valency classes. For example, *heeilted* ‘wave’ is a transitive verb, *teâuddjed* ‘fill (with smth)’ has a NOM_LOC frame (see Section 3.4), and the second argument of *lee’d miöllstes* ‘be content’ takes the postposition *diõtt* ‘because of’ (see Section 3.6).

3.4. NOM_LOC

This class includes nine predicates that have their first argument in the nominative and second argument in the locative, as shown in (16).

- (16) *Peâtt* *vuõ'lj-i* *meâdda* *suu*
 Peter.SG.NOM leave-PST.3SG away 3SG.GEN
šõdd-âm-siid-âst.
 born-PST.PTCP-village-SG.LOC
 ‘Peter left his native city.’

Three of them, namely *mõõnnâd* ‘go out’, *vue'lgged* ‘leave’ and *luâšttõõttâd* ‘descend’, have a meaning of physical movement away from a specific point, and they can emphasize it using a spatial adverb like *meâdda* ‘away’ in the example above. One more predicate, *põõllâd* ‘fear, be afraid’, seems to have a similar semantic component on a metaphorical level. The verb *põõllâd* ‘fear, be afraid’ also belongs to a subgroup within this valency class comprising meanings related to feelings and mental activities, along with *niõggeed* ‘dream, see in one’s sleep’ and *tu'kķeed* ‘like, love’ illustrated in (17).

- (17) *Peâtt* *tu'kķad* *čee-st.*
 Peter.SG.NOM like.PRS.3SG tea-SG.LOC
 ‘Peter likes tea.’

The verb *tu'kķeed* ‘like, love’ was considered the most natural translational equivalent for three predicates in the questionnaire and is, therefore, counted as three different predicates. The reason for this is that the contexts in which it appears are different. Apart from the example in (17), where the second argument is a non-specific inanimate stimulus, it is also used to translate the sentence ‘Peter likes this shirt’ with a specific inanimate second argument, and the sentence ‘Peter is fond of Mary’ with a specific animate argument. In all these cases, the valency frame remains the same.

The last predicate in our data with the NOM_LOC valency frame is *teâuddjed* ‘fill (with smth)’ shown in (18), which does not seem to have any semantic links to the rest of the class.

- (18) *Skääll* *teâuddj-i* *čää'zz-est.*
 bucket.SG.NOM become.filled-PST.3SG water-SG.LOC
 ‘[There was a bucket near the house during the rain.] The bucket filled with water.’

3.5. NOM_ILL

This class consists of eight predicates with the first argument in the nominative and the second argument in the illative, as illustrated in (19).

- (19) *Peätt* *la'ddj-i* *ridd-u.*
 Peter.SG.NOM reach-PST.3SG bank-SG.ILL
 ‘[Peter was crossing the river in a boat.] Peter reached the bank.’

Most predicates in this group belong to two different but neighbouring classes in Say’s (2018c: 569–570) analysis, which means that in some languages they do pattern together in terms of argument expression. The first of these subclasses in Skolt Saami comprises predicates whose second argument denotes the target or direction at which a physical action is aimed, namely *la'ddjed* ‘reach’, *pue'tted* ‘come’, *kuõskkâd* ‘touch’, and *kõskkâd* ‘flash (or hit) with lightning’. The second subclass encompasses verbs that refer to interactions between humans where the second participant can also be perceived as a kind of target or recipient. These verbs are *ceä'lkked* ‘tell’, *taa'ppjed* ‘lose, be beaten (in a game)’, and *va'stteed* ‘answer’; see (20).

- (20) *Peätt* *vastti-i* *u'čteel-a.*
 Peter.SG.NOM answer-PST.3SG teacher-SG.ILL
 ‘[The teacher asked Peter what his parents’ names are.] Peter answered the teacher.’

The meaning of the last verb with the NOM_ILL valency frame, *vaikkted* ‘influence’, illustrated in (21), does not cluster typologically with the meanings listed above. In other words, according to Say’s cross-linguistic study (2018c: 569–570), the verbs meaning ‘influence’ do not often encode their arguments in the same way as the verbs meaning ‘reach’, ‘come’, ‘tell’, ‘answer’, etc. However, the semantics of targeted action appears to pertain to this predicate as well.

- (21) *Áä'kk* *vaaikat* *muštt-u.*
 age.SG.NOM influence.PRS.3SG memory-SG.ILL
 ‘Age influences memory.’

3.6. NOM_GEN *diõtt* ‘because of’

This class features six predicates and is the largest among those that use postpositions for argument encoding. The first argument is in the nominative, and the second argument is expressed by a postpositional phrase with the postposition *diõtt* ‘because of’, as shown in (22).

- (22) *Peâtt* *õõmtõõžž-i* *skîânk*
 Peter.SG.NOM be.surprised-PST.3SG gift.SG.GEN
diõtt.
because.of
 ‘[Maria gave Peter an expensive gift.] Peter was surprised at this gift.’

The predicates belonging to this class are *õõmtõõššâd* ‘be surprised’, *lee’d kuärgast* ‘be glad’, *lee’d miõllstes* ‘be content’, *neu’rrõõvvâd* ‘take offence’, *käinnad* ‘be shy or embarrassed (about smth)’, and *niõggeed* ‘dream’. All of them are emotion predicates, and the second participant can be seen as either stimulus or reason. Cross-linguistically, most of these verbs also tend to form a class of their own, or pattern together with the verbs belonging to the class outlined in the next section.

3.7. NOM_GEN + *ool* ‘on(to)’

This class consists of five verbs with the first argument in the nominative and the second argument expressed by a postpositional phrase with the postposition *ool* ‘on(to)’, as shown in (23).

- (23) *Ståkkán* *pâššn-i* *påå’rd*
 glass.SG.NOM get.caught-PST.3SG table.SG.GEN
ool.
onto
 ‘[The table was recently covered with paint. Peter didn’t know this and placed his glass on it.] The glass got stuck to the table.’

For two of the verbs, *kõrmmled* ‘attack’ and *pâššned* ‘get stuck’, the second participant has a meaning of a physical goal. The rest are emotion predicates that have the stimulus as their second participant. They are *suttõõttâd* ‘get upset’, *gåårâtted* ‘get irritated’, and *lee’d suttâm* ‘be angry’; see (24).

- (24) *Peâtt* *lij* *suttâm* *Mää’rj*
 Peter.SG.NOM be.PRS.3SG angry Marja.SG.GEN
ool.
onto
 ‘Peter is angry with Maria.’

3.8. LOC_NOM

As mentioned in Section 3.1, this is the only class in Skolt Saami with P locus of non-transitivity. Its first argument is in the locative, while the second argument receives nominative encoding; see (25).

- (25) *Peätt-ast* *lij* *nuõpp-kõpp*.
 Peter-SG.LOC be.PRS.3SG cold(disease)-disease.SG.NOM
 ‘Peter has the flu’.

The class consists of four predicates. Two of the meanings are conveyed simply by the verb *lee'd* ‘be’, namely the meaning ‘have’ in the context of owning (e.g., ‘Peter has a car’), and in the context of having a specific disease, as illustrated in (25) above. The two other predicates use additional material, that is, *lee'd teänab* ‘remain’ and *lee'd nokk* ‘be enough’; see (26).

- (26) *Áá'n* *Peätt-ast* *lie* *nokk* *tie'gğ*.
 now Peter-SG.LOC be.PRS.3PL. enough money.PL.NOM
 ‘[P. has found a new job.] Now Peter has enough money.’

Typologically, these predicates often belong to a larger class featuring some other verbs with non-volitional first participant, such as ‘like’ or ‘need’ (Say 2018c: 570). In Skolt Saami, however, the class in question is limited to verbs with some kind of possessive meaning.

3.9. ACC_ESS

From this section on, we list classes that contain only a single predicate. In the first one of these classes, the first argument is in the accusative, and the second bears essive case marking, as shown in (27).

- (27) *Tän* *tiing* *kääččēet* *kompas-ân*.
 PROX.SG.ACC thing.SG.ACC call.PRS.4SG compass-ESS
 ‘[A boy saw his father’s compass and asked what this was. The father answered him:] ‘This thing is called a compass.’

The predicate *kääččēet* ‘call’ is the only member of this class. This is the only predicate in Skolt Saami with A&P locus of non-transitivity.

3.10. NOM_NOM

In this class, both the first and the second argument appear in the nominative case, as shown in (28).

- (28) *Tät* *kopp* *mähss* *õhtt*
 PROX.SG.NOM cup.SG.NOM cost.PRS.3SG one.NOM
eu'rr.
 euro.SG.NOM
 ‘This cup costs 1 euro.’

The predicate *mähssed* ‘cost’ is the only member of this class.

3.11. NOM_GEN *vuâkka* ‘like’

In this class, the first argument is in the nominative, and the second argument is expressed by a postpositional phrase with the postposition *vuâkka* ‘like’, as displayed in (29).

- (29) *Pä'rnn* *lij* *ää'jj-es*
 Boy.SG.NOM be.PRS.3SG grandfather.SG.GEN-POSS.3SG
vuâkka.
like
 ‘[Miika is Peter’s grandfather] The boy resembles his grandfather.’

The only meaning belonging to this class is ‘resemble’, and it is conveyed by a combination of the verb *lee'd* ‘be’ and the postposition.

3.12. NOM_GEN *tuâkka* ‘behind’

In this class, the first argument is in the nominative, and the second argument is expressed by a postpositional phrase with the postposition *tuâkka* ‘behind’, as shown in (30).

- (30) *Peätt* *kuâddj-i* *Mää'rj*
 Peter.SG.NOM remain-PST.3SG Marja.SG.GEN
tuâkka.
behind
 ‘[Peter and Maria started out to school together, but] Peter fell behind Maria.’

The only meaning belonging to this class is ‘fall behind’, and it is conveyed by a combination of the verb *kuâddjed* ‘remain’ and the postposition.

3.13. NOM_GEN *rââst* ‘across’

In this class, the first argument is in the nominative, and the second argument is expressed by a postpositional phrase with the postposition *rââst* ‘across’, as shown in (31).

- (31) *Peâtt* *mõõn-i* *čuõkku* ***rââst.***
 Peter.SG.NOM go-PST.3SG road.SG.GEN **across**
 ‘[Peter saw that the shop is on the other side of the road.] Peter crossed the road.’

The only meaning belonging to this class is ‘cross’, and it is conveyed by a combination of the predicate *mõõnnâd* ‘go’ and the postposition.

3.14. NOM_GEN *vuâlla* ‘under’

In this class, the first argument is in the nominative, and the second argument is expressed by a postpositional phrase with the postposition *vuâlla* ‘under’, as shown in (32).

- (32) *Stokk* *vuâja-i* *čää'žž* ***vuâlla.***
 log.SG.NOM sink-PST.3SG water.SG.GEN **under**
 ‘[This tree has very heavy wood. I threw a log in the river.] The log sank in the water.’

The predicate *vuâjjad* ‘sink’ is the only member of this class.

3.15. NOM_GEN *vuâstta* ‘towards’

In this class, the first argument is in the nominative, and the second is expressed by a postpositional phrase with the postposition *vuâstta* ‘towards’, as shown in example (1) and repeated here in (33) for convenience.

- (33) *Peâtt* *pääčča-i* *lââ'dd* ***vuâstta.***
 Petja.SG.NOM shoot-PST.3SG bird.SG.GEN **towards**
 ‘Peter shot at the bird. [He missed.]’

The predicate *pääččad* ‘shoot’ is the only member of this class.

3.16. NOM_GEN *ârra* ‘towards’

In this class, the first argument is in the nominative, and the second argument is expressed by a postpositional phrase with the postposition *ârra* ‘towards’, as shown in (34).

- (34) *Peâtt* *lij* *ķiičče-men*
 Peter.SG.NOM be.PRS.3SG watch-PROG.PTCP
põõlv-i ***ârra.***
 cloud-PL.GEN **towards**
 ‘Peter is looking at the clouds [trying to determine if it will rain].’

The predicate *küččeed* ‘look’ is the only member of this class.

3.17. NOM_ *ko* ‘than’ NOM

In this class, the first argument is in the nominative, and the second argument is introduced by the conjunction *ko* ‘than’, as shown in (35).

- (35) *Ij,* *muu* *päi'dd* *lij*
 NEG.3SG 1 SG.GEN shirt.SG.NOM be.PRS.3SG
jee'resnallšem ko tuu päi'dd.
 different than 2SG.GEN shirt.SG.NOM
 ‘[Maria told Peter that they have identical shirts. Peter answered her,] No, my shirt is different from yours [, it is darker].’

The predicate *lee'd jee'resnallšem* ‘be different’ is the only member of this class.

3.18. NOM_ *mâte* ‘like’ NOM

In this class, the first argument is in the nominative, and the second argument is introduced by the conjunction *mâte* ‘like’, as shown in (36).

- (36) *Muu* *küöđ* *â'psse*
 1 SG.GEN hand.PL.NOM smell.PRS.3PL
mâte be'nsin.
 like gasoline.SG.NOM
 ‘[I was fixing my car.] My hands smell like gasoline.’

The predicate *â'psse* ‘smell’ is the only member of this class.

3.19. Summary of valency classes

Table 1 below summarizes all 17 valency classes of bivalent predicates attested in Skolt Saami based on the questionnaire introduced in Say et al. (2018). The structure of the table is based on Ovsjannikova (2018: 53). The classes are listed in descending order of the number of predicates they comprise. The locus of non-transitivity (first argument, second argument, or both) is given in bold.

Table 1. Classes of bivalent predicates in Skolt Saami.

<i>Argument encoding</i>	<i>Examples of predicates</i>	<i>Number of pred.</i>

1 st	2 nd		<i>Semantic role(s) of the non-transitive argument</i>	
NOM	ACC	<i>kâ'dded</i> 'kill', <i>kõõškeed</i> 'throw', <i>â'kkõõššâd</i> 'miss'	Patient and others	67
NOM	COM	<i>tuärrad</i> 'fight', <i>seäkkned</i> 'become mixed', <i>siõrrâd</i> 'play'	Companion, Instrument	11
NOM	LOC	<i>vue'lğğed</i> 'leave', <i>tu'kķeed</i> 'like, love', <i>teâuddjed</i> 'fill (with smth)'	Source, Stimulus, Instrument	9
NOM	ILL	<i>la'ddjed</i> 'reach', <i>va'stteed</i> 'answer', <i>taa'ppjed</i> 'lose, be beaten'	Goal, Recipient, Patient	8
NOM	GEN <i>diõtt</i>	<i>õõmtõõššâd</i> 'be surprised', <i>lee'd miõllstes</i> 'be content'	Stimulus, Reason	6
NOM	GEN <i>ool</i>	<i>kõrmmled</i> 'attack', <i>pâššned</i> 'become stuck', <i>lee'd suttâm</i> 'be angry'	Goal, Stimulus	5
LOC	NOM	<i>lee'd teänab</i> 'remain', <i>lee'd nokk</i> 'have enough'	Possessor, Experiencer	4
ACC	ESS	<i>kääčceed</i> 'call'	Theme (1 st and 2 nd)	1
NOM	NOM	<i>mä'hssed</i> 'cost'	Theme	1
NOM	GEN <i>vuâkka</i>	<i>lee'd</i> 'be' (> 'resemble')	Standard of comparison	1
NOM	GEN <i>tuâkka</i>	<i>kuâddjed</i> 'remain' (> 'fall behind')	Landmark	1
NOM	GEN <i>rââst</i>	<i>mõõnnâd</i> 'go' (> 'cross')	Goal	1
NOM	GEN <i>vuâlla</i>	<i>vuäjjad</i> 'sink'	Goal	1
NOM	GEN <i>vuâstta</i>	<i>pääčcad</i> 'shoot'	Goal	1

NOM	GEN <i>arra</i>	<i>ķiičceed</i> ‘look’	Stimulus	1
NOM	<i>ko</i> NOM	<i>lee'd jee'resnallšem</i> ‘be different’	Standard of comparison	1
NOM	<i>mâte</i> NOM	<i>â'pssed</i> ‘smell’	Standard of comparison	1

4. Skolt Saami from a Uralic perspective

As mentioned in the introduction, analysing the organization of bivalent verbs classes across languages allows us to uncover different genetic and areal tendencies, as well as specific manifestations of contact influence. Uralic languages are particularly interesting in this respect, since they do bear certain family similarities, but at the same time show clear signs of contact influence at different levels. Therefore, it seems promising to compare the behaviour of bivalent predicates in Skolt Saami to their behaviour in other Uralic languages examined within the project so far. At this point, data are available for 13 Uralic languages, at least one from each major traditionally recognized subgroup within the Uralic language family (cf. e.g., Salminen 2002), namely Skolt Saami and South Saami (Saamic), Estonian, Standard Finnish and Ingrian dialect of Finnish (Finnic), Erzya Mordvin and Moksha Mordvin (Mordvinic), Hill Mari (Mari), Zyrian Komi, Permyak Komi and Udmurt (Permic), Hungarian (Ugric), and Forest Enets (Samoyedic). In this section, we provide preliminary observations concerning the place of Skolt Saami among these languages with respect to the organization of bivalent verb classes. We will both compare Skolt Saami to the languages of other branches and briefly comment on the similarities and differences between the two Saamic languages. A more detailed comparison of Skolt Saami and South Saami is presented in Blinova & Shagal (2020).

Since valency classes cannot be directly equated across languages for a number of reasons, in order to draw meaningful cross-linguistic comparison, Say (2014, 2018c) introduced a spectrum of quantitative metrics describing the way valency class systems are organized in different languages. The first metric, which captures the coarsest distinctions between languages, is the transitivity ratio, i.e., the share of transitive verbs among the predicates included in the questionnaire. Skolt Saami has a transitivity ratio of 0.56, which is one of the highest values in the whole family. Outside the Saamic branch, the only language with the same ratio is Forest Enets, while in all other Uralic languages examined in the project, transitivity ratios are no more than 0.5. Earlier studies suggested an important role of areality in the distribution of transitivity ratios; see Say (2018c) for Northern Eurasia in general, and Ovsjannikova (2018) for Siberia. Indeed, Forest Enets is close in this respect

to other Siberian languages, such as Evenki (0.59), Sakha (0.53), Chukchi (0.59), and Nanai (0.60). For Skolt Saami, on the other hand, the areal explanation does not seem to work in a straightforward way. Finnic languages spoken roughly in the same region show the lowest values in the whole Uralic family (0.37 in Ingrian Finnish, 0.34 in Estonian, and 0.30 in Standard Finnish), and other languages of the region are also fairly intransitive, e.g., Russian (0.42) and Lithuanian (0.44). The geographically closest unrelated language with a high transitivity ratio examined in the project is Norwegian (0.54), but it can be seen as the northernmost outlier of the western European high-transitive area featuring such languages as Dutch (0.61), English (0.63), and Swedish (0.64). The transitivity ratio of South Saami is even higher (0.71), presumably due to its close contact with Swedish, but from the Uralic perspective the value attested in Skolt Saami is still remarkable.

The high transitivity ratio in Saamic languages is exceptional from one more point of view. Say (2014: 139) shows that in a sample of European languages transitivity ratios correlate negatively with the size of the case inventory for nouns. In other words, the more cases a language has, the lower the transitivity ratio. Skolt Saami, however, does not quite fit the tendency, with a transitivity ratio of 0.56 and nine nominal cases. Neither does South Saami, with a transitivity ratio of 0.71 and eight cases. The transitivity ratio of Skolt Saami is close to that of German and Ancient Greek (0.55), which both have four cases, while Norwegian and French with transitivity ratios of 0.54 and 0.56 respectively do not have cases in nouns at all. At the same time, languages with a comparable number of cases tend to be significantly less transitive. For example, Ossetic (nine cases) has a transitivity ratio of 0.39, and the transitivity ratio of Kalderaš Romani (seven cases) is 0.41 (Say 2014: 165). Furthermore, Standard Finnish, with 15 cases and a transitivity ratio of 0.3, supports the observation and contrasts Skolt Saami which it is both genealogically and geographically close to.

A possible answer to this puzzle might lie partly in the historical development of Saamic languages. Naturally, the majority of predicates that are transitive in Skolt Saami are transitive in all or almost all Uralic languages, in particular highly transitive predicates, such as ‘eat’, ‘make’, or ‘bend’. However, those verbs that are transitive in Skolt Saami but intransitive in most other Uralic languages do have a notable thing in common. Almost all of them are either transitive or have a nominative–partitive valency frame in Finnic languages, or at least in Standard Finnish. The first case can be illustrated by the predicate ‘forfeit’, which requires abessive encoding of the second argument in most other Uralic languages for which data are available (Zyrian Komi, Udmurt, and Erzya), but is transitive both in Skolt Saami and in Standard Finnish, as shown in examples (37) and (38).

- (37) *Peâtt* *mõõn-t-i*
 Peter.SG.NOM go-CAUS-PST.3SG
põõrt-âs.
 house.SG.ACC-POSS.3SG
 ‘[There was a fire in the village. As a result] Peter lost his house.’
- (38) *Pekka* *mene-tt-i*
 Pekka.SG.NOM go- CAUS-PST.3SG
talo-nsa.
 house.GEN/ACC-POSS.3SG
 ‘[There was a fire in the village. As a result] Pekka lost his house.’

The predicates ‘think’, ‘miss’, ‘make fun (of smb)’, ‘help’, ‘govern’, ‘follow’, ‘avoid’, ‘obey’, ‘wave’, ‘move’ and ‘need’ show a whole range of different intransitive encoding strategies across Uralic languages, including valency frames with A locus of non-transitivity. The predicate ‘believe’ requires dative encoding of the second argument in almost all subgroups of the family, and yet it is transitive in Skolt Saami and has a NOM_PART valency frame in Standard Finnish, as illustrated in (39) and (40):

- (39) *Peâtt* *âskk* *suu*.
 Peter.SG.NOM believe.PRS.3SG 3SG.ACC
 ‘Peter believes her.’
- (40) *Pekka* *usko-o* *Matti-a*.
 Pekka.SG.NOM believe-PRS.3SG Matti-SG.PART
 ‘Pekka believes Matti.’

In other words, the class of transitive verbs in Skolt Saami corresponds quite neatly to the classes of transitive verbs in Finnic languages, and it also includes a number of predicates that in Finnic languages show a NOM_PART valency frame. However, as already mentioned in Section 3.1, the use of the partitive in Skolt Saami is very limited (Feist 2015: 254–255), and although the partitive case is reconstructed for Proto-Saami (Sammallahti 1998: 70–71), in many Saamic languages (including South Saami) it was lost altogether and a number of its functions were taken up by the accusative. If we take this into account, the drastic difference in transitivity ratio between Skolt Saami and Finnic languages is not that striking. The only two verbs that are transitive in Skolt Saami and have intransitive valency frames elsewhere in the Uralic family (at least as the primary strategy) are *na’ddjõõttâd* ‘trust’ and *õõmšed* ‘marvel, admire’. This number, however, is small enough not to invalidate the outlined overall pattern.

Another metric used in Say (2018c: 584–590), which focuses on the distribution of intransitive predicates into classes, is relative entropy of

intransitives, or $H_{\text{rel}(\text{intr})}$. It allows us to compare the relative elaborateness of the intransitive class system in different languages and is calculated as the actual entropy of intransitives divided by its theoretical maximum for a given number of predicates. Actual entropy of intransitives is calculated using the following formula.

$$(41) \quad H(x) = - \sum_{i=1}^k p(x_i) \times \log(p(x_i))$$

Here, k is the number of intransitive classes in a given language, and $p(x_i)$ is the ratio of the i th class of verbs relative to the overall number of intransitive verbs considered for this language (see Say 2014: 149–151). The theoretical maximum of the entropy of intransitives equals $\log(x)$ where x is the number of intransitive verbs attested in a given language (e.g., $\log(65) \approx 4.17$ for a language with 65 intransitive verbs); this would happen if all the intransitive verbs formed separate classes, each with one member. The opposite situation where all intransitive verbs belong to one class would yield an entropy of 0. In other words, the more disorder in the grouping of verbs, the higher the entropy value.

For Skolt Saami, 53 verbs have been deemed intransitive, and they fall into 16 classes based on their valency frames. The class sizes are (11, 9, 8, 6, 5, 4, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1). Actual entropy of intransitives in this case is ≈ 2.33 . Theoretical maximum of entropy for 53 verbs equals $\log(53) \approx 3.97$. Consequently, relative entropy of intransitives for Skolt Saami equals 0.59, or 59% ($2.33/3.97$). In the sample of languages analysed in Say (2018c), the values of relative entropy fall within the interval from 0.31 to 0.77 (Say 2018c: 586–590). Thus, Skolt Saami shows a relatively elaborate system of intransitive classes, and the predicates are distributed relatively evenly across these classes. One notable aspect of $H_{\text{rel}(\text{intr})}$ metric is that its value often differs significantly for members of one language family or areal. Within Uralic languages, the value of $H_{\text{rel}(\text{intr})}$ puts Skolt Saami together with Permian (0.60–0.63) and Mordvinic (0.60–0.62), while Standard Finnish shows a much lower value (0.41). At the same time, Skolt Saami is rather close in $H_{\text{rel}(\text{intr})}$ to Russian (0.63), one of its major contact languages. The connection between Skolt Saami and Russian will be touched upon once again later in this section.

In South Saami, which has 31 intransitive verbs distributed across 15 classes with the sizes of (7, 7, 4, 2, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1), the actual entropy of intransitives is exactly the same as in Skolt Saami, that is ≈ 2.33 . Since the theoretical maximum of entropy for 31 intransitive verbs is lower, $\log(31) \approx 3.43$, the value of $H_{\text{rel}(\text{intr})}$ in South Saami is higher, namely 0.68. However, the small number of intransitive verbs attested in South Saami (and therefore the low theoretical maximum of the entropy) might be partly explained by the fact that no translational equivalents were obtained for a number of predicates that cross-linguistically tend to be intransitive (see Say 2018c: 560–564). Some of the lacking predicates are likely to belong to the larger attested classes, e.g., the predicates meaning ‘get mixed’ or ‘agree’ can

be expected to pattern with other NOM_COM predicates. If this is the case, it would decrease the South Saami $H_{\text{rel}(\text{intr})}$ and bring it closer to Skolt Saami.

While areal and genetic factors are not always reflected in the values of relative entropy for individual languages, a positive correlation is observed between this metric and the number of cases in a given language. Yet, Say (2018c: 589–590) hypothesizes that the reason for this lies on a deeper level: languages are more likely to have a high number of valency classes if they use numerous grammaticalized spatial markers for argument encoding. For Skolt Saami (and South Saami as well), this hypothesis may account for the relatively high value of $H_{\text{rel}(\text{intr})}$. The language is, indeed, rather rich in nominal cases (it has nine), although for many other Uralic languages the number is even higher. In addition, spatial markers appear in several classes of intransitive verbs as means of coding arguments. Almost all of the postpositions involved in argument encoding (e.g., *ool* ‘on(to), towards’, *tuákka* ‘behind’, *vuälla* ‘under’) carry some spatial meaning which may or may not be realized in a literal sense in a specific context.

Finally, the most fine-grained level of analysis is comparing the distribution of predicates into individual valency classes across languages.⁶ As shown in Say (2014: 152–158) based on the data from 29 European languages, at this level it is possible to identify numerous low-level convergences between languages, both genetic and purely areal. This implies that distributions of verbs into valency classes change fast and can be easily transferred in contact situations. For Skolt Saami, a particularly salient class in this respect is that with a LOC_NOM frame (see Section 3.8), containing predicates meaning ‘have’, ‘have enough’, ‘have (a disease)’, and ‘remain’. In many other Uralic languages, these predicates fall into different classes. Moreover, these classes can also differ in terms of their locus of non-transitivity. For instance, in Udmurt, the meaning ‘have’ is conveyed by means of a possessive construction, with the genitive on the first argument, and a possessive marker on the second, as shown in (42). The predicate ‘have (a disease)’ has the first argument in the nominative, and the second argument in the instrumental, as shown in (43).

(42)	<i>Petâ-lèn</i> Peter-SG.GEN ‘Peter has a car.’	<i>mašina-ez</i> car.SG.NOM-POSS.3SG	<i>van’</i> . COP.PRS

⁶ For analysing this aspect of valency class organization, Say (2018c: 590–597) uses a metric based on contingency tables for pairs of languages, and ultimately on comparing entropies of distributions. In this paper, however, we only provide some preliminary observations concerning the place of Skolt Saami in the Uralic language family. Quantitative analysis of Saamic languages in the Uralic context is currently in the works.

(43)	<i>Petâ</i> Peter.SG.NOM 'Peter has the flu.'	<i>gripp-en</i> flu-SG.INS	<i>vis-e.</i> be.sick-PRS.3SG

In Finnic languages, on the other hand, the argument encoding used with these predicates is more uniform. There is, indeed, some minor variation across languages as to how they treat these four predicates, but they all have A locus of non-transitivity, and in most cases the first argument is encoded by the adessive; see examples (44) and (45) from Standard Finnish.

(44)	<i>Peka-lla</i> Pekka-SG.ADE 'Pekka has a car.'	<i>on</i> be.PRS.3SG	<i>auto.</i> car.SG.NOM
(45)	<i>Peka-lla</i> Pekka-SG.ADE 'Pekka has the flu.'	<i>on</i> be.PRS.3SG	<i>flunssa.</i> flu.SG.NOM

From an areal point of view, it is also remarkable that in Russian, all the four predicates allow for a similar frame, using a locative possessive construction with the preposition *u* 'at' (although this frame was not deemed primary for all these predicates in the project). Unlike Skolt Saami, Finnic and Russian, South Saami uses a transitive verb *utnedh* 'have' to encode possessive relations. The verb itself is of Saamic origin, but its use in the possessive construction is clearly a result of North Germanic influence (Kowalik 2016: 52–54). Overall, the behaviour of the predicates expressing possession and similar concepts in the aforementioned neighbouring languages can be seen as an illustration of Say's (2014: 160) thesis that valency frames can spread independently of morphosyntactic devices, that is, in this case pattern borrowing does not imply matter borrowing (see Matras & Sakel 2007 for these notions).

The correspondences between Skolt Saami transitive verbs and verbs with transitive and NOM_PART valency frames outlined above are also indicative of a low-level areal signal. At the same time, Skolt Saami has at least one class with respect to which it patterns with both genealogically and geographically distant Uralic languages, namely the NOM_COM class described in Section 3.3. While this class appears slightly larger in Skolt Saami than in most other Uralic languages, the predicates 'encounter', 'fight', 'get to know', 'speak', 'get mixed', 'agree' and 'have a quarrel' share a valency frame also in Permic and Mordvinic (although morphosyntactic devices encoding the second participant, of course, differ, e.g., comitative in Zyrian Komi and Permyak Komi, instrumental in Udmurt, postposition *marto* in Erzya). In Standard Finnish, on the other hand, only some of these verbs use the frame with the second participant encoded by the postposition *kanssa* 'with', while the verb *tavata* 'encounter' is transitive, and the verbs *tutustua* 'get to know' and

sekoittua ‘get mixed’ have a nominative–illative frame. Interestingly, the latter three verbs all have their second participant encoded by the postposition *kans* ‘with’ in Ingrian Finnish. A possible explanation here is that Permic and Mordvinic languages, as well as Skolt Saami and Ingrian Finnish, all have been in contact with Russian, which uses similar encoding of the second participant (preposition *s* ‘with’) with all the verbs in question. In South Saami, which has not experienced a strong Russian influence, the meaning ‘encounter’ is encoded using a NOM_COM frame, but the meaning ‘get to know’ is expressed by a transitive construction, while for the meaning ‘get mixed’ there is currently no available data. Therefore, even in this case, the influence of contact is probable in the development of an individual valency class.

5. Conclusions

In this paper, we investigated bivalent predicates in Skolt Saami based on a questionnaire introduced in Say’s (2018a) typological project and reported our preliminary observations regarding the place of Skolt Saami within the Uralic language family and in a wider cross-linguistic context. Overall, Skolt Saami has a relatively wide range of ways to encode arguments of two-place predicates. Out of 130 stimulus sentences presented in the questionnaire, 120 yielded valid responses suitable for further analysis. The verbs can be split into 17 groups in total (see Table 1 in Section 3.19). The transitive class (NOM_ACC) is the largest, encompassing 67 verbs. Among intransitive classes, six have several members, and ten consist of a single member each. One class with four members has A locus of non-transitivity (LOC_NOM), one single-member class shows A&P locus (ACC_ESS), and the rest have P locus, which is cross-linguistically a fairly common picture (Say 2018c: 565–566). Six classes of intransitive verbs only use cases for argument encoding, eight classes use postpositions, and two single-member classes use conjunctions.

Both genetic and areal patterns can be identified in the way bivalent predicates are arranged into classes in Skolt Saami. Its transitivity ratio is 0.56, which is very high compared to other members of the Uralic family. The only non-Saamic language with a comparable value is Forest Enets (also 0.56), which is among the most distant relatives of Skolt Saami, both genealogically and geographically. However, the large gap between the transitivity ratios of Skolt Saami and its neighbouring Finnic languages can be at least partly explained by the fact that many of the verbs that are transitive in Skolt Saami, in Finnic languages have their second argument in the partitive, which is almost never used for argument encoding in modern Saamic languages. Skolt Saami also patterns with Finnic languages in the organization of certain specific valency classes, such as the LOC_NOM class comprising verbs denoting some sort of possession. Interestingly, in this particular case Skolt Saami also shows clear similarities with Russian, a language with which it has been in close contact for centuries. Russian is also close to Skolt Saami in its

relative entropy of intransitives, along with some other Uralic languages that it is known to have influenced. However, a relatively high value of $H_{\text{rel}(\text{intr})}$ in Skolt Saami (0.59) can also be explained by language-internal reasons, namely by the fact that the language employs numerous spatial markers, such as locative cases and postpositions, to code argument relations.

The data collected for this study already allow for quite an extensive analysis. Yet, it would be possible to draw more stable conclusions and possibly confirm some issues if more speakers filled out the questionnaire. Another way to develop the findings and put them into a regional perspective would be to compare results of this study to data on a wider variety of Saamic languages. Within this and other studies, applying Say's (2018c: 593) Mutual Information profiling would allow us to better assess similarities and differences between languages in regard to their valency systems. Systematic comparison of Saamic languages to their non-related contact languages, such as Russian and North Germanic languages, would shed more light on the areal aspect of some observed phenomena.

References

- Bickel, Balthasar, Taras Zakharko, Lennart Bierkandt & Alena Witzlack-Makarevich. 2014. Semantic role clustering: An empirical assessment of semantic role types in non-default case assignment. *Studies in Language* 38(3). 485–511.
- Blinova, Alena & Ksenia Shagal. 2020. Dvuxmestnye predikaty v saamskix jazykax v areal'noj perspective. In Ksenija P. Semenova (ed.), *Malye jazyki v bol'shoj lingvistike: Sbornik trudov konferencii*, 25–33. Moskva: Buki Vedi.
- Blokland, Rogier & Michael Rießler. 2011. Komi-Saami-Russian contacts on the Kola Peninsula. In Cornelius Hasselblatt, Peter Houtzagers & Remco van Pareren (eds), *Language contact in times of globalization* (Studies in Slavic and General Linguistics 38), 5–26. Leiden: Brill.
- Dowty, David. 1991. Thematic proto-roles and argument selection. *Language* 67. 547–619.
- Feist, Timothy. 2015. *A grammar of Skolt Saami* (Suomalais-Ugrilaisen Seuran Toimituksia 273). Helsinki: Suomalais-Ugrilainen Seura.
- Haspelmath, Martin. 2001. The European linguistic area: Standard Average European. In Martin Haspelmath, Ekkehard König, Wulf Oesterreicher & Wolfgang Raible (eds), *Language typology and language universals: An international handbook* (Handbooks of Linguistics and Communication Science 20/2), 1492–1510. Berlin: Walter de Gruyter.
- Haspelmath, Martin. 2015. Transitivity prominence. In Andrej Malchukov & Bernard Comrie (eds), *Valency classes in the world's languages*, 131–147. Berlin: Mouton de Gruyter.
- Kowalik, Richard. 2016. *Predicative possession in South Saami*. Stockholm: Stockholm University. MA thesis.
- Kroik, David. 2016. *Differential object marking in South Saami*. Umeå: Umeå University. Licentiate thesis.
- Malchukov, Andrej L. 2006. Transitivity parameters and transitivity alternations: constraining co-variation. In Leonid Kulikov, Andrej L. Malchukov & Peter de Swart (eds), *Case, valency and transitivity* (Studies in Language Companion Series 77), 329–359. Amsterdam: John Benjamins.

