

Faculty of Arts
University of Helsinki

POSTPOSED DEMONSTRATIVES IN FINNIC AND NORTH RUSSIAN DIALECTS

Chingduang Yurayong

DOCTORAL DISSERTATION

To be presented for public discussion with the permission of the Faculty of Arts of
the University of Helsinki, in Auditorium PIII, Porthania building,
on the 14th of November, 2020 at 10 o'clock.

Helsinki 2020

Postposed demonstratives in Finnic and North Russian dialects

Copyright © Chingduang Yurayong 2020

Doctoral Programme in Language Studies and
Department of Finnish, Finno-Ugrian and Scandinavian Studies
Faculty of Arts
University of Helsinki

Cover image: «se» – a computer graphic based on Novgorod birch bark documents
№202, №220, №292, №419, and №913
Designed by Puritip Suriyapatarapun

The Faculty of Arts uses the Urkund system (plagiarism recognition) to examine all
doctoral dissertations.

ISBN 978-951-51-6691-3 (paperback)
ISBN 978-951-51-6692-0 (PDF)

Printed by Unigrafia
Helsinki 2020

ABSTRACT

The present study investigates the use and development of demonstratives that follow head word, postposed demonstratives, which are characteristic of eastern Finnic and North Russian dialects. Some previous studies regard these postposed demonstratives as definite articles, while other recent studies identify additional functions related to information structure and discourse. Given that postposed demonstratives are not a feature common to all East Slavic languages, several studies propose that this characteristic feature of North Russian could have resulted from language contact with the Uralic-speaking population who adopted Russian as their second language, particularly Finnic speakers.

The main goal of the present study is to answer three research questions:

- 1) How do postposed demonstratives function as grammatical markers?
- 2) What does the development of demonstratives tell us about the history of Finnic and Slavic languages?
- 3) Do postposed demonstratives result from a Finnic substratum in North Russian dialects?

For this purpose, the present study examines spoken language data comprising thirteen Finnic and two North Russian varieties which have been in contact during the latest millennium, as well as Novgorod birch bark documents from the 11th–15th centuries. The typological analysis identifies properties and functions of postposed demonstratives from various perspectives: word order, host attachment, syntactic and pragmatic functions. The analysis also combines results with geographical data, which shows the correlation between the speaking areas and linguistic similarities among varieties.

The results achieved in the present study justify the following conclusions. First, postposed demonstratives function as grammatical markers with a basic function to organise information structure. At the same time, the properties of information-structural uses as topic and focus markers have secondarily extended to contexts of use in which postposed demonstratives co-occur with definite referents, and are used to code the speaker's evaluation. The functional extension is particularly common in North Russian dialects and adjacent Finnic varieties in the east. Second, the development of demonstrative systems from Proto-Finnic to modern Finnic languages is influenced by later contacts among Finnic sub-branches that share areal features. Based on these isoglosses, the Finnic demonstrative system can be classified into four groups: 1) western Finnic (Livonian, South Estonian, and North Estonian), 2) central Finnic (Votic and Ingrian), 3) Karelian Finnic (Olonets Karelian and Northern Lude), and 4) eastern Finnic (Southern Lude and Veps). Third, the postposed demonstrative “-to” and its variants in North Russian dialects do not result from the Finnic substratum, but from the adstratum. Through mutual reinforcement with the Veps demonstrative “se”, the indeclinable “-to” inherited from the Central dialect of Middle Russian has developed further properties to inflect and co-occur more often with definite referents. Such a developed pragmatic use later also spread to Lude and Olonets Karelian.

Keywords: demonstrative, Finnic languages, North Russian dialects, substratum, language contact, areal typology, definiteness, information structure

TIIVISTELMÄ

Tämä tutkimus käsittelee käyttökontekstin ja kehityksen näkökulmasta pääsanan jälkiasemassa käytettäviä demonstratiiveja, ts. pääsanan jälkeisiä demonstratiiveja, jotka ovat itäisten itämerensuomalaisten ja pohjoisvenäjän murteiden tunnusomainen piirre. Osa aiemmista tutkimuksista pitää pääsanan jälkeisiä demonstratiiveja määräisinä artikkeleina, kun taas toiset tuoreemmat tutkimukset tuovat esiin myös muita funktioita, jotka liittyvät informaatiorakenteeseen ja diskurssiin. Sen perusteella, että tämä piirre ei ole yhteinen kaikille itäslaavilaisille kielille, muutamat tutkimukset esittävät selityksen, jonka mukaan tämä pohjoisvenäjän ominaispiirre juontuisi kielikontakteista uralilaisen, erityisesti itämerensuomalaisen väestön kanssa, joka on omaksunut venäjän toisena kielenä.