- Matras, Yaron & Jeanette Sakel. 2007. Introduction. In Yaron Matras & Jeanette Sakel (eds), *Grammatical borrowing in cross-linguistic perspective* (Empirical Approaches to Language Typology 38), 1–13. Berlin: Mouton de Gruyter.
- Moshnikoff, Satu. 1992. *Maaddârää'jji mainnâz*. Utsjoki: Girjegiisä Oy.
- Narrog, Heiko & Shinya Ito. 2007. Re-constructing semantic maps: the comitative-instrumental area. *Sprachtypologie und Universalienforschung* 60 (4). 273–292.
- Ovsjannikova, Maria. 2018. Kodirovanie učastnikov pri dvuxmestnyx predikatax v lesnom dialekte èneckogo jazyka. *Ural-Altaic Studies* 28(1). 49–68.
- Rießler, Michael. 2009. Loanwords in Kildin Saami, a Uralic language of northern Europe. In Martin Haspelmath & Uri Tadmor (eds), *Loanwords in the world's languages: A comparative handbook*, 384–413. Berlin: Mouton de Gruyter.
- Salminen, Tapani. 2002. Problems in the taxonomy of the Uralic languages in the light of modern comparative studies. In Tat'jana B. Agranat & Ol'ga A. Kazakevič (eds), *Lingvističeskij bespredel. Sbornik statej k 770-letiju A. I. Kuznecovoj*, 44–55. Moskva: Izdatel'stvo Moskovskogo universiteta.
- Sammallahti, Pekka. 1998. *The Saami languages: An introduction*. Karasjok: Davvi Girji.
- Say, Sergey. 2014. Bivalent verb classes in the languages of Europe: A quantitative typological study. *Language Dynamics and Change* 4. 116–166.
- Say, Sergey (ed.). 2018a. *Valentnostnye klassy dvuxmestnyx predikativ v raznostrukturnyx jazykax*. Saint Petersburg: ILI RAN.
- Say, Sergey. 2018b. Valentnostnye klassy dvuxmestnyx predikativ: teoretičeskij kontekst, zadači issledovanija i struktura sbornika. In Sergey Say (ed.), *Valentnostnye klassy dvuxmestnyx predikativ v raznostrukturnyx jazykax*, 5–24. Saint Petersburg: ILI RAN.
- Say, Sergey. 2018c. Markirovanie aktantov dvuxmestnyx predikativ: predvaritel'nye itogi tipologičeskogo issledovanija. In Sergey Say (ed.), *Valentnostnye klassy dvuxmestnyx predikativ v raznostrukturnyx jazykax*, 557–616. Saint Petersburg: ILI RAN.
- Say, Sergey (ed.). 2020-. *BivalTyp: Typological database of bivalent verbs and their encoding frames*. Saint Petersburg: Institute for Linguistic Studies, RAS. (Available online at <https://www.bivaltyp.info>, Accessed on 22 June 2021.)
- Say, Sergey, Dmitry Gerasimov, Sergey Dmitrenko, Natalia Zaika, and Viktor Xrakovskij. 2018. Valentnostnye klassy dvuxmestnyx predikativ: tipologičeskaja anketa i instrukcija issledovatelju. In Sergey Say (ed.), *Valentnostnye klassy dvuxmestnyx predikativ v raznostrukturnyx jazykax*, 45–46. Saint Petersburg: ILI RAN.
- Shagal, Ksenia & Alena Blinova. 2020. Bivalent patterns in Skolt Saami. In Sergey Say (ed.), *BivalTyp: Typological database of bivalent verbs and their encoding frames*. St. Petersburg: Institute for Linguistic Studies, RAS. (Data first published on August 23, 2020; last revised on August 23, 2020.) (Available online at <https://www.bivaltyp.info>, Accessed on 16 June 2021.)
- Shagal, Ksenia & Sergey Say. 2018. *Bivalent verb classes in Uralic languages*. Talk given at Kielitieteen päivät XLV, Helsinki, 8 May 2018. Available online at https://bivaltyp.info/docs/Shagal_Say_2018_Bivalent_Uralic.pdf (Accessed on 24 February 2021).

On the origin of South Saami *edtjedh* ‘shall, will’

Torbjörn Söder, Uppsala University

1. Introduction

The South Saami verb *edtjedh* ‘shall, will’ belongs to a set of auxiliaries, the members of which typically accompany a lexical verb to add modal meaning to a clause. Other auxiliaries of this category express meanings such as obligation: *byöredh*, *sjödtedh*; ability and possibility: *maehtedh*; desire: *sijhtedh*; necessity: *tjoeredh*, *tjoeveridh*; permission: *åadtjodh*, *fihkedh*.

Like the other auxiliaries, *edtjedh* typically appears in combination with a verb in the infinitive, which it most commonly precedes. Its meaning usually applies to an action or event, which is posterior to the time of reference of the sentence. Depending on context, the verb may express different modal meanings. It may express, as in example (1), a deontic meaning:

- (1) *Datne edtjh darjodh guktie manne.*
you shall.PRS.1SG do.INF how I
‘You shall do as I do.’ (SaaS, Mattsson Magga 2016: 119)

In this article, however, the capacity of the verb to express future, as either intention or purpose, is in focus. It is in this capacity the verb displays interesting correspondence with the supine, which has been attested in Ume, Pite, Lule and North Saami (Korhonen 1974: 197), and the potential, which has been attested in all Saami languages (Korhonen 1967: 97).

In Saami linguistics, the supine refers to an infinite verb form which typically occurs in combination with a finite verb to express purpose, as in example (2), where the supine *boražit* (< *borrat* ‘eat’) denotes the purpose of the departure. The potential refers to specific finite verb forms, which can express uncertainty, but which also, as in example (3), may express future.

- (2) *Sii vulge dohko boražit.*
 they leave.PST.3PL there eat.SUP
 ‘They went there in order to eat.’ (SaaN, Svonni 2019: 128)
- (3) *Vuobditjá má idet bijlav?*
 sell.POT.2SG probably tomorrow car.ACC
 ‘You will sell the car tomorrow, right?’ (SaaL, Tuolja & Angéus Kuoljok 2013: 151)

The correspondence between *edtjedh*, the supine and the potential is not only semantic, but there also exists a clear phonetic correspondence. This article paper promotes the idea that *edtjedh* shares a common origin with both the supine and the potential, and that *edtjedh* has evolved through a process of degrammaticalization, as the Finno-Saamic¹ suffix **-ηsí²* in South Saami and adjacent varieties transformed into an auxiliary. Degrammaticalization is a relatively uncommon phenomenon in general, but it is not unheard of in Saami. Saami has even played an important role in the degrammaticalization debate (Nevis 1986; Norde 2009: 207–209; Ylikoski 2016b). Already in 1992, Bergsland suggested the etymological affinity between *edjedh* and the supine (Bergsland 1992: 11). Ylikoski also discussed the possible affinity, including the potential, in a lecture held in Levanger in 2019³.

The article is structured as follows: Section 2 and 3 outline the semantic features of *edtjedh* pertaining to intention and purpose respectively. Section 4 discusses morphological and phonetic features of the verb. Section 5 is a survey of intentional and purposive expressions in North, Lule, Pite and Ume Saami. Section 6 presents a discussion about the affinity between the Supine and the potential, whereas Section 7 contains a discussion about the origin of *edtjedh* and some concluding remarks.

2. *Edtjedh* in expressions with intentional meaning

In the framework of this article, intention refers to a commitment to perform an action, as in *I will go to town*. In this meaning *edtjedh*, as in example (4), appears in combination with an infinitive in a main clause (Bergsland 1994: 72), but there are also instances where the verb alone constitutes the predicate

¹ For a discussion about the notion Finno-Saamic, see Aikio (2012).

² Korhonen (1981: 252), in addition, reconstructs a form with palatal dental affricate: **-ηcí*, whereas Sammallahti (1998: 213) and Aikio (2012: 68) have **-ηsí*.

³ I would like to thank Jussi Ylikoski for sharing parts of his manuscript from the lecture and for drawing my attention to Bergsland’s article. I’m also grateful to Torgny Hedström and Olle Kejonen for valuable comments.

of the clause (5). The commitment not to perform an action is expressed by negating *edtjedh* (6).

- (4) *Maam* *edtjebe* *gaskebeajjan* *byöpmédidh?*
 what.ACC will.PRS.1PL dinner.ILL eat.INF
 ‘What will we have for dinner?’ (SaaS, NDB s.v. *edtjedh*)
- (5) *Gåabph* *edtjh* *datne?*
 where will.PRS.2SG you
 ‘Where are you going?’ (SaaS, NDB s.v. *edtjedh*)
- (6) *Im* *edtjh* *dejnié* *vajmedh.*
 NEG.PRS.1SG shall.CONNEG that.COM lament.INF
 ‘I shall not lament that.’ (SaaS, Kuvsje)

In this respect, the usage of the verb very much resembles the intentional usage of the Norwegian auxiliary *skulle* ‘shall, will’ (Faarlund et al. 1997: 604), and, in Swedish, the auxiliary *skola* (cf. SAG II 283). This is illustrated by the Norwegian translations of the examples above, where *skal* (< *skulle*) appears as counterpart to forms of *edtjedh*, both when combined with a main verb or functioning as the predicate alone (5). In the latter example, an implicit verb of motion could be assumed, just as the case is in Norwegian and Swedish (SAG IV 906; cf. Faarlund et al. 1997: 527). The status of *edtjedh* as a verb of intention is also highlighted in the translation of the Norwegian verb *skulle* ‘shall, will, should, ought to’ in Mattsson Magga’s Norwegian–South Saami dictionary (NSO s.v. *skulle*). The dictionary mentions *edtjedh* in combination with Norwegian clauses that express intention: *jeg skal dra (dra i vei)* ‘I will leave’ (7), *vi to skal dra, la oss dra* ‘let us two go, let’s go’ (8), *han sa han skulle hjelpe til* ‘he said that he would help’ (9). The deontic and epistemic meanings of the Norwegian verb, on the other hand, have translations that include *galkedh* ‘should, must’ and *byöredh* ‘ought to, should’ (ibid.). In his grammar of northern South Saami, Nilsson-Mankok (1975: 12–13) deviates partly from this analysis, ascribing the verb *edtjedh* ‘shall, will, must, must have the possibility to’ both intentional and deontic meaning, whereas *galkedh*⁴ ‘shall, will’ is treated as a temporal auxiliary with future reference.

⁴ The verbs have the forms *etjet* and *galket* in Nilsson-Mankok’s orthography.

- (7) *Edtjem* *vuelkedh.*
 shall.PRS.1SG leave.INF
 ‘I shall leave.’ (SaaS, NSO s.v. *skulle*)
- (8) *Edtjien* *vuelkedh.*
 shall.PRS.1DU leave.INF
 ‘We two shall leave/Let’s leave.’ (SaaS, NSO *ibid.*)
- (9) *Dihte jeehti, satne edtja viehkiehtidh.*
 he say.PST.3SG he.himself will.PRS.3SG help.INF
 ‘He said that he will help.’ (SaaS, NSO *ibid.*)

3. *Edtjedh* in expressions with purposive meaning

Besides the intentional usage, *edtjedh* also commonly appears in constructions that express the desired result of an action. This purposive meaning is illustrated in the sentence *I go to town in order to buy food*, where the subordinate construction (*in order to buy food*) denotes the purpose of the subject’s trip to town. In its purposive meaning, the verb *edtjedh* has a similar function as the subordinating conjunction *in order to*, which precedes the infinitive phrase in the English example above. In this function, the verb typically initiates the subordinate construction and combines with the main verb in the infinitive, as in examples (10) and (11). It should be pointed out that the two finite verb forms in the examples are not instances of conjunctionless coordination, which is common in South Saami (cf. example 12).

- (10) *Dihte vöölki edtja viehkiehtidh.*
 he leave.PST.3SG will.PRS.3SG help.INF
 ‘He went away in order to help.’ (SaaS, NSO s.v. *for*)
- (11) *Jih dle vuelkieh edtjieh minnedh*
 and then leave.PRS.3PL will.PRS.3PL go.INF
håagkodh.
 fish.INF
 ‘And they leave in order to go fishing.’ (SaaS, Bergsland 1994: 82)
- (12) *Dah tjåadtjoeh vuartasjeh.*
 they stand.PRS.3PL watch.PRS.3PL
 ‘They stand and watch.’ (SaaS, Jaahkenelkien 1993: 30)

The ability of *edtjedh* to express both intention and purpose is quite salient. Already Lagercrantz (LW 3) points out that *edtjedh* has these two functions and gives several examples of both, using the German verb *sollen* ‘shall, will’ as equivalent in the former function and the subordinate conjunctions *um zu* ‘in order to’ and *damit* ‘idem’ in the latter.⁵

In South Saami, negative purpose may be also expressed with an asyndetic clause. Such constructions include the imperative forms of the negative verb (Bergsland 1994: 82). The negative verb has two imperative forms, based on the stems *aellie-* and *olle-*, the former of which expresses prohibition and the latter a warning (Bergsland 1994: 44). Both may express negative purpose, as in (13) and (14). Purpose may also be expressed by means of the subordinate conjunction *juktie* (15).

- (13) *Tjoerh* *geehtedh* *aellieh* *tjaajenh*.
must.PRS.2SG watch.out.INF NEG.IMP.2SG get.lost.CONNEG
‘Watch out lest you get lost.’ (SaaS, Bergsland 1994: 82)
- (14) *Manne* *raajem* *varke* *olkese,* *olles*
I run.PRS.1SG quickly out NEG.IMP.3SG
dihite *båvvesh*.
it perish.CONNEG
‘I run out quickly [to take care of the calf] lest it dies.’ (SaaS, Mattsson Magga 2016: 114)
- (15) *Bööremes* *liegkedidh* *juktie* *barkose*
good.SUPERL rest.INF in.order.to work.ILL
madtjeles *sjidtedh*.
prepared become.INF
‘One had better rest in order to get prepared for work.’ (SaaS, NDB s.v. *juktie*)

4. The verb *edtjedh* and its forms

The verb *edtjedh* (Table 2) displays features of a parisyllabic, or, to be more specific, a disyllabic verb. South Saami disyllabic⁶ verbs are usually divided

⁵ For a generative approach on *edtjedh* and “as the exponent of subjunctivity” in null subject clauses see Vinka (2022).

⁶ I will henceforth use the term disyllabic, since it is the far most common type of parisyllabic verbs and since it is more relevant in the discussion about *edtjedh*.

into six groups (I–VI). The division is based on the quality of the vowel in the second syllable and refers to rows I–IV in the umlaut table (see Table 1). Verbs of group I, II and IV regularly undergo umlaut, as opposed to verbs of the other groups. The second syllable vowel of a disyllabic verb corresponds to one of the vowels in the second column from the right in table. The far-right column shows how these vowels are realized in trisyllabic feet. The indexes 4–6 refer to vowels that, for historical reasons, induce umlaut in different ways. The columns 1–7 refer to umlaut series and present different qualities that a first syllable vowel of a word possibly can have. The South Saami umlaut has occurred as an instance of regressive assimilation: the quality of the first syllable vowel is dependent on the quality of the second syllable vowel.

Table 1. Umlaut in South Saami, based on Magga & Magga (2012: 27; superscripts refer to the groups).

	Umlaut series							Vowel of the second syllable	Vowel of the second syllable of trisyllabic feet
	1	2	3	4	5	6	7		
I	<i>i</i>	<i>e</i>	<i>u</i>	<i>ie</i>	<i>ae</i>	<i>äe</i>	<i>ue</i>	<i>ie</i>	<i>e</i> ¹ or <i>i</i> ¹
II	<i>ä</i>	<i>a</i>	<i>å</i>	<i>ea</i>	<i>aa</i>	<i>åa</i>	<i>åa</i>	<i>oe</i>	<i>o</i>
III	<i>ä</i>	<i>a</i>	<i>å</i>	<i>ea</i>	<i>aa</i>	<i>åa</i>	<i>ua</i>	<i>a</i>	<i>e</i> ³ or <i>i</i> ³
IV	<i>ĩ</i>	<i>a [â]</i>	<i>o (å)</i>	<i>ĩe</i>	<i>aa</i>	<i>åa</i>	<i>oe</i>	<i>e</i> ⁴ or <i>i</i> ⁴	<i>e</i> ⁴ or <i>i</i> ⁴
V	<i>y</i>	<i>o</i>	<i>o</i>	<i>yö</i>	<i>åå</i>	<i>åå</i>	<i>åå</i>	<i>e</i> ⁵	<i>e</i> ⁵
VI	<i>i</i>	<i>i [ĩ]</i>	<i>u</i>	<i>ee</i>	<i>ee [ææ]</i>	<i>öö</i>	<i>öö</i>	<i>e</i> ⁶ or <i>i</i> ⁶	<i>e</i> ⁶ or <i>i</i> ⁶

The initial vowel of the infinitive *ed tjedh* (phonotactically trisyllabic), the present tense 1DU *ed tjien*, 2DU *ed tjeden*, 1PL *ed tjebe*, 2PL *ed tjede*, 3PL *ed tjieh* would suggest that the verb, like the verb *ved tedh* ‘give’ in Table 3, belongs to group I. However, the present tense 1SG *ed tjem*, 2SG *ed tjh*, 3SG *ed tja* show that the verb, unlike verbs in group I, is not subject to umlaut, and the present tense 2DU *ed tji jidien*, 3DU *ed tji jægan*, 1PL *ed tji jibie*, dissimilarly to group I verbs, do not have *-ie-* in the second syllable (cf. Table 2 and 3). Thus, the verb *ed tjedh* appears to be irregular. NDB defines it as belonging to verb group IV (s.v. *ed tjedh*), which seems correct judging from the personal suffixes. The South Saami verb handbook *Verbh* (Frändén et al 2007: 11, 26–27), on the other hand, lists it as an irregular verb, due to its lack of umlaut. Other verbs of group IV, with a first syllable vowel belonging to column 2 (see Table 1), regularly display an *a* in the first syllable of the infinitive, and of all present tense forms, except for the 1st person dual and the 3rd person plural forms (Magga & Magga 2012: 23), cf. *ed tjedh* and *barkedh* in Table 2 and 4.

Table 2. The present tense forms of the verb *edtjedh* ‘shall, will’.

	singular	dual	plural
1 st person	<i>edtjem</i>	<i>edtjien</i>	<i>edtjebe/edtjijibie</i> ⁷
2 nd person	<i>edtjh</i>	<i>edtjeden/edtjijidien</i>	<i>edtjede</i>
3 rd person	<i>edtja</i>	<i>edtjijægan</i>	<i>edtjieh</i>

Table 3. The present tense forms of the verb *vedtedh* ‘give’.

	singular	dual	plural
1 st person	<i>vadtam</i>	<i>vedtien</i>	<i>vedtebe/vedtiejibie</i>
2 nd person	<i>vadtah</i>	<i>vedteden/vedtiejidien</i>	<i>vedtede/vedtiejidie</i>
3 rd person	<i>vadta</i>	<i>vedtiejægan</i>	<i>vedtieh</i>

Table 4. The present tense forms of the verb *barkedh* ‘work’.

	singular	dual	plural
1 st person	<i>barkem</i>	<i>berkien</i>	<i>barkebe/barkijibie</i>
2 nd person	<i>barkh</i>	<i>barkeden/barkijidien</i>	<i>barkedede/barkejidie</i>
3 rd person	<i>barka</i>	<i>barkijægan</i>	<i>berkieh</i>

The standard orthographic form of the infinitive *edtjedh* suggests the pronunciation [et̪ʰ:ət̪], with a short initial vowel followed by a long affricate. In records where non-standard forms of the verb appear, there are instances where the phonetic forms of the verb differ quite significantly from the standard. Lagercrantz (LW 2–3), Hasselbrink (SW s.v. *addjeb*) and Halász (1891 s.v. *ač̣a*) report a variation between front and back vowels in the first syllable and a variation of quantity in the stem consonant. In *Sámien lukkeme-gärjá* from 1957, written in the former South Saami orthography, the verb is given an initial back vowel, e.g., *ažžem* ‘I will’ (Bergsland & Hasselbrink 1957: 21). Hasselbrink mentions that the initial vowel often is reduced to [ə]. In Andersson’s *Maam dah soptsestamme* (1992), a collection of South Saami tales from the 1940s, the forms of the verb are in the majority of cases even transcribed without the initial vowel at all, as in examples (16) and (17), which underlines the reduced nature, or even lack, of the initial vowel. This feature appears regardless of whether a vowel or a consonant precedes the auxiliary. Halász (1891: s.v. *č̣a*) also reports forms lacking an initial vowel.

⁷ Disyllabic verbs may have a trisyllabic stem, which includes *-j-*, in present tense 1PL, 2DU and 2PL. Forms appearing in the table have been retrieved from NDB.

- (16) *Nov tjem manne dijem vuesiehtidh.*
 so will.PRS.1SG I you.ACC.PL show.INF
 ‘So I will show you.’ (SaaS, Andersson 1992: 31)
- (17) *Manne tjoekim dellie dahkoe jih*
 I ski.PST.1SG then there and
kroehkim tjim veartasjidh, mij
 dig.PST.1SG will.PST.1SG see.INF what
lij dennie boernesne.
 be.PST.3SG that.ATTR.INE storage.pit.INE
 ‘Then I skied there and dug in order to see what was in the storage pit.’
 (SaaS, Andersson 1992: 27)

Lagercrantz mentions that the verb is “structurally weakened (Ger. *strukturell abgeschwächt*)” (LW 2) and that its consonant center is not realized as a full geminate, which would otherwise be expected in a position after a short vowel. Lagercrantz and Hasselbrink mention that the stem consonant either appears as a half-long or as a short fricative and indicate that this variation is due to the verb’s position and prominence in the clause, seemingly implying that it is shorter when it appears as an enclitic. The view that the verb may appear as an enclitic finds support in Lagercrantz (1923) and in SW, where it is transcribed as an enclitic attached to pronouns. Lagercrantz (61) gives the form *mannatjeb*⁸ ‘I will’ and Hasselbrink *mannadtjem* (s.v. *addjeb*). In the standard South Saami orthography, the sequence is rendered *manne edtjem* I will.PRS.1SG ‘I will’ (See also Korhonen 1967: 107). The enclitic-like nature of the verb also appears in combination with the negative verb, where the two verbs may form a phonological word: *immetjh* (LW 182), *immotj(h)*⁹ (Collinder 1943: 47), cf. standard SaaS *im edtjh* NEG.PRS.1SG shall.CONNEG ‘I shall not’. See Halász 1891: 94 s.v. *ča*, for more examples of the enclitic use of *edjedh*.

Interestingly, both Lagercrantz and Hasselbrink mention that the verb lacks an infinitive. The usage of the infinitive of modal auxiliaries is *per se* limited, since auxiliaries typically appear in finite forms and combine with the infinitive of another verb. However, for the modal auxiliaries *galkedh* ‘shall, should’ and *tjoeveridh* ‘must’, both Lagercrantz and Hasselbrink mention

⁸ The transcriptions of Lagercrantz and Hasselbrink have been adapted to the standard South Saami orthography. The final bilabial stop in Lagercrantz’s example, as opposed to *-m* in the standard orthography, reflects a feature of northern South Saami.

⁹ Collinder’s example is from Härjedalen. The transcription has been adapted to the standard orthography. His example does not reflect the final *h* of standard South Saami connegatives, hence the parenthesis.

infinitives (SW s.v. *galgedh*, *tjuoveridh*; LW: 236, 113). However, on page 143 in SW Hasselbrink mentions the infinitive *addjedh*. In modern South Saami dictionaries, the headword is always the infinitive *edjedh* (NDB s.v.; ÅDB s.v.; SS s.v.). A search for the lemma *edjedh* in the SIKOR database indicates that the usage of the infinitive still today is very limited: out of more than 10 000 occurrences, less than 10 are in the infinitive.

5. Intentional and purposive expressions in North, Lule, Pite and Ume Saami

5.1 Intentional expressions

In North, Lule and contemporary Pite Saami, there seem to be no attested cognates to *edjedh* among the auxiliary verbs. Instead, the North Saami verbs *áigut* ‘will, intend’ and *galgat* ‘have to, must, shall’, and its cognates in Lule, *ájggot*, *galggat* (LLO s.v.) and Pite, *ájggot*, *gallgat* (BB s.v.), appear to be the closest equivalents to the South Saami verb in its intentional meaning, cf. examples (18–20). Magga (1982: 21), who labels the North Saami verb *áigut* a dynamic epistemic auxiliary, mentions that the verb *galgat* may paraphrase *áigut* in some contexts, as in example (17), adding that, for *galgat*, a deontic reading often is possible along with the dynamic epistemic reading (Magga 1982: 28; cf. Nickel & Sammallahiti 2011: 430; Koskinen 1998: 43).

- (18) *Mun áiggun gilvit nisuid duon*
 I intend.PRS.1SG sow.INF wheat.ACC.PL that.ATTR.ILL
beldui.
 field.ILL

‘I will sow wheat on that field.’ (SaaN, NDS s.v. *áigut*)

- (19) *Gánda áigu vuolgit dan skiippa*
 boy intend.PRS.3SG leave.INF that.GEN ship.GEN
lusa borjjastit.
 to sail.INF

‘The boy intends to go to the ship to sail away.’ (SaaN, Magga 1982: 21¹⁰)

¹⁰ Magga’s example is from Qvigstad (1927: 340).

- (20) *Son maid galggai beassat oastit*
 she also will.PST.3SG get.INF buy.INF
juoidá.
 something.ACC
 ‘She would also get to buy something.’ (SaaN, Magga 1982: 28¹¹)

In Ume Saami the corresponding verbs are *ájgguot* ‘intend’, *galggat* ‘will’ (BG s.v.). In Ume Saami the verbs *ussjadit* ‘think, mean, intend’ and *adtjat* ‘shall, will’ (idem s.v.) also express intention. The latter is cognate with *edtjedh* in South Saami (21).

- (21) *Ij adtjh vuan.*
 NEG.PRS.3sg shall.CONNEG PTCL
 ‘He should not, should he?’ (SaaU, BG s.v. *adtjat*)

In addition to the verbs mentioned above, Lule Saami has the verb *gehkat* ‘intend’ (LLO s.v. *kāhkat*) as a verb of intention. In South Saami *galkedh*, the cognate of North Saami *galgat*, may also express intention (22). South Saami also has the noun *aajkoe* ‘intention’, which is etymologically related to North Saami *ájgut* etc., and similarly to Ume Saami *ussjadit* the South Saami cognate *ussjedidh* ‘think’ may also be used to express intention (23).

- (22) *Manne galkem oktegh vaedtsedh sjaltan.*
 I shall.PRS.1SG alone walk.INF village.ILL
 ‘I shall go to the village alone.’ (SaaS, Mattsson Magga 2016: 10)

- (23) *Datne aaj ussjedh sjaltan vuelkedh?*
 you also think.PRS.1SG village.ILL go.INF
 ‘Are you also thinking of going to the village?’ (SaaS, Mattsson Magga 2016: 21)

In Lagercrantz’s records of Pite Saami from the early 20th century, there is another verb with intentional meaning (24). The verb lacks an attested infinitive in his records. He refers to it by the form *atjav* (1926: 24), which is the first person singular of the present tense. Apart from the German equivalent *sollen* ‘shall, will’, he also formulates translations such as “muss ich” (‘must I’) and “ich möchte” (‘I would like’) in his examples. Lehtiranta (1992) consistently uses the Finnish equivalent *täytyä* ‘must’ in his

¹¹ Magga’s example is from Jernsletten & Hatle (1975: 70).

translations of the verb¹², which, judging from Lagercrantz’ translations, seems a little simplistic. The form and function of the verb indicate an etymological relation to *edjedh* and *adtjat* in South and Ume Saami (see Section 5.2).

- (24) *Mij atjap muätkiev giesset.*
 we will.PRS.1SG isthmus.ACC pull.INF
 ‘We will pull [the boat] across the isthmus’ (SaaP, Lagercrantz 1926: 24)

Before leaving the discussion, it is worth mentioning that lexical verbs in the present tense also may express intention, as in example (25) from North Saami; for the future reference of the potential, see Section 6.

- (25) *Goas don finat min geahčen?*
 when you visit.PRS.2SG we.GEN at
 ‘When will you come and visit us?’ (SaaN, Nickel & Sammallahti 2011: 270)

5.2 Purposive expressions

North, Lule, Pite and Ume Saami altogether display several ways of expressing purpose. The constructions may include the supine (26–27, 32, 35), a verb of motion and an infinitive (30–31), an auxiliary verb (33, 34) or a subordinate conjunction (36–39). For this paper, constructions with the supine and constructions including an auxiliary are of the most interest.

- (26) *Sii vulge dohko boražit.*
 they leave.PST.3PL there eat.SUP
 ‘They went there (in order) to eat.’ (SaaN, Svonni 2019: 128)

- (27) *Ulmütja bāhtin suv gehtjatjit*
 people.PL come.PST.3PL he.ACC watch.SUP
ja gulldalittjat.
 and listen.SUP
 ‘People came in order to look at and listen to him.’ (SaaL, Tuolja & Angéus Kuoljok 2013: 147)

¹² In one example, which Lehtiranta has from Halász (1893: 276), he gives the Finnish equivalent *jotta* ‘in order to’ (Lehtiranta 1992: 95).

In Lule Saami, where the supine is a common feature, the supine is, in the standard language, marked by the suffix *-tjit* [tʃit] in parisyllabic and contracted stems, and *-ittjat* [itʃ:at] in imparisyllabic stems. In the spoken language, however, the quality of the final and penultimate vowel of imparisyllabic verbs vary, e.g., *málestattjit* ~ *málestattjat* ~ *málestittjat* (LLO 1886; cf. also Kejonen, this volume). In Lule Saami, the usage is not restricted to the purposive use, as in example (27), but the supine may also appear in constructions that convey other non-purposive meanings (Tuolja & Angéus Kuoljok 2013: 147; cf. Spiik 1989: 106). In example (28) the supine functions as a modifier to the adjective and corresponds to expressions in English like *be* adj *at* + continuous form. It may also appear in constructions together with a finite form of the verb *liehket* ‘be’. These constructions have future reference and may appear in intentional expressions (Ylikoski 2016a: 221–226), as in example (29).

- (28) *Áhpaddje* *lij* *tjuorbbe* *tjielggittjit*.
 teacher be.PST.3SG bad explain.SUP
 ‘The teacher was bad at explaining.’ (SaaL, Tuolja & Angéus Kuoljok 2013: 147)

- (29) *Dánna* *la* *gus* *dân* *tjähkkåhittjat?*
 here be.PRS.2SG Q you sit.SUP
 ‘Do you intend to sit here?’ (SaaL, *ibid.*)

In North Saami, where the use of the supine is limited to the western dialects (Collinder 1949: 51; Svonni 2019: 128; Nickel & Sammallahti 2011: 68; for the negative supine, see below), the supine seems to be restricted to purposive expressions, as in example (26). In standard North Saami the supine suffix is rendered *-žit* [tʃit] in parisyllabic and contracted stems, and *-eažžat* [itʃ:at] in imparisyllabic stems (Svonni 2019: 129). In many North Saami dialects, the supine of other verbs than the negative verb is not in use (Korhonen 1974: 197–198, 204). Instead, the infinitive, as in example (30), has the same function (cf. examples 50–51). This type of construction is also possible in South Saami (31).

- (30) *Son* *bodii* *sisa* *biepmu* *viežžat*.
 he come.PST.3SG in food.ACC fetch.INF
 ‘He came in to pick up food.’ (SaaN, Nickel & Sammallahti 2011: 301)

- (31) *Dihte jååktan vöölki ryöjnesjidh.*
 he yesterday leave.PST.3SG herd.INF
 ‘He went off to herd [the reindeer].’ (SaaS, Magga & Mattsson Magga 2012: 183)

From Pite and Ume Saami there are some records of the supine. According to Lehtiranta the usage is rare (1992: 103) and he mentions the possibility that forms identified as the supine *giehtjatjit* (< *gähtjtat* ‘see’) actually could be the infinitive of diminutive verbs, cf. *giehtjatjit* ‘see a little’. The rareness, or the possible lack, is also emphasised by the fact that Wilbur’s grammar on contemporary Pite Saami (2014) does not mention the supine at all (cf. however Korhonen 1974: 197; Sjanggo 2015: 18, 21). Lehtiranta mentions a construction (32) which indicates that the supine in Pite, as in Lule Saami, not only appears in purposive constructions, when in use.

- (32) *Giesátjit lin māj dajt.*
 pull.SUP be.PRS.1DU we.two those.ACC
 ‘We two must pull those.’ (SaaP, Lehtiranta 1992: 103)

Korhonen (1974: 198) defines the southern border of the supine as the border between Pite and Ume Saami, meaning that the single mention of the supine *guulahtjit* in (Schlachter 1958 XVIII), with its deviating pre-aspirated affricate as marker, indicates that the supine was not in use when Schlachter made his recordings in the 1950s. Calleberg’s recordings of Ume Saami from the early 1920s also display deviating forms. From Maskaur (ULMA 2860) he records the supine *lūčkätit* (< *lūčkät* ‘read’, cf. standard SaaU *luhkkat*) with the Swedish translation ‘för att läsa’, and from Malmesjaur (ULMA 2966) he noted *lūčkätum* followed by Swedish “läst” in parenthesis. The former translation, ‘in order to read’, is in line with the semantic properties of the supine in Lule and North Saami, whereas *läst* renders a form which according to standard Swedish linguistic terminology is called *supinum* (≠ supine; SAG I 228), which functionally correspond to the perfect participle in Saami. On the other hand, Ume Saami, in line with South Saami, has the verb *adtjat* ‘shall, will’, which typically initiates purposive subordinate constructions (33). The verb is attested with a full paradigm (*ätšäü* ‘in order for me to’ etc.) in Calleberg’s recordings from Malmesjaur (ULMA 2966).

- (33) *Adtjuv mænna dájđđiet.*
 will.PRS.1SG I know.INF
 ‘In order for me to know.’ (SaaU, BG s.v. *adtjat*)

A cognate verb is also reported from Pite Saami in the early 20th century. Lagercrantz (LW 3; cf. Lagercrantz 1926: 122) mentions, with reference to his records from Arjeplovve (Swe. Arjeplog; see example 34), that the verb, like its South Saami cognate, displays variation between a longer and a shorter stem consonants. According to Lagercrantz’s transcription, the forms presented for Pite Saami all have half short stem consonants (rendered as a single affricate in example 34). The past tense form *ittij* (shall.3SG) in example (34) only contrasts with the corresponding form of the negation verb in that the stem consonant is shorter, cf. *ittij* NEG.PST.3SG (cf. LW 183). The form *ačai* (PST.3SG) in Halász 1893 (276) translated by means of the Hungarian conjunction *hogy* ‘that’ (idem 277), refers to the same word (cf. Lehtiranta 1992: 92).¹³

- (34) *Dä härrá budij ittij suv*
 Then priest come.PST.3SG shall.PST.3SG he.ACC
vuojjelit.
 chase.away.INF
 ‘Then the priest came in order to chase him away.’¹⁴ (SaaP, LW 3)

Lagercrantz also treats the verb *atjav* (PRS.1SG) in connection to what he calls “optative clauses” (1926: 22–23). He states that the suffix *-atja-*, which appears in such constructions with an optative particle (Ger. *Hilfswort*), originate (Ger. *zurückgeht*) in the verb *atjav*. Lehtiranta (1992: 88), on the other hand, identifies the examples given by Lagercrantz, e.g., *råkiertitjip-mi* ‘plane down’, as forms of the potential with enclitic pronouns. According to Lagercrantz, an optative meaning may also be expressed without an optative particle, with forms such as *mujjtetatjav* ‘I will probably remember it, I would like to remember it’¹⁵ (< *mujjtet* ‘remember’). Such forms are according to Lehtiranta (ibid.) the result of agglutination, involving the verb *atjav*.

In addition to the supine discussed above, there is also a negative supine, which expresses the purpose of an action to avoid something (Nickel & Sammallahti 2011: 301). In North Saami, the negative supine is based on the stem *ama-*, to which the possessive suffixes are attached, thus yielding a paradigm of nine different forms (Svonni 2019: 129). The negative supine

¹³ In the Pite Saami text collection *Alvos giehto* (Halász 2023), a selection of tales from Halász’s *Svéd-lapp nyelv 5*, the editors Inger Fjällås och Peter Steggo have substituted forms of the verb *atja-* with forms of the verb *gallga-* ‘shall’, which indicates that the former verb nowadays is considered obsolete in Pite Saami (Kejonen 2024: 201).

¹⁴ Lagercrantz’s German translation is ‘Dann kam der Priester, (“er sollte”, d.h.) um ihn zu verjagen’.

¹⁵ Lagercrantz’s German translation is ‘ich werde wahrscheinlich, ich möchte: mich daran erinnern’. The examples given by Lagercrantz are only in the first person singular.

combines with an infinitive (31). The third person singular form, *amas*, has in some North Saami varieties been lexicalised as a subordinating conjunction (Korhonen 1981: 311–312; Sammallahti 1998: 87); in combination with the negative verb it is used to express negative purpose (36). Constructions based on *ama-* do not appear in the whole of the North Saami speech area; instead, combinations of the subordinate conjunction *vai* ‘in order to’, which is discussed below, and the negative verb are used (37).

- (35) *Leage* *várrugas* *go* *váccát*
 be.IMP.2SG careful when walk.PRS.2SG
amat *gahččat.*
 lest.2SG fall.INF
 ‘Be careful when you walk lest you fall.’ (SaaN, Nickel & Sammallahti 2011: 301)

- (36) *Vácce* *jaska,* *amas* *mánát*
 walk.IMP.2SG silently lest child.PL
eai hear.CONNEG
 NEG.PRS.3PL *gula!*
 ‘Walk silently lest the children hear.’ (SaaN, idem 204)

- (37) *Sii* *mannet* *stobu* *sisá* *vai*
 they go.PRS.3PL cottage.GEN into in.order.to
eai *galbmó.*
 NEG.PRS.3PL get.cold.CONNEG
 ‘They went into the cottage in order not to get cold.’ (SaaN, Svonni 2019: 129)

The only cognate of North Saami *ama-* in Lule Saami is possibly the subordinate conjunction *abma~abme* ‘lest’, which according to Harald Grundström (LLO s.v.) only appears in the literary language. In Pite Saami and in Ume Saami, the cognate subordinate conjunctions *amá* ‘so that you don’t’ (BB s.v.) and *ammá~abmá* ‘idem’ (BG s.v.; Schlachter s.v. *abmaa*), as in example (38), appear with a similar function. Lehtiranta (1992: 84) also has *vaj amaah*¹⁶ ‘in order not to’ from Pite Saami, where the subordinate conjunction *vaj* ‘so that’ is combined with *amá*.

¹⁶ Lehtiranta’s orthography.

- (38) *Ammá dâdna galgh bâhtiet.*
 lest you will.PRS.2SG come.INF
 ‘Lest you will come.’ (SaaU, BG s.v. *ammá*)

In Lule Saami purposive constructions may include the subordinate conjunctions *vaj* ‘in order to’, as in example (39) (Spiik 1989: 94; Tuolja & Kuoljok 2013: 232), or *ahte* ‘idem’ (LLO s.v.). The Lule Saami conjunction *vaj* has cognates with the same function in North, *vai* (Nickel & Sammallahti 2011: 403), and Ume Saami, *vaj* (BG s.v.; Schlachter s.v. *vai*). In contemporary Pite Saami, the conjunction *ahte* ‘in order to’ appears in purposive constructions (Wilbur 2014: 190). In Ume Saami, the cognate *ahttie* and in North Saami *ahte* may have purposive meaning (BG s.v.; Schlachter s.v. *àhtè*; Nielsen s.v. *âttě*). In addition, Lule Saami has *vaják* ‘in order to’ (Korhonen 2006: s.v.). Lehtiranta, who cites recordings made by Israel Ruong in 1939 and 1944, reports for Pite Saami both the form *vaj* (1992: 170; cf. Halász 1893: 136) and *vajaah* (1992: 181), of which the latter corresponds to Lule Saami *vaják*.