Tutkimuksen tavoite on vastata seuraaviin kolmeen tutkimuskysymykseen:

- 1) Miten pääsanan jälkeiset demonstratiivit toimivat kieliopillisena merkitsijänä?
- 2) Mitä demonstratiivien kehitys kertoo itämerensuomalaisten ja slaavilaisten kielten historiasta?
- 3) Ovatko pääsanan jälkeiset demonstratiivit pohjoisvenäjässä syntyneet itämerensuomalaisesta substraattista?

Tavoitteen saavuttamiseen tämä tutkimus tarkastelee puhekielen aineistoa, joka koostuu kolmestatoista itämerensuomalaisesta ja kahdesta pohjoisvenäjän varietetista, jotka ovat olleet kontaktissa keskenään viimeisen vuosituhaten aikana, samoin Novgorodin tuohiasiakirjoja 1000–1400-luvulta. Typologinen analyysi määrittelee pääsanan jälkeisten demonstratiivien ominaisuuksia ja funktioita sanajärjestyksen, pääsanaan liittämisen sekä syntaktisten ja pragmaattisten funktioiden näkökulmista. Analyysi yhdistää tulokset myös maantieteelliseen tietoon, joka osoittaa puhuma-alueiden ja kielellisen samankaltaisuuden vastaavuussuhteen.

Tutkimustulokset johtavat seuraaviin johtopäätöksiin. Ensiksi pääsanan jälkeiset demonstratiivit toimivat kieliopillisena merkitsijänä, jonka peruskäyttö liittyy informaatiorakenteen jäsentämiseen. Samalla informaatiorakenteen käyttö topiikki- ja fokusmerkitsijänä on toissijaisesti laajentunut käyttökonteksteihin, joissa pääsanan jälkeiset demonstratiivit esiintyvät määräisten tarkoitteiden kanssa ja niitä käytetään merkitsemään puhujan arvioinnin. Kyseinen laajennus on havaittavissa erityisesti pohjoisvenäjän murteissa ja vierekkäisissä itämerensuomalaisissa kielissä idässä. Toiseksi demonstratiivijärjestelmän kehitykseen kantasuomesta nykykieliin ovat vaikuttaneet myöhäisemmät kontaktit niissä kieliharoissa, joissa esiintyy yhteisiä areaalipiirteitä. Näiden isoglossien perusteella itämerensuomalainen demonstratiivijärjestelmä voidaan luokitella neljään ryhmään: 1) läntisiin (liivi, eteläviro ja pohjoisviro), 2) keskisiin (vatja ja inkeri), 3) karjalaisiin (aunuksenkarjala ja pohjoislyydi), ja 4) itäisiin (etelälyydi ja vepsä). Kolmanneksi pohjoisvenäjän pääsanan jälkeinen demonstratiivi ”-to” variantteineen ei ole syntynyt itämerensuomalaisesta substraattista vaan adstraattista. Keskinäisenä vahvistuksena vepsän demonstratiivin ”se” kanssa keskivenäjän keskisestä murteesta periytynyt taipumaton pääsanan jälkeinen demonstratiivi ”-to” on alkanut taipumaan sekä esiintymään useammin määräisten tarkoitteiden jäljessä. Sellainen kehittynyt pragmaattinen käyttö myöhemmin levisi myös lyydiin ja aunuksenkarjalaan.

Avainsanat: demonstratiivi, itämerensuomalaiset kielet, pohjoisvenäjän murteet, substraatti, kielikontaktit, areaalitypologia, määräisyys, informaatiorakenne

SAMMANDRAG

Denna forskning behandlar användningen och utvecklingen av demonstrativer som följer sitt huvudord, dvs. efterställda demonstrativer, som är kännetecknande för östliga östersjöfinska och nordryska dialekter. Några tidigare forskningar ser dessa efterställda demonstrativer som bestämda artiklar, medan senare forskningar också identifierar andra funktioner kopplade till informationsstruktur och diskurs. Eftersom efterställda demonstrativer inte är kännetecknande för alla östslaviska språk, föreslår flera forskningar att detta fenomen i nordryska kan vara ett resultat av språkkontakt med den uraliska, särskilt östersjöfinska, befolkning som antagit ryska som andraspråk.