- (39) *Rabá uvsav, vaj mân*
 open.IMP.2SG door. ACC so.that I
besav âlgus.
 can.get.PRS.1SG out
 ‘Open the door so that I can get out.’ (SaaL, LLO s.v. *vai*)

5.3 The word *atja* in 17th and 18th century sources

In the grammar section of the Saami-Swedish-Latin dictionary *Lexicon Lapponicum* (LL), Öhrling treats the words *atja*, which also appears in the form *attja* (s.v.), and *abma* as subordinate conjunctions (LXXIX; cf. Korhonen 1981: 297). The words appear in combination with both personal verbal inflectional suffixes and possessive suffixes, but *atja*, as far as the singular and plural forms are concerned, combines to a greater extent than *abma* with verbal suffixes¹⁷. For *abma* only the forms in the second person singular, *abmah*, and plural, *abmabet* and *ammabet*, display verbal suffixes. For the second person singular the form *ammata*, possibly with a possessive suffix, is also mentioned. The word *abma* is clearly cognate with the Ume Saami conjunction *ammá~abmá*, and corresponding words in North, Lule and Pite Saami, and *atja*, with South Saami *edjtjedh*, Ume Saami *aditjat* and the

¹⁷ The comparison was made in relation to the possessive suffixes in *Lexicon Lapponicum* (XLVII–XLIX).

Pite Saami cognate (see Sections 5.1 and 5.2). The function and meaning of *atja*, as seen in (40), emphasize its functional and semantic relation to *edjtjedh* and *adtjat*.

- (40) *Hålå* *munji*, *atjab* *tetet*.
 speak.IMP.2SG I.ILL will.PRS.1SG know.INF
 ‘Tell me, so that I will know.’ (LL: LXXX)

Qvigstad (1947: 31–32), in his analysis of Olaus Stephani Graan’s *Manuale Lapponicum* (ML) from 1669, discusses the word *uggu-*, which he relates to *edjtjedh*¹⁸. From the examples given by Qvigstad, it appears that the word can express both purpose (41) and intention (42)¹⁹.

- (41) *Uggus* *sii* *denehied* *kalgeh*.
 will.IMP.3PL they serve.INF shall.PRS.3PL
 ‘In order that they serve.’ (Qvigstad 1947: 32)
- (42) *Kåbbagiam* *tij* *sijtewet* *uggum*
 who.of.two.ACC you.PL want.PRS.2PL will.PRS.1SG
mon *tiies* *luewes* *vuebde?*
 I you.ILL.PL free give.INF
 ‘Who of the two do you want me to release?’ (ibid.)

Graan’s *uggu-* must be cognate to *uggie-/uggiu-*, which occurs in Lars Rangius’s translation of the gospel of Mark from 1713. Wilson identifies the form as a verb, meaning ‘must, shall, will’ (2008: 108). Graan’s <gg> (Qvigstad 1947: 19) and Rangius’s <ggi> (Wilson 2008: 107) correspond to <dtj> in modern standard Pite, South and Ume Saami orthography. In both Graan’s and Rangius’s texts, the stems frequently combine with an *-s* when referring to a third person subject. This is also a feature of *atja* in *Lexicon Lapponicum*, where the form *atjas* appears (LXXIX). The *-s* is most probably identical with the third person singular possessive suffix, which, besides appearing in possessive constructions (cf. SaaP *áhttje-s* ‘her/his father’), also combines with verbal stems to form imperatives of the third person (cf. Sammallahti 1998: 79). In Saami third person imperatives can express an optative or permissive meaning (Wiklund 1915: 114; Korhonen 1981: 256; Nickel & Sammallahti 2011: 278; Svonni 2019: 106), and it seems the forms

¹⁸ The reference is made to the entry *ečem* ‘I will, shall’ in LW (2).

¹⁹ Qvigstad’s examples are from page 230 and 264 in ML. Qvigstad has omitted some words of the original text in the former example.

ending in *-s* (*uggus*, *uggies*, *atjas*) have a similar function in Graan's and Rangius's texts. Rangius also uses the third person imperative marker with other verbs, of which the forms *galges* (cf. SaaU *galggat* 'shall, will'), *allos* and *ollos* (both imperatives of the negative auxiliary) are the most frequent (Wilson 2008: 103). The use of *allos* and *ollos* corresponds to that of imperatives of the negative auxiliary in South Saami (see examples 13 and 14 in Section 3).

Fjellström (GL 56) mentions that the auxiliary *uggiub/agieb* (1SG) functions as a causal conjunction corresponding to Latin *ut* and *quo*, both of which the meaning is 'in order to, so that', as in example (43).

- (43) *Hweset* *agiab* *hweidnet!*
 show.IMP.2SG shall.PRS.1SG see.INF
 'Show so that I see!' (GL 56)

Fjellström gives a paradigm of all grammatical persons in singular and plural of both stem variants, *uggiu-* and *agie-/agia-*. The two different stems display variation concerning both vowel quality and consonant quantity. He does not mention how the stem variants relate to his division of Saami into a southern (Ume Saami) and northern (Pite Saami) dialect (Larsson 2018: 34), but one can establish that the latter variant corresponds to forms in Pite and Ume Saami which have been attested in the 20th century (see Section 5.2). Information from Halász 1891 possibly indicates that forms with an initial close back vowel may have a southern distribution. His examples are from Stensele (44), in the southern Ume Saami area, and from Frostviken (45), in the northern South Saami area (cf. the negated form *immoťj(h)* in Section 4). From the information in Halász's examples, it is impossible to discern any personal suffix, hence the glossing. Judging from the translations, *uč'*, like its cognates, has future reference, possibly with intentional meaning.

- (44) *Koapp* *uč'* *mennet?*
 where will go.INF
 'Where will you go?' (Halász 1891: s.v. *uč'*)
- (45) *Tuh* *uč'* *waccet.*
 they will walk.INF
 'They will walk.' (ibid.)

The *Lexicon Lapponicum* also displays forms that phonetically relate to Fjellström's *uggiu-*: *udťjub* 1SG, *udťjubet* 2PL, *udťjus* 3SG (LL s.v.). It also

presents the forms *ádtjob* and *ádtjos* (s.v.) with reference to *udtjus* and not to the following entry *ádtjot* ‘receive’. Fjellström defines forms ending in *-s* (*uggius*, *agies*) as third person plural forms, as opposed to Rangius, who has *uggius* both for third person singular and plural (Wilson 2008: 103).

6. The supine and the potential

Since the potential mood shares both morphological and semantic features with the supine (Korhonen 1981: 298; Sammallahti 1998: 222), the potential is of interest for the topic of this article. The potential in Saami, which is historically related to the conditional mood in Finnish (Sammallahti 1998: 83), is usually described as a mood which, in the view of the speaker, expresses possibility, assumption or uncertainty of an event (Nickel & Sammallahti 2011: 273; Svonni 2019: 103). Even though widely attested in Saami (Korhonen 1974: 97; Bartens 1980: 1–99), it is quite low in frequency, and in some varieties, if in use at all, it only appears in the copula (< PS **lea-* ~*li-* ‘be’), which also may function as a temporal auxiliary and an existential verb. In the *Lexicon Lapponicum* Öhrling uses the Swedish term *konjunktiv* ‘subjunctive’ when referring to the category. To the clause *Mon etsitjab* (I love.POT.1SG), he gives the Swedish translation ‘Jag må älska’ (I may love, I must love’; LL: LXVII), which reflects a way of expressing optative meaning in Swedish. Halász, following Lönnrot (1856: 156–57), labels the potential “Indefinitivus”, meaning that the term better signifies the category than the subjunctive, since “its value is indefinite” (Halász 1881: 52) and it commonly appears in imperative- or concessive-like constructions.

In Standard North Saami the marker is *-ž-* [tʃ] for parisyllabic verbs, *bođežan* ‘I will probably come’ (< *boahtit* ‘come’), and *-žž-* [tʃ:] for imparisyllabic verbs, *muitalažžá* ‘he probably will tell’ (< *muitalit* ‘tell’), which phonologically corresponds to the Lule *-tj-/-ttj-* and Pite Saami marker *-tj-/-dtj-* (cf. Sjaggo 2015: 20–33). Concerning contemporary Pite Saami, Wilbur (2014: 152–153) states that the potential mood is very little in use²⁰. According to Bartens (1980: 303) the use of the potential is limited south of Pite Saami. Schlachter (1958: XVIII) states that forms of the conditional and the potential mood are very difficult to obtain from his Ume Saami informant Lars Sjulsson. However, Calleberg (ULMA 2966) reports full paradigms of present tense potential from Malmesjaur in his Ume Saami recordings from 1922. In South Saami, instances of the potential mood are mostly attested (as

²⁰ In a personal comment Wilbur (2021-09-13) told me that this may be due to the limited dataset of his study.

in 46) with the verb *lea-* ‘be’ (SW: 143), but Bergsland (1994: 91) mentions the possibility to use the mood also with other verbs (47). Lagercrantz (1923: 7), from whom Bergsland probably copied his example, has several other examples of forms in the potential with different verbs. The use of the potential mood, however, seems to be very rare in modern South Saami. An indication of this is that NDB treats *ledtjie* as an adverb meaning ‘maybe, might, perhaps’ and not as the potential present of the third person singular.

- (46) *Dagke dihte ledtjie baaateme.*
 maybe he be.POT.3SG come.PTCP
 ‘Maybe he has come.’ (SaaS, Bergsland 1994: 91)

- (47) *Vööjnetje dle maam vööjnetje.*
 see.POT.3SG then what.ACC see.POT.3SG
 ‘He will see what he will see.’ (SaaS, *ibid.*)

The semantic correspondence between the potential mood and the supine has its basis in the categories’ future reference. This circumstance is most obvious if one compares the supine with potentials of the present tense. Events or actions denoted by the potential typically occur, or might occur, at a point in time that follows the time of utterance (cf. examples 48 and 49). By the use of a potential perfect construction, it is possible to place the action before the time of utterance, as in example (46). This phenomenon, however, has more to do with the properties of the perfect tense to put actions in the past, than of properties of the potential. The semantic component of uncertainty, however, remains in the perfect.

- (48) *Baláv da muv lastutja .*
 fear.PRS.1SG that.PL I.GEN willow.bark.PL
luvátji.
 get.wet.POT.3PL
 ‘I am afraid that my willow bark might get wet.’ (SaaL, Tuolja & Angéus Kuoljok 2013: 148)

- (49) *Na gal don dal oaččut*
 well surely you now get.PRS.2SG
dan beannot čuohtái, go juo
 it.ACC one.and.a.half hundred.ILL if already
válddežat.
 take.POT.2SG

‘You nevertheless get it for 150, if you would take it.’ (SaaN, Nickel & Sammallahiti 2011: 273)

The supine, at least in its most widely attested use (see examples 26 and 27 in Section 5.2), has future reference too, but, unlike the potential, it typically denotes a planned action in the future. As such, it does not explicitly denote uncertainty, but since the future always brings certain amount of uncertainty, one can ascribe uncertainty to supine constructions too. Dixon states that future time, because of the uncertainty of the future, may be shown by modalities within irrealis (2012: 7). Korhonen (1974: 200–201) illustrates this uncertainty of a supine construction (50), taken from Qvigstad (1929: 500), by comparing it to a corresponding infinitive construction (51). According to Korhonen, example (51) quite objectively states that the boy went off to fetch his mother, whereas in example (50), there exists the possibility that the boy never fetched his mother even though it was the purpose to begin with.

- (50) *De vulggii bárdni eatni-s*
 then leave.PST.3SG boy mother.ACC-PX.3SG
vieččážit.
 fetch.SUP.
 ‘So the boy went off in order to fetch his mother.’ (SaaN, Korhonen 1974: 200–201)

- (51) *De vulggii bárdni eatni-s*
 then leave.PST.3SG boy mother.ACC-PX.3SG
viežžat.
 fetch.INF
 ‘So the boy went off to fetch his mother.’ (SaaN, idem 201)

Formally, it is perhaps possible to deduct such a semantic discrepancy between the two sentences, especially if the sentences are in the past tense. The corresponding sentences of the present tense, on the other hand, would present a scenario where both the infinitive and the supine have future reference, and, accordingly, a similar portion of uncertainty that the future brings about. Hence, it seems that the infinitive and the supine in many contexts can be exchanged with each other, and, this would perhaps partly explain why the supine today is lacking in many North Saami dialects, even where it has been attested earlier (cf. Korhonen 1974: 198–199). The impression from Lule Saami is that the supine is the default choice in constructions such as that found in example (27). Importantly, however, at

some point in time, the semantic properties of the supine must have motivated its systematic use, both where the use is present today and where it has disappeared.

Bartens, in his thorough investigation of the semantic properties of the potential in Saami (1980: 1–99), points out that the potential in many cases has foremost a temporal meaning and that the modal component in many cases is very weak, if existing at all. As far as the present tense potential is concerned, as in Bartens's examples (idem 24–25), it is often included in statements with future meaning: declaration of intent, promise, announcement and prediction (Ger. *Absichtserklärung*, *Versprechen*, *Ankündigungen*, *Voraussage*). These types of statements put the potential semantically closer to the supine than the meaning described in Saami grammars (cf. examples 48 and 49). Sammallahti also describes the future reference of the potential stating that it “is sometimes used as kind of future form which is otherwise lacking in Saami” (1998: 82–83).

The etymologies of the potential and the supine clearly support a common origin. Sammallahti (1998: 215, 222; cf. Korhonen 1981: 298) points out the Finno-Saamic **-ŋsi-tAk* as the origin of supine, the former component being identical with the suffix, which developed into the potential marker in Saami and the conditional marker in Finnish. The component **-tAk* is identical with the Finno-Saamic infinitive marker. Accordingly, the supine could be described as the infinitive of the potential.

Regarding semantics, it has been suggested that the potential originates in a derivational suffix which has been ascribed frequentative-continuative (Lehtinen 2007: 92) and diminutive (Korhonen 1981: 252) properties. It is not within the scope of this article to discuss earlier semantic components of the suffix; however, the identification of a common derivational suffix, implies that the relation of the potential and supine is, or at least at some point was, that of a finite and non-finite verb forms based on the same suffix.

7. Discussion and concluding remarks

The auxiliary *ed tjedh* in South Saami and its cognates in Ume and in Pite Saami do not seem to have cognates among auxiliary verbs in any other Saami variety. From a functional and semantic point of view, however, the verbs show striking similarities with the supine. The supine, which is best preserved in Lule Saami and in western North Saami varieties, is typically used in purposive constructions (see examples 26–27), which is one of the main functions of the South Saami verb *ed tjedh* as well (see examples 10–11). Besides the functional and semantic correspondence, there is also a

conspicuous phonetic correspondence, since both *edjtjedh* and the supine are formed with a postalveolar affricate as its core. A closer analysis of the potential shows similar semantic, phonetic and functional correspondences.

Although there are semantic discrepancies between the potential, the supine and the verb *edjtjedh*, they are, in my opinion, not distinct enough to rule out a historical cognation (see Section 6). However, to include *edjtjedh* in the etymology, one must first explain the circumstance that *edjtjedh* is a verb in its own right, and that the other two possible cognates, the supine and the potential, are morphological markers. Since suffixes mark the supine and the potential in Saami and the conditional in Finnish, it seems more plausible that *edjtjedh* developed from an original suffix than vice versa. This assumption implies that *edjtjedh* appeared as an auxiliary verb through the process of degrammaticalization, or, more specifically, through a process of debonding (see Norde 2009: 186), when the originally bound morpheme transformed into an auxiliary.

A possible first step in the process of degrammaticalization could have been that the suffix, in an inflectional form, started to adhere to categories other than verb stems. From Pite Saami there is evidence of the cognate verb *atjav* attaching to infinitives (see Section 5.2), and in South Saami it is attested how *edjedh* merges with pronouns (see Section 4). These two instances give an indication of how the suffix may have separated from verb stems. When separated from the lexical verb stem, the constituent starts acquiring the syntactic properties of an auxiliary. The next step in the development towards a verb would be that the form starts to appear as a separate word, thus as an auxiliary in its own right. The evidence from Pite Saami shows that cognates in the form of the auxiliary verb, of the supine and of the potential existed side by side (see Section 5.3).

Another circumstance which points to the verb's origin in a suffix is the variation concerning the quality of the initial vowel (see section 4). The variation indicates a reflection of different vowels, which, depending on the quality of the verb stem, precede the consonant of the suffix. A similar variation, for instance, appears today in the following Lule Saami first person potential forms: *javl-a-tjav* < *javlla-* 'say', *váld-i-tjav* < *vállde-* 'take', *jáhtål-i-ttjav* < *jáhtål-* 'move, migrate'. From *Lexicon Lapponicum*, the first person singular potential forms *ets-i-tjab* (LXVII) < *etse-* 'love' (cf. SaaS *iehtsie-*, SaaU *eähttsie-* 'idem') and *jukk-a-tjab* (LXXVI) < *jukke-* 'drink' (cf. SaaS *jovke-*, SaaU *juvkka-* 'idem') may serve as an illustration of this vowel variation as well.

The stress in Saami falls on the first syllable. That implies that a verb like *vööjnetje* 'he will see' (cf. example 43), yields an unstressed vowel in second

syllable. When a constituent formed of the two last syllable nuclei and the intermediate consonant detaches from the rest of the word, which this article also suggests has happened, the initial vowel inherits some properties of the unstressed vowel. The outcome of such a process would explain the attested prosodic weakness and the reduced quality, or even complete lack, of the first vowel. The tendency of South Saami auxiliaries to be unstressed and pronounced as monosyllabic words, e.g., *gem* instead of *galkem* ‘I shall’ (Bergsland 1994: 82), may also have contributed to the occurrence of forms without an initial vowel (see example 14).

In earlier stages of South Saami, potential forms of the parasyllabic verb *vuejnedh* ‘see’ and the imparisyllabic verb *byöpmédidh* in the present singular could, somewhat speculatively, be assumed to have been **vööjne-tjem* 1P, **vööjne-tjh* 2P, *vööjne-tje* 3P (cf. example 35) and **byöpméd-adtjem* 1P, **byöpméd-adtjh* 2P, **byöbmed-adtja* 3P. In the process of degrammaticalization, the verb adopted the morphology of a parasyllabic verb from the imparisyllabic potential forms, and on basis of the potential inflection, the verb came to be associated with verb group IV (see Tables 1 and 2), with which it had most inflectional correspondence.

Since the potential in Saami has the same origin as the conditional in Finnish, it seems plausible that they have developed from a category that had modal properties. It is of course not possible to exactly pinpoint the category’s modal properties, which quite plausibly also could have been quite varied, but judging from the semantic properties of the potential and supine in Saami and the conditional in Finnish, it would possibly have had a modal meaning which could be labelled *irrealis*. In Saami this category developed into the potential and the supine, the supine being the potential suffix plus the infinitive suffix. In the course of its development, the semantic properties of the conditional must have interplayed with those of the potential, at which point the two categories gained their modal status (cf. Korhonen 1981: 254).

The semantic properties of expressions of intention and of purpose are similar. They both involve an aim and have future reference. The possibility to express either intention (Section 2) or purpose (Section 3) with *edjtjedh* is essentially conditioned by syntax. Typically, when in a main clause, it combines with the other constituents of the clause to express intention, and, when in a subordinate construction, it combines with other constituents of the construction to express purpose. It is its function similar to that of a purposive subordinator, which is the most conspicuous, since, with the exception of Ume and (earlier) Pite Saami, no other Saami languages display a verb with similar function, and, interestingly, neither do the most influential contact languages Norwegian and Swedish. However, purposiveness is found in constructions

with the supine, which, initially, possibly has motivated its function as a purpose marker, when it started to appear as a separate verb. The rise of *ed tjedh* as a purpose marker may also have caused the displacement of the supine, which would explain why the supine is poorly attested in South Saami. In addition, one should note in this context that expressions of negative purpose in South Saami, also may be initiated by finite constructions (see examples 13 and 14). The Norwegian *skulle* and the Swedish *skall* could possibly have affected the semantic extension of the verb so that it can also express deontic meaning (see example 1).

Even though South Saami has a lot in common with the other Saami languages, it also possesses features in several linguistic domains that make it stand out from them (see, e.g., Magga 2012). The auxiliary verb *ed tjedh*, may not be unique for South Saami, but, owing to its high frequency and its central role as a modal auxiliary verb, it is quite emblematic for the language. In addition, the verb with its historical background provides an interesting example of degrammaticalisation, which is a quite rare feature (see Norde 2009).

Abbreviations

ACC—accusative; ATTR—attributive; COM—comitative; CONNEG—connegative; DU—dual; GEN—genitive; Ger. —German; ILL—illative; IMP—imperative; INE—inessive; INF—infinitive; NEG—negation verb; Nor. —Norwegian; PL—plural; POT—potential; PPTCP—past participle; PRS—present tense; PST—past tense; PTCL—particle; PX—possessive suffix; Q—question participle; SG—singular; SUP—supine; SUPERL—superlative; Swe.—Swedish.

Unpublished sources

- ULMA 2860 = Calleberg, Carl Axel 1923: Maskaurdialektens grammatik. Renskriven av U. Rehnlund 1930. SOFI.
 ULMA 2966 = Axel Calleberg 1922: Malmesjaurdialektens grammatik, renskriven i okt. 1930 av Uno A. Rehnlund. SOFI.

References

- Aikio, Ante. 2012. An essay on Saami ethnolinguistic prehistory. In Riho Grünthal & Petri Kallio (eds), *A Linguistic Map of a Prehistoric Northern Europe* (Mémoires de la Société Finno-Ougrienne 266), 63–117. Helsinki: Suomalais-Ugrilainen Seura.
- Andersson, Ando. 1992. *Maam dah soptsestamme*. Jokkmokk: Sámi girjjit.
- Bartens, Hans-Hermann. 1980. *Die Verwendung von Potential und Konditional im Lappischen* (Mémoires de la Société Finno-Ougrienne 177). Helsinki: Suomalais-Ugrilainen Seura.
- Bergsland, Knut. 1992. Language contacts between Southern Sami and Scandinavian. In Ernst Håkon Jahr (ed.), *Language Contact: Theoretical and Empirical Studies*, 5–15. Berlin: Mouton de Gruyter.
- Bergsland, Knut. 1994. *Sydsamisk grammatikk*. Karasjok: Davvi Girji.

- Bergsland, Knut & Gustav Hasselbrink. 1957. *Sámien lukkeme gárjá: sydlapsk läsebok med grammatik och ordlista*. Gardvik: P.J. Fjellström.
- BB = Bidumsáme Báhkogirje
(https://saami.uni-freiburg.de/psdp/pite-lex/simple_search.php?searchMode=off&lang=Swe)
- BG = Barruk, Henrik. 2018. Báhkuogirjje: ubmejesámien–dárueon, dárueon–ubmejesámien. [Gagnef]: Kanaanstiftelsen.
- Collinder, Björn. 1943. Lappisches Wörterverzeichnis aus Härjedalen. Uppsala: Lundequistska bokh.
- Collinder, Björn. 1949. *The Lappish Dialect of Jukkasjärvi: A morphological survey*. (Skifter utg. av Kungl. Humanistiska Vetenskaps-Samfundet I Uppsala. 37:3.) Uppsala & Leipzig.
- Dixon, Robert M. W. 2010. *Basic linguistic theory. Vol. 1, Methodology*. Oxford: Oxford University Press.
- Faarlund, Jan Terje, Svein Lie & Kjell Ivar Vannebo. 1997. *Norsk referansegrammatikk*. Oslo: Universitetsforl.
- Frändén, Märit, Björn Lundqvist & Karin Wilson. 2007. *Verbh: En sydsamisk verbhandbok*. Uppsala: Uppsala Publishing House AB.
- GL = Fjellström, Pehr 1738. *Grammatica lapponica* [...]. Holmiæ.
- Halász, Ignác. 1881. *Svéd-lapp nyelv. 1, Nyelvtan és olvasmányok*. Budapest: MTA.
- Halász, Ignác. 1891. *Svéd-lapp nyelv. 4, Déli-lapp szótár*. Budapest: MTA.
- Halász, Ignác. 1893. *Svéd-lapp nyelv. 5, Népköltési gyűjtemény: a Pite lappmark Arjepluogi egyházkerületéből*. Budapest: MTA.
- Ignác Halász. 2023. *Alvos giehto. Förskräckliga berättelser från Arjeplog. Pite-samiska berättelser upptecknade av Ignác Halász*. Årråsit ulgusvaddedumbidumsámen: Inger Fjällås ja PeterSteggo. Gåvå: Stina Renberg. Árjepluovve/Arjeplog: Silver-museet,
- Jernsletten, Johan & Sverre Hatle. 1975. *Gin'na*. Oslo: Universitetsforl.
- Jaahkenelkien, Anna. 1993. *Goltelidh jih soptsestidh*. Daasta berterne: Aarborte
- Kejonen, Olli. 2024. Pite Saami Tales of Terror. [Review of Halász 2023.] *ARV. Nordic Yearbook of Folklore*. 80: 200-202.
- Korhonen, Mikko. 1967. *Die Konjugation im Lappischen. Morphologisch-historische Untersuchung. I: Die finiten Formkategorien*. (Mémoires de la Société Finno-Ougrienne 143.) Helsinki: Suomalais-Ugrilainen Seura
- Korhonen, Mikko. 1974. *Die Konjugation im Lappischen. Morphologisch-historische Untersuchung. II. Die nominalen Formkategorien*. (Mémoires de la Société Finno-Ougrienne 155.) Helsinki: Suomalais-Ugrilainen Seura.
- Korhonen, Mikko. 1981. *Johdatus lapin kielen historian* (Suomalaisen Kirjallisuuden Seuran Toimituksia 370). Helsinki: Suomalais-Ugrilainen Seura.
- Korhonen, Olavi 2006. Báhkogirje: julevsámes dárruj, dáros julevsábmáj = Ordbok: lulesamisk svensk, svensk lulesamisk. Jokkmokk: Samernas utbildningscentrum
- Koskinen, Arja. 1998. *Toiminnan välttämättömyys ja mahdollisuus. Pohjoissaamen modaalisten ilmausten semantiikkaa ja syntaksia* (Mémoires de la Société Finno-Ougrienne 231.) Helsinki: Suomalais-Ugrilainen Seura.
- Kuvsje (<http://kuvsje.oahpa.no/index.php/2012/10/02/komitativ>)
- Lagercrantz, Eliel. 1923. *Sprachlehre des Südlappischen nach der Mundart von Wefsen*. (Kristiania Etnografiske Museum. Bulletin 1). Kristiania.

- Lagercrantz, Eliel. 1926. *Sprachlehre des Westlappischen nach der Mundart von Arjeplog*. (Mémoires de la Société Finno-Ougrienne 55). Helsinki: Suomalais-Ugrilainen Seura.
- Larsson, Lars-Gunnar. 2012. *Grenzen und Gruppierungen im Umesamischen* (Veröffentlichungen der Societas Uralo-Altaica 83). Wiesbaden: Harrassowitz.
- Larsson, Lars-Gunnar. 2018. *Per Holmberger och sockenlapparnas språk*. (Acta Bibliothecae R. Universitatis Upsaliensis 52). Uppsala: Uppsala universitet.
- Lehtinen, Tapani. 2007. *Kielen vuosituhanet: suomen kielen kehitys kantauralista* (Tietolipas 215). Helsinki: Suomalaisen Kirjallisuuden Seura.
- Lehtiranta, Juhani. 1992. *Arjeplogsaamen äänne- ja taivutusopin pääpiirteet* (Mémoires de la Société Finno-Ougrienne 212). Helsinki: Suomalais-Ugrilainen Seura.
- LL = Lindahl, Erik & Johan Öhrling. 1780. *Lexicon Lapponicum*. Holmiæ: Lange.
- LLO = Grundström, Harald. 1948–1954. *Lulelappsk ordbok. Lulelappisches Wörterbuch*. 1–4. (Skifter utg. genom Landsmåls- och folkminnesarkivet I Uppsala. Ser. C:1). Uppsala & København:
- LW = Lagercrantz, Eliel. 1939. *Lappischer Wortschatz*. I-II. (Lexica Societatis Fenno-Ugricae 6). Helsinki: Suomalais-Ugrilainen Seura.
- Lönnrot, Elias. 1856. Ueber den Enare-Lappischen Dialekt. In *Acta societatis scientiarum Fennicae* IV. 133–280.
- Magga, Ole Henrik & Lajla Mattsson Magga. 2012. *Sørsamisk grammatikk*. [Karášjohka:] Davvi Girji.
- Magga, Ole Henrik. 1982. *Modalverb og infinitive innen verbalet: prosjektrapport*. Sámi instituhtta, Guovdageaidnu.
- Magga, Ole Henrik. 2012. Om ord og setninger i sørsamisk. In Cornelius Hasselblatt, Hans-Hermann Bartens & Eberhard Winkler (eds), *Lapponicae investigationes et Uralicae. Festschrift zum 65. Geburtstag von Lars-Gunnar Larsson* (Veröffentlichungen der Societas Uralo-Altaica 82), 111–123. Wiesbaden: Harrassowitz.
- Mattsson Magga, Lajla. 2016. *Mojhtesh*. Indre Billefjord: Idut.
- ML = Graan, Olaus Stephani. 1669. *Manuale lapponicum* [...]. Stockholm.
- NDB = Nedtedigibaakoeh (baakoeh.oahpa.no)
- NDS = Neahttagisánit (<https://sanit.oahpa.no>)
- Nevis, Joel A. 1986. Decliticization and deaffixation in Saame: Abessive taga. In Brian D. Joseph (ed.), *Studies on language change* (Ohio State University Working Papers in Linguistics 34), 1–9. Columbus: The Ohio State University Department of Linguistics.
- Nickel, Klaus Peter & Pekka Sammallahti. 2011. *Nordsamisk Grammatikk*. Kárášjohka: Davvi girji.
- Nielsen, Konrad. 1932. *Lapp dictionary: based on the dialects of Polmak, Karasjok and Kautokeino: A-F*. Vol. 1. Oslo: Universitetsforl.
- Nilsson-Mankok, Erik. 1975. *Systematik i sydsamiskan, Vilhelmina-Vefsen*. Vilhelmina.
- Norde, Muriel. 2009. *Degrammaticalization*. Oxford: Oxford University Press.
- NSO = Mattsson Magga, Lajla 2009. *Norsk-sydsamisk ordbok [Daaroen-áarjelsaemien baakoegærja]*. Indre Billefjord: Idut.
- Qvigstad, Just Knud. 1927. *Lappiske eventyr og sagn*. Volume 1. Oslo: Aschehoug.
- Qvigstad, Just Knud. 1929. *Lappiske eventyr og sagn*. Volume 4. Oslo: Aschehoug.

- Qvigstad, Just Knud. 1947. Sproget i Graans Manuale Lapponicum. In *Lapponica*. (Studia septentrionalia 3.) Oslo. 18–39.
- SAG I = Teleman, Ulf, Staffan Hellberg, Erik Anderson & Lisa Holm. 1999. *Svenska akademiens grammatik. Volym 1, Inledning, Register*. Stockholm: Svenska akademien.
- SAG II = Teleman, Ulf, Staffan Hellberg, Erik Anderson & Lisa Holm. 1999. *Svenska akademiens grammatik. Volym 2, Ord*. Stockholm: Svenska akademien.
- SAG IV = Teleman, Ulf, Staffan Hellberg, Erik Anderson & Lisa Holm. 1999. *Svenska akademiens grammatik. Volym 4, Satser och meningar*. Stockholm: Svenska akademien.
- Sammallahti, Pekka. 1998. *The Saami languages: an introduction*, Kárášjohka: Davvi girji
- Schlachter, Wolfgang. 1958. *Wörterbuch des Waldlappen-dialekts von Malå und Texte zur Ethnographie* (Lexica Societatis Fenno-Ugricae 14.) Helsinki: Suomalais-Ugrilainen Seura.
- SIKOR = Giellatekno - Saami Language Technology, UiT The Arctic University of Norway and The Divvun group at UiT The Arctic University of Norway, 2015, SIKOR North Saami free corpus, Common Language Resources and Technology Infrastructure Norway (CLARINO) Bergen Repository, <http://hdl.handle.net/11509/100>.
- Sjaggo, Ann-Charlotte. 2015. *Pitesamisk grammatik: en jämförelse med lulesamiska*. Tromsø: Tromsø universitet, Senter for samiske studier, Solentro.
- Spiik, Nils Eric. 1989. *Lulesamisk grammatik*. (2nd ed.). Jokkmokk: Sameskolstyrelsen.
- SW = Hasselbrink, Gustav. 1981–85. *Südlappisches Wörterbuch I-III* (Skrifter utgivna genom Dialekt- och folkminnesarkivet i Uppsala. Ser. C: 4). Uppsala: Lundequistska bokhandeln.
- Svonni, Mikael. 2019. *Davvisámi giellaoahppa*. Giron: Ravda Lágádus.
- Tuolja, Karin & Susanna Angéus Kuoljok. 2013. *Giellaj hilá*. (2nd ed.). Jokkmokk: Samernas utbildningscentrum.
- Wilbur, Joshua. 2014. *A grammar of Pite Saami*. (Studies in Diversity Linguistics 5.) Language Science Press: Berlin.
- Wilson, Karin. 2008. *Markusevangeliet i Lars Rangius samiska översättning från 1713*. (Acta Academiae Regiae Gustavi Adolphi 101.). Uppsala: Kungl. Gustav Adolfs Akademien för svensk folkkultur.
- Wiklund, K.B. 1915. *Lärobok i lapska språket* (2nd ed.). Stockholm: Björck & Börjesson.
- Vinka, Mikael. 2022. Two types of null subjects in South Saami. In Gréte Dalmi, Egor Tsedryk & Piotr Ceglowski (eds), *Null Subjects in Slavic and Finno-Ugric: Licensing, Structure and Typology*, 307–346. Berlin, Boston: De Gruyter Mouton.
- Ylikoski, Jussi. 2016a. Future time reference in Lule Saami with some remarks on Finnish. *Eesti ja soome-ugri keeleteaduse ajakiri* 7(2). 209–244.
- Ylikoski, Jussi. 2016b. Development of adpositions, adverbs and a free lexical noun from inflectional and derivational suffixes. *Finnisch-Ugrische Mitteilungen* 40. 113–173.

A comparative study of user attitudes to Sámi orthographies

Riitta-Liisa Valijärvi, Uppsala University/University College London & Lily Kahn, University College London

1. Introduction

The purpose of this paper is to analyse the perceptions of three different Sámi orthographies – North, South and Skolt – among L1 and L2 users of these systems. We have chosen these three orthographies as they are used in three different comparable Nordic countries, while having different national histories and maximally different Latin-script orthographic conventions and sound systems. South Sámi, for example, lacks consonant gradation, whereas North Sámi has a complex system of consonant gradation and Skolt Sámi has a rich consonantal inventory including a palatalised versus non-palatalised phonemic distinction. North Sámi (*davvisámegiella*) is spoken by around 25,000 people in Northern Norway, Sweden and Finland, and has the most widely used orthography. South Sámi (*åarjelsaemien gïele*) is spoken by around 300–500 people in certain municipalities in central Sweden and Norway. Skolt Sámi (*nuõrttsää'mkiõll*) is spoken by around 300 people mostly in Sevettijärvi in Inari municipality in Finland. We have not included Kildin Sámi and its Cyrillic script in this initial study due to limitations in scope and lack of access to the speaker community. For a historical overview and analysis of Kildin Sámi orthography see Rießler (2013), Siegl and Rießler (2015) and Ivaniščeva (2016).

Orthography refers here to a standardised way of writing a language and representing its sounds in writing; in this study the term will be used interchangeably with *writing system*. The latter term can also refer to the categories of alphabets, syllabaries, and logographic systems (cf. Coulmas 2003: 35). The history of different Sámi orthographies dates back to missionary work in the Sámi-speaking area (Wiklund 1922; Korhonen 1976: 56, 1981: 53–57). Orthographies were developed for Bible translations and linguists carried on this work in line with a top-down approach. As early as

1648 Johannes Torneaus tried to combine elements from the Pite and Torne varieties of Sámi in his *Manuale Lapponicum* (Korhonen 1976: 56–57). Such a unified language planning approach has not been attempted since. Recent trends in Sámi orthography development have been more bottom-up with community involvement. Examples of the latter are the launches of an Ume Sámi orthography (Påve 2016) and Pite Sámi orthography (e.g., Sunna et al. 2019; Wilbur in this volume). Sámi languages have been perceived as separate languages especially after the creation of early orthographies (Aikio-Puoskari 2002: 102), which could be interpreted to suggest that having a distinct orthography for one's own variety of Sámi is a particularly important marker of language identity (see also Jones and Mooney 2017: 1).

The current North Sámi orthography has been in the making since Knud Leem's grammar in 1748 and dictionaries in 1768–81 (see Korhonen 1976: 57–58). Rasmus Rask introduced a phonemic principle for North Sámi orthography in 1832 (Rask 1832), which was used by Stockfleth in his grammar of 1840 (Stockfleth 1840). Nielsen launched an orthography in the late 1920s that did not remain the standard, and in 1947 a common orthography for the Sámis in Sweden and in Norway was discussed. The task of uniting the orthography was given to the Norwegian linguist Knut Bergsland and the Swedish (Pite Sámi) linguist Israel Ruong (Korhonen 1976: 57). The new orthography was adopted in 1950 by the Swedish and Norwegian Boards of Education to be used in schoolbooks (Bergsland 1952: 37–40). At the time, the Finnish North Sámi kept using their own orthography devised by the Finnish linguists Paavo Ravila in 1930s and Erkki Itkonen in 1950s (Korhonen 1981: 53–57). A new common North Sámi orthography for all three countries was adopted in 1979 and modified in 1985; it was seen as necessary for the survival of the language (Aikio 1987: 457). Some dialectal and country-specific variation is allowed in North Sámi publications.

An orthography for more southern varieties of Sámi was also launched relatively early, when Per Fjellström introduced a written standard which he used in the publication of a grammar and a dictionary in 1738 as well as a translation of the New Testament in 1755 (Korhonen 1976: 56). Contemporary South Sámi, which is used in Norway and Sweden, employs an orthography that was initially adopted in 1976. While the Sámi Language Board currently recommends a single unified orthography for users in both Norway and Sweden (cf. Magga and Magga 2012: 12), in practice South Sámi publications from Norway tend to use Norwegian graphemes while their counterparts from Sweden tend to use Swedish graphemes (see below for details).

In contrast to North Sámi and the more southern varieties, the first Skolt Sámi orthography was developed only in the early 1970s. Linguists Mikko Korhonen, Jouni Moshnikoff and Pekka Sammallahti chose to base it on the *Suõ'nn'jel* (Suonikylä) dialect (Kulonen et al. 2005: 399; Moshnikoff 2014). It was initially planned to be similar to the existing North Sámi writing system, but by the time the North Sámi orthography was revised and unified in the late 1970s, Skolt no longer resembled it (Wesslin and Alajärvi 2016).

There are few studies that focus exclusively on Sámi orthographies, and only one previous study that investigates users' attitudes to them. Sorjonen (2019) studied perceptions of Sámi orthographies among Finnish Sámi in the context of revitalisation. She identified three main challenges: learning the orthography, producing special characters, and the approval of written standards by the language community. Sámi orthographies are also mentioned in official reports (Moshnikoff and Moshnikoff 2006; Aikio-Puoskari 2016; Pasanen 2016) and in relation to the assessment of revitalization efforts (Olthuis et al. 2013; Pasanen 2015). For example, according to Pasanen's (2016: 43) report, a lack of reading skills in all the Sámi varieties spoken in Finland (i.e., North Sámi, Inari Sámi, and Skolt Sámi) is an obstacle to taking advantage of Sámi-language services. Similarly, Moshnikoff and Moshnikoff (2006: 4–5) mention that speakers of Skolt Sámi note difficulties in reproducing the special orthographic symbols in print, and that as a result the quality of Skolt Sámi publications suffers. In addition to these official reports, there is debate in the media about Sámi orthographies. For example, Wesslin and Alajärvi (2016) reported in the Finnish news that many Skolt Sámi speakers do not know how to read and write their mother tongue. The writing system is especially difficult for the older generation as it was devised and taught in schools only from the 1970s; many feel that the spoken language is very different from the written language (Wesslin and Alajärvi 2016). Moreover, Sámi languages are generally considered to be too different from each other to have one unified orthography (Korhonen 1981: 53), although some people see this as a consequence of a divisive politics (e.g., Lindqvist 2016). Nils-Henrik Sikku from the Sámi Writers' Association *Bágo* says that the different orthographies are a taboo subject (Lindqvist 2016).

Against this background, the research questions in our pilot survey study are: How do speakers of the chosen Sámi languages (North, South and Skolt) perceive their orthographies? What are the challenges in using Sámi writing systems from the speakers' point of view? Would the speakers want to change their orthographies? What are the theoretical implications of speakers' preferences? What are the discursive practices and ideologies surrounding the

use of the selected orthographies? What do the respondents use their orthographies for, i.e., what do they read and write?

2. Theory

Theoretically, our study is based on the sociolinguistics of language policy and practice. Spolsky (2004, 2011: 148–150) divides language policy into three components: language practices, language beliefs or ideology, and language management. In our study the first component refers to the languages our respondents write or read in. The second component, beliefs, refers to the attitudes the respondents have toward the orthography they use or are expected to use. Language management refers to attempts to modify the status or use of a language, attempts to modify the specific standard, and attempts to modify or increase the number of speakers. We ask the speakers about the ways in which they would change the orthography and how they feel about its use, which seeks to explore the underlying question of the relationship between the degree of a language's endangerment and its orthography. Graphisation, i.e., receiving a written form, makes many maintenance and revitalisation activities possible, including curriculum development and teaching materials (Lüpke 2011: 312). An orthography can increase the prestige and status of a language and may help create a sense of one unified identity (Jones and Mooney 2017: 1).

We also make use of different dimensions of orthography development (e.g., Sebba 2007, 2009; Grenoble and Whaley 2006; Cahill 2014; Hinton 2014; Lüpke 2011; Jones and Mooney 2017). The following factors have been identified as being key in creating successful orthographies. When devising an orthography, one should consult the speaker community when making decisions about the orthography. In addition to speaker input and community involvement, one should aim to find a balance between the ease of learning and ease of reproduction of an orthography vs the maximum representation of speech. If one chooses to represent all phonemes or even allophones in a language, the result may end up being an orthography with many diacritics and additional denotations that the user cannot remember or acquire, or type with computer or phone. The rendering of phonological distinctions can in this respect be characterised in terms of overrepresentation vs underrepresentation. In the case of Sámi languages, it applies to ways of marking consonant and vowel length, aspiration, and consonant gradation. This relates to the concept of orthographic depth (cf. Ziegler et al. 2010: 551–552). Shallow orthographies have a one-to-one phoneme to grapheme match and are predictable. Deep orthographies, on the other hand, are more arbitrary in the

way they are pronounced and are often based on etymologies or morphophonemics. One also needs to take into consideration users' knowledge of majority-language orthography or legacy orthography: it might be worth basing graphemes and orthography conventions on these as the users will already be used to them. A case in point is Lule Sámi, whose orthography users are most familiar with Swedish orthography. Sometimes a so-called 'ideological distancing' (Jones and Mooney 2017: 25–26) may be preferred by the speaker community. This means that they may consciously choose to use graphemes or even a script that is different from the one used by the majority language or a different variety of the language. This is the case for example with the Inuktitut syllabary which is maximally different from the Latin script used by the colonial majority language English, and is based on a Cree syllabary instead. One should also decide which dialect one should base the orthography on. One solution to this would be to allow optional spelling or graphemes for dialectal variation rather than force everyone to use one standard.

We can exemplify some of the concepts above with some of the present-day Sámi orthographies used in Norway, Sweden, Finland, and Russia. An example of majority-language influence is that fact that Inari Sámi (spoken in Finland) uses the Finnish ⟨ä⟩ and ⟨y⟩, word-initial ⟨k⟩, ⟨p⟩, and ⟨t⟩ for unaspirated word-initial stops, and double vowels to indicate vowel length, whereas in South Sámi (Norway, Sweden) the Norwegian/Swedish graphemes ⟨æ⟩/⟨ä⟩, ⟨ø⟩/⟨ö⟩, and ⟨å⟩ are used, with readers and writers tending to choose the variant that they prefer based on their country of residence (so that South Sámi written in Norway typically uses the Norwegian graphemes, while South Sámi written in Sweden typically uses the Swedish ones). Similarly, Kildin Sámi (spoken in Russia) is written using the majority (Cyrillic) alphabet. An example of ideological distancing is the fact that North Sámi (spoken in Norway, Sweden, and Finland) is free from country-specific graphemes and instead uses graphemes that are distinctly Sámi and are not found in the majority languages, e.g., ⟨č⟩, ⟨š⟩ and ⟨ŋ⟩. This can be contrasted with South Sámi where these same sounds are written as they would be in Norwegian or Swedish: ⟨tj⟩, ⟨sj⟩ and ⟨ng⟩, respectively. All the Sámi orthographies are relatively shallow, but Skolt Sámi is a good example of an orthography that chooses to represent fine phonological distinctions which are not indicated in other Sámi orthographies; for example, Skolt Sámi distinguishes orthographically between e.g., ⟨a⟩ /a/ and ⟨â⟩ /æ/, between ⟨ʒ⟩ /d͡ʒ/ and ⟨ž⟩ /d͡ʒ̥/, between ⟨g⟩ /g/, ⟨ǵ⟩ /j͡g/, and ⟨g̃⟩ /ɣ/, between ⟨k⟩ /k/ and ⟨k̃⟩ /c͡c̥/, and between o /o/ and ⟨ö̃⟩ /ə/.

3. Methodology and data

3.1. Survey design

The method in this study is primarily qualitative content analysis. Our data were collected via an online survey which was created using SurveyMonkey. The survey consisted of 10 multiple-choice questions supplemented by some open questions regarding reading and writing practices among speakers of North Sámi, South Sámi, and Skolt Sámi. The questions were produced in three different versions, one in Norwegian (for North Sámi speakers living in Norway), one in Swedish (for South Sámi speakers living in Sweden), and one in Finnish (for Skolt Sámi speakers living in Finland).¹ Participants were told that they could answer the questions in the language of their choice (Norwegian, Swedish, Finnish, or a Sámi language). Participants were asked how often they read and write in the relevant Sámi language and given a choice of ‘often (many times a week or on a daily basis), sometimes (no more than once or twice a week), rarely (no more than once or twice a month), or never’. In each case they were given the option to add comments in a free-text box. They were also asked how often they write in the relevant variety of Sámi, with the same options for responses and space for additional comments. Participants were then asked what types of materials they read and write in their Sámi language most frequently, with a range of options listed including novels, non-fiction books, textbooks, homework, websites, text/messaging, emails, etc. They were also given space to add comments in a free-text box. Participants were also asked how they felt about reading and writing in their Sámi language, with a free-text box for responses. Finally, participants were asked whether there was anything that they would change about the orthography of their Sámi language, with another free-text box for responses. At the end of the survey, participants were asked to provide their age, gender, occupation, first language, and languages regularly spoken. The three versions of the survey were disseminated mainly via Facebook pages for the communities in question (the North Sámi page *Ártegis sagat*, the South Sámi *Numhtie naemhtie soptsestibie*, and the Skolt Sámi community page *Kolttakulttuurisäätiö*).² They surveys were completed between August and

¹ Producing alternative versions of each questionnaire for Sámi speakers in the other relevant countries (i.e., Swedish- and Finnish-language versions of the North Sámi questionnaire for speakers living in Sweden and Finland, a Norwegian-language version of the South Sámi questionnaire for speakers living in Norway, and a Russian-language version of the Skolt Sámi questionnaire for speakers living in Russia) was beyond the scope of this study, but is a desideratum for a larger-scale research project on Sámi orthographies.

² We would like to thank Olle Kejonen, Richard Kowalik, and Timothy Feist for their help in disseminating the survey.

October 2019. All participants gave permission for their responses to be used in the study.

3.2. Survey responses

North Sámi

We received 25 responses from North Sámi participants. Survey respondents were aged between 18 and 72, with most speakers in their 30s and 40s. The respondents were relatively equally balanced in terms of gender, with 11 men and 14 women. They represented a range of different occupations: consultant/advisor, language worker, student, (Sámi) teacher, carpenter, pensioner, artist, IT consultant, reindeer herder, psychologist, self-employed, journalist, and translator. The majority of respondents (16) listed Sámi or North Sámi as their first language, with another 2 respondents listing Sámi and Norwegian, 1 listing Sámi and Swedish, 3 listing Norwegian, 1 listing Finnish, and another 2 unspecified. All respondents reported that they speak North Sámi every day or all the time.

South Sámi

We received 11 responses for our South Sámi survey. While fewer than the number of North Sámi responses, this is in keeping with the much smaller South Sámi speaker population. Like North Sámi, the South Sámi respondents represented a wide range of ages, between 19 and 71. The gender division was less even than that of North Sámi, with 3 men and 8 women. A large selection of different types of professions were represented, with respondents citing the following occupations: teacher, administrator, pensioner, economist, self-employed, reindeer herder, South Sámi language professional. The respondents reported a much wider range of first languages than their North Sámi counterparts, with the following responses: Norwegian (2), Norwegian and South Sámi (1), Swedish (2), Sámi (2), South Sámi (2), North Sámi (1), and Swedish/English bilingual (1). Most respondents reported that they speak South Sámi daily, with a minority reporting that they speak it less often.

Skolt Sámi

We received 14 responses for our Skolt Sámi questionnaire, which is substantial considering the small total speaker population. As in the case of North and South Sámi, there was a considerable age range (25–79). Like in the South Sámi survey, there were more female than male respondents with 1 man, 12 women, and 2 undeclared. There was again a wide range of self-reported professions: student, professional, care assistant, researcher, customer service assistant, sales assistant, occupational therapist, consultant,

pensioner, secretary, and chef. There was a greater range of first languages, with respondents reporting Skolt Sámi (4), Karelian and Finnish (1), Finnish (5), Finnish and Skolt Sámi (2), Finnish, with Skolt Sámi acquired as an adult (1). One respondent did not provide an answer for this question. There was also a wider range of answers regarding frequency of use: some respondents used the language every day, while others said ‘whenever there’s someone to speak to’, ‘a couple of times a week’, ‘a couple of times a month’, and ‘rarely’.

4. Analysis of data

North Sámi

The North Sámi respondents reported very high frequencies of reading, with 84% saying daily or many times per day, 12% sometimes, and just 4% rarely. Most respondents read multiple types of texts, i.e., internet news, social media, emails, text messages, work documents, children’s books, and textbooks. Of these, internet news, social media, and emails were the most common. When asked how it feels to read in North Sámi, the answers were generally positive. Some respondents expressed great enthusiasm, e.g., ‘it’s the best thing there is’ and ‘great’. There were several responses focusing on ease and intuitiveness, e.g., ‘it’s easy and often very poetic as I’ve had it in school’, ‘easy, good, and simple orthography’, and ‘it’s easy’. Others highlighted the ‘naturalness’ of reading in the language, e.g., ‘good, simple, and natural’, ‘it is very natural to read one’s own language. I am used to reading Sámi’, and ‘almost as natural for me as speaking North Sámi’. One respondent described the reading experience as ‘fun and educational’. Only a small minority of respondents suggested that they found North Sámi difficult to read. However, they also framed their responses in a positive way, saying e.g., ‘I prefer to read in North Sámi although it can be hard’ and ‘I think it’s nice but it is slower than in Norwegian [...]’. Only one respondent expressed a purely negative opinion, describing the experience of reading Sámi merely as ‘tough’.

The North Sámi respondents reported similarly high writing frequencies, with 84% saying daily or many times a day, 8% saying sometimes and only 4% saying rarely. As with reading, respondents reported writing a wide range of texts, i.e., e-mails, text messages, comments on social media, lists, reports, work documents, stories, and poetry. Of these, emails, text messages, and comments on social media were the most common.

When asked how it feels to write in North Sámi, respondents were generally very positive. Some respondents reported excitement and great enthusiasm, e.g., ‘Great! It’s soooooo great. The best way to communicate with others’ and ‘I think it’s exciting to write in my mother tongue’. Another

commented on it being ‘fun’ to write in North Sámi. Other respondents commented on the ‘naturalness’ of writing in North Sámi, with one explaining ‘It’s easier because it comes so naturally. It’s easy to write one’s own language as it’s so natural. When I write Norwegian I have to think more about how I should write’ and another saying ‘it’s natural for me as it’s my mother tongue but I still don’t write 100% correctly’. This latter comment reflects a wider concern about making mistakes when writing in North Sámi; another respondent observed that ‘It’s easy but it bothers me that I still make mistakes writing’. Along similar lines, two respondents noted simply that it is ‘difficult’ to write in North Sámi. Others addressed specific difficulties, with one saying that ‘sometimes it’s difficult especially with diphthongs’, and another referring to the speed of writing: ‘It goes a bit slower than in Norwegian’. Only one respondent commented on keyboard limitations, saying ‘I like to write North Saami. I do it as often as I can if I’m not hindered by the keyboard [...]’. This suggests that most participants did not find the special North Sámi characters particular obstacles to writing the language, but keyboard problems were not completely unknown.

When asked for their opinions about the North Sámi orthography, many respondents again gave positive answers. Some commented on its logical and phonetic nature, e.g.: ‘it’s logical, reminds me of maths’, while another responded that ‘It works well and is logical, the orthographic rules are closer to the spoken language than Norwegian’. One respondent observed that the current North Sámi orthography is ‘easier than the earlier orthography’, while another expressed concern for the prospect of the system being changed again, stating that ‘it’s very well suited for the language. It shouldn’t be changed. We Sámi writers should instead learn to use it and children must have teachers who can use it’. Other respondents were less positive: one touched on the lack of orthographic representation for different North Sámi varieties, saying, ‘the Kautokeino dialect, it doesn’t suit those of us from Varanger’. Another expressed the opinion that the special North Sámi characters ‘should’ve been thrown in the bin long time ago. It just makes a mess, especially within IT’. One respondent specifically addressed the issue of distancing, saying ‘I don’t think that a small language like Sámi benefits from having different letters than the big neighbouring languages’.

When asked whether they would change anything in the North Sámi orthography, 69% said no, and 26% said yes. One respondent who did not want anything to be changed commented that ‘it has been changed too many times already’, pointing to the importance of continuity and stability. One respondent who did want to see changes said that they would remove the special signs ⟨á⟩, ⟨š⟩, ⟨č⟩, ⟨ž⟩, ⟨ŋ⟩, ⟨t̥⟩, ⟨đ⟩. Another commented that they would

‘mark length of vowels, mark the vowels [sic] today marked with *h* in front, e.g., *hl*, *hr*’, with a third saying ‘when I teach Sámi, of course sometimes one would wish that consonant gradation was *hht* : *ht* instead of *htt* : *ht* and *sst* : *st* instead of *st* : *stt* so that it corresponded better with the way it’s pronounced’. One respondent said that they would change the orthography ‘to include Varanger dialect’. Conversely, another observed that, despite the competing demands of speakers of different dialects of North Sámi, on balance it would not be practical for the orthography to take these into account: ‘The orthography is a compromise between all the different dialects so if I start to wish for changes based on my dialect then others might want thousands of small changes based on their dialect and then it would be chaos’.

South Sámi

The South Sámi respondents generally reported a high frequency of reading, although somewhat lower than their North Sámi counterparts: 73% reported reading daily or many times a day, with only 9% saying sometimes. A relatively high percentage, 19%, reported reading only rarely. Most respondents reported reading multiple types of texts, i.e., social media, textbooks, text messages, children’s books, newspapers, online news, popular science, professional literature, and novels. The most popular among these were social media, textbooks, and text messages. The inclusion of textbooks in this list differs from the North Sámi data, and reflects the slightly different demographic among the South Sámi cohort which contains more L2 speakers.

When asked how it feels to read, the South Sámi respondents, like their North Sámi counterparts, generally reported positive feelings. However, the reactions were less emphatic: while one respondent commented that it was ‘great; sometimes fantastic’, most said simply ‘good’. One focused on the educational and symbolic value of reading in South Sámi, noting that ‘it’s fun, the language feels important. It’s educational to get to see new words, how other people put together sentences, etc.’ Their use of the term ‘important’ echoes one of the North Sámi comments, as well as their observation that an orthography elevates the language’s importance. Another respondent expressed a more neutral viewpoint, saying, ‘There’s no difference in reading South Sámi and Swedish but texts about Sámi culture, traditional and reindeer herding are better for reading in South Sámi or North Sámi’. The South Sámi respondents, in contrast to their North Sámi counterparts, expressed some concern with the quality of written language, e.g.: ‘good if there aren’t too many grammatical and spelling mistakes, and full of Norwegianisms’. On a similar note, another respondent referred to difficulties stemming from lack of fluency in the language itself, noting that ‘it’s difficult when you don’t know

so much but everyday expressions and sentences are fine'. This also reflects the greater number of L2 speakers of South Sámi in general, and among the survey respondents in particular.

The South Sámi respondents reported a somewhat lower frequency of writing than their North Sámi counterparts, with only 65% writing on a daily basis or multiple times per day. 19% said that they write sometimes, with another 19% reporting that they write only rarely. Respondents reported writing a wide variety of textual genres, namely notes, text messages, social media comments, poetry, newspaper articles, stories, language learning materials, and essays, with notes, text messages, and social media comments being the most common.

When asked how it feels to write in South Sámi, some respondents answered 'good' or 'fun', with one commenting enthusiastically that 'to express yourself in your own language is a feeling that cannot be described'. One respondent said, 'thanks to the Sámi spellchecker from the University of Tromsø it is easy to write now'; adding 'I also write a dialect which has become more accepted recently'. However, some respondents reported more negative experiences of writing, but, as in the case of reading discussed above, these were tied to grammatical difficulties rather than stemming directly from orthographic challenges. One commented, 'It feels good most of the time but because you don't do it often you get rusty', while another said that 'it's difficult to try to change the form of verbs and nouns'.

When asked for their opinions about the South Sámi orthography, some respondents responded positively that it is 'logical, actually completely fantastically constructed' or 'easy, intuitive but the excess of *h* can sometimes go overboard'. One respondent said 'I've learnt it, it's easy to use', suggesting that their comfort with the written language is attributable directly to the fact that they have invested time in acquiring proficiency in it. One speaker felt that the system is too complex but that this complexity is justified, saying 'there are many who think there are too many vowels and too many rules but they are important to distinguish words'. Another echoed this sentiment, saying 'crazy long but if you know it, it's not a problem'. A third acknowledged the tensions inherent in the development of the system: 'It's been a long process with many compromises. I accept that those who worked on the writing system came up with this. But I of course pronounce the words as in my own dialect but then you have to understand which dialect you're supposed to be speaking'. One respondent expressed slightly more dissatisfaction with the complexity of the system, remarking that '17 letters in one word is too much'.

Finally, when asked whether they would change anything about the South Sámi orthography, 91% said no, with only 9% saying yes. Of that 9%, the only respondent to comment further said that they would like to see the South Sámi orthography move ‘closer to the North Sámi orthography’.

Skolt Sámi

Skolt Sámi respondents reported more variation in reading practices than their North and South Sámi counterparts: only 43% said that they read daily or many times a day, with another 43% saying sometimes and the remaining 14% saying rarely. Most respondents reported reading multiple genres, namely social media, children’s books, online news, textbooks, memoirs, books, and emails. Of these, the most common were social media, children’s books, and online news. This differs slightly from the North and South Sámi data, which did not include children’s books as a commonly read text type. Some Skolt Sámi respondents commented on the reasons for infrequent reading, with one mentioning a lack of materials, while another cited difficulty in the skill of reading, saying ‘I haven’t learnt the letters’ and a third mentioned insufficient command of the language itself.

Regarding their feelings about reading in Skolt Sámi, a few respondents gave positive answers, but less enthusiastically than their North and South Sámi counterparts, e.g., ‘no problem’ and ‘good’. One respondent echoed the theme of identity seen in some of the North and South Sámi responses: ‘Great. Language is part of your identity’. Another expressed satisfaction associated with the achievement of learning to read the language, saying ‘I’ve learnt in school so it [feels] good’. However, many respondents expressed a more ambivalent opinion, with some commenting on it being difficult, slow, or cumbersome. In one case a respondent cited challenges stemming from the fact that the language had only developed a standardised orthography relatively recently (i.e., in the early 1970s; cf. Moshnikoff and Moshnikoff 2006). In another case, a respondent attributed their challenges with reading to the soft sign (i.e., the diacritic ‘*’*), which is added before a dental consonant to mark palatalization), calling it an ‘unnecessary invention’ and saying that it ‘makes it hard to recognize the text and it all looks like a mess’.

Skolt Sámi respondents reported significantly lower rates of writing than their North and South Sámi peers, with only 29% saying that they wrote daily or many times a day, 29% saying sometimes, 29% rarely, and 14% never. This is the only respondent group to contain a percentage who never write. The Skolt Sámi respondents reported writing a wide range of different materials, namely text messages, emails, social media comments, homework, translation work, *leu'dd* (traditional songs), stories, prose texts, letters, and jokes with

Skolt friends. Some of these (*leu'dd* and jokes), were not mentioned in the other groups' responses. As in the case of North and South Sámi, the most commonly written genres were text messages, e-mails, and social media comments.

When asked how it feels to write in Skolt Sámi, there were only a few positive responses, in marked contrast to the other groups. While two respondents said that it was 'good' or 'nice', others gave more nuanced answers, e.g., 'quite easy as I've been practicing for years', highlighting the fact that years of study were a factor in their positive attitude. Two respondents expressed the feeling that writing in Skolt Sámi has no function ('there isn't actually a need to write' and 'it's not meaningful to write on one's own'), a sentiment not observed in the North and South Sámi groups. Other respondents suggested that writing in Skolt Sámi is a struggle, citing lack of tuition, e.g., 'the letters weren't studied' and 'limited knowledge' and the complexity of the orthography itself, e.g., 'I would write more if the written language wasn't so shockingly complicated. Having studied [the language] I don't dare to write spoken language anymore' and 'damned hard and difficult. Separate keyboard has to be installed and you put them on the work phone and the work computer. Softening signs ruin everything, you never know where you're supposed to put them. There's no joy or ease to writing like this'. Another respondent raised the same issue of keyboard incompatibility: 'I don't know how to write and I don't have the letters on the phone and computers'. These difficulties are much greater than those mentioned by the North and South Sámi respondents.

When asked specifically about their opinions of the Skolt Sámi orthography, some respondents expressed positive feelings but highlighted the need to study it explicitly in order to achieve mastery of it: 'the alphabet is clear if you know it and learn it', 'I've learnt it as a child so I don't find it hard', and 'I think it's functional: if you know how to pronounce the language, writing and reading go without problems'. Other respondents expressed varying degrees of dissatisfaction with particular elements of the orthography. One identified a particular set of consonants that they felt should be written more like their Finnish equivalents: 'In my opinion the letters *cc*, *c*, *č* are not needed but it would be written as *ts*. For example, water = *čääcc* = *tšääts*'. Two respondents commented on the soft signs and apostrophes being frustrating and difficult to remember (though one noted that they are important for the language's phonology). One respondent remarked more generally that the orthography is 'far too complicated', and another described it as 'difficult [and] messy'. The difficulty of writing on a standard keyboard was raised again, with one respondent saying that it is 'challenging on a computer and on

the phone because the keyboards are designed for different kinds of alphabets. It's especially difficult if you want to write Finnish/Scandinavian languages and Skolt at the same time.' Overall, the opinions on Skolt Sámi orthography diverge significantly from those given for North and South Sámi.

When asked whether they would change anything about their orthography, interestingly, only 28% of Skolt Sámi respondents replied positively, with 71% saying that they would not. This percentage resembles that of North Sámi, but the comments differ widely between the two groups, with the North Sámi respondents much less likely to suggest specific changes. One Skolt Sámi respondent suggested that there should be a symbol ⟨ts⟩, as in Finnish, instead of the special symbol ⟨č⟩ (which, interestingly, is shared with North Sámi; the respondent would prefer a digraph shared with the majority national language, Finnish). Another commented on several areas for reform relating to special consonant symbols: 'I would leave the "threes" [⟨ʒ ʒ⟩ /dʒ/ and ⟨ʒ ʒ⟩ /dʒ/] out as no one has been able explain their exact pronunciation. And tone down the apostrophe hysteria. To my understanding there are three different ones: does anyone else apart from linguists and language fanatics feel the need for them?' Another respondent said 'at least get rid of soft signs'. One respondent gave a general comment in the same vein: 'simplify the alphabet so the written language would be more approachable'.

5. Conclusions

Our study has revealed both similarities and differences in the way the three groups of respondents employ and perceive their orthographies. Out of the three groups of respondents, North Sámi speakers use the language the most both in speech and writing on a daily basis. North Sámi respondents generally reported positive feelings about the orthography, calling it logical, natural, and an important symbolic marker of identity. This is consistent with Sorjonen's (2019) findings regarding the importance of orthography as a positive marker of identity among Finnish Sámi, as well as with those of Jones and Mooney (2017) for endangered language users more generally. The distancing principle underpinning the orthography was not identified as particularly problematic, with only a small number of respondents expressing a desire for the special graphemes to be removed. Just one respondent mentioned difficulties with keyboard incompatibility, so this does not appear to have been a major concern. There was a general feeling that the orthography has already been changed too many times and should now remain stable, with more effort going into learning and using it. Only a minority of respondents mentioned having some difficulties in writing correctly.

South Sámi respondents reported reading and writing somewhat less frequently than their North Sámi counterparts, but were also happy with their orthography overall. Like the North Sámi respondents, they mentioned the important symbolic value of the orthography for their linguistic identity. In contrast to North Sámi, South Sámi does not reflect a distancing model but rather its graphemes are based on those of the majority languages Swedish and Norwegian. Thus, the data show that both distancing and non-distancing models can be popular among readers and writers of Sámi languages. Interestingly, one South Sámi respondent reported a desire for their language to be spelt like North Sámi, suggesting a drive towards greater symbolic unity between the different Sámi languages rather than the need for a close resemblance to the majority national languages. (Interestingly, such an orthography was once developed; see Bergsland and Hasselbrink 1957). North Sámi and South Sámi respondents are aware of the fact that orthography is a compromise between different dialects and are generally satisfied with this situation. In contrast to North Sámi respondents, the South Sámi respondents reported greater levels of difficulty with the language itself (since a higher proportion consists of L2 speakers) and a lower quality of literary outputs in the language (with grammatical mistakes and Norwegianisms). Again, in keeping with Sorjonen's (2019) findings regarding Finnish Sámi, our South Sámi respondents associated knowledge of the orthography with linguistic competence, equating the ability to write with mastery of grammar and vocabulary.

Skolt Sámi respondents read and write less often in their language than South and North Sámi respondents. As in the case of the other two groups, some respondents highlighted the importance of the orthography as a marker of linguistic identity. However, the Skolt Sámi cohort differed markedly from the North and South Sámi ones in that they expressed high levels of dissatisfaction with the orthography. Many respondents considered it to be overly complicated unless you know it well and have used it since childhood (cf. Sorjonen 2019), which older speakers had not had the opportunity to do. In contrast to the North Sámi speakers, who were generally at ease with the distancing orthographic model used for their language, many of the Skolt Sámi respondents expressed frustration at the special symbols used in their orthography. This could be due to a perceived overrepresentation, as the Skolt Sámi orthography makes finer phonemic distinctions than North Sámi. While only a minority in the North Sámi group mentioned a desire for the abolition of the special symbols, this was a very prominent issue raised by the Skolt Sámi group, suggesting that they perceive a higher level of complexity and/or have a lower level of training and familiarity with the orthography. The

complexity of the Skolt Sámi system might explain the expressed preference for greater identification with the majority national language, as opposed to a dislike of a distancing model *per se*.

It is to be hoped that this study could lead to further comparative research on Sámi orthographies. The number of respondents in our study was relatively low and not all Sámi languages or orthographies were included. Furthermore, there might have been a bias toward respondents who are keen writers and readers or even activists as the survey was conducted via Facebook pages rather than oral interviews. Some professional groups, such as language teachers or translators, or dialect areas, such as North Sámi in Finnmark, may have been overrepresented among our respondents. For this reason, we would welcome a further investigation of the following research questions: How would a quantitative study among a larger number of respondents correlate with our findings? Are Swedish North Sámi equally happy with the orthography or do they feel it is not fit for their purpose as it is not based on their dialect? How widely are the new Ume and Pite Sámi orthographies used? How do Kildin Sámi or Lule Sámi speakers use and view their orthographies? Does the launch of digital resources, such as the Skolt Sámi keyboard, change the use and evaluation of the orthography?

References

- Aikio, Samuli. 1987. Yksi mieli, yksi kieli: Pohjoissaamen kirjakielen yhtenäistyminen. *Virittäjä* 91 (4). 457–490.
- Aikio-Puoskari, Ulla. 2002. Kielten ja kulttuurien risteysasemalla kysymyksiä saamen kielestä ja identiteetistä. In Sirkka Laihiala-Kankainen, Sari Pietikäinen & Hannele Dufva (eds), *Moniääninen Suomi: kieli, kulttuuri ja identiteetti*, 90–109. Jyväskylä: University of Jyväskylä.
- Aikio-Puoskari, Ulla. 2016. *Gullos sámegiella! Saamen kielten elvyttämisen parhaita käytäntöjä ja kansallisen politiikan linjauksia Suomessa, Ruotsissa ja Norjassa*. Inari: Saamelaiskäräjät.
- Austin, Peter K. & Julia Sallabank (eds). 2011. *The Cambridge handbook of endangered languages*. Cambridge: Cambridge University Press.
- Bergsland, Knut. 1952. Hvordan den nye samiske rettskrivningen ble til. *Sameliv/Sami ællin, Samisk Selskaps Årbok* 1951–52. 28–50.
- Bergsland, Knut & Gustav Hasselbrink. 1957. *Sámien lukkeme gárjá: sydlapsk läsebok med grammatik och ordlista*. Gardvik: P. J. Fjellström.
- Coulmas, Florian. 2002. *Writing systems: an introduction to their linguistic analysis*. Cambridge: Cambridge University Press.
- Cahill, Michael. 2014. Non-linguistic factors in orthographies. In Michael Cahill and Keren Rice (eds), *Developing orthographies for unwritten languages*, 9–25. Dallas, TX: SIL International.
- Grenoble, Lenore & Lindsay J. Whaley. 2006. *Saving languages: an introduction to language revitalization*. Cambridge: Cambridge University Press.
- Hinton, Leanne. 2014. Orthography Wars. In Michael Cahill and Keren Rice (eds), *Developing orthographies for unwritten languages*, 139–168. Dallas, TX: SIL International.

- Ivaniščeva, Ol'ga. 2016. Language policy experiments: creation of a Kola-Saami writing system in the 1930s. *Linguistica Uralica* 52 (4). 289–303.
- Jones, Mari C. & Damien Mooney. 2017. *Creating orthographies for endangered languages*. Cambridge: Cambridge University Press.
- Korhonen, Olavi. 1976. Linguistic and cultural diversity among the Saamis and the development of standard Saamish. *Linguistics* 14 (183). 51–66.
- Korhonen, Mikko. 1981. *Johdatus lapin kielen historiaan* (Suomalaisen Kirjallisuuden Seuran Toimituksia 370). Helsinki: Suomalaisen Kirjallisuuden Seura.
- Kulonen, Ulla-Maija, Irja Seurujärvi-Kari & Risto Pulkkinen (eds). 2005. *The Saami: A Cultural Encyclopaedia*. Helsinki: Suomalaisen Kirjallisuuden Seura.
- Lindqvist, Marit. 20 March 2016, updated 04 May 2016. *Våga skriva på samiska!* <https://svenska.yle.fi/artikel/2016/03/20/vaga-skriva-pa-samiska> (accessed 29 March 2020).
- Lüpke, Friederike. 2011. Orthography development. In Peter K. Austin & Julia Sallabank (ed.), *The Cambridge Handbook of Endangered Languages*, 312–336. Cambridge: Cambridge University Press.
- Magga, Ole Henrik & Lajla Mattsson Magga. 2012. *Sørsamisk grammatikk*. Karasjok: Davvi Girji.
- Moshnikoff, Jouni & Satu Moshnikoff. 2006. *Selvitys koltansaamen nykytilasta ja tarpeellisista toimenpiteistä*. Institute for the Languages of Finland, <https://www.kotus.fi/files/715/koltansaamenselvitys.pdf> (accessed 1 April 2020).
- Moshnikoff, Satu. 2014. Koltansaamen kirjakielen kehitys ja kouluopetus. In Pigga Keskitalo, Veli-Pekka Lehtola & Merja Paksuniemi (eds), *Saamelaisten kansanopetuksen ja koulunkäynnin historia Suomessa*, 270–273. Turku: Siirtolaisinstituutti.
- Olthuis, Marja-Liisa, Suve Kivelä & Tove Skutnabb-Kangas. 2013. *Revitalising Indigenous Languages: How to Recreate a Lost Generation*. Clevedon, UK: Multilingual Matters.
- Pasanen, Annika. 2015. *Kuávsui já peeivičuová. 'Sarastus ja päivänvalo'. Inarinsaamen kielen revitalisaatio* (Uralica Helsingiensia 9). Helsinki: Suomalais-ugrilainen seura ja Helsingin yliopisto.
- Pasanen, Annika. 2016. *Saamebarometri 2016. Selvitys saamenkielisistä palveluis-ta saamelaisalueella - Sámi giellabaromehter 2016. Čielggadus sámegielat bálvalusain sámeguovllus*. Oikeusministeriö <http://urn.fi/URN:ISBN:978-952-259-530-0> (accessed 3 April 2020).
- Påve, Marja. 5 April 2016. *Umesamiskan godkänns som eget skriftspråk*. <https://sverigesradio.se/artikel/6404213> (accessed 30 March 2020).
- Rask, Rasmus. 1832. *Ræsonneret lappisk sproglære efter den sprogart, som bruges af fjældlapperne i Porsangerfjorden i Finmarken*. København: J.H. Schu-bothes Boghandling.
- Rießler, Michael. 2013. Towards a digital infrastructure for Kildin Saami. In Erich Kasten & Tjeerd de Graaf (ed.) *Sustaining Indigenous Knowledge: Learning Tools and Community Initiatives for Preserving Endangered Languages and Local Cultural Heritage*, 195–218. Fürstenberg/Havel: Kulturstiftung Sibirien.
- Siegl, Florian & Michael Rießler. 2015. *Uneven Steps to Literacy. The History of Dolgan, Forest Enets and Kola Saami Literary Languages*. In Heiko F. Marten, Michael Rießler, Janne Saarikiv & Reetta Toivanen (eds), *Cultural and linguistic minorities in the Russian Federation and the European Union*. (Multilingual Education 12), 189–229. Cham: Springer.

- Sebba, Mark. 2007. *Spelling and Society: the Culture and Politics of Orthography around the World*. Cambridge: Cambridge University Press.
- Sebba, Mark. 2009. Sociolinguistic approaches to writing systems research. *Writing Systems Research* 1 (1): 35–49.
- Spolsky, Bernard. 2004. *Language Policy*. Cambridge: Cambridge University Press.
- Spolsky, Bernard. 2011. Language and Society. In Peter K. Austin & Julia Sallabank (eds), *The Cambridge handbook of endangered languages*, 141–156. Cambridge: Cambridge University Press.
- Sorjonen, Jenna. 2019. *Kuin ympäröivä luonto vai vieraaksi tehty insinöörikieli?: Suomen saamelaiskielten kirjakieliä koskevistä käsityksistä ja niiden roo-lista kielen elvytyksen kontekstissa*. MA thesis, University of Helsinki <http://hdl.handle.net/10138/301221> (accessed 3 April 2020).
- Stockfleth, Nils Vibe. 1840. *Grammatik i det Lappiske Sprog, saaledes som det tales i Norsk-Finmarken*. Christiania: Grøndahl.
- Sunna, Anna, Anne Marit Päiviö & Anna-Karin Niia. 2019. *Nu har pitesamiskan eget skriftspråk*. 19 August. <https://sverigesradio.se/sida/artikel.aspx?programid=2327&artikel=7283047> (accessed 29 March 2020).
- Wesslin, Sara & Martta Alajärvi. 2016. *Moni äidinkielen kolttsaamelainen ei osaa lukea ja kirjoittaa omaa äidinkieltään*. 15 March. https://yle.fi/uutiset/osasto/sapmi/moni_aidinkielen_kolttsaamelainen_ei_osaa_lukea_ja_kirjoittaa_omaa_aidinkieltaan/8744739 (accessed 29 March 2020).
- Wiklund, Karl Bernhard. 1922. De första lapska böckerna. *Nordisk tidskrift för bok- och biblioteksväsen*, 9:13–28.
- Ziegler, Johannes C., Daisy Bertrand, Dénes Tóth, Valéria Csépe, Alexandra Reis, Luís Faisca, Nina L. Saine, Heikki Lyytinen, Anniek A. Vaessen, & Leo P. Blomert. 2010. Orthographic Depth and Its Impact on Universal Predictors of Reading: A Cross-Language Investigation. *Psychological Science* 21 (4): 551–559.

Areal variation in structure types of Saami place names

Taarna Valtonen, University of Oulu and University of Helsinki

1. Introduction

In my doctoral thesis and in a subsequent study (Valtonen 2014), I studied the place names of four Saami communities, each of them speaking a different Saami language. Back then, my main intention was to do a basic analysis of structure and vocabulary of Saami place names and to study how the connection with other cultures and languages is reflected in place names. Even then, I realized that the research materials that I collected and the basic analysis that I conducted would provide suitable material for other linguistic analyses as well.

In this article, I return to the same materials and some of the results reached in that study (Valtonen 2014). However, I will also deal with some new materials and with completely new questions. The materials concerned here are place name samples from South Saami (southern and northern dialects), North Saami (Torne and the eastern inland dialect), Aanaar or Inari Saami (northern dialect) and Skolt Saami (southern dialect). The main objective is to show how place names can be used to study areal variation of Saami languages.

In addition to interdisciplinary toponomastics, i.e., place name studies, I will base my analysis on results gained in studies of cognitive linguistics. The long-term linguistic behavior of the users of a language is reflected in the typology of place name systems. The place names, in turn, have qualities that make it possible to connect linguistic behavior to time (often) and space (always). In this paper, I will first use onomastic and cognitive linguistic theories to argue and explain why place names are such good indicators of areal differences. In the latter part, I will give several examples of the results that can be obtained with this method.

The results provide an independent interpretation of the areal variation of Saami languages and help to recognize historical areal innovations or locally

loaned features as well as areal interdependences. These results can be compared with the results of dialectology or studies of areal variation based on reconstruction of historical evolution of related languages (e.g., Rydving 2013 based on vocabulary; Sammallahti 1998 based on inflectional morphophonology). This approach has also been inspired by Asbjörn Nesheim's (1967) studies on areal distribution of culture-specific words. By comparing and combining these independent results, we can achieve much more information about the prehistorical, early historical, and historical times of Saami cultures and languages.

2. Cognitive linguistics and place names

To give a place a name is universal behavior of human beings: some sort of place names are found in all known human languages (Ainiala et al. 2016, 13, 64). Another universal phenomenon is human beings' natural tendency to categorize everything and to interpret and convert everything new to fit already-existing categories (Lakoff 1987, xvii, 5–8). Place names are one way to categorize the surroundings of human communities: named places are defined as entities and only somehow meaningful places are named (see, e.g., Ainiala et al. 2016, 17–19).

Members of a category share certain features that make them recognizable as members of that category. They might be prototypical, i.e., be easily recognized as members due to some features that are shared by most of other members of that category, or they might be atypical, i.e., they share only some features with other members of that category. Often-given examples of prototypical and atypical members of a category are sparrow and penguin, respectively, which represent the category of birds. Individual place names, as with any other words, can be understood as prototypical or atypical representatives of their category. A typology based on the most common features of a sample of linguistic material is a structured description of categories, i.e., typologically determined categories are representations of prototypes.¹ (Sjöblom 2006, 71–73.)

According to Valtonen (2014), a prototypic Saami place name is a compound of two stems—in onomastics, known as name elements—of which the latter (generic part) is a terrain-describing appellative in nominative singular and the first (specific part), a substantive in nominative singular, e.g.,

¹ About prototype theory in general, see Dabrowska & Divjak 2015, 68; Lakoff 1987, 5–9.

SaAn *Kedgi|vyeppee*, ‘Stone Bay’.² Approximately one-third of the Saami place names that I have studied belong to this category. Second in rank are place names with another place name in genitive singular as a specific part, e.g., SaSk. *Čiŋŋlõsjääura|vääraž*, ‘Deep Lake’s Mountain’. Atypical Saami place names are, for instance, names with a particle, such as a sound-describing onomatopoeic lexeme, as a specific part, e.g., SaSk. *Šarr|čuä’lmm*, ‘Šarr [onomatopoeic lexeme describing the babbling sound of a brook] Strait’.

It is not a surprise that the place names of North Germanic and Finnic languages, the nearest neighboring languages of the Western and some of the Eastern Saami languages, also share the same prototypes. It is, however, noteworthy that the place names of the Russian language often follow another pattern: the terrain appellative is not a mandatory (generic) part of a Russian place name but an additional qualifier that can be added when needed. The Russian place names of Kola peninsula are often derivatives in which a derivational suffix is used instead of a generic part. (Kert 2005, 40; Senkevich-Gudkova 1963, 27.)

As a category of words, place names have some qualities that make them special. First, place names comprise a distinguished category of words (proper nouns) since their referent is always one definite place.³ This means that place names are linguistic utterances with a definite spatial connection. If a linguistic phenomenon is attested in a place name, this phenomenon is also tied to the referent because place names are always given in the same place or in the vicinity of the named place. It also means that when the referent was named, the language community that gave the place name used this phenomenon in their daily language and that they were physically there where the named place is.⁴ (Ainiala et al. 2016, 15–19; Coates 2016, 525–526, 538–539.)

A second quality that makes place names special is that place names are parts of systems that are called Idealized Cognitive Models or ICMs (see Lakoff 1987, 68–74), also known as schemas (see, e.g., Dabrowska & Divjak 2015, 67; Langacker 1987, 132–135; Tuggy 2012, 83–85). Schemas are organized cognitive memory structures of mind. They help us to construct in

² For a thorough description and definitions of the terminology used in this article and generally in Finnish and Saami onomastics see, e.g., Ainiala et al. 2016; Helander 1991; Kiviniemi 1990; Valtonen 2014.

³ The focus of this paper is aimed at the place names with true-life referents. This means that I do not take in account, e.g., the viewpoints of literary onomastics, a sub-discipline that is concentrated on names with imaginary referents (see, e.g., Ainiala et al. 2016, 255–264).

⁴ In my material there are no planned official place names, only names that have been given spontaneously by the language community. The method I use in this article is not suitable for analyzing planned place names because they have different naming principles.

our minds models of the objective world and to organize in simplified and abstracted form the information that we perceive with our senses (Lakoff 1987, xi–xii, xiv–xv, 58, 68; Tuggy 2012, 83–84). Individual place names are nodes or links in these multidimensional cognitive networks that comprise a schema (Lakoff 1987, 69–70; Sjöblom 2006, 73). Because place names have manifold connections in a place name schema as well as with other schemas, they carry and accumulate rich language- and culture-specific meanings (see also Ainiala et al. 2016, 17–20).

Place name schemas are coherent systems that follow language- and culture-specific rules, i.e., rules that regulate the functions of and interaction between schemas. These rules are called place name grammar (see, e.g., Ainiala et al 2016, 36, 63; Zilliacus 2002, 142–143). They regulate the naming process, e.g., which places are named and what kind of linguistic structures and themes are accepted for a name. Place name grammar also tells how names should be used in varying linguistic and cultural contexts. Place name grammar is constantly but very slowly changing due to language's internal evolution, changes in the world or influences acquired from other languages and cultures. The place name grammar of genetically related languages is often quite similar, but similarities are also born in close contact between unrelated languages or, at least theoretically, independently by chance (see Valtonen 2019, 8–16).

A third quality that makes place names special is that place names of major referents are durable, often hundreds, but sometimes even thousands, of years old, and conservative (see, e.g., Ainiala 1997, 233–234; Pamp 2007, 88–89). They can preserve linguistic phenomena that were used during the naming period but have otherwise become extinct in a language. In addition to the fact that individual place names can be very old, the place name schemas that provide prototypical place name models or molds that are used when new place names are created are even more durable than individual names and change very slowly. This means that place names and place name systems are also a great repository of old local linguistic innovations and loans. (Ainiala et al. 2016, 16–17, 21–22, 36, 79; Coates 2016, 525–529, 534–536; Kiviniemi 1990, 121, 207–209; Sjöblom 2006, 72–73.)

These three qualities make place names excellent material for diachronic areal studies. When a place name is given, it is based on the local linguistic and cultural model, i.e., place name system or schema, which regulates its content, structure, and use. It is from that moment connected to a certain location, i.e., the referent. The place name and its linguistic features can be preserved in a name connected to this location even though the linguistic context around the named place would be changing. Sometimes place names

have preserved loans, such as loan words or loaned constructions, from other languages or local linguistic innovations. These loans and innovations are stored in the place name system and can later be attested as areal differences. (See Fig. 1.)

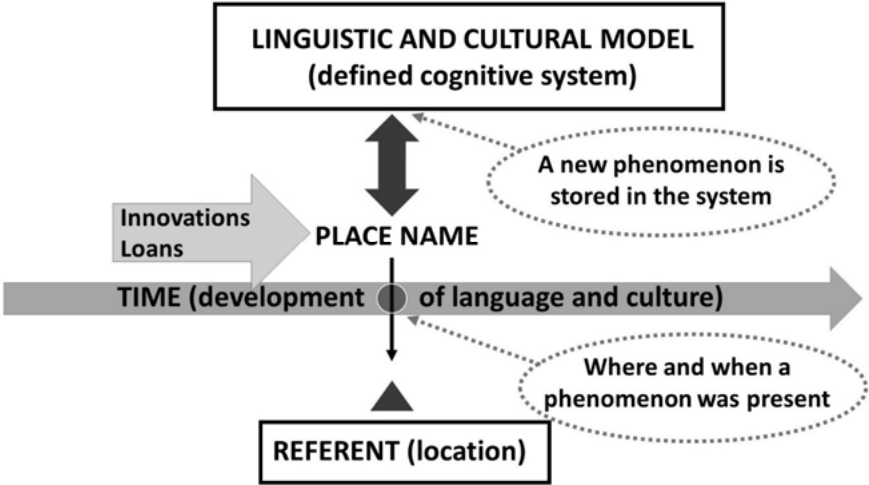


Figure 1. The theoretical model that describes place names’ ability to capture and preserve linguistic innovations and loans and to connect them with permanent locus.

3. Method and research material

A sample of place names can be divided into typological categories according to many features, such as the naming principles, i.e., the themes of place names (see, e.g., Ainiala et al. 2016, 72–75; Kiviniemi 1990, 119–162), type of referents (see, e.g., Aalto 2019), and vocabulary or structure of names (see, e.g., Valtonen 2014). All these categorizations open one important aspect of a place name system and add to our understanding of its nature. However, when it comes to contrastive studies based on several place name systems, the systematic categorization based on structure gives the most precise results. The reason is that the grammatical rules that linguistic constructions are based on represent such a high abstraction level that they are almost subconscious in nature and followed automatically. This means that grammar is not in the acknowledged core of the naming process but only provides unconsciously understood linguistic limits for the process. In addition, it has been noted that it is not common that influences coming from other languages and cultures would be borrowed to a morphological system. However, place names are as

subject to normal phonetic change as any other word of a language. (Coates 2016, 539; Lakoff 1987, 6, 13; Sankoff 2013, 11–13.)

By comparing the structure of Saami place names, it is possible to learn which constructions are attested, which are common, and which are uncommon. Such knowledge helps, for instance, to create convincing place name etymologies, since constructions that are common now were often common already hundreds of years before and vice versa. The conservative natures of Saami place name systems can be attested by studying old maps and other old documents.⁵ For instance, in many Saami languages the genitive singular and genitive plural are distinguished by only one sound at the end of the stem. If a historical place name or name borrowed from a Saami language to another language can be reconstructed as either of those, it is more convincing to propose genitive singular since genitive plural is extremely rare in present-day and historically recorded place names of Saami languages (Valtonen 2014, 272). Typology is a tool for source criticism that can provide limits of plausible interpretations (for principles of source criticism in etymological onomastic studies, see Aikio, A. 2007.) Due to spatial connection, the typology based on structure can also be used to study areal variation among place name systems of Saami languages. In such studies the typological similarities can be seen on distribution maps as concentrations or blank spaces that represent the original area of past innovations and contacts.

Valtonen (2014, 200–222) presents a method for dividing place names of Saami languages into categories based on their syntactic structure, morphology, morphophonology, and, in some special categories, also semantics. The method is simple. First, the place names are divided into syntactic parts (i.e., name parts), to a specific part and to a generic part that most often in Saami languages are stems of a compound (for dividing rules, see Ainiala et al. 2016, 36–37, 71–82; Valtonen 2014, 200–202). The division is marked with a pipe (|) in the examples. The further analysis is based mainly on specific parts since the generic parts are most often in the nominative,⁶ which makes them uninteresting from a comparative point of view.

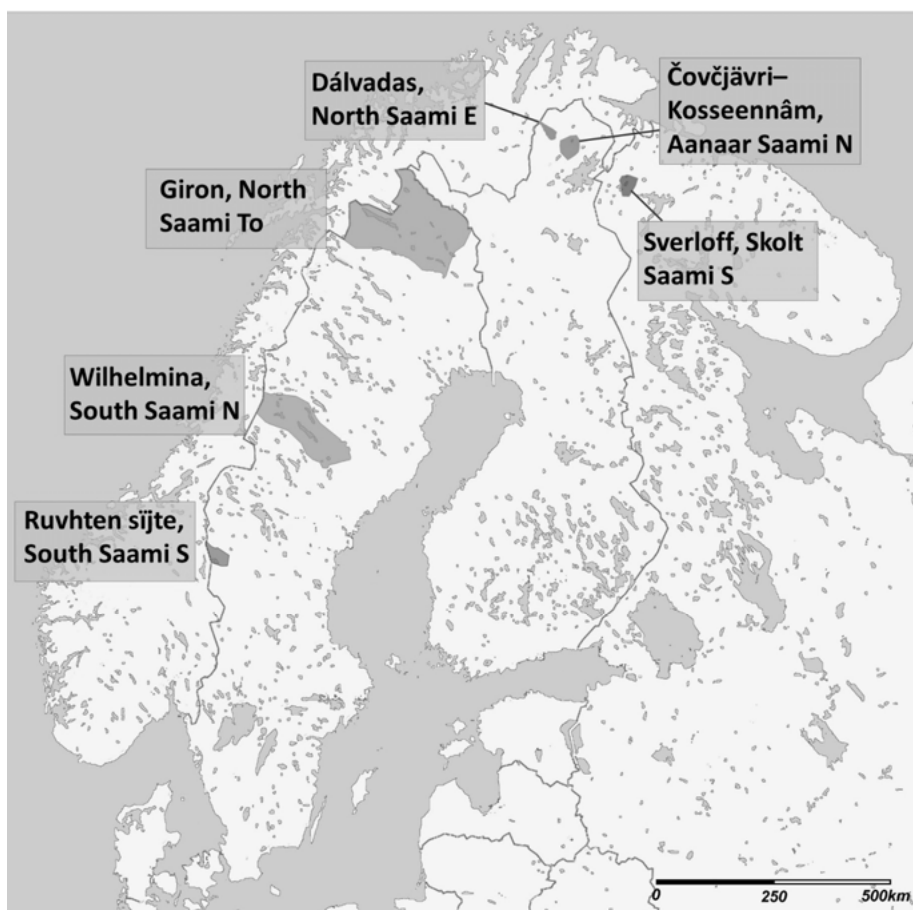
The place names in Valtonen (2014) are further categorized according to the morphology, morphophonology and semantics of the specific part. In addition, there are separate structure groups for place names with only one

⁵ E.g., old taxation documents, Jonas Persson Gedda and Anders Olofsson Holm's map and other documents from Ume Lappmark of 1671 (see Norstedt 2011), Olof Tresk's map and description of Torne and Kemi Lappmark of 1642–1643 (see Tresk 1928), and Peter Schnitler's maps and documents of the border region between Denmark(-Norway) and Sweden of 1742–1745 (see Schnitler 1742–1745).

⁶ For a special type in which the basic form is in genitive case only found in South Saami, see Valtonen 2014, 208, 210–212.

name part (typically the generic part), place names with a specifying modifier as the specific part, place names with loaned name elements (hybrid place names) and place names that are unanalyzable because the meaning is unclear (Valtonen 2014, 203–222). The two latter types are not discussed in this article, but shortening of unstressed middle name elements (reduction) and the number of name elements, i.e., the number of stems in a place name, are taken in account as structural phenomena. When counting the number of name elements, each stem (i.e., a name element) with its case endings or derivational suffixes is counted as one. In compound words, each separable stem is counted as one. The method of counting the number of name elements is adopted from Kiviniemi (1990, 91; see also Valtonen 2014, 267).

In this article, I will present examples of analyses that can be made with the above-described method. It is worth keeping in mind that the main objective of this article is to introduce the method, not the results. My material consists of six separate place name samples. Four of them are the materials that were analyzed in Valtonen (2014). These are 168 South Saami (southern dialect) place names (Ruvhten sįjte, Herjedaelie/Härjedalen, Sweden); 222 North Saami (eastern inland dialect) names (Dálvadas, Ohcejohka/Utsjoki, Finland); 561 Aanaar or Inari Saami (northern dialect) names (Čovčjävri and Kosseennâm areas, Aanaar/Inari, Finland) and 655 Skolt Saami (southern dialect) names (Sverloff family area, Suõ’nn’jelsijdd/Suonikylä/Songelsk, Peäccam/Petsamo/Pechenga, Russia). For the purposes of this article, I have additionally analyzed two other samples: 1747 North Saami (Torne dialect) place names of the Giron/Kiruna municipality (Sweden) and a sample of 569 South Saami (northern dialect) names of the Vualtjere/Vilhelmina municipality (Sweden). (See Map 1.) In addition, I will make some minor notes on place names of Kildin Saami, Ume Saami and (historical) Kemi Saami.



Map 1. Studied areas, languages, and dialects.

The major difference between the materials in Valtonen (2014) and the materials analyzed for the purposes of this article is that the first mentioned are based on detailed archive studies, but the two newer materials are map names uploaded from the internet service *Kartsök och Ortnamn* provided by the Lantmäteriet, i.e., the Swedish Mapping, Cadastral and Land Registration Authority.⁷ In archive collections, it is possible to take into account the natural variation of a place name system, but on maps, only one name variant can be displayed. This means that the materials from Giron and Vuoltjere do not represent the whole spectrum of variation. However, it is worth remembering that place names printed on maps of Nordic Countries are real place names used by real language communities. It is equally important to understand that

⁷ Since 2019 when the present material was uploaded, the internet service *Kartsök och Ortnamn* (kso.etjanster.lantmateriet.se/#) has been replaced by a new, very similar internet service called *Ortnamn* at ortnamn.lantmateriet.se/.

map names present only a sample of the real place name heritage. Note also that the areas of Giron and Vualtjere displayed on Map 1 represent whole municipalities even though most Saami place names can be found in the western or the more remote parts of those communities. Other studied areas, Ruvhten sįjte, Dálvadas, Čovčjávri–Kosseennâm, and Sverloff family area, are displayed as the members of the communities have defined them (see Valtonen 2014, 78–193).

Place names of Saami languages share many features with place name systems of other North European languages. Especially closely related are the place name systems of Finnic languages, especially Karelian, but Saami place name systems also share many features with place name systems of North Germanic languages. It is noteworthy that, when compared with the Russian place name system(s), the differences are more distinctive, as described above. Since the major structure-type categories of place name systems of Saami languages are mostly similar to the structure-type categories of the closest neighbor languages or we are lacking detailed comparable information, one should be very careful when comparing these categories between Saami languages. From the viewpoint of source criticism, it would be impossible to tell whether the differences are caused by differences between Saami languages or only by influence from the majority languages. Because of this, I will examine the morphological differences by analyzing only the marginal constructions that are not typical in neighbor languages. However, this means that some of these structure types that I present are very rare. Because of this, the results that I present should be read as preliminary approximations, not as definitive truth.

In this article, I have chosen to take as examples place names with a specific part that is a personal name, an action nominal (verbal noun), an attributive adjective or a specifying modifier. In addition, I will present some areal differences of other constructions: place names with shortening of unstressed middle name elements (reduction) and number of name elements. The purpose of this article is to present the method, and I want to emphasize that numerally and spatially, more comprehensive materials can provide more convincing results than what I can present here. More comprehensive data would also make it possible to analyze a wider selection of structure types without fear that the borrowed and local linguistic innovations of the naming systems would be as easily confused with each other.

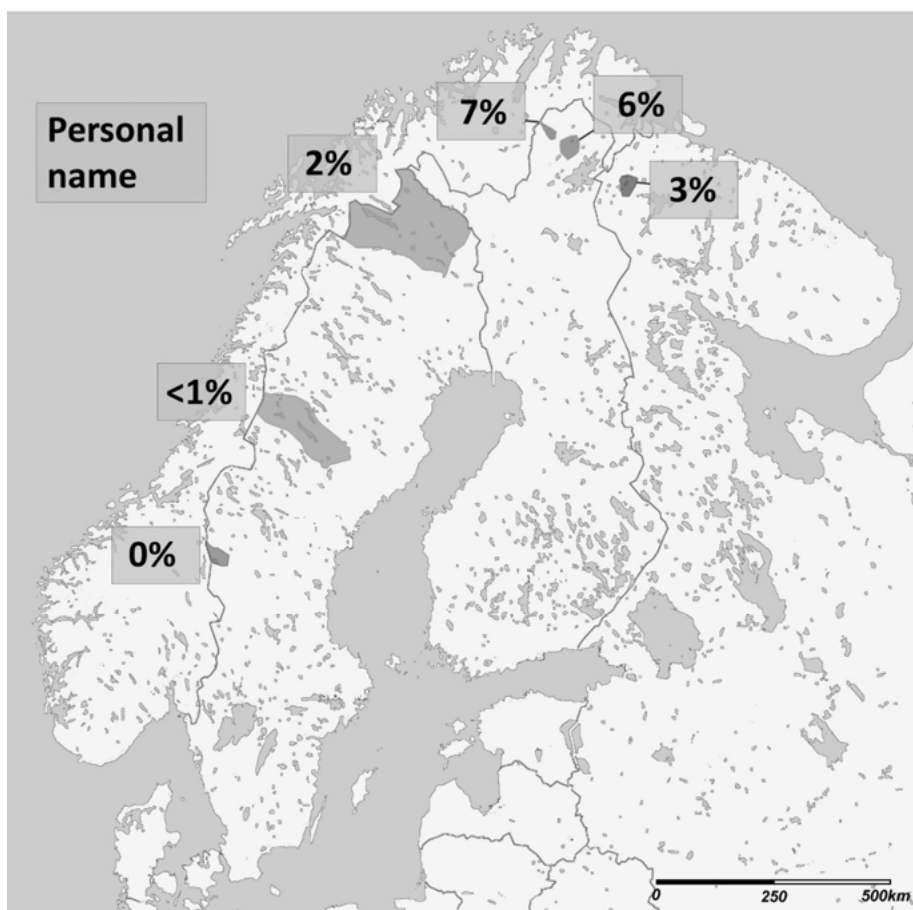
4. Results

4.1 Personal name as a specific part

Place names with a personal name as a specific part⁸ have a surprising areal distribution. They are common in Dálvadas (7% of all the names of this material) and Čovčjävri–Kosseennâm (6%), less common in Giron (2%) and Sverloff family area (3%), but nonexistent in Ruvhten sijte and only one place name, *Meerken|jaevrie*, ‘Meerke’s Lake’, is found in Vualtjere. The eastern inland dialects of North Saami and Aanaar Saami seem to belong to the core area of this phenomenon, whereas the phenomenon seems to be clearly less typical in the surrounding Torne Saami and Skolt Saami languages. In South Saami, on the other hand, this structure type does not appear to be among the prototypical place name types at all. (See Map 2.) Because of the limited material I am dealing with, it is impossible to say where the borderline between the northern and southern area goes. However, it is noteworthy that in the map of Ume Lappmark, drawn by Jonas Gedda and Anders Holm in 1671, there are some place names that seem to include a personal name, e.g., *Ouls Jauri* ~ *Oulus Jaur* ~ *Ohlss Jauri*, ‘Åvl(o)s’s Lake’ and *Clemetz-Jauri*, ‘Clemet’s ~ Clemets(en)’s Lake’ (Norstedt 2011, 199, 206). This area is only about 100 kilometers north of the northernmost South Saami area.

According to Nicolaisen (1986), in the languages of the world, on average, 10 percent of place names include a personal name. Kiviniemi (1986) writes that from 10 to 15 percent of Finnish place names have a personal name as one name element. From this perspective, all of the studied Saami place name systems differ from the average, but South Saami material is something truly special. To cross-check the result, I searched the 28 most common South Saami personal names from Swedish and Norwegian place name databases that include thousands of place names. The result was two South Saami place names that contain a personal name. *Meerken|jaevrie*, mentioned above, was one of them. Since personal names are common in the place names of the North Germanic languages (see, e.g., Nyström 2021, 59–101; Rygh 1898), the best explanation seems to be that the South Saami place name grammar includes a culture-specific rule that personal names do not belong to place names. At the present, it is impossible to say whether it is South Saami or Aanaar and North Saami eastern inland dialect that represents an innovation.

⁸ In this paper, family names, which are quite rare, have been included in this group. Personal names as specific parts are always in genitive, typically in genitive singular.



Map 2. Personal name as a specific part.

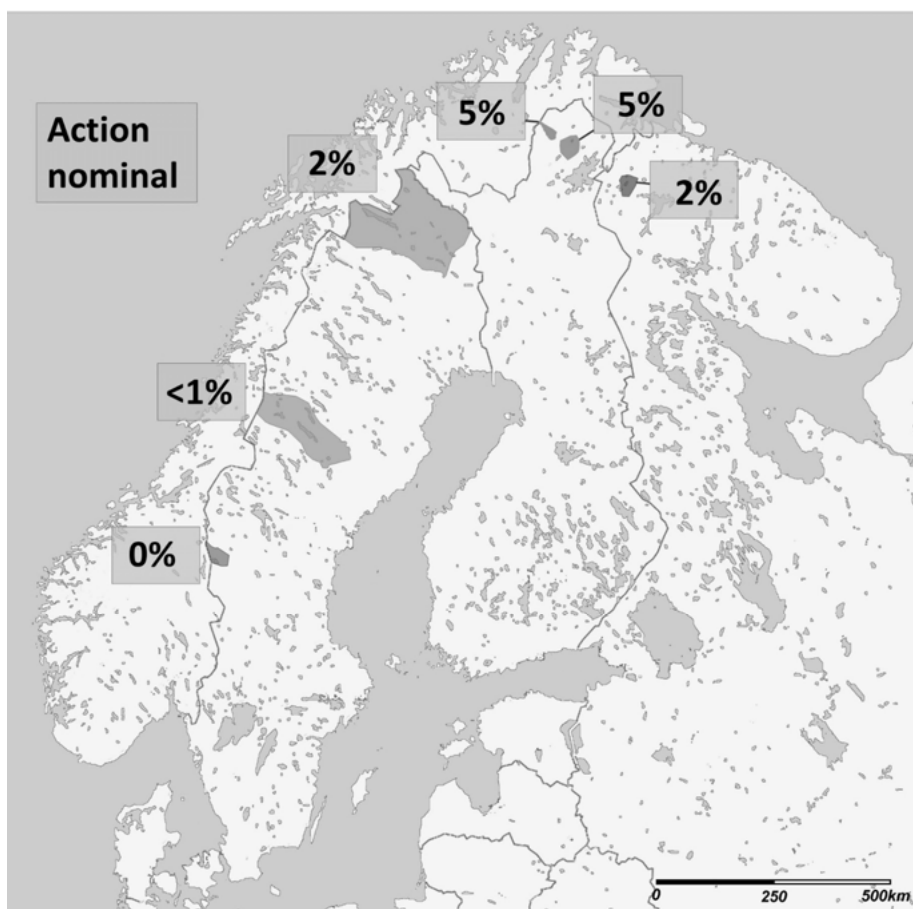
4.2 Action nominal as a specific part

Place names with action nominals, also known as verbal nouns, as a specific part have an interesting areal distribution. In South Saami, they are very rare: absent in the Ruvhten sįjte's material and less than 1 percent in Vualtjere's material, one example being *Orremeljohke*, 'Living River'. In Giron's and Sverloff family area's materials, there are 2 percent of such names. However, 5 percent of Dålvasdas's and Čovčjävri-Kosseennâm's place names have an action nominal as a specific part, making it a rather common construction.

These results, shown in Map 3, indicate the following three areal differences: 1) South Saami with marginal use of action nominals; 2) North Saami Torne dialect and Skolt Saami with minor use of action nominals; and 3) North Saami eastern inland dialect and Aanaar Saami with quite commonly used action nominals—hence, the same situation as with personal names as a

specific part. Gedda and Holm's map of Ume Lappmark from 1671 presents evidence that action nominals have been used for a long time, at least in place names in an area close to present-day South Saami areas. In Máláge/Malå, some 100 kilometers northeast of Vuoltjere, Holm reports a place name *Pwelem Kidde/Brändwallen*—in modern Ume Saami orthography, *Buallám|geäddie*, 'Burned Dwelling Place' (Norstedt 2011, 207).

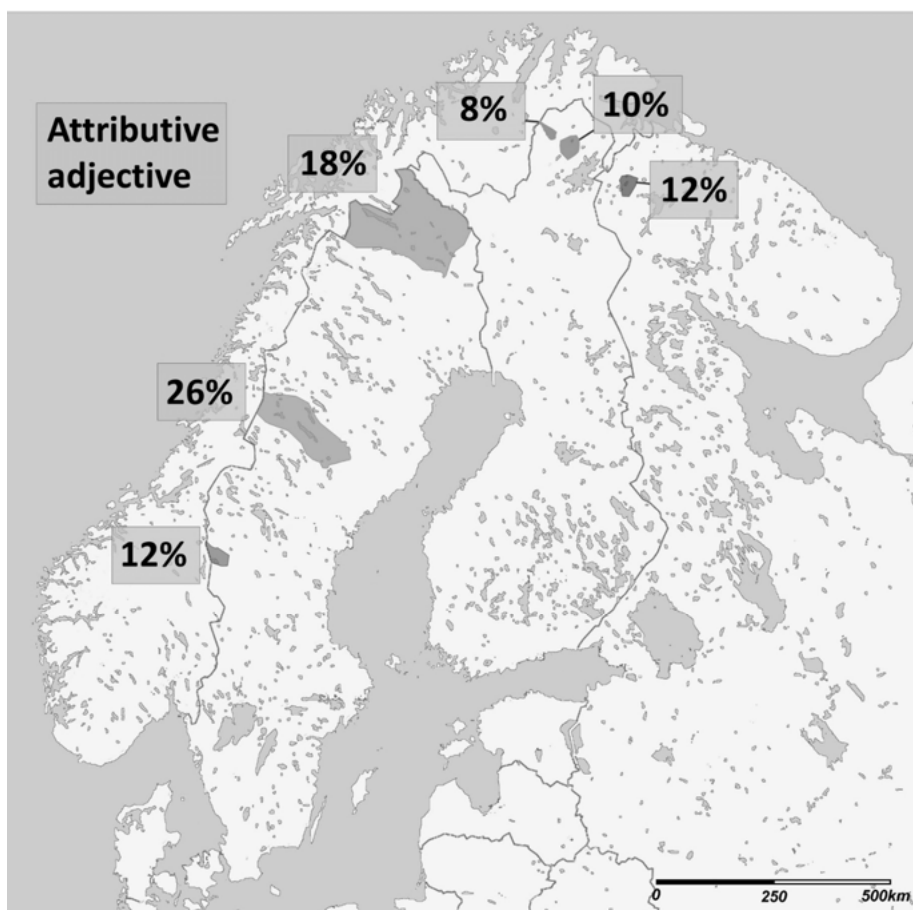
In addition, there are among Giron's, Dálvadas's, Čovčjävri–Kosseennâm's, and Sverloff family area's materials place names that have an action nominal and its object as a specific part. One such name is *Gazzagaikun|johka*, 'Hoof Breaking River' belonging to the only place name cluster in Giron of this type. By contrast, in Dálvadas, only 2 percent and in Čovčjävri–Kosseennâm 1 percent of place names have this construction. In Sverloff family area, the construction is not common. In Dálvadas, Čovčjävri–Kosseennâm and Sverloff family area, there are also place names constructed by an action nominal and its subject as a specific part, such as *Vue'zzkõckâm|jääuraž*, 'Reindeer Calf Drowning Lake'. These are, however, quite rare. In Čovčjävri–Kosseennâm area, there are in addition some examples of place names with an action nominal and its subject and object as specific parts, such as *Maati |Kähvivuoššâm|jávrááh*, 'Matti's Coffee Making Lakes'. Such North Saami place names are also found in Ohcejohka/Utsjoki, e.g., *Ville |Gámalahppinvárri*, 'Ville's Shoe Losing Mountain'. These findings are in good agreement with the distribution described above and support the idea that Aanaar Saami and the eastern inland dialect of North Saami might represent an innovation area.



Map 3. Action nominal as a specific part.

4.3 Attributive adjective or specifying modifier as a specific part

Even though attributive adjectives are common as specific parts in compound place names of Saami languages, as well as in neighboring languages, I have decided to present these results because they reveal some phenomenon or phenomena that seem to have nothing to do with the majority languages. In the materials representing the Eastern Saami language, i.e., Aanaar and Skolt Saami languages, a moderate 10 and 12 percent of place names have an attributive adjective as a specific part. In Dálvadas, the number is a bit lower, 8 percent. But in Giron, 18 percent of place names have an attributive adjective as a specific part. However, the situation in South Saami is even more surprising: 26 percent of Vualtjere's place names, but only 12 percent of Ruvhten sįjte's place names, have an attributive adjective as a specific part. (See Map 4.)

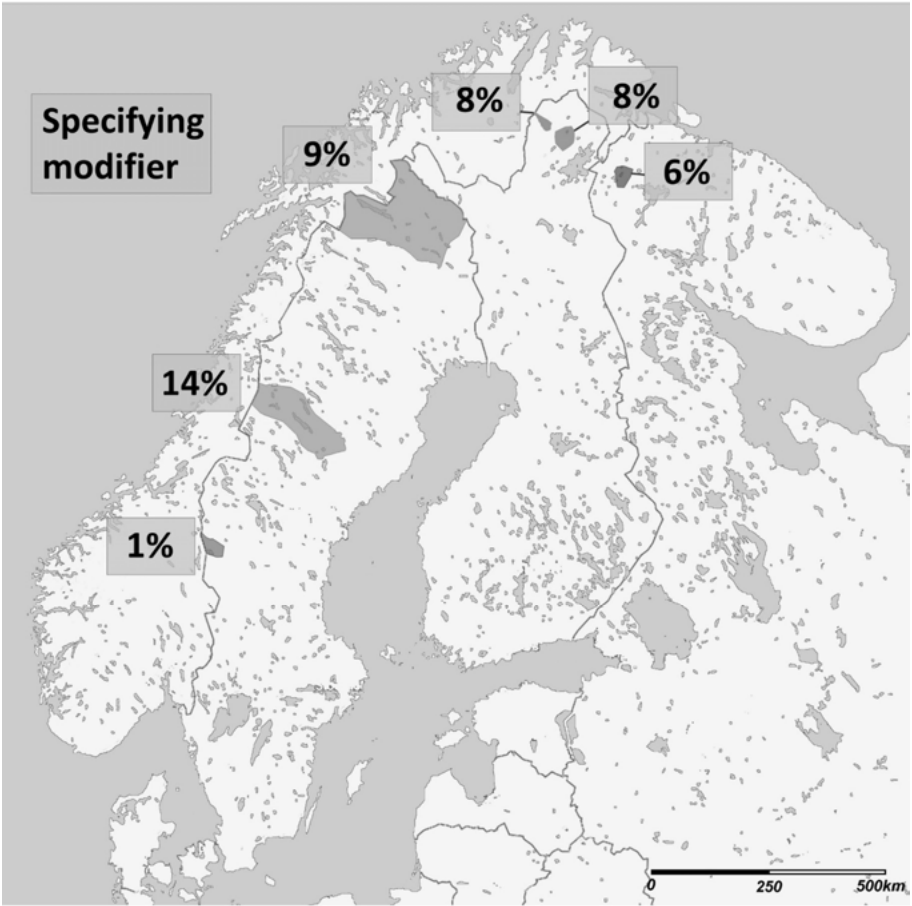


Map 4. Attributive adjective as a specific part.

The situation in Giron can hardly be explained as a Meänkieli influence, since according to Eero Kiviniemi (1990, 150), only 8 percent of Finnish place names have an attributive adjective as a specific part. However, it is important to look at another map (Map 5) that presents the percentage of place names with a specifying modifier as a specific part. This is a structure type that has as a generic part a whole place name that is distinguished from other, similar place names by a specifying modifier, such as South Saami (N) *Ohtje| Sudtsjaevrie*, ‘Small Ice-Free Lake’ and *Stoere| Sudtsjaevrie*, ‘Big Ice-Free Lake’ or Aanaar Saami *Irjánâšrohe| Lyemekaavâ*, ‘The Late Irjánâš’s Cloudberry Bay’.

As can be seen from these examples, some of these names have an attributive adjective as a specific part, but the number of such names varies from one language to another. In Ruvhten sǣjte, all the place names belonging to this type have an attributive adjective as their specific part, in Vualtjere all

but one name, and in Giron 96 percent of such names. But in Dálvadas, Čovčjävri–Kosseennâm and Sverloff family area, only approximately half of the specifying modifiers are attributive adjectives. This makes the differences between Giron’s and Dálvadas’s place name systems even more evident: the rules that specify what kind of lexemes can be used as a specifying modifier are important in this context. There are, however, also differences between South Saami materials and Giron, since comparative and superlative forms of adjectives are not attested in the South Saami materials but are quite common in Giron’s material. It seems that place names that include adjectives form several structure-type categories, which differ from one Saami language to another. This topic should be studied further.



Map 5. Specifying modifier as a specific part.

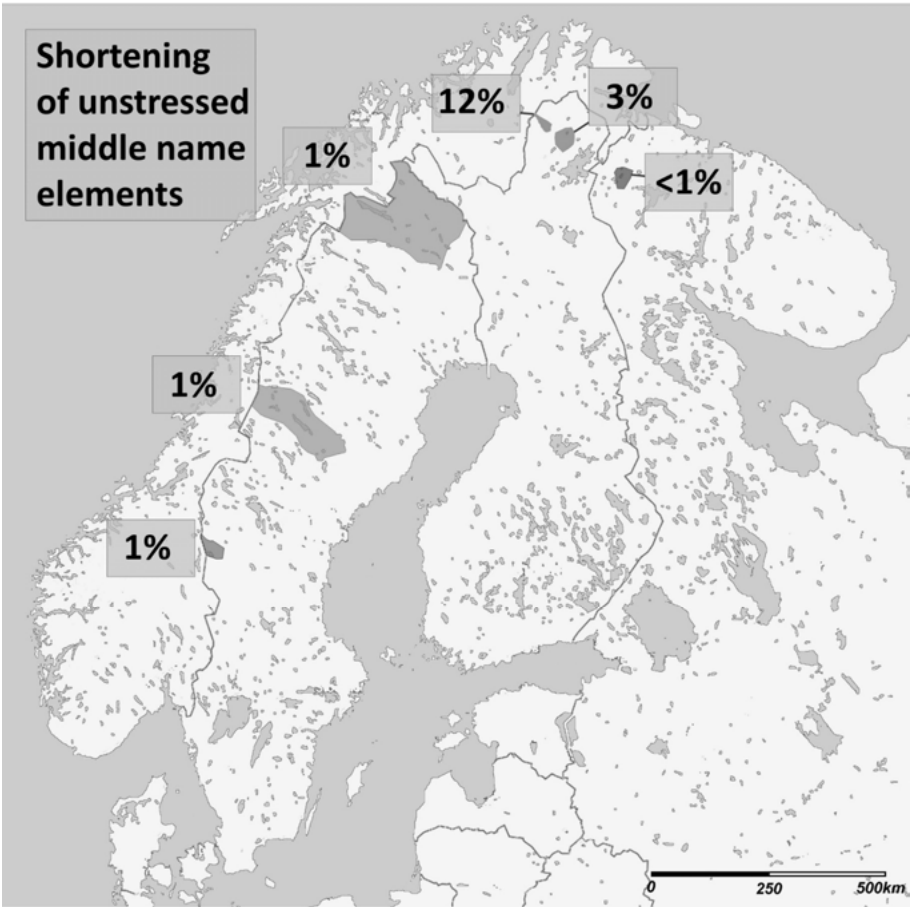
4.4 Shortening of unstressed middle name elements (reduction)

There are in each of the studied languages morphophonological shortening processes that also concern place names. These processes are typically manifested in the unstressed middle name elements of place names with three or more name elements. In North Saami (at least western and eastern inland and sea dialects), the reduction process is systematic, though somewhat irregular, and can be attested also in compounds other than place names, especially often in the eastern inland dialect (for a full description of the process see, e.g., Aikio, A. 2007, 173–174; Aikio, S. 2017, 14–15; Helander 1991, 82; Sammallahti 2019, 219–220; Valtonen 2014, 254–256). Some examples from Dálvadas are *Beahcelljohka* < *Beahcelahjohka* < **Beahceleagijohka*, ‘Pine River Valley’s River’ and *Njiljohmoh|várri* < **Njiljogamohki|várri* ‘Njil River’s Curve’s Mountain’.

The typicality can be clearly attested also in the results of the present study: up to 12 percent of place names of Dálvadas have this type of shortening. In other studied languages, the process is clearly less common and can be attested regularly only in my Aanaar Saami material (3%), e.g., *Hiävušvär|luobâl* < *Hiävušvääri|luobâl*, ‘Horse Mountain’s Lake’. It is, however, good to keep in mind that the Čovčjävri and Kosseennâm areas belong to the northern dialect, which has adopted many features from nearby North Saami dialects. How typical and systematical this phenomenon is, should be tested with material from the eastern dialect. The phenomenon is almost nonexistent in my Skolt Saami material, probably mainly due to the linguistic processes. As a result of these, the second syllable vowels have been shortened to extra-short, and the historically disyllabic stems have become monosyllabic, causing the shift of the syllable boundary as a rule.

However, the materials from Giron, Vualtjere and Ruvhten sįjte also include a very low amount of shortened place names (1% each). The low amount is not a big surprise in the South Saami materials, since South Saami has in general preserved the historically long forms and end-syllables. The clearest examples that can be found are of type *Viglenjuen|baahka* < *Viglenjuenien|baahka*, ‘Vigle Mountain’s Slope’ (in Ruvhten sįjte) and attested only in archive records. On maps, only such variants as *Baeltehk|jaevrie* < **Baelhtehke(n)|jaevrie*, ‘Cliffy(‘s) Lake’ can be seen. On the other hand, the small number of such cases in Giron is very surprising, since reduction of unstressed middle name elements is often presented as a common North Saami phenomenon (see, e.g., Aikio, A. 2007, 173–174; Aikio, S. 2017, 14–15; Helander 1991, 82; Sammallahti 2019, 219–220; Valtonen 2014, 254–256). The phenomenon is present in the Giron material, and these cases are consistent with the morphophonology of similar cases in

Dálvadas material, but the number is very low. In this case, it seems that a clear difference in the place name grammar is found inside North Saami. (See Map 6.)



Map 6. Shortening of unstressed middle name elements.

3.5 Number of name elements

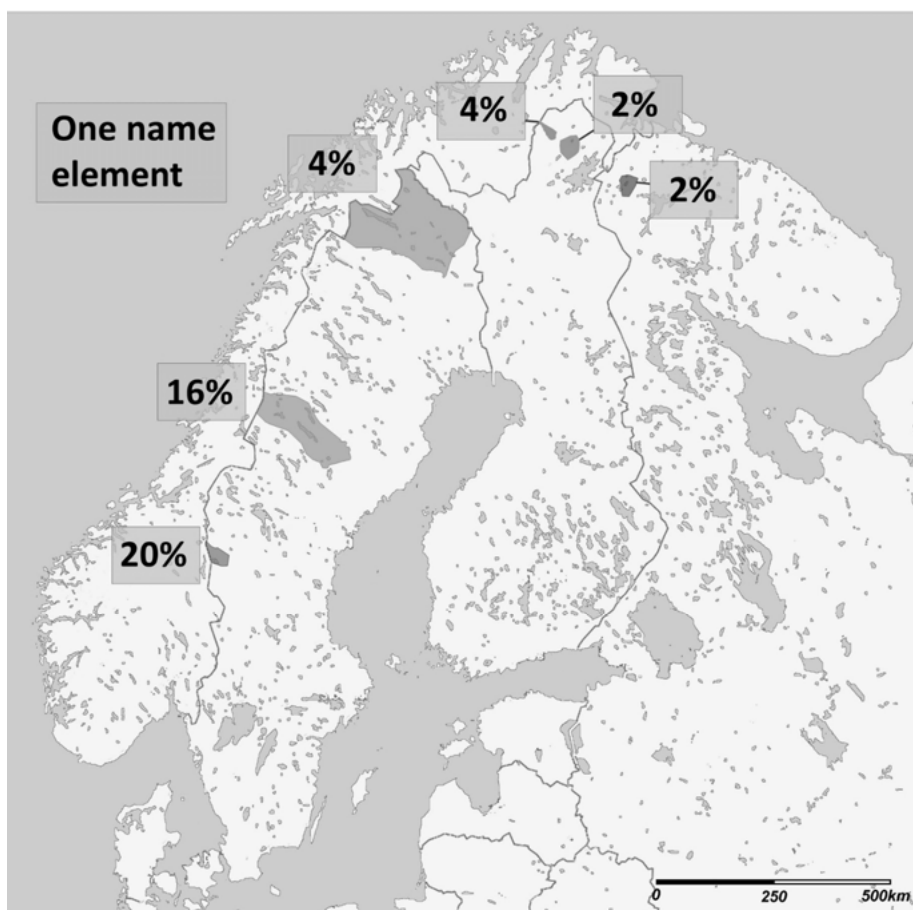
How many name elements there can be in a place name is one of the structural features, which is guided by place name grammar as much as the above-mentioned choices of specific parts or shortening processes. The shortest place names in my research material comprise a single name element, i.e., of one stem. Into this group I have also counted combinations of stem and derivational suffix. They are, however, very rare in other materials except in Vualtjere, where the number is a bit higher than in other materials due to commonness of place names belonging to the *-Vre*-reduction construction

type, e.g., *Gierkere* < *Gierkie|vaerie*, ‘Stone Mountain’ and *Saaljere* < *Saalje|jaevrie*, ‘Goat Willow Lake’ (for detailed description, see Valtonen 2014, 208, and sources mentioned there).

As can be seen in Map 7, South Saami seems to differ significantly from other Saami languages with 20 and 16 percent of place names with one name element. In addition, in the thoroughly studied material of Ruvhten sįjte, 13 percent of the names belong to a unique South Saami structure type that I have referred to in Valtonen (2014) as freely varying single- or two-part names and that have either one or two syntactic name parts, such as *Bavla* ~ *Bavlan|gāetieh*, ‘Hollow ~ Hollow’s Dwellings’ in Ruvhten sįjte. In everyday speech, these names are often used as a single-part variant, but the two-part variant is thought to be the complete form. Because the “complete” forms are most often chosen for maps, this phenomenon cannot be attested in my material from Vualtjere, but it is attested in place name materials in all South Saami areas (see Valtonen 2014, 210–212, for a full description of this type and Valtonen 2014, 208–209, for similar but rare processes in North Saami).

The more northern materials seem to be divided into two areas: the Giron (4%) and Dálvadas (4%) vs. the Čovčjävri–Kosseennâm (2%) and Sverloff family area (2%). However, the difference is not as great as compared to South Saami. It is interesting that the smallest amounts are in the east, since single-part (one name element) place names, especially stem + suffix constructions, are common in the Russian place names of Kola Peninsula and Murmansk Oblast, e.g., *Лопарская* (*Loparskaya*) ← *лопар*, ‘a Saami (person)’ + possessive/adjective suffix -ск- + feminine ending -ая (see also Kert 2005, 40; Senkevich-Gudkova 1963, 27–28). This seems to indicate that at least the Eastern Saami languages belonging to the mainland group (see Sammallahti 1998, 26–31) have not had intense influence from the Russian language.⁹ According to my preliminary observations based on a place name sample of some hundred lexemes (Itkonen 1958 and the map of Aarsjogk sįjtt [Varzino] published by Sámiid Vuorká-Dávvirat in 2019; see also Kert 2002, 127), the situation in Kildin Saami seems to be the same as in the mainland group.

⁹ I have also gone through the place names belonging to the northern dialects of the Skolt Saami published by Itkonen (1958) and Tanner (1928), and they do not differ significantly from the place names of the southern dialects. Since I have not made similar detailed calculations for the materials, I cannot give figures for them. Akkala Saami place names are very poorly known, and name collections are not readily available. The few place names presented by Senkevich-Gudkova (1963) resemble the place names of Skolt Saami.



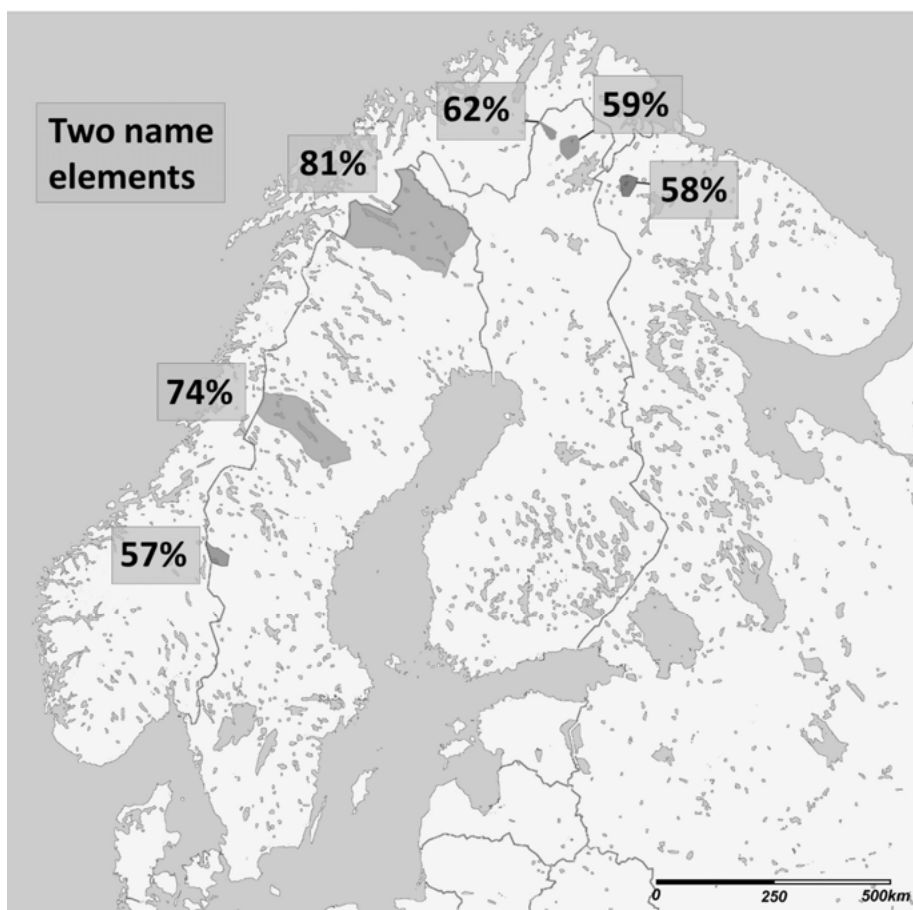
Map 7. Place names with one name element.

There are two name elements in most place names in all Saami languages. In most cases, this means that there is one stem as a specific part and one stem as a generic part. However, there are some areal differences. The results of Dálvadas, Čovčjävri–Kosseennâm and Sverloff family area are very similar: 62, 59 and 58 percent, respectively, but in Giron the situation is different: 81 percent of the names have two name elements. The results are almost equally high in Vualtjere (74%), but it is likely that a significant part of these names belong to the group of freely varying single- or two-part names. In Ruvhten sįjte, the result (57%) seems to follow the pattern of the northernmost area, but if the freely varying single- or two-part names (13%), as well as the few names (2%) that vary between two or three name elements, were added, the result would be 72 percent. Due to differences in the nature of the studied materials, the idea of a comparable situation in the South Saami area is, however, speculative only. (See Map 8.)

Because the materials from languages and areas between South Saami and North Saami Torne dialect are not included in this analysis, it also remains open whether the Giron result represents an independent innovation area or is connected to the South Saami results. The idea of an independent innovation based on linguistic contacts is supported by the language situation: in Giron there have been speakers of Meänkieli at least since the 17th century, and the languages and their speakers have been interacting in many ways for centuries. This includes place names that most often have been borrowed from North or Lule Saami to Meänkieli (e.g., Me. *Vakkojärvi* < SaN. *Vähkujávri*), but there are also place names borrowed from Meänkieli to Saami (e.g., SaN. *Murhávárri* < Me. *Murhavaara*) and reverse loans (e.g., SaN. *Alajávri* < Me. *Alajärvi* < SaN. *Vuollijávri*;¹⁰ see, e.g., Swedell 2008; 2009, 28–31).¹¹ According to results from my former study (Valtonen 2014, 266–270), the Finnish place names are on average one name element shorter than the place names of North Saami (eastern inland dialect), Aanaar Saami and Skolt Saami (see also Kiviniemi 1990, 91). The place names with three or more name elements are clearly rarer in place names of Giron than in place names of Dálvadas, Čovčjávri–Kosseennâm and Sverloff family area.

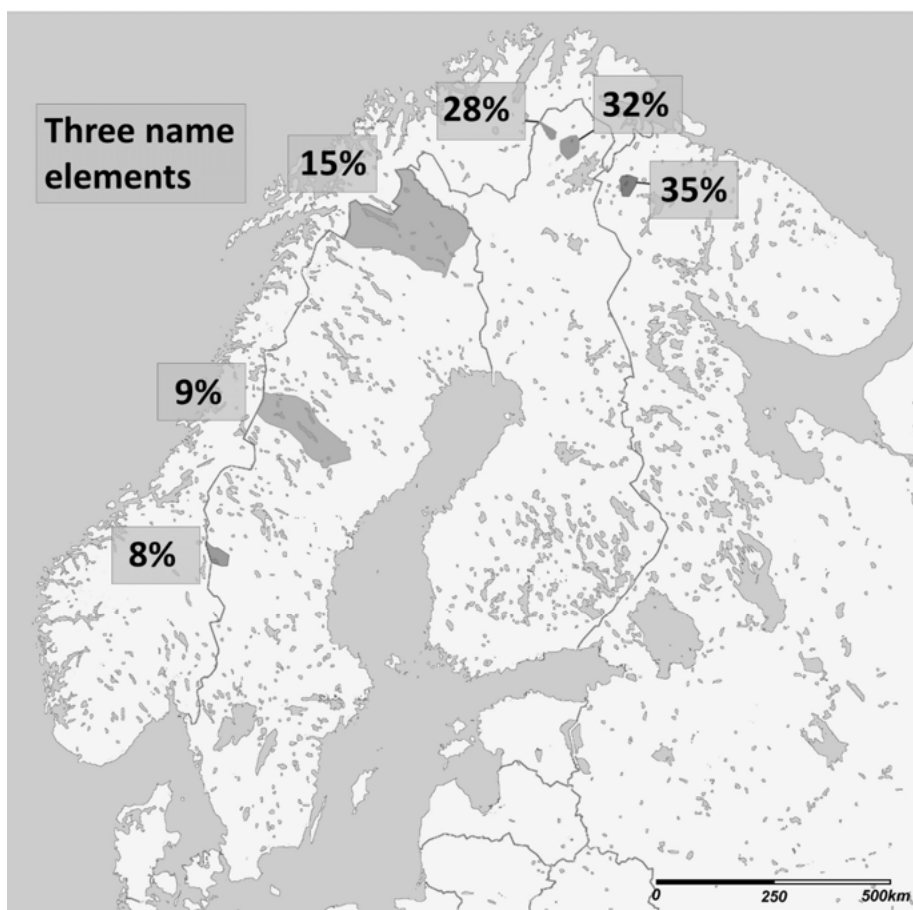
¹⁰ These place name examples are taken from the digital version of the place name archive Namnarkivet i Uppsala administered by Institutet för språk och folkminnen (Isof), see <https://ortnamnsregistret.isof.se/place-names>.

¹¹ In her autobiographical book *Nådevalpar* (1971), Saami writer Sara Ranta-Rönnlund, born in Giron in 1903, describes well the trilingual way of life in the municipality and especially the intertwined lives of the Saami-speaking reindeer herders and Meänkieli-speaking peasants. Her book includes several passages in which multilingual place name practices are included.



Map 8. Place names with two name elements.

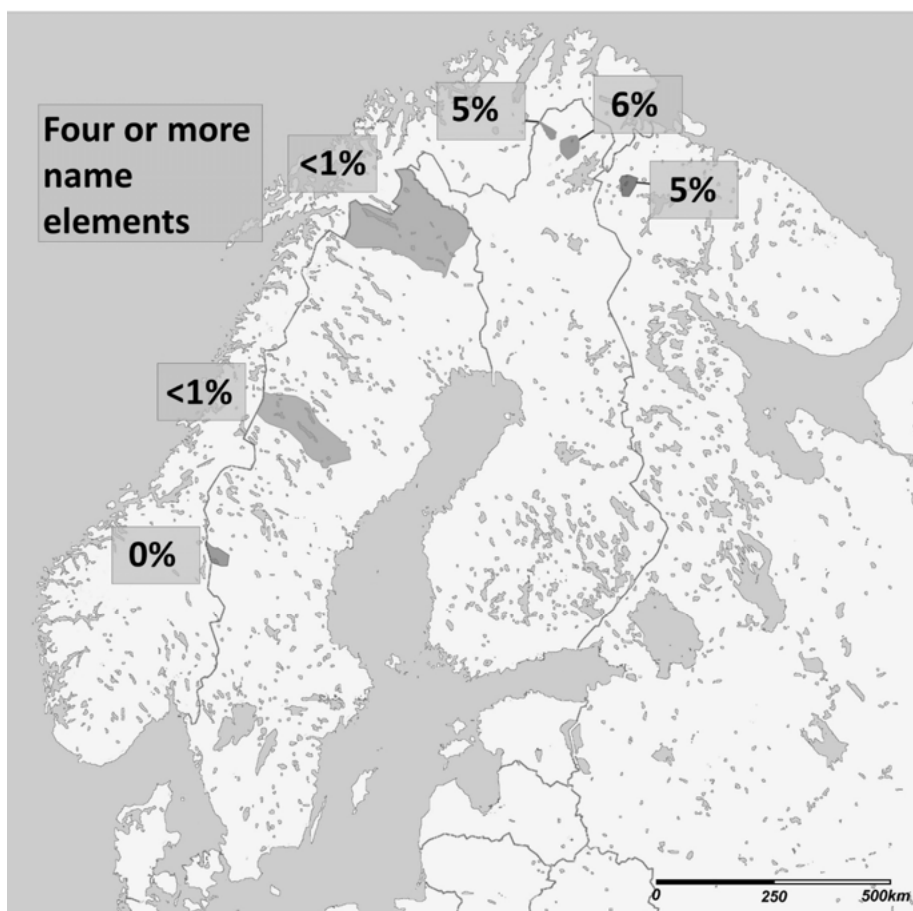
Place names with three name elements are a common feature in my place name samples of eastern inland North Saami, Aanaar Saami and Skolt Saami, since approximately a third of the place names have three name elements. Giron shows a clear difference with a result of only 15 percent. The South Saami materials show even lower numbers: 9 percent in Vualtjere and 8 percent in Ruvhten sijte. (See Map 9.)



Map 9. Place names with three name elements.

When it comes to the place names that have four or more name elements, the place names materials of Dálvadas, Čovčjävri–Kosseennâm and Sverloff family area show a clear contrast with the Giron, Vualtjere and Ruvhten sijte materials. In the more eastern languages, 5 or 6 percent of the names have four or more name elements, whereas in other languages, such names are very rare or nonexistent. (See Map 10.) The longest place name in my material is a six-stem Skolt Saami place name from Sverloff family area: *Hoattjääuraluutvuäjnjää'lm|jeä'gğ*, 'Useless Lake's Bay's Brook Mouth's Bog'. It is also noteworthy that the Finnish place names of eastern Finnish Lapland that have a clear Kemi Saami substrate are typically longer than Finnish place names in other parts of the country. This gives a hint that the Kemi Saami areas have belonged to the northern and eastern area where long names have been a typical phenomenon. On the other hand, it seems that based

on the before-mentioned sources, the Kildin Saami place names are not as often as long as place names of the more western Eastern Saami languages.



Map 10. Place names with four or more name elements.

5. Discussion

To compare the studied place name materials, I have calculated a relative equality value for each region and for each linguistic phenomenon. This means that the area with the highest percentage of occurrence of each studied linguistic phenomenon is taken as a point of comparison and gets a value of 100. Other areas get a relative equality value between 0 and 99 according to the number of occurrences. The higher the value, the more that phenomenon occurs in that material. I have further divided these relative equality values into six categories: 0; 1–19; 20–39; 40–59; 60–79; and 80–100.

Table 1. The number on the upper line next to each studied linguistic phenomenon measures occurrence as a percentage of the studied place name material, and the number on the lower line is the relative equality value. The color of the cells describes the categorization of the relative equality values, the darkest being the highest and the lightest the lowest (see also Map 11).

Phenomenon / Area	Ruvhte n sįjte	Vualtjer e	Giron	Dálvadas	Čovčjávri –Koss- eennâm	Sverlof f family area
Personal name as specific part	0% 0	<1% <14	2% 29	7% 100	6% 86	3% 43
Action nominal as a specific part	0% 0	<1% <20	2% 40	5% 100	5% 100	2% 40
Attributive adjective as a specific part	12% 46	26% 100	18% 69	8% 31	10% 38	12% 46
Specifying modifier as a specific part	1% 7	14% 100	9% 64	8% 57	8% 57	6% 41
Shortening of unstressed middle name elements (reduction)	1% 8	1% 8	1% 8	12% 100	3% 25	<1% <8
One name element	20% 100	16% 80	4% 20	4% 20	2% 10	2% 10
Two name elements	57% 70	74% 91	81% 100	62% 77	59% 73	58% 72
Three name elements	8% 23	9% 26	15% 43	28% 80	32% 91	35% 100
Four or more name elements	0% 0	<1% <17	<1% <17	5% 83	6% 100	5% 83

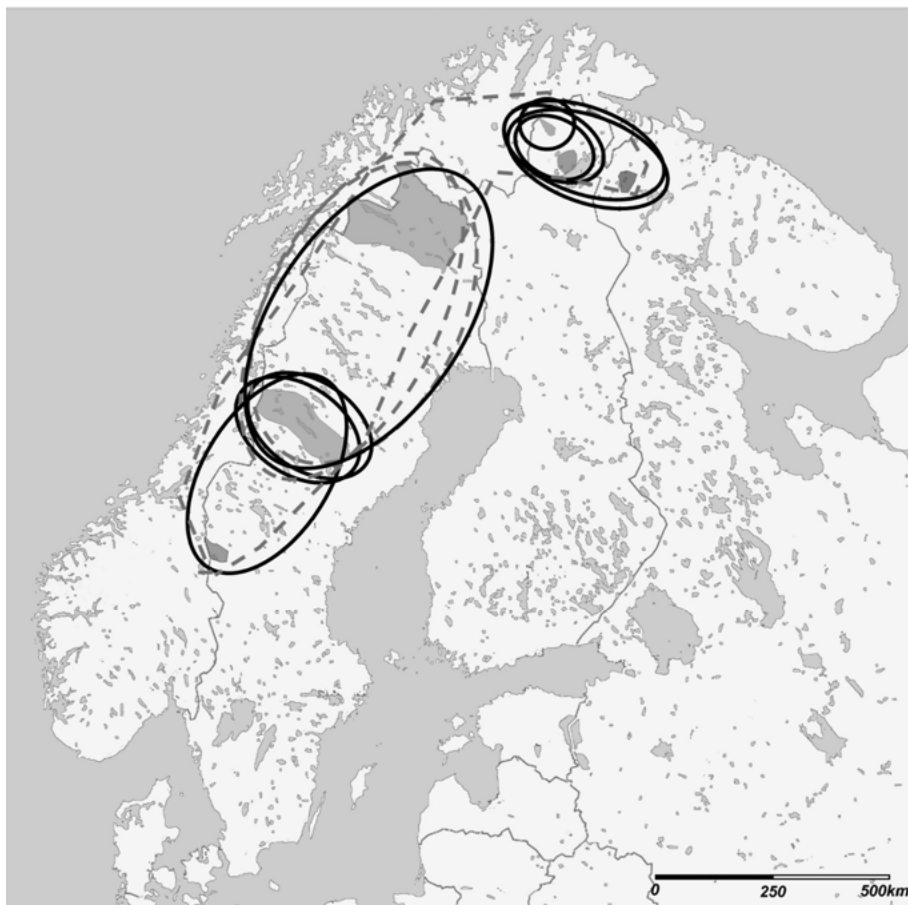
I have summarized the percentage and relative equality values of each material, and each studied linguistic phenomenon in Table 1 above. In the table, the six categories are distinguished by different shades of gray, the darkest being the highest and the lightest the lowest.

The highest (and sometimes the lowest) relative equality values can be, in my opinion, interpreted to represent an innovation center (or a borrowing area) of a linguistic phenomenon. These innovation centers (areas with relative equality values between 80 and 100) are marked on Map 11 with boundaries drawn with solid black lines. The innovation areas together with the areas where the innovation seems to be diffusing (areas with relative equality values between 60 and 100) are marked with a dashed gray line. Map 11 with its marked areas resembles an isogloss map, but the lines must not be interpreted as typical isoglosses that represent the boundary between existence and nonexistence of a certain linguistic phenomenon. However, large differences in values of neighboring materials can be interpreted as indicating (at least gradual) differences in the prototypical place name schemas represented as differences in place name grammars. At the same time, it is worth remembering that these results are partly based on quite small numbers of place names and unsystematically selected structural features, and because of this, the statistical results are not representative but rather indicative.¹² However, the main objective of this article, to show how place names can be used to represent areal variation of Saami languages, is fulfilled.

Despite the poor statistical representativeness, I want to reflect on some of the results. In Map 11, the only clearly common structural feature is that place names with two name elements are common in all studied Saami languages. The place names of the South Saami area, especially in Vualtjere, seem to have several structural features that are typical for them. Attributive adjectives and specifying modifiers as specific parts are clearly more typical in Vualtjere's place name system than in other studied materials. Place names with only one name element are most typical in Ruvhten sįjte's place name material, but very common also in Vualtjere. Vualtjere's place name system shares some typical structural features with Giron's material, but these are not shared with Ruvhten sįjte's material. As Ruvhten sįjte's material is very small in number compared to other studied materials, it is plausible to think that a more representative material would produce a more similar result than in Vualtjere, but there could certainly be some differences as well. It would be interesting to compare the South Saami materials with Ume Saami place name material, since based on language historical morphophonological evidence,

¹² Compare, for example, *Atlas Linguarum Europae*'s material, which aimed to capture the list of 549 stems in 34 different Saami dialects (see Rydving 2013, 93–107).

these languages have been thought to belong to the same subgroup of Saami languages (see, e.g., Korhonen 1981, 19; Sammallahti 1998, 22)— earlier, even being dialects of the same language (see, e.g., Rydving 2013, 36, 38–41).



Map 11. The innovation centers (areas with relative equality values between 80 and 100; see Table 1) are marked with solid black lines, and the innovation areas and the areas where the innovation seems to be diffusing or is otherwise more typical (areas with relative equality values between 60 and 100; see Table 1.) are marked with a dashed gray line.

The most distinctive difference on Map 11 is shown as a gap between the place name materials of the studied North Saami dialects: Giron's place name system seems to be more closely related to the South Saami systems and Dálvadas's place name system with the systems of the studied Eastern Saami languages. Giron's place name system shares with both South Saami place name systems two structural features: the low number of personal names as

specific parts and the rarity of place names with more than three name elements. As both types are very common in Dálvadas's material, the difference is striking. Giron's material also shares with Vualtjere's material the high number of attributive adjectives and specifying modifiers as specific parts, even though the number is clearly higher in Vualtjere. The difference with Dálvadas's material in the number of attributive adjectives as specific part is high. The number of place names with two name elements is highest in Giron, but they are very common in Vualtjere as well. This feature separates these two materials from others, although the difference is not radical. The reduction (shortening of unstressed middle name elements) is rare in Giron, even though it is very common in Dálvadas.

The division between the Torne Saami dialect and the inland dialects of North Saami has not been highlighted in language historical studies, but as Rydving (2013, 187) notes, Torne Saami is understudied and poorly documented. It is also worth noting that my materials do not include any place name materials of the western inland dialect, and because of this, it is impossible to know if these assumed differences are between Torne Saami and the western inland dialect or between western and eastern inland dialects. There are, however, at least three reasons that can explain the differences between Giron's and Dálvadas's materials. First, it might relate to the long common history of Saami and Meänkieli in Giron, which I have discussed above. Second, Giron's place name system may share old common features with Lule Saami name systems, as well as some substrate influences. The existence of a Lule Saami substrate in Giron can be attested on the lexical level in some loaned (opaque) appellative name elements (e.g., *moarhmmá* < SaLu *moarmes* : *moarmmá*-, '(reindeer) fetus; May') and terrain appellatives (e.g., SaN southern Torne dialect *čavil*, 'mountain ridge' < SaLu *tjávelk*). These are more common in the municipality's southwestern border area. The second hypothesis could be tested by analyzing other Torne Saami place name materials, especially material from the Suomanjárga/Käsivarsi region in Finland that has no connections with the Lule Saami language. It would be important to analyze also place name materials of the Lule Saami language.

The third reason that might explain the differences between Giron's and Dálvadas's materials is that Dálvadas's material seems to share several features with more eastern place name materials. Personal names and action nominals as specific parts are most typical in Dálvadas's and Čovčjävri–Kosseennâm's place name systems. Place names with three and four or more name elements are most typical in Dálvadas's, Čovčjävri–Kosseennâm's, and Sverloff family area's place name systems. On the other hand, only

Dálvadas's place name material includes lots of reduction, but it is also known to be typical in the western inland dialect of North Saami.

It is interesting to compare my findings with the results that Håkan Rydving (2013) got when he studied the lexicon of Saami languages based on 34 samples collected during the project *Atlas Linguarum Europae* (ALE). His results show that the South Saami dialect of Gåebrie has the highest lexical correspondence to the other South Saami materials (Rydving 2013, 140, 175–176). This result is consistent with my findings that the South Saami forms an area of its own that also shares common place name systems. The close connection between Aanaar Saami and North Saami of Ohcejohka, as well as the connection between Aanaar Saami and the Skolt Saami dialect of Njeä'llem/Nellim¹³ in Aanaar/Inari, is also recognized by Rydving (2013, 161–166, 177).

However, in Rydving's results the lexical correspondence of the Torne Saami dialect of Gárasavvon is as high to the materials representing the eastern and the western inland dialects of North Saami. It has the highest correspondence, as can be expected, to the Torne Saami dialect of Čohkkiras. Even the correspondence to the Lule Saami material of Jåhkåmåhkke is almost as high as to the North Saami inland dialects. (Rydving 2013, 151–154, 176.) The area with the highest lexical correspondences is almost the same as the area where the North Saami of Guovdageaidnu and Gárasavvon were forcibly relocated during the end of the 19th and the beginning of the 20th centuries (Rydving 2013, 187–188). It is probable that this fast and radical historical process is reflected in the lexical correspondences that Rydving studied, but it might not have had similar effect on place names due to their more conservative nature.

Rydving also writes about the betweenness of the Aanaar Saami language in relation to the Eastern-Western Saami division. He suggests that it could be seen as an independent transition language that cannot be classified in either category. This idea is supported by the complex morphophonology in nominal and verbal inflection and derivation that seems to have gained its complexity mainly during the separate development of Aanaar Saami (see, e.g., Valtonen et al. 2022). If one must choose, then Rydving would place Aanaar Saami with the Western Saami, based on lexical evidence. This choice is supported by the historical data. The Aanaar Saami and Ohcejohka Saami traded their goods and fished regularly with the Sea Saami in the Várjavuonna/Varangerfjorden region by the Arctic Ocean, where they also had connections with the speakers of the northern Skolt Saami dialects, including the communities that now are

¹³ The Skolt Saami of Njeä'llem moved to the area after the second World War from their traditional territories situated by the Arctic Ocean.

living in Njeä'llem. For centuries, the Aanaar Saami have also moved permanently to Várjahuonna when there has been a shortage of resources in Aanaar. (e.g., Nahkiaisaja 2003, 174, 178–179, 183; Valtonen 2014, 164–165.) It is obvious that these connections have been very important in transmitting linguistic innovations but also in maintaining old connections.

However, there is also a third possibility: that the border between Western and Eastern Saami languages, nowadays most often situated between North and Aanaar Saami, is not as fundamental and old as present research maintains. This view has been presented previously by Mikko Korhonen (Korhonen 1981, 21; see also Bergsland 1995, 13). There are several good morphophonological arguments to suggest that one, at the present less well recognized, linguistic border has passed between western and eastern inland dialects of North Saami.¹⁴ It can be argued that the present border between North and Aanaar Saami is a result of eastward migration from West Finnmark caused by economical, institutional, and political changes that started in the 16th century (see, e.g., Aikio, S. 2013). Migrating populations spread new lexicons and morphophonological innovations, but the older linguistic structures remained as a substratum.

6. Conclusions

At the epilogue of his investigation *Words and Varieties*, Håkan Rydving wishes that his study “will at the same time initiate an intensified discussion about Saami linguistic geography, its methods, criteria, material and theoretical perspectives [...] and] contribute to the evaluation of different criteria, be they [...] language historical, political or communicative” (Rydving 2013, 188). I agree with Rydving’s wish, because more research and

¹⁴ There are several morphophonological differences that support the idea that the border between Eastern and Western Saami languages may have moved further east. East of North Saami western inland dialect, for example: 1. The stressed single vowels have become long vowels before short and relatively short consonants (e.g., North Saami eastern inland dialect: /mâānnah : mâānan/ and Aanaar Saami: moonnâđ : moonâm vs. North Saami western inland dialect: /mannah : manan/ (‘to go’ INF : SGLPRES); 2. Single intervocalic stops /k/ and /p/ have changed after stressed syllables to spirant v ~ g (~ Ø) (e.g., North Saami eastern inland dialect: /soahkii : soagii ~ soa.ii; suuhpii : suuvii/ and Aanaar Saami: suáhi : suávi; supe : suve vs. North Saami western inland dialect: /soahkii : soakii; suhpai : supii/ (‘birch’ and ‘aspen’ NOM : GEN-ACC); 3. Cases and other grammatical features are often (but not always) marked also by modifications to the root (also other than consonant gradation) and not just with suffixes (e.g., North Saami eastern inland dialect: /meāhcii : miehcii : meāhcäij/, Aanaar Saami: mecci : meeci : miäcán, and Skolt Saami: meä’cc : mie’cc : meäcca vs. North Saami western inland dialect: /meahccii : meahcii : meahccaaj/ (‘forest; area without permanent settlement’ NOM : GEN-ACC : ILL). In addition to these, there are the well-known features of the sea dialect that agree with the Eastern Saami languages. (See Korhonen 1981, 19–22; Sammallahhti 1998, 11–14.)

new openings are needed to question the present, partly unsatisfying conclusions. In this process, extensive systematic analysis of the structure types of place names of all Saami languages and major dialects can provide one additional independent perspective, as I have demonstrated in this article.

Acknowledgments

Writing of this article has been possible due to the funding granted by the Academy of Finland (proj. nr. 308882) and the Kone Foundation.

References

- Aalto, Tiina. 2019. *Ukon jäljet paikoissa ja paikannimissä: Tutkimus ukko-elementin sisältävistä nimistä*. Helsinki: Helsingin yliopisto. <<http://hdl.handle.net/10138/304967>>
- Aikio, Ante. 2007. The study of Saami substrate toponyms in Finland. In Ritva Liisa Pitkänen & Janne Saarikivi (eds), *Borrowing of Place Names in the Uralian Languages* (Onomastica Uralica 4), 159–197. Helsinki & Debrecen: Onomastica Uralica research program.
- Aikio, Samuli. 2013. Aanaar viestärist nuortás. *Anarâš juovlámáánu* 2013, 3–7.
- Aikio, Samuli. 2017. *Davvisámi báikenamat*. Kárášjohka: Čálliidláágáduš.
- Ainiala, Terhi. 1997. *Muuttuva paikannimistö*. Helsinki: Suomalaisen Kirjallisuuden Seura.
- Ainiala, Terhi, Minna Saarelma & Paula Sjöblom. 2016. *Names in Focus: An Introduction to Finnish Onomastics*. Helsinki: Finnish Literature Society.
- Bergsland, Knut. 1995. *Bidrag til sydsamenes historie*. Tromsø: Sámi dutkamiid guovddáš / Senter for samiske studier.
- Coates, Richard. 2016. Names and historical linguistics. In Carole Hough (ed.), *The Oxford Handbook of Names and Naming*, 525–539. Oxford: Oxford University Press.
- Dabrowska, Ewa & Dagmar Divjak. 2015. *Handbook of Cognitive Linguistics*. Berlin: De Gruyter Mouton.
- Helander, Kaisa Rautio. 1991. Sámegiel báikenamaid strukruvra ja dehálemos čoagginrávvagat. In Hans Ragnar Mathisen (ed.), *Sámi kulturmuittut. Báikenammačoaggima giehtagirji – Samiske kulturminner. Håndbok i stedsnavnregistrering*, 72–93. Romsa: Romssa Sámi Searvi.
- Itkonen, T. I. 1958. *Koltan- ja kuolanlapin sanakirja. Wörterbuch des Kolta- und Kolalappischen. I–II* (Lexica Societatis Fenno-Ugricae 15). Helsinki: Suomalais-Ugrilainen Seura.
- Kert, G. M. 2002. Saami toponymy of the Kola Peninsula as a research object In István Nyirkos (ed.), *History of the Study of Toponyms in the Uralian Languages* (Onomastica Uralica 2), 121–132. Debrecen/Helsinki: Onomastica Uralica research program.
- Kert, G. M. 2005. Saami village names in the Kola Peninsula. In Sándor Maticsák, (ed.), *Settlement Names in the Uralian Languages* (Onomastica Uralica 3), 35–44. Debrecen/Helsinki: Onomastica Uralica research program.
- Kiviniemi, Eero. 1986. De vanligaste fornarnamn i finska ortnamn. In Jørn Sandnes & Ola Stemshaug (eds), *Personnamn i stadnamn* (NORNA-rapporter 33), 199–206. Trondheim: Tapir forlag.
- Kiviniemi, Eero. 1990. *Perustietoa paikannimistä* (Suomi 149). Helsinki: Suomalaisen Kirjallisuuden Seura.

- Korhonen, Mikko. 1981. *Johdatus lapin kielen historiaan* (Suomalaisen Kirjallisuuden Seuran Toimituksia 370). Helsinki: Suomalaisen Kirjallisuuden Seura.
- Lakoff, George. 1987. *Women, Fire, and Dangerous Things*. Chicago: University of Chicago Press.
- Langacker, Ronald W. 1987. *Foundations of cognitive grammar, Volume 1: Theoretical prerequisites*. Stanford, CA: Stanford University Press.
- Nahkiaisola, Tarja. 2003. Uudisasuttajien aika (1750–1876). In Veli-Pekka Lehtola (ed.), *Inari – Aanaar: Inarin historia jääkaudesta nykyaikaan*, 164–215. Inari: Inarin kunta.
- Nesheim, Asbjörn. 1967. Eastern and western elements in Lapp culture. In *Lapps and Norsemen in Olden Times*, 104–168. Oslo: Universitetsforlaget.
- Nicolaisen, W.F.H. 1986. Personal Names as Place Names. In Jørn Sandnes & Ola Stemshaug (eds), *Personnamn i stadnamn* (NORNA-rapporter 33), 207–216. Trondheim: Tapir forlag.
- Norstedt, Gudrun. 2011. *Lappskaftelanden på Geddás karta: Umeå lappmark från 1671 till 1900-talets början*. Obbola: Thalassa Förlag.
- Nyström, Staffan (ed.). 2021. *Namn och namnforskning: Ett levande läromedel om ortnamn, personnamn och andra namn* (Version 2, 2021-02-25). <<http://uu.diva-portal.org/smash/get/diva2:1531251/FULLTEXT01.pdf>>
- Pamp, Bengt. (1988) 2007. *Ortnamnen i Sverige*. Lund: Studentlitteratur.
- Ranta-Rönnlund, Sara. 1971. *Nådevalpar: Berättelser om nomader och nybyggare i norr*. Stockholm: Askild & Kärnekull.
- Rydving, Håkan. 2013. *Words and varieties: lexical variation in Saami* (Suomalais-Ugrilaisen Seuran Toimituksia 269). Helsinki: Suomalais-Ugrilainen Seura.
- Rygh, Oluf. 1898. *Norske Gaardnavne: Oplysninger samlede til Brug ved Matrikelens Revision: Forord og Indledning*. Kristiania: W. C. Fabritius & Sønners Bogtrykkeri.
- Sammallahti, Pekka. 1998. *The Saami Languages: An Introduction*. Kárášjohka: Davvi Girji.
- Sammallahti, Pekka. 2019. *Láidehus sámegiela jietnadatoahpa dutkamii*. Oulu: Giellagas-instituut. <<http://urn.fi/urn:isbn:9789526222578>>
- Sankoff, Gillian. 2013. Linguistic Outcomes of Bilingualism. In J. K. Chambers & Natalie Schilling (eds), *The Handbook of Language Variation and Change* [online], 1–19. Chichester: John Wiley & Sons. <<https://doi.org/10.1002/9781118335598.ch23>>
- Schnitler, Peter. 1742–1745. = *Major Peter Schnitlers grenseeksaminasjonsprotokoller 1742–1745* 1-3 bd. (bd. 2 ed. by Just Qvigstad & Karl Bernhard Wiklund 1929, bd. 1 ed. by Kristian Nissen & Ingolf Kvamen 1962, bd. 3 ed. by L. I. Hansen & T. Schmidt 1985). Oslo: Riksarkivet.
- Senkevich-Gudkova, Viktorija V. 1963. Strukturno-slovoobrazovatel’nye elementy v toponimike Kol’skogo poluostrova. In A. R. Rot (ed.), *Tezisy dokladov i soobshchenija k Vsesojuznoi konferentsii po voprosam finno-ugorskogo jazykoznanija*, 26–28. Užgorod: Užgorodskii gosudarstvennyj universitet, Institut jazykoznanija akademii nauk SSSR, MV i USSR.
- Sjöblom, Paula. 2006. *Toiminimen toimenkuva: Suomalaisen yritysnimistön rakenne ja funktiot* (Suomalaisen Kirjallisuuden Seuran toimituksia 1064). Helsinki: Suomalaisen Kirjallisuuden Seura.
- Swedell, Ulla. 2008. Ortnamnens vittnesbörd om språkmöten i norra Norrland. In Ulla Börestam & Satu Gröndahl & Boglárka Straszer (eds), *Revitalisera mera! En artikelsamling om den språkliga mångfalden i Norden tillägnad Leena Huss*, 197–208. Uppsala: Uppsala universitet.

- Swedell, Ulla. 2009. Gällivares bebyggelsenamn ur minoritetsperspektiv. In Lars-Erik Edlund & Susanne Haugen (eds), *Namn i flerspråkiga och mångkulturella miljöer: Handlingar från NORNA:s 36:e symposium i Umeå 16–18 november 2006*, 27–36. Umeå.
- Tanner, Väinö. 1928. Petsamon aleen paikannimiä I: Lappalaisia paikannimiä. *Fennia* 49(2).
- Tresk, Olof. 1928. *Kartor över Kemi & Torne Lappmarker 1642 och 1643*. Stockholm.
- Tuggy, David. 2012. Schematicity. In Dirk Geeraerts & Hubert Cuyckens (eds), *The Oxford Handbook of Cognitive Linguistics*, 82–116. Oxford: Oxford University Press. <DOI: 10.1093/oxfordhb/9780199738632.013.0004>
- Valtonen, Taarna. 2014. *Mielen laaksot – Mielen vuemieh – Miela vuomit – Miela vyemeh – Miöl vue'm: neljän saamen kielen paikannimien rakenne, sanasto ja rinnakkaisnimet vähemmistö–enemmistö-suhteiden kuvastajina* (Suomalais-Ugrilainen Seuran toimituksia 271). Helsinki: Suomalais-Ugrilainen Seura. <<https://www.sgr.fi/fi/files/original/278c81d5c0fb20cdf608fb19b94b6af0.pdf>>
- Valtonen, Taarna. 2019. Báikenamat giella- ja kulturkontávttaid govvideaddjin – ovdamearkkat njealji sámegielas. *Dutkansearvvi diedalaš áigečála* 1/2019, 1–38.
- Valtonen, Taarna, Jussi Ylikoski & Luobbal Sámmol Sámmol Ánte (Ante Aikio). 2022. Aanaar (Inari) Saami. In Marianne Bakró-Nagy, Johanna Laakso & Elena Skribnik (eds), *The Oxford guide to the Uralic languages* 178–195. Oxford: Oxford University Press.
- Zilliacus, Kurt. 2002. *Forska i namn*. Helsingfors: Svenska Litteratursällskapet i Finland.

Bisumsáme tjállemvuoge birra - A review of the official Pite Saami orthography

Joshua Wilbur, University of Tartu

1. Background and context

Pite Saami is a critically endangered¹ Uralic language spoken today by a handful of individuals stemming from Árjepluovve municipality² in Swedish Lapland, and from adjacent areas in Norway and Sweden. Historically, the language community was never particularly large, but was robust and vibrant, at least until the local North Germanic-speaking settlers gained a majority.³ This ultimately led to language shift among many Pite Saami speakers to the respective North Germanic language variants. Nowadays, only a small portion of the ethnic Pite Saami population speaks Pite Saami as a first language (Valijärvi & Wilbur 2011, Wilbur 2014: 6–7), and all speakers are bilingual in Pite Saami and Swedish/Norwegian.

During the course of the second half of the 20th century, the number of speakers decreased slowly but steadily for a variety of reasons mainly

¹ The online version of The Atlas of the World's Languages in Danger (Moseley 2010: accessed on 2020-03-03), created and published by UNESCO, lists Pite Saami as “critically” endangered, which is defined as the state in which “the youngest speakers are grandparents and older, and they speak the language partially and infrequently.” However, according to my own familiarity with the Pite Saami language community, the UNESCO definition for the attribute “severely” endangered is in fact more fitting. According to this definition, a severely endangered “language is spoken by grandparents and older generations; while the parent generation may understand it, they do not speak it to children or among themselves”. In fact, the UNESCO characterization of the degree “definitely endangered” is also partially suitable, as this says “children no longer learn the language as mother tongue in the home”. While these UNESCO categories are clearly meant to generalize and be applicable for many languages from across the globe in exceptionally varied situations, the sheer number of speakers of Pite Saami (around 35 individuals) clearly makes the attribute “critically” quite appropriate, whatever the actual demographics of the various generations of speakers.

² In Swedish, Árjepluovve is known as Arjeplog.

³ Wallström (1943: 20–21) discusses evidence that the initial generations of Swedish settlers in the area spoke some Pite Saami (around the turn of the 18th century).

connected to the detrimental influence of the majority culture.⁴ This adversely affected the situation that the Pite Saami language was in, a fate shared by other smaller Saami languages and indeed many other minority languages around the globe. However, the first two decades of the 21st century have witnessed a slow reversal in this trend, with an increase in awareness, respect and urgency concerning the language, as well as a change in attitude towards multilingualism.

One result of this change was the creation and official approval of a standard orthography for Pite Saami in August 2019, which is the focus of the current paper. In this, I take a critical look at this new orthographic standard by attempting to analyze it concerning its linguistic integrity and soundness as well as the potential it holds to affect the vitality of the Pite Saami language in the future. In doing so, I have tried to keep in mind as a guiding principle the following passage from an insightful chapter on orthography development in the context of language revitalization in Grenoble & Whaley (2005: 148):

“Although linguists, in their desire to develop orthographies that accurately capture the phonemic system of a language, may balk at involving speaker intuitions which may obscure the system, an orthography needs to be designed with potential readers in mind, and so needs to be suited to their needs and perceptions in ways that a linguistic account need not.”

This discussion is structured as follows. To begin with, Section 2 introduces the new orthography and the document that presents it, including a review of the document itself. Then, in Section 3, I provide a history of how Pite Saami has been written in the past, as well as how the current standard came about. Section 4 is a critical review of how successful the orthography is, i.e., how well it fulfills some ideals for what a standard orthography should be like (e.g., as outlined in chapter 6 of Grenoble & Whaley 2005). While this discussion highlights theoretical aspects, such as how well graphemes correspond to phonemic realities of the language, it also considers practical aspects such as how likely the orthography will be adopted by actual non-linguist users of the language. Finally, in Section 5, I look at what the future of Pite Saami may hold considering a written standard now exists, and what potential this orthography has in supporting the language community, especially in revitalization efforts.

⁴ It is beyond the scope of this paper to go into the history and effects of colonization on the Pite Saami language.

2. Bisumsáme tjállemvuohke - an overview

On the 20th of August 2019, the first officially recognized Pite Saami orthography was signed into life during a ceremony at Silvermuseet in Árjepluovve, Sweden. The document is entitled *Pitesamisk ortografi: Beslut om pitesamisk ortografi*,⁵ and at the time of writing this article and to the best of my knowledge, this document is only available via a link in an article in the North Saami language version of the Giellagáldu website; the accessibility of the document will be discussed briefly below. As dictated by the Saami parliamentary council (SPR) on the 15th of May 2018, this orthographic standard was developed by a working group organized and funded by *Sámi Giellagáldu*, a collaborative working group of the Saami parliaments in Norway, Sweden and Finland focussing on standardizing writing systems and new terminology for the Saami languages, as well as supporting revitalization efforts and language support for speakers.⁶ In its essence, the document contains guidelines specifying sounds in the Pite Saami language, and how to use a version of the Latin alphabet to represent these sounds in writing. The meta-language of the document is Swedish, but it contains many Pite Saami examples, which are of course written in the new orthography. The working group consisted of four members from both Sweden and Norway: two linguists specializing in Saami languages, and two fluent Pite Saami speakers (one of whose native languages is Pite Saami), in addition to a secretary. The group met four different times, and developed the document collaboratively during those meetings as well as off-site between meetings (Peter Steggo, personal contact).

The document is organized into the following parts:

1. *Uppdraget* (p. 1): a brief presentation of the working group's purpose and work
2. *Ordens struktur* (p. 1): a short description of prosodic structures relevant to words and morphology
3. *Pitesamiskans alfabet* (pp. 1–31): the main part of the document; here, the spelling rules are presented
4. *Dialektvariation* (pp. 31–33): a section on spoken dialectal variation and to what extent variants are accommodated by the orthography
5. *Särskilda skrivregler* (p. 34): a comment on incorporating loan words

⁵ “Pite Saami orthography: resolution on a Pite Saami orthography” (my translation).

⁶ Cf. the group's website <http://www.giella.org/about> (accessed on 2020-04-28).

6. *Slutkommentar* (p. 34): a statement from the working group about their stance concerning the document
7. *Textexempel* (p. 34): an example text (originally written and published by Lars Rensund in 1986) written according to the new orthographic standard

The rest of the current section provides details about the main part of the document (no. 3 above), as well as briefly about the final comment (no. 6 above). For more on dialectal variation (no. 4 above), see Section 4.4.1.

The main section of the document begins with the letters used for writing the Pite Saami language. Specifically, Pite Saami words are spelled using the graphemes⁷ presented in Figure 1. The document does not specify alphabetic order for Pite Saami, but it seems reasonable to assume that the order presented in the document (and repeated in Figure 1) can be considered alphabetic as this nearly completely aligns with the Swedish/Norwegian alphabets⁸ as well as with Lule Saami, which was frequently used as a point of comparison during orthography development. In this alphabet, non-standard Swedish/Norwegian characters are included immediately following the corresponding base letter, such that, e.g., <á> follows <a>, <đ> follows <d>, <ŋ> follows <n> and <t> follows <t>; the other two graphemes with diacritics, namely <â> and <ä>, are listed last, which is consistent with the Swedish/Norwegian alphabets as well. The document also notes that “in loan words, foreign words and names, other letters can also occur” (p. 1, my translation), and lists <c>, <q>, <x>, <y> and <z> as examples (surprisingly, <ö> is not listed, although this is potentially a common non-Pite Saami grapheme in loan words). Despite a lack of information concerning the status of upper-case letters, it seems safe to assume that the common corresponding upper-case letters are included in the alphabet as well, although it might have been useful to at least include the upper-case variants for the less common graphemes <á>, <đ>, <ŋ> and <t>.

a á b d đ e f g h i j k l m n ŋ o p r s t t u v â ä

Figure 1. The graphemes used in the Pite Saami orthography for native words.

⁷ The technical term “grapheme” refers to the smallest building blocks of the writing system, especially when it is important to distinguish these from the actual sounds the graphemes may represent. Graphemes are marked by the less-than and greater-than symbols, e.g., <a>, <á>, . The term is roughly equivalent to the non-technical term “letter”.

⁸ Note that the order of the graphemes <â> and <ä> follow the Swedish standard, not the Norwegian standard.

Following this, the document presents the *språkljud* (“speech sounds”). It begins by describing the main ambiguities present in the written standard, specifically the fact that certain length distinctions are not marked in the orthography (cf. Sections 4.1 and 4.3). Note that recommendations are provided for how dictionaries and teaching materials can mark these distinctions, even if the standard orthography does not.

The document then presents a detailed inventory of Pite Saami vowels and consonants, and how these are represented in the written standard. The presentation of the vowels begins with tables dividing the vowels into groups based on place of articulation for monophthongs. In these tables, each vowel grapheme is presented with short descriptions of its pronunciation and distribution in the three relevant prosodic slots of the first, second and third syllable positions; examples of words are also provided. The document only uses the letters from the Pite Saami alphabet, and does not present the sounds using a phonetic alphabet such as UPA or IPA; length distinctions are marked when necessary by double vowels in supplementary “pronunciation spellings” provided in parenthesis (e.g., *deedij* ‘know.3SG.PST’, which is officially spelled *dedij*). A table describing “orthographic diphthongs” is also provided (p. 5). This is followed by a table presenting paradigmatic vowel alternations in the first syllable; these are dependent on both the quantity of the consonant center (grade I, II or III) and the second syllable vowel. Finally, a number of orthographic rules pertaining to vowel alternations are presented; for each rule (except for rule no. 5), there are examples presented in the correct orthographic form and the pronunciation, along with a Swedish translation. These rules deal mainly with the paradigmatic vowel alternations outlined in the preceding table, and with the consequences of not marking certain vowel lengths.

The next part of the document deals with consonants. The presentation of the consonants begins with tables dividing the consonants into groups based on manner of articulation. In these tables, each consonant grapheme (including grapheme combinations representing a single sound such as <sj>) is presented with short descriptions of its pronunciation and distribution in relevant prosodic slots such as word-initial position, the consonant center, etc.; examples are also provided. These are followed by a brief discussion of word-initial consonant graphemes, including those in loanwords, and a table of examples. Next, morphophonemic consonant alternations (i.e., consonant gradation) in the consonant center (referred to as the *stamkonsonant* “stem consonant”) are presented as nine alternation patterns (or “series”) which are divided into four main “groups”. After an overview of what consonant gradation entails and a brief introduction of some technical terms (e.g., “grade

III”, “strong grade”, “weak grade”, etc.), the gradation groups and patterns are summarized. The document then provides details in the form of two tables for each pattern. For each pattern, the first table presents an inventory of the actual known instantiations of consonant alternation sets (e.g., bbd:bd, bbj:bj, bbl:bl as three instantiations of pattern 4) and necessary comments, while the second table provides examples for each instantiation (e.g., *rabbda:rabda* ‘edge’, *libbja:libja* ‘decoration’, *gäbbne:giebne* ‘pot’) with a Swedish translation. Finally, brief discussions of consonant graphemes bridging the second and third syllables as well as consonant graphemes in word-final position are discussed in the last two subsections. Each of these sections also includes a table of examples.

After a few pages presenting variation in spoken Pite Saami and how this is dealt with in spelling (cf. Section 4.4.1), and a very short section dealing with how to adapt loanwords to the spelling rules, the document presents a *slutkommentar* “final statement” before the text example is provided on page 34. This statement is brief, yet worth detailing here because it seems to provide insight into the process of developing this orthography. The first paragraph points out that the members of the working group are in agreement about the “principles” of the spelling system. The second paragraph highlights the fact that the group considered a variety of factors concerning the language, linguistics, pedagogy and language politics in their work, while also emphasizing that not all variation can be accounted for, and ultimately that the document will provide a foundation on which to further develop the Pite Saami orthography. The third paragraph diplomatically and anonymously points out that not all members agree with all details in the descriptions of the language’s sounds.

Having outlined the document above, a brief critique of the document as such is in order as well (a contextualized critique of the actual orthographic system is provided from Section 4 onward). Overall, the document is well structured and detailed enough to cover most situations in which Pite Saami will be written, while also providing clear guidelines for teachers and even students. The language used in the descriptions is sufficiently specific and technical to present the spelling system effectively while being accessible at least to users who are somewhat familiar with Swedish linguistics terminology. This is wise because entirely avoiding technical jargon in such a document would have made it overly complicated without significantly improving its readability for the target audience. In addition, many examples are provided to support the technical descriptions; this further increases the chances that users without any linguistics background will still be able to understand the document, although it will certainly require some extra effort

on their part. It would have been useful to actually indicate what alphabetical order is, and future editions of the orthographic system could benefit from descriptions of how to punctuate and capitalize written Pite Saami. Finally, the “final statement” seems to be a kind of caveat, indicating that the document is not the only possible way to do things, and that disputes may exist. While I do not know for certain about the motivation behind this statement, its inclusion makes the document – despite its consistent, determined and rigid pedagogical approach to standardizing the language in written form – seem refreshingly human, and ultimately accepting that it is prone to flaws. Although it could be interpreted as a sign of conflict within the working group, I prefer the interpretation that its authors are also quite human, that language is a flexible phenomenon impossible to represent adequately to written form, and that even an authoritative document like this is subject to revision and improvement. Especially in the context of such a small and highly endangered language, this is an important fact.

Unfortunately, a final note is necessary, although hopefully this situation will be quickly remedied. Specifically, it seems unnecessarily difficult to actually obtain a digital copy of the orthography document since (to the best of my knowledge) the document is solely available via a link “hidden” in an article in the North Saami language version of the Giellagáldu website.⁹ The link at the beginning of the article leads to a scanned version of the signed orthography document. The document itself has an exceptionally cumbersome url.¹⁰ Furthermore, as far as I know, no DOI¹¹ exists which provides a permanent link to this document, and the continued availability of the links provided here cannot be guaranteed. It is unfortunate that this document, which is an important milestone and needed reference for the Pite Saami language community, seems to be so difficult to find, and not readily discoverable via a Swedish or Norwegian interface.¹² It even seems to be out of Google’s reach; at least my attempts to do a google search for “pitesamisk ortografi” on both google.no and google.se did not lead to any hits leading to the document or even to anything on www.giella.org.¹³

⁹ This article about the new Pite Saami orthography can be found at:
<http://www.giella.org/articles/6YKrmI8IbgWVUc6KS7dyKp>

¹⁰ This is provided here as a matter of principle and in the name of transparency:
https://assets.ctfassets.net/kkxfrpcp7iyu/01JD6UHiBbZ5J26a1ufEsf/051217eecfe9fe7cf374df0bb33a37b5/Pitesamisk_ortografi_23.08.19.pdf

¹¹ DOI is a “digital object identifier” which is intended to uniquely identify a specific digital object in the long term.

¹² Based on my own estimation, very few potential users of the new orthography read North Saami, which is the language used on the only webpage that links to the actual document.

¹³ These searches were carried out on 28-04-2020.

3. The road to the current standard

The new standard Pite Saami orthography is hardly the first attempt at representing Pite Saami in written form.¹⁴ Over the last 130 years, a number of methods have been used, as is summarized in Table 1. In general, linguists collecting Pite Saami texts during the first half of the previous century and earlier each used their own variant of the Uralic Phonetic Alphabet (UPA); this included Hungarian linguist Ignász Halász, Finnish linguist Eliel Lagercrantz, Pite Saami linguist Israel Ruong, Swedish linguist Bo Wickman, and Finnish linguist Juhani Lehtiranta.

Israel Ruong also used an adapted version of the Swedish alphabet in a number of the texts he wrote down and which are archived at *Institutet för språk och folkminnen*¹⁵ (ISOF) in Uppsala. Carl Johansson collected a bit more than 60 pages of text in Pite Saami, and these can also be found in the ISOF archive, and are written in a modified version of the Swedish alphabet. J. K. Qvigstad used his own adaptation of a North Saami orthography (for instance *dab'ča* ‘whetstone’ and *nad̥da* ‘handle’). The texts in Lehtiranta’s publication from 1992 are transcribed in a less technical alphabet which could possibly be best described as a modified Swedish-Finnish alphabet (with spellings such as *åällulij̥t* ‘jaw.ACC.PL’ and *haavka* ‘pike.GEN.SG’), in addition to his UPA transcription.

Lars Rensund is certainly the most prolific Pite Saami author, and published a number of shorter texts (both prose and poetry) throughout his life in small booklets and in Saami periodicals. In Rensund’s representation of Pite Saami, only standard Swedish letters and orthographic principles are used.

As professor for Saami studies at Umeå University, Olavi Korhonen developed a modified version of the standard Lule Saami alphabet to write Pite Saami; it was used in teaching at the university, and is – to my knowledge – unpublished.¹⁶ The first digital collection of written Pite Saami was likely the blog maintained during 2007 and 2008 by Peter Steggo.¹⁷ The blog contains a number of heritage Pite Saami texts in a transcription based essentially on the spelling system developed by Korhonen (Peter Steggo, p.c.).

¹⁴ The following outline of the history leading to the new orthography is based on my own familiarity not only with the literature but – more importantly – with the main individuals involved and events that transpired over the last decade. Bibliographic references are provided whenever needed or relevant, but when not otherwise indicated, the information stems from my own observations and familiarity with the situation.

¹⁵ The Institute for Language and Folklore.

¹⁶ For instance, Korhonen used his orthography to transcribe a Pite Saami text (originally published by Lars Rensund) in a booklet titled *Samiska textprov: med glossor och kommentarer rörande vokalism och konsonantism* from 1992, and revised in 1994. I am grateful to Peter Steggo for telling me about this, and showing it to me.

¹⁷ <http://arbbe.blogspot.com/> (accessed 2020-04-21).

Table 1. A summary of orthographies used to write Pite Saami previous to the current standard, in approximate chronological order.

author(s)	period	UPA	orthography <i>adaptation of</i>		
			Swedish	saN	saL
I. Halász	1890s	✓			
E. Lagercrantz	1920s–1960s	✓			
J.K. Qvigstad	1920s			✓	
I. Ruong	1930s–1970s	✓	✓		
B. Wickman	1943	✓			
C. Johansson	?–1940s		✓		
J. Lehtiranta	1992	✓	✓		
L. Rensund	1950s–1990s		✓		
O. Korhonen	1990s				✓
P. Steggo	2007–2008				✓
Wordlist project ¹⁸	2008–2012				✓
J. Wilbur	2008–...				✓
A.-C. Sjaggo	2015				✓
O. Utne	2018–...				✓

To my knowledge, the current orthographic standard had its genesis in the project *Insamling av pitesamiska ord*,¹⁹ which was initiated in 2008 by members of the Arjeplog Saami association with funding from the European Union’s regional development fund. A number of individuals were involved at various periods throughout the project, and they all deserve to be mentioned here for the role they played in creating what ultimately formed the foundation of the official orthographic standard which exists today. They are (in alphabetical order): Nils-Henrik Bengtsson, Marianne Eriksson, Inger Fjällås, Eva-Karin Rosenberg, Gry Helen Sivertsen, Valborg Sjaggo, Dagny Skaile and Peter Steggo. The project resulted in a lexical database (as a Microsoft Excel table) of Pite Saami words with Swedish and Norwegian translations, as well as some grammatical information concerning morphological alternations and part of speech; it contains approximately 6800 entries. Based on my own knowledge of the project, which I consulted with regularly (as detailed below), the group used Lule Saami spelling conventions as a foundation, but altered these as needed, both to more accurately represent Pite

¹⁸ See below for more about the wordlist project.

¹⁹ “Collection of Pite Saami words”, hereafter referred to simply as “the wordlist project”.

Saami sounds and linguistic structures²⁰ as well as to clearly differentiate Pite Saami from Lule Saami.²¹ As shown in Table 1, all published Pite Saami texts since this project began have used some variant of their orthographic standard, up until and including the officially recognized standard from 2019. This includes publications by Ann-Charlotte Sjaggo, and the lexical app still under development by Olve Utne, as discussed below.

By some serendipitous stroke of fate,²² my own work with the Pite Saami language coincided with the work being done by the wordlist project. As a beginning PhD student with a scholarship to document the language,²³ my first visit to Árjepluovve commenced in June 2008, and I was pleased to find that this wordlist project not only was taking place (an indication of local interest in the language), but that they even had regular working meetings in the basement of the place I was staying (Samegårdén). I mention this not as an egocentric record of my own activities, but because this led to occasional informal symbiotic collaboration, with me learning about aspects of Pite Saami lexicography, and my own qualifications in general linguistics at least somewhat helping the group in their work with keeping a lexical database and abstracting phonological and grammatical structures from their own native speaker intuitions. In addition, my exposure to the group's work gave me a preliminary written standard to use in my own linguistic documentation work on the language.

This symbiosis led to a workshop specifically aimed at providing the Pite Saami language community with input from external experts in other Saami languages and orthography development, as well as a stage to discuss challenging aspects and the best way forward. The two-day workshop took place in May 2011 in Árjepluovve, and was made possible by the Saami Documentation and Revitalization Network (SaamiDocNet) with funding from NordForsk. The most tangible outcome of the workshop was the creation of a preliminary wordlist using the working orthography in summer 2011. That version contained a random selection of approximately one third of the lexical entries from the wordlist at that point, and was distributed to members of the language community, who were in turn asked to provide their own feedback and suggestions on spelling in general or even for specific entries.

²⁰ For instance, Pite Saami *luakktal/luokta* 'bay' in nominative singular and genitive singular, vs. Lule Saami *luokta/luovta*.

²¹ For instance: consistently doubling the initial consonant grapheme in a word-internal consonant cluster, e.g., Pite Saami *dåbbdåt* vs. Lule Saami *dåbddåt* 'know'.

²² ...or divine intervention?

²³ Five years of funding for my work on documenting Pite Saami were graciously provided by the Hans Rausing Endangered Language Project. While this funding was not directly intended to support orthography development, it provided me with the opportunity to work with the Pite Saami language community, and indirectly support them in precisely this undertaking.

However, to my knowledge, no significant feedback ever reached the wordlist project group. The project continued its work until funding ran out in 2012, at which point there seemed to be a general implicit consensus on how to spell Pite Saami words, although this was not specified in a written way (such as in the form of rules or guidelines), but could only be derived by studying spellings in the wordlist.

However, this was sufficient to allow this working version of a Pite Saami orthography to be used not only in my own documentary and linguistic work, but also by others interested in studying and writing the language. For instance, Ann-Charlotte Sjaggo published *Pitesamisk grammatik: en jämförande studie med lulesamiska* in 2015, a work which basically follows the orthographic rules at that stage of development.

In the meantime, interest in actively promoting²⁴ the Pite Saami language was growing on the Norwegian side, particularly at *Duoddara Ráfe pitesamisk senter* in Beiern, mainly due to Stig Morten Kristensen and Lennart Ranberg. Their political initiative ultimately helped move the Norwegian Saami parliament to decide to provide *Duoddara Ráfe* with funding to turn the wordlist into a full-fledged paper publication in book form. It then became my responsibility to turn this project into reality, and over the course of four months in 2016, I was able to dedicate myself full-time to this undertaking. Four months is an incredibly short time period to complete a lexicography project,²⁵ but it seemed more important to produce something visible and tangible, even if the result was not completely consistent or matured, rather than continuing work on Pite Saami on a seemingly endless lexical project, and thus surely postponing any potential usefulness of such work. Thanks to this funding and some additional funding from Arjeplog municipality to help cover printing costs, *Pitesamisk ordbok samt stavningsregler*²⁶ (Wilbur 2016b) was published in September 2016. The book was premiered during Saami Week in Árjepluovve at the beginning of October that year. It consists, as the title clearly indicates, of two parts: a dictionary with Pite Saami entries and Swedish and English translations, and the set of orthographic rules implemented in the first part. The dictionary part (Bengtsson et al. 2016) contains of an extensive selection of the entries from the wordlist which I had edited in a partly successful attempt to apply the spelling rules consistently, and to include relevant grammatical and lexical information for each entry. The orthographic rules (Wilbur 2016a) are presented as correspondences between graphemes and sounds, but do not cover morphological stem

²⁴ Perhaps ‘fighting for’ would be a more accurate description.

²⁵ Anyone even slightly familiar with lexicographic projects understands that these are typically projects that span a lifetime.

²⁶ “Pite Saami dictionary and spelling rules”.

alternations in nearly as much detail as the official document with the new orthography standard which was approved in 2019.

Although the set of rules in the Pite Saami dictionary were not sanctioned by any official political entity, they served as a reference point for further work promoting the language. One specific instance of this concerns language technology. In collaboration with Giellatekno (the research group for Saami language technology at the University of Tromsø), this initial orthography was used in developing language technology tools,²⁷ the lexical database *Nähttagibáhko* with inflected forms,²⁸ and even the initial spell checker (still under development).²⁹ In addition, several other online resources exist which implement these rules, including an extended version of the Pite Saami dictionary (Wilbur 2016b),³⁰ and of those initial spelling rules.³¹ Finally, a further lexical resource is being developed by Olve Utne; this is a mobile app with Pite Saami lexical items, example sentences and translations into Norwegian, Swedish, English and German; a beta version is available via Facebook.³²

With the continuing implementation of this preliminary Pite Saami orthography as a background, as well as a revival in interest in Pite Saami, in 2018 *Sámi Giellagáldu* (on behalf of the Saami parliamentary council) decided to implement a committee to draw up an official Pite Saami orthography. The committee consisted of Inger Fjällås, Ole Henrik Magga, Bruce Morén-Duolljá and Peter Steggo, as well as Ann-Charlotte Sjaggo as secretary. The committee met face-to-face four times in 2018 and 2019, and together composed the document that is now the basis for the official Pite Saami orthography. The orthography became officially recognized at a signing ceremony at Silvermuseet in Arjeppluovve on 20th August 2019. As a result, the Pite Saami language has achieved official status in relevant political realms in both Sweden and Norway.³³

²⁷ Specifically, these tools are based on automatic wordform generators and analyzers as well as syntactic parsers; see giellatekno.uit.no/cgi/index.sje.eng.html (accessed 28-04-2020).

²⁸ See bahkogirje.oahpa.no/ (accessed 28-04-2020).

²⁹ A beta version can be downloaded from divvun.org/proofing/proofing.html (accessed 28-04-2020).

³⁰ This is added to, corrected and updated regularly, and is based mainly on the wordlist from 2012 (Bengtsson et al. 2016); see saami.uni-freiburg.de/psdp/pite-lex/ (accessed 28-04-2020).

³¹ See <https://sjelex.keeleressursid.ee/stavningsregler/> (accessed 28-04-2020).

³² See the Facebook group *BidumBágo*.

³³ See Section 4.5 for more on revitalization efforts.

4. An assessment of the standard orthography

It seems reasonable that most speakers of the world's majority languages take for granted the existence of a writing system. However, speakers of small, minority languages are likely more appreciative of and have a better understanding for what it means to develop an orthographic system, as its existence is hardly a given. For some of those minority language speakers, it could even be that the mere existence of a standard orthography is reason enough to be satisfied with that orthography, and its symbolic import outweighs its actual usefulness. For instance, the fact alone that Pite Saami now has an officially recognized orthography has suddenly promoted the language to a legitimate political entity, which hopefully will have nothing but positive outcomes (cf. Section 4.5 on revitalization, as well as the outlook summary in Section 5). Nonetheless, it is worth exploring the new orthography with a critical eye in order to both highlight its strengths and discover potential challenges that could be addressed in future revisions.

In the following sections, I try to do just that. In addition to examining the rules applied in the orthography document³⁴ based on my own familiarity with Pite Saami linguistic structures (especially phonological, morphological and morphophonological ones), I also apply some general criteria for assessing orthographic systems for any language. To do this, I refer mainly to Chapter 6 in Grenoble & Whaley's *Saving languages* (2005), which presents a number of criteria, many of which are especially useful in the context of endangered languages. With that chapter as a point of reference, in the following sections, I explore the official document presenting the new Pite Saami orthography with respect to its consistency in dealing with symbol-sound correspondences (Section 4.1), how successfully it makes morphemes uniquely recognizable (Section 4.2), aspects of "diacritic overload" and the ease of entering it on a keyboard (Section 4.3), how well it deals with representing variation (Section 4.4), and, last but not least, how well it stands to support revitalization efforts (Section 4.5). My own background is in descriptive linguistics, and as a result, it would be most natural for me to focus on how well the orthography represents Pite Saami as a linguistic system; however, I have tried to go beyond this abstract linguistic framework and consider the orthography in its actual context, as the first standardized writing system for a critically endangered language for which the vast majority of native speakers are in the oldest generations; the orthography will thus be most useful if it is accessible to learners of the language. This approach is highlighted, and indeed inspired by Grenoble & Whaley, who point out that "orthographies for communities

³⁴ See Section 2 for more details about the document itself.

creating revitalization programs should be designed primarily with beginning readers in mind” (2005: 142; see also Section 4.5).

4.1 Symbol-sound correspondences

Ideally, the underlying phonemic structure of any given word is represented unambiguously in a language’s orthographic system. In other words, each grapheme should reliably correspond to a single sound (Grenoble & Whaley 2005: 140).³⁵ To a large extent, the new Pite Saami orthography adheres to this principle, as the set of examples in Table 2 illustrates. Here, the symbols (graphemes) on the left correspond uniquely to the sounds (phonemes) on the right, and vice versa. In other words, when readers see <á>, then they know this refers to a long -a- (the underlying sound /a:/, although its actual pronunciation may vary depending on its context).

Table 2. A selection of symbol-sound pairs for Pite Saami.

<u>grapheme symbol</u>	<u>sound (phoneme) in IPA</u>
<a>	/a/
<á>	/a:/
	/p/
<d>	/t/
<đ>	/ð/
<e>	/e/
<f>	/f/
<g>	/k/

Similarly, when writers want to represent the underlying sound /a:/, then they have to use the grapheme <á>. Keeping these relations consistent should make both reading and writing significantly easier to learn, which is an especially important factor for a critically endangered language like Pite Saami, which would only be further hindered by unreliable symbol-sound correspondences (such as those prevalent for instance in the standard English orthography).

However, the standardized orthography does not adhere to this principle for all letters and sounds, as there are a number of exceptions. In a few cases, one symbol can refer to multiple sounds, while in a few other cases, a single

³⁵ Aside from the fact that, as Grenoble & Whaley point out (2005: 140), an alphabetic orthography is the best choice of orthographic systems concerning sound-to-symbol correspondences, using a different system such as a semi-syllabic system (e.g., for Hindi) or a logographic system (e.g., Chinese) would have likely met with significant resistance by speakers and learners who are only used to alphabetic systems. This is sufficient reason to not consider further whether or not an alphabetic system is the best choice for Pite Saami, and this will therefore not be discussed in any more detail here.

sound has multiple means of representation, as discussed in the following two sections.

4.1.1 One symbol, multiple sounds

Although the orthography marks the length distinction for the vowels represented by <a> and <á>, this is not the case for the graphemes <â> and <u>, which do not differentiate length distinctions that do in fact exist for these vowel sounds: /ɔ/~ɔ:/ and /u/~u:/, respectively. The working group behind the orthography document is fully aware of this drawback, as the document points out that these length distinctions are not visible in spelling (p. 7). Furthermore, a footnote on p. 2 implies that the working group does not feel entirely responsible for allowing such ambiguities in vowel length. Specifically, the footnote points out that the working group presented the Pite Saami language community with suggestions about how to mark vowel length, but the community instead recommended using the system ultimately presented in the document. In principle, the working group could have ignored the language community's preference; however their choice to listen to the language community was likely a wise decision (see Section 4.5), even if this decision is not ideal from a purely linguistic point of view.

Similarly, a distinction between extra long geminates (in grade III)³⁶ and the other geminates (in grade II) is not captured in the orthography. The orthography document deals with this ambiguity by using an apostrophe to mark extra long (quantity III) geminates in the gradation pattern tables on pages 17 and 18, e.g., <ff'> (as opposed to the shorter geminates in quantity II which are presented as a doubled consonant, e.g., <ff>), but points out that the apostrophe is not used in spelling (p. 17).

In Table 3, three sets of minimal pairs or near minimal pairs are presented to illustrate how these length distinctions are relevant for the pronunciation, but not marked in the orthographic representations. Note that, while the third example in this table presents the length ambiguity of the specific compound grapheme <ff>, this is true for all quantity III vs. quantity II geminate contrasts (such as <jj>, <mm>, <ŋŋ> and <ssj>). The theoretical ambiguity of <h> is not included because it is never truly ambiguous, as discussed below.

These phonemic contrasts are underspecified in the written form, which is unfortunate because it can be somewhat problematic for beginning learners of the language. However, even if it is not allowed in the official orthography, lexical and pedagogical resources can indicate these contrasts, much as the orthography document does.

³⁶ Note that I have chosen to use <::> (a doubled IPA length symbol) to transcribe grade III geminates.

Table 3. Symbols which correspond to more than one sound (here, the geminate grapheme <ff> is exemplary for any geminate), and examples.

symbol	sound	example		
		orthographic	phonemic	gloss
å	/ɔ/	båhtjet	/pɔ ^h tjet/	‘milk.INF’
	/ɔ:/	båhtet	/ pɔ: ^h tjet /	‘come.INF’
u	/u/	gulliv	/kul::iv/	‘hear.1SG.PRS’
	/u:/	gulliv	/ku:l::iv/	‘fish.1SG.PST’
ff	/f:/	jáffo	/ja:f:o/	‘flour.GEN.SG’
	/f::/	jáffo	/ja:f::o/	‘flour.NOM.SG’

Although learners will simply have to remember these distinctions when reading and writing, well planned dictionaries and teaching materials should make finding and understanding these distinctions easier, and thus prevent these orthographic ambiguities from becoming an unsurmountable problem.

Note that the grapheme <h> is also ambiguous because it can correspond to an unvoiced glottal fricative /h/ (as in *hiejos* ‘poor.ATTR’) or the preaspiration of a following plosive or affricate in the consonant center (as in *båhtet* ‘come.INF’). However, this ambiguity is never problematic because these two contexts are mutually exclusive. In other words, <h> can only indicate preaspiration when preceding a plosive or an affricate in the consonant center; in all other cases, it corresponds to an unvoiced glottal fricative.

4.1.2 One sound, multiple symbols

Each singleton, unaspirated plosive sound can be represented using two ways, depending on the location of the sound in a word. Specifically, in word-final position, the graphemes <p>, <t> and <k> are used, while in all other cases, the graphemes , <d> and <g> are used, respectively, as illustrated in Figure 2.

$$\begin{array}{c} \textit{buorep} \text{ /pu}^{\widehat{}}\text{orep/ ‘better.NOM.SG’} \\ \text{vs.} \\ \textit{buoreba} \text{ /pu}^{\widehat{}}\text{orepa/ ‘better.NOM.PL’} \end{array}$$

Figure 2. Example of using multiple symbols (and <p>) for one sound (/p/).

Note that, while <p> and represent the same underlying sound, namely /p/, the pronunciation of some speakers exhibits allophony that aligns with this graphemic distinction such that this phoneme is postaspirated word-finally (when <p> is used); the same is true for the other two plosives as well. This “double” representation may seem unnecessary, but, with this allophony in

mind, this “allography” is not only acceptable, but potentially an advantage for learners, as it actually corresponds to a pronunciation difference which is probably at least salient to non-native speakers. In addition, it may seem familiar to some learners, as similar rules exist in other Saami orthographies.

Technically speaking, there is another situation in which single sounds are represented by multiple symbols. This is the case when a combination of graphemes, which otherwise represent single sounds individually, correspond to a single sound when used in combination. For instance, the combination <nj> (the graphemes <n> and <j>) represents a palatal nasal (/ɲ/). Here are some other examples, but this list is not exhaustive: <sj> refers to a palatal sibilant (/ʃ/), <vv> refers to a labio-dental fricative in either quantity II (/v:/) or quantity III (/v::/), <hp> refers to a preaspirated bilabial plosive (/ʰp/), and <ts> refers to a alveolar affricate (/t͡s/).

From a purely theoretical point of view, this situation should be avoided since it is better to have a single symbol represent a single sound. However, the alternative would be introducing new graphemes for these sounds (for instance using <ʃ> instead of <sj>), thereby adding to an alphabet which otherwise aligns nicely with extant alphabets used in the area, and possibly making the use of modern technology more difficult as well (as discussed in more detail in Section 4.3). However, the orthography’s grapheme combinations will likely be familiar to most users since these are used in Norwegian/Swedish and/or Lule Saami orthographies as well, so learning them should not be overly difficult for most users. More importantly, none of the grapheme combinations used to represent single sounds can occur in other contexts in which the sounds represented by the graphemes taken individually are intended. Thus, confusion as to what these grapheme combinations represent is highly unlikely, once their symbol-sound correspondences have been internalized by orthography users.

With the above exceptions in mind, the Pite Saami orthography quite successfully adheres to the ideal of a one-symbol-one-sound principle. However, with the exception of the lack of visible length distinctions on certain vowels and geminates, even these exceptions are likely no big deal for both learners and native speakers.

4.2 Uniquely recognizable morphemes

Ideally, each and every morpheme should have a unique form in an orthography so as to be unambiguously recognizable (Grenoble & Whaley 2005: 141). This is especially useful for cases of homophony; Pite Saami has many such cases (e.g., the syncretism of nominative plural and genitive singular forms). The Pite Saami orthography clearly fails to reach this ideal; a

few examples of ambiguous orthographic forms are listed in Table 4. In the first example in this table *bissin*, even the lexical morpheme in the stem is ambiguous (it could mean ‘fry’ or ‘wash’), while all cases exhibit ambiguous grammatical morphemes.

Table 4. A few examples of homophony and ambiguous spellings.

<i>ambiguous</i>	
<i>orthographic form</i>	<i>possible readings</i>
bissin	‘fry.3PL.PST’
	‘fry.1DU.PRS’
	‘wash.3PL.PST’
	‘wash.1DU.PRS’
lä	‘be.3SG.PRS’
	‘be.3PL.PRS’
	‘be.2SG.PRS’
	‘be.CONNEG’
luokta	‘bay.NOM.PL’
	‘bay.GEN.SG’

However, due to the complex, non-linear morphophonology of the Pite Saami language, the entire concept of “uniquely recognizing” morphemes is problematic, so this is not exactly the orthography’s fault, especially since an alphabet-based orthography is necessarily linear (while Pite Saami morphology is strikingly non-linear).

The alternatives, such as having a logographic system (e.g., Chinese) or accepting a highly inconsistent correlation between graphemes and sounds (e.g., English), would be much worse choices for Pite Saami. While fluent literate speakers have considerably less difficulties with spellings for words that sound the same (an example from English for this is *sea* vs. *see*), this is only after the years of work involved in learning a spelling system (Grenoble & Whaley 2005: 141).³⁷ However, the Pite Saami language community most likely does not have any truly fluent users of the orthography (at the same level of literacy as for instance Swedish or Norwegian). As a result, the main target user group for the orthography consists of orthography learners (both fluent speakers and language learners); for this group, which is vital to revitalization efforts, it is much more important to have consistent and reliable symbol-to-sound relationships than symbol-to-morpheme relationships. With

³⁷ This is perhaps one of only a few positive aspects of English orthography, and certainly an advantage of a logographic system like Chinese.

this in mind, this seeming ‘shortcoming’ of the orthographic standard is understandable, forgivable, and in fact the best choice given the alternatives.

4.3 Avoiding diacritic overload and minimizing keystrokes

Having (wisely) chosen an alphabet system, the next integral part of orthography development is selecting a set of symbols to represent the phonemes of the given language. While using symbols that language users are already familiar with is clearly an advantage at least concerning the ease of learning, it is often the case that any “loaning” alphabets do not suffice in covering the target language’s sounds. Rather than creating entirely new characters, more practical technical solutions include multiple graphemes to refer to a single sound (cf. Section 4.1.2), adding diacritics,³⁸ and using additional markings on familiar letter symbols to represent some sounds. All three solutions are implemented to some extent in the Pite Saami orthography; this section discusses the advantages and disadvantages of these solutions from both technical and practical perspectives.

Starting with the use of diacritics, as Grenoble & Whaley point out, there may be technical reasons to avoid diacritics in general, and to specifically avoid what they call “diacritic overload” (2005: 146), which is when more than one diacritic on a single character base is used (such as for instance <á̊>, which has both an acute accent and a ring on the base <a>). The case is in fact the same for letters with additional marks, such as <ŋ> (a base <n> with a hook on the bottom right). Specifically, such letters may require the user to make more keystrokes or even unfamiliar keystrokes when typing, or may not be accessible at all (especially on mobile devices), which is a possible deterrent to using a new writing system. Furthermore, there can even be sociological reasons to avoid such letters because users may perceive diacritics (especially when there is diacritic overload) or non-standard letters as too strange, foreign or visually unpleasant, thus decreasing the new orthographic standard’s chances of finding acceptance.

The new Pite Saami orthographic system successfully restricts the use of diacritics to three characters, strictly speaking: <á>, <ä> and <å>. The latter two are familiar to most users, since they are part of the Swedish alphabet as well.³⁹ In addition, there are three other Pite Saami graphemes which, while strictly speaking do not host any diacritic, have marking in addition to the base character; these are <ŋ>, <đ> and <ɞ>. These latter three are not found in the

³⁸ Diacritics are supplementary marks written above or below a base symbol, such as the two dots (called dieresis) on top of the letter a in <ä>.

³⁹ Norwegian users are surely familiar with <ä>, even if they use <æ>; for more, see Section 4.4.2.

respective Nordic alphabets, but are used in other Saami alphabets, so may be familiar to some users as well. Thus, there are a total of six graphemes which are derivations from a base letter, at least two of which are familiar to potential Pite Saami orthography users.⁴⁰ Finally, a few grapheme combinations exist that represent a single sound (cf. Section 4.1.2) such as <nj>, <sj>, <tj> and <ts>.⁴¹

This set of “deviant” graphemes and grapheme combinations seems quite reasonable in size, especially considering that many of these are used in alphabets that users are at least familiar with, if not in fact fully literate in. However, using these solutions comes with two caveats.

The first caveat is only relevant for users who actively want to type Pite Saami, and essentially only pertains to <á>, <ŋ>, <đ> and <ɛ>. Specifically, it is these graphemes that have diacritics or other additional markings which may be difficult to find on standard keyboard layouts, especially those not shared by Nordic alphabets, and particularly when using mobile devices, on which accessing uncommon graphemes is even less straightforward for untrained users than on laptop and desktop computers.⁴² This could result in users of the orthography not adhering to the standard simply because they do not know how to implement it digitally, or that their devices do not support these characters at all without adding a supplemental app. For mobile devices running iOS and Android systems, Divvun (together with Giellatekno at the University of Tromsø, specializing in language technology for Saami and other minority languages), offers a separate app called “Sami keyboards”⁴³ to fulfill specifically this lack in the standard software. While, at least in the iOS version, the installed app does not offer a specifically Pite Saami keyboard, the Lule Saami keyboard has all necessary graphemes, and in fact only <đ> and <ɛ> are not accessible directly via a single keystroke.⁴⁴ Another workaround for iOS is using the Finnish layout, which has all of the graphemes relevant for Pite Saami, although these are only accessible by pushing and holding the respective base letters. In addition to the potential

⁴⁰ Note that the number of words which may be spelled using the graphemes <đ> or <ɛ> is quite small. Furthermore, words using these two letters also have acceptable variant spellings using <ɕ>/<d> or <s>, respectively, which makes them potentially avoidable as well, if necessary.

⁴¹ This is a bit simplified since it actually disregards geminate length (e.g., <nnj> or <dtj>). Furthermore, <hp> (and other <h>-plosive combinations) are also combinations representing a single sound (preaspirated plosives); however, the phonological status of preaspiration is debatable, so I leave these out here for simplicity’s sake.

⁴² Note that finding non-standard graphemes on typewriters may be even more problematic, but due to my own lack of knowledge about typewriters and to the fact that I am assuming that typewriters are uncommon today, especially among younger potential users of the Pite Saami orthography, I will not go into this aspect any further.

⁴³ Cf. divvun.no/keyboards/mobileindex.html (accessed 2020-04-21).

⁴⁴ With ‘keystroke’, I also mean “finger-stroke” on mobile devices.

difficulties in finding these special symbols on a keyboard, accessing them sometimes requires extra effort, i.e., using more than a single keystroke. While this may not sound like much “effort” (indeed, it is merely lifting a finger), minimizing keystrokes is clearly preferred by many users, especially in informal settings, as exhibited by the standardization and even internationalization of such keyboard abbreviations as ‘lol’, ‘btw’, ‘omg’, etc.

With this in mind, it is quite relevant that an orthography be easily implemented in digital technologies, especially for orthographies like the one for Pite Saami, which are partially aimed at winning over younger user groups who are technologically savvy and lazy on their keyboards. With the few exceptions noted above, the Pite Saami orthography should be easily adoptable when using current technologies, although strict adherence to using all graphemes may prove difficult for the reasons mentioned. Nonetheless, the alternative would be restricting the alphabet to only those symbols already available on relevant standard keyboard layouts (i.e., Swedish and Norwegian), and this would come at a higher price. Not only would the dialectal variation accessible via <đ> and <ţ> not be possible, which would thus possibly alienate some users,⁴⁵ but vital contrastive linguistic distinctions accessible through the graphemes <á> and <ŋ> would not be visible in the orthographic representations of the language. This lack of transparency would be difficult for fluent speakers, but especially for learners who rely on orthographic representations corresponding unambiguously to linguistic realities (as discussed in Section 4.1).

The second caveat is perhaps more important for that very reason: in a few cases, contrastive length distinctions are not unambiguously coded in the orthography. Specifically, this concerns the graphemes <â> and <u>, as well as the way grade III geminates and grade II geminates are represented. The details of these underrepresented length distinctions are presented in Section 4.1.1 and summarized in Table 3 above, as well as reasons why this can be considered a weakness in the orthography. Choosing to allow such ambiguities likely stems from an attempt by the orthography developers to limit ‘diacritic overload’ and an overly redundant, ‘foreign’-looking spellings. Technical alternatives could include using the acute accent to mark length on the vowels (based on the model of the <a> vs. <á> length distinction), and thus use <ú> and <ấ>. For the latter, one could certainly argue that this is a case of ‘diacritic overload’, which has practical, aesthetic and technical problems (as discussed above). However, I cannot really find any good reason to not accept the former, especially considering <á> is acceptable; perhaps this

⁴⁵ Cf. Section 4.4 for more on this variation.

choice was made simply because users have seen <á> in other, more common Saami spelling systems (basically North Saami and Lule Saami), but not <ú>.

Another possibility for long <â> would be to double the grapheme, i.e., <ââ>. Even though this approach is used for long consonants (e.g., <bb> in *tjábbe* ‘beautiful’), I have heard speakers disapprove of this for vowels because it “looks too Finnish”, for instance.⁴⁶

As for choosing to suppress the length distinctions between grade III geminates and grade II geminates, the main alternative could be using an apostrophe between the single elements of a quantity III geminate, such as <f'f>. Indeed, the document presenting the orthography uses precisely this approach in a number of tables presenting grade alternations, although the document specifically points out that this is not accepted in the orthography.⁴⁷ Note that this approach is also adopted in Lule Saami, which may have been one source of inspiration for this restriction.

Table 5. Examples of hypothetical symbol-sound correlations, and the resulting spellings which are unambiguous concerning length distinctions.

<i>symbol</i>	<i>sound</i>	<i>example</i>		
		<i>orthographic</i>	<i>phonemic</i>	<i>gloss</i>
â	/ɔ/	bâhtjet	/pɔ ^h tjet/	‘milk.INF’
â	/ɔ:/	bâhtet	/ pɔ: ^h tjet /	‘come.INF’
u	/u/	gulliv	/kul::iv/	‘hear.1SG.PRS’
ú	/u:/	gúlliv	/ku:l::iv/	‘fish.1SG.PST’
ff	/f:/	jáffo	/ja:f:o/	‘flour.GEN.SG’
f'f	/f::/	jáf'fo	/ja:f::o/	‘flour.NOM.SG’

In summary, the new standard orthography successfully minimizes the number of keystrokes needed to type words and avoids diacritic overload. The resulting drawbacks, as discussed above, are minimal.

4.4 Representing variation

Within the Pite Saami language area, there is a certain amount of linguistic variation between speakers; this can correlate with a variety of factors, including geographic, demographic, sociological, and even individual ones. As Grenoble & Whaley (2005: 151–158) discuss quite thoroughly, a shared standard orthography can serve as an identity building factor which can

⁴⁶ They were presumably referring to the fact that long vowels are marked in Finnish orthography using a doubled vowel grapheme.

⁴⁷ For example, “in these tables, the apostrophe (') indicates grade III, but it is not used in the spelling system” (my translation, p. 17).

successfully contribute to revitalization efforts for highly endangered languages. Creating a successful standard involves a balancing act between representing enough variation so as not to alienate users, while simultaneously ensuring that the standard is consistent and not overly variable. In other words, allowing too much variation risks spoiling the creation of a single, unifying standard. While dealing with variation mainly corresponds to the representation of spoken language (dealt with in Section 4.4.1), even written variation can be a factor, as discussed briefly in Section 4.4.2.

4.4.1 Spoken variation

The document describing the new Pite Saami orthography dedicates more than two pages to a number of specific cases of variation in spoken Pite Saami and how these are each dealt with in the written standard (pp. 31–33). Each description of a type of variation presents the relevant variants in the spoken language using orthographic representations, and indicates clearly whether the variation is accepted in the standard.⁴⁸ In cases where more than one variant is acceptable, the document points out that it is important to stick to one variant consistently, and not mix the variants within a text. For instance, the word for “rock.NOM.SG” can be pronounced and spelled three ways: *gärrge*, *gäddge* and *gäddge*; these three spellings are all acceptable, “as long as one is consistent” (p. 32, my translation). On the other hand, the word for ‘bay.NOM.SG’ has two main pronunciations: *luakkta* or *luäkkta*; here, only the former spelling is accepted in written form. A few more examples are presented in Table 6.

The orthography document successfully and consistently points out that all these cases of variation are acceptable in *spoken* Pite Saami. This is very important, as it emphasizes that the document does not attempt to prescribe how people speak, but it is instead focused on how to represent Pite Saami in *written* form. Although a variety of spoken variants are acceptable in written form as well (cf. Table 6 and the document itself), and speakers who identify with such variants should have no complaints about the orthographic system, speakers who identify with spoken variants that do not have an exact written representation will hopefully also understand that some sacrifices need to be made in the name of standardization, and that these compromises only concern the written forms.

⁴⁸ Note that only the description of variation between short <ä> and long <â> in the second syllable (p. 32) lacks this recommendation, but that is simply because the written language does not distinguish between short and long <â> even when it does not depend on dialectal variation.

Table 6. Variant forms of a selection of Pite Saami words, and whether these are acceptable written forms (✓) or not (X).

<i>accepted</i>		
<i>variant</i>	<i>written form?</i>	<i>gloss</i>
gäddge	✓	‘rock.NOM.SG’
gärrge	✓	
gäddge	✓	
muossá	✓	‘maternal_aunt.NOM.SG’
muottá	✓	
mannat	✓	‘go.INF’
mannad	X	
várest	✓	‘mountain.ELA.SG’
váres	X	
luakkta	✓	‘bay.NOM.SG’
luäkkta	X	
biegga	✓	‘wind.GEN.SG’
biekka	X	

Although the document does not force individual writers to adhere to its guidelines, to what extent speakers of non-written variants follow the official orthography will be an important test of its viability and success.

4.4.2 Variation in writing traditions

The grapheme <æ> is not included as an alternative to <ä> in the document describing the Pite Saami orthography. This is different from the Lule Saami orthography, which allows for orthographic variation such that <ä> is used in Sweden and <æ> is used in Norway (in alignment with the fact that <ä> is used in the Swedish alphabet, and <æ> in the Norwegian alphabet). Since this official document does not mention <æ> as being an option, it is therefore technically not an official option. Although the majority of Pite Saami native speakers today are from the Swedish side, it could be wise to allow variation between <ä> and <æ> with the goal of maximizing grass-root support for revitalization efforts on both sides of the border. Perhaps this was simply an oversight of the committee that created the document, but perhaps it was intentional; in either case, it seems to me that not including the possibility to have this specific variation runs the risk of alienating potential users (such as teachers, learners and authors) who are used to using Norwegian-based input methods in modern technology, as well as in handwriting.

On the other hand, this is a relatively insignificant aspect of the orthography, and does not affect the effectiveness and utility of the standard as a whole.

With this in mind, I would not expect it to significantly impact revitalization efforts in a negative way, although the symbolism of only allowing a grapheme associated specifically with Sweden may hold more weight than I can imagine as an outsider. And, ultimately, people are free to write as they choose in unofficial instances, so it is possible that some Norwegian users may simply use <æ> anyway, perhaps out of principle, or simply out of habit.

4.5 Supporting revitalization efforts

When designing or even reviewing a writing system, it is essential to contextualize the orthographic standard because, although some generalizations can be made, ultimately, every orthography attempts to fulfill its purpose for linguistic, political and cultural situations that are at least as unique as the linguistic diversity that drives speakers to develop unique orthographies to begin with. For the Pite Saami context, something that Grenoble & Whaley point out seems particularly relevant, namely that “orthographies for communities creating revitalization programs should be designed primarily with beginning readers in mind” (2005: 142). The extent to which the Pite Saami language community is actively working on revitalization programs will not be covered here, but it is clear that the language’s current situation requires a significant amount of revitalization for the language to remain vital in the coming years. Specifically, to summarize the situation (cf. Section 1), most native speakers are in the older generations and only a small number of younger speakers are potentially learning the language natively; this imbalance in the distribution of the age of speakers indicates that new speakers are needed, and the needs of this group of beginning readers and writers alike should be prioritized over the needs of the larger group of fluent older speakers in order to help maximize revitalization efforts. The more easily accessible the language is via its orthography, the better are the chances that revitalization efforts will prove successful; the writing system is surely a small part of the effort, but a language like Pite Saami needs all the help it can get.

Having thus identified the importance of focusing on new readers and writers, the current official orthographic standard fulfills this criterion quite successfully, although not entirely. As mentioned in previous sections, a few weaknesses of varying gravity exist which will likely make learning the language via reading and writing somewhat difficult. Specifically, these are:

- a few cases of using one symbol to represent multiple sounds, or multiple symbols to represent one sound (cf. Section 4.1)

- the difficulty of making morphemes uniquely recognizable in an inherently morphophonologically complex language (cf. Section 4.2)
- the relative inconvenience involved in accessing a few characters on standard keyboard layouts (cf. Section 4.3)
- variation in spelling reflecting dialectal variants (cf. Section 4.4)

However, as discussed in the respective sections on these points of criticism, none of these taken alone is crucial enough to significantly affect the overall success of the orthography. Furthermore, these points are different enough from one another, and small enough in number, that these relative shortcomings taken together should not prohibit the orthography from being useful either, even for beginning readers and writers. It may require some extra effort, particularly in these respects, but in comparison to learning orthographies for other languages, the Pite Saami writing system in fact quite reliably corresponds to linguistic realities. Although the availability of resources for learning orthographies used for large languages such as those for English or French (just to name two particularly inaccessible systems for learners) make it nearly pointless to attempt to compare these with Pite Saami, the fact remains that millions of learners successfully navigate English and French reading and writing; in comparison, learning to read and write Pite Saami should be a walk in the park. In addition, dictionaries and teaching materials can help compensate as well, especially concerning the underspecified length distinctions by marking these in a way not used in the official orthography.

Although the focus here has been on beginning language users, it should also be pointed out that the Pite Saami orthography should be even more accessible to native speakers learning to read and write in their native language. This is because they have the advantage of possessing native speaker intuition to guide them. Especially pertaining to the underspecification of length distinctions, native speakers will know instinctively what the correct length is, even if it is not represented unambiguously in writing. Hopefully the flexibility of the orthography in allowing a number of dialectal variants will satisfy native speakers' desire to feel that each individual spoken Pite Saami variant is sufficiently represented. Indeed, successful acceptance on behalf of native speakers can have important symbolic and practical affects in revitalization efforts.

5. Outlook and potential

In summary, the new orthographic system should prove to benefit rather than inhibit revitalization efforts. While its medium-term and long-term effects remain to be seen, it is hopefully an effective part in the complex process of promoting the status of Pite Saami and, even more importantly, increasing the number of Pite Saami speakers. Overall, it minimizes linguistic shortcomings and maximizes the needs of the language community, particularly those of language learners. The writing system forms a reliable, detailed standard that can be easily implemented in producing much needed teaching materials. Official recognition of this standard will hopefully lead to tangible results felt in the Pite Saami community because this administrative political act has put Pite Saami officially on the map as its own politically legitimated entity, whereas before the standard was official, Pite Saami was often just considered a dialect of Lule Saami, if it was visible at all. This will hopefully make it easier for specifically Pite Saami language projects, such as those desperately needed to promote revitalization, to receive funding. Furthermore, it should require local and regional governmental agencies to officially support using the Pite Saami language in relevant public spheres (such as in day-care and care of the elderly). This document has hopefully triggered a chain of events critical to the Pite Saami language's continued survival as a spoken language. Namely, this improvement in social status will promote an increase in language use, which will ultimately boost the number of speakers such that at some point, Pite Saami's future will no longer be endangered.

At the time of writing, the Pite Saami orthography is only eight months old, but already the first ever Pite Saami book has been published, and it uses this new standard. In late February 2020, *Biejve Nájjda* was released (Somby 2020); it is a translation by Pite Saami speakers Inger Fjällås and Peter Steggo of the Saami children's book *Solens dattar* by Marry Áilonieida Somby. Despite this new book and a number of online lexical resources that use the new standard (as outlined in Section 3), the Pite Saami language is still lacking any published, easily accessible teaching materials, such as a primer. The official recognition of a generally successful orthographic standard has already removed one of the hurdles encountered in producing such needed materials in support of revitalization efforts. This orthography has the potential to do even more in paving the way to a more promising future for the Pite Saami language.

Finally, to come back to the scholarly source that has informed my evaluation of the new Pite Saami orthography, I would like to emphasize one more passage from Grenoble & Whaley (2005) in highlighting the fact that it is not me as an outsider and academic with a supposedly critical and distanced

eye, and in fact not even a political entity like *Sámi Giellagáldu* who will in the end determine the success or failure of this spelling system. Instead, “[r]egardless of how linguistically and technically sound an orthography might be, its initial (and continued) acceptance by the people for whom it is designed is critical in determining its eventual effectiveness and use” (p. 137). In is my sincere hope that the Pite Saami language community will make the most out of the opportunity given to them with the creation of an official *bisumsáme tjállemvohke*.

Abbreviations

1—first person; 2—second person; 3—third person; ACC—accusative; CONNEG—connegative; DU—dual number; ELA—elative; GEN—genitive; INF—infinitive; IPA—International Phonetic Alphabet; NOM—nominative; PL—plural number; PRS—present tense; PST—past tense; saL—Lule Saami; saN—North Saami; SG—singular number; SPR—Saami parliamentary council; UPA—Uralic Phonetic Alphabet.

References

- Bengtsson, Nils-Henrik, Marianne Eriksson, Inger Fjällås, Eva-Karin Rosenberg, Gry Helen Sivertsen, Valborg Sjaggo, Dagny Skaile, Peter Steggo, & Joshua Wilbur. 2016. Pitesamisk ordbok. In Joshua Wilbur (ed.), *Pitesamisk ordbok samt stavningsregler* (Samica 2), 13–121. Freiburg: Albert-Ludwigs-Universität Freiburg.
- Grenoble, Lenore A. & Lindsay J. Whaley. 2005. *Saving languages. An introduction to language revitalization*. Cambridge: Cambridge University Press.
- Moseley, Christopher (ed.). 2010. *Atlas of the World's Languages in Danger*. 3rd ed. Paris: Unesco Publishing.
- Sjaggo, Ann-Charlotte. 2015. *Pitesamisk grammatik. En jämförande studie med lulesamiska* (Senter for samiske studiers skriftserie 20). Tromsø: Septentrio Academic Publishing.
- Somby, Marry Áilonieida. 2020. *Beijve Näjjda*. Trans. by Inger Fjällås & Peter Steggo. Snåsa: Snåsa kommune.
- Valijärvi, Riitta-Liisa & Joshua Wilbur. 2011. The past, present and future of the Pite Saami language. Sociological factors and revitalization efforts. *Nordic Journal of Linguistics* 34. 295–329.
- Wallström, Sigvard. 1943. *Studier i Övre Norrlands språkgeografi med utgångspunkt från Arjeplogmålet*. Uppsala: Almqvist & Wiksell.
- Wilbur, Joshua. 2014. *A grammar of Pite Saami* (Studies in Diversity Linguistics 5). Berlin: Language Science Press.
- Wilbur, Joshua. 2016a. Stavningsregler. In Joshua Wilbur (ed.), *Pitesamisk ordbok samt stavningsregler* (Samica 2), 123–197. Freiburg: Albert-Ludwigs-Universität Freiburg.
- Wilbur, Joshua (ed.). 2016b. *Pitesamisk ordbok samt stavningsregler* (Samica 2). Freiburg: Albert-Ludwigs-Universität Freiburg.

Acta Universitatis Upsaliensis

Studia Uralica et Altaica Upsaliensia

Prior to 1990 the series was called *Studia Uralica et Altaica Upsaliensia*. The following books were published under this title.

1. BJÖRN COLLINDER: Sprachverwandtschaft und Wahrscheinlichkeit. Ausgewählte Schriften neu veröffentlicht zum 70. Geburtstag des Verfassers 22. juli 1964, zusammen mit einer Bibliographie der Werke von Björn Collinder 1921–1964. 1964.
2. SEUNG-BOG CHO: A phonological study of Korean. With an historical analysis. 1967.
3. NILS ERIK HANSEGÅRD: Recent Finnish Loanwords in Jukkasjärvi Lappish. 1967.
4. ÖDON LAVOTHA: Studien zu dem ungarischen Potentialsuffix *hat/het*. 1967.
5. VALEV UIBOPUU: Similarkomparative Konstruktionen im Finnischen und Estnischen insbesondere in der modernen Schriftsprache. Syntaktisch-stilistische Untersuchungen. 1970.
6. NILS ERIK HANSEGÅRD: Finnish Loanwords in Jukkasjärvi Lappish Classified according to their Meaning. Supplement to *Recent Finnish Loanwords in Jukkasjärvi Lappish* by the same author. 1971.
7. EIVOR NYLUND TORSTENSSON: De finska sjönamnen i Korpilombolo socken. Summary: *Finnish lake-names in Korpilombolo, Sweden*. 1973.
8. VALTER TAULI: Standard Estonian Grammar. Part I. Phonology, morphology, wordformation. 1973.
9. MILJA ANDERSSON: Huomioita suomen mainoskielestä. Zusammenfassung: *Bemerkungen zur finnischen Werbesprache*. 1974.
10. OSCAR LAZAR: The Formation of Abstract Nouns in the Uralic Languages. 1974.
11. SIIRI SAHLMAN-KARLSSON: Specimens of American Finnish. A field study of linguistic behaviour. 1976.
12. AARAND ROOS: Morfologiska tendenser vid språklig interferens med estniska som bas. Zusammenfassung: *Morphologische Tendenzen bei sprachlicher Interferenz mit Estnisch als Grundsprache*. 1980.
13. RAIMO RAAG: Lexical Characteristics in Swedish Estonian. 1982.
14. VALTER TAULI: Standard Estonian Grammar. Part II. Syntax. 1983.
15. LARS-GUNNAR LARSSON: Studien zum Partitivgebrauch in den ostseefinnischen Sprachen. 1983.
16. ANNE NESSER: ”Se oli kurja käya koulusa”. Bruket av inre och yttre lokalkasus i sverigefinska skolbarns språk. 1986.
17. TUULI FORSGREN: Finska participattribut i svensk översättning. Summary: *Finnish Attributive Participles translated into Swedish*. 1986.
18. NORE JOHANSSON: Tornedalsfinska ord ur dialektgeografisk synvinkel. Zusammenfassung: *Wörter des finnischen Dialekts von Tornedalen. Eine dialektgeographische Untersuchung*. 1987.
19. Proceedings of the Second Scandinavian Symposium on Aspectology. Ed. LARS-GUNNAR LARSSON. 1989.

Acta Universitatis Upsaliensis

Studia Uralica Upsaliensia

Editors: Rogier Blokland and Torbjörn Söder

20. MÁRIA DUGÁNTSY: Erzä-mordwinische rituelle Klagegesänge. 1991.
21. LEENA MARJATTA HUSS: Simultan tvåspråkighet i svensk-finsk kontext. 1991.
22. TOIVO TIKKA: Vepsän suffiksoituneet postpositiot. Kieliopillisiin sijoihin liittyvä suffiksoituminen. 1992.
23. MANJA IRMELI LEHTO: Ingrian Finnish: Dialect Preservation and Change. 1996.
24. TIINA SÖDERMAN: Lexical Characteristics of the Estonian North Eastern Coastal Dialect. 1996.
25. PEEP NEMVALTS: Case Marking of Subject Phrases in Modern Standard Estonian. 1996.
26. Lapponica et Uralica. 100 Jahre finnisch-ugrischer Unterricht an der Universität Uppsala. Vorträge am Jubiläumssymposium 20.–23. April 1994. Herausgegeben von LARS-GUNNAR LARSSON. 1996.
27. SATU M. GRÖNDAHL: Den ofullkomliga traditionen. Bilden av Ingermanlands kvinnliga runotradition. 1997.
28. SUSANNA ANGÉUS KUOLJOK: Nominalavledningar på *ahka* i lulesamiskan. 1997.
29. VIRVE RAAG: The Effects of Planned Change on Estonian Morphology. 1998.
30. STEFAN M. PUGH: Systems in Contact, System in Motion: The Assimilation of Russian Verbs in the Baltic Finnic Languages of Russia. 1999.
31. LEENA HUSS: Reversing Language Shift in the Far North: Linguistic Revitalization in Northern Scandinavia and Finland. 1999.
32. SUSANNA FAHLSTRÖM: Form and Philosophy in Sándor Weöres' Poetry. 1999.
33. TORBJÖRN SÖDER: "Walk This Way": Verbs of Motion in Three Finno-Ugric Languages. 2001.
34. LEELO KEEVALLIK: From Interaction to Grammar: Estonian Finite Verb Forms in Conversation. 2003.
35. ANDRÉ HESSELBÄCK: Tatar and Chuvash Code-copies in Mari. 2005.
36. ANNIKA JANSSON: Sami Language at Home and at School: A Fieldwork Perspective. 2005.
37. Inheriting the 1990s. The Baltic Countries. Ed. BAIBA METUZĀLE-KANGERE. 2010.
38. BOGLÁRKA STRASZER: "Ungerska för rötternas skull". Språkval och identitet bland andragenerationens ungrare i Sverige och Finland. 2011.
39. GÁBOR TILLINGER. Samiska ord för ord. Att mäta lexikalt avstånd mellan språk. 2014.
40. TÜNDE BLOMQVIST: Mellan två stolar. Författarskap i Sverige med ungerskspråkig bakgrund 1945–2015. 2017.
41. HANS-HERMANN BARTENS: Die Erzähltraditionen der Pite- und Umesaamen. 2020
42. ROGIER BLOKLAND, MICHAEL RIEBLER & TORBJÖRN SÖDER (eds.) Saami Linguistics in Uppsala. The SAALS 4 Symposium. 2025.