Målet med denna studie är att svara på följande tre frågor:

- 1) Hur fungerar efterställda demonstrativer som grammatiska markörer?
- 2) Vad berättar utvecklingen av demonstrativer om de östersjöfinska och slaviska språkens historia?
- 3) Är efterställda demonstrativer i nordryska ett resultat av ett finskt substratum?

För att besvara dessa frågor undersöker denna studie talspråkskorpusar omfattande tretton östersjöfinska och två nordryska varianter, samt Novgorods näverdokument från 1000–1400-talen. Den typologiska analysen identifierar egenskaper och funktioner hos efterställda demonstrativer ur olika perspektiv: ordföljd, fastsättning på huvudordet, syntaktiska och pragmatiska funktioner. Analysen kombinerar också resultaten med geografisk information, vilket visar korrelationen mellan lokaliseringen av språksamhällen och språkliga likheter mellan varianterna.

Resultaten föreslår följande tre slutledningar. För det första fungerar efterställda demonstrativer som grammatiska markörer vars grundfunktion är organisationen av informationsstruktur. Samtidigt har de som informationsstrukturella markörer för topik och fokus sekundärt också utvidgats till kontexter där de förekommer efter bestämda referenter och används för att markera talarens utvärdering. Denna funktionella utvidgning påträffas särskilt i nordryska dialekter och intilliggande östersjöfinska varianter i öster. För det andra har utvecklingen av demonstrativsystemet från uröstersjöfinska till moderna språk påverkats av senare kontakter mellan östersjöfinska undergrenar som delar areala drag. Utgående från dessa isoglosser kan det östersjöfinska demonstrativsystemet klassificeras i fyra grupper: 1) västliga (liviska, nordestniska och sydestniska), 2) centrala (votiska och ingriska), 3) karelska (olonetskarelska och nordludiska), och 4) östliga (sydludiska och vepsiska). För det tredje är den efterställda demonstrativen ”-to” och dess varianter i nordryska dialekter inte ett resultat av ett östersjöfinskt substratum utan av ett adstratum. Genom den gemensamma förstärkningen med den vepsiska demonstrativen ”se” har den oböjliga demonstrativen ”-to”, som ärvts från den centrala dialekten av mellanryska, utvecklat ytterligare böjningsegenskaper samt börjat förekomma oftare efter bestämda referenter. Sådan avancerad pragmatisk användning spreds senare också till ludiska och olonetskarelskan.

Nyckelord: demonstrativ, östersjöfinska språk, nordryska dialekter, substratum, språkkontakt, arealtypologi, bestämdhet, informationsstruktur

РЕЗЮМЕ

Настоящая работа исследует развитие и употребление характерного для восточных прибалтийско-финских и севернорусских диалектов типа указательных местоимений, так называемых постпозитивных демонстративов, отличающихся позицией после определяемого слова. В ранней научной литературе постпозитивные демонстративы считались определёнными артиклями, тогда как дальнейшие исследования позволили установить функции, связанные с актуальным членением и дискурсом. Так как постпозитивные демонстративы не встречаются в других восточнославянских языках, предлагалось объяснить их наличие в севернорусских диалектах контактным влиянием прибалтийско-финских языков.

Основные задачи работы заключаются в следующих вопросах:

- 1) Какими грамматическими функциями обладают постпозитивные демонстративы в исследуемых языках?
- 2) Что можно сказать об истории исследуемых языков в свете данных о развитии демонстративов?
- 3) Связано ли происхождение постпозитивных демонстративов в севернорусских диалектах с прибалтийско-финским субстратом?

Материалом для исследования послужили речевые корпуса тринадцати прибалтийско-финских и двух севернорусских вариантов, а также новгородские берестяные грамоты XI–XV веков. В работе анализируются с типологической точки зрения такие характеристики постпозитивных демонстративов, как линейный порядок, тип связи с определяемым словом, синтаксические и прагматические функции. Результаты типологического анализа связываются с географическими данными с тем, чтобы проследить взаимосвязанность географической близости и языковых сходств между исследуемыми вариантами.

Результаты настоящей работы следующие. Во-первых, основная грамматическая функция постпозитивных демонстративов связана с актуальным членением предложения, однако в севернорусских диалектах и соседних восточных прибалтийско-финских вариантах наблюдаются и вторичные функции, сочетающийся с определённым референтом и употребляющийся как маркеры оценки говорящего. Во-вторых, развитие праприбалтийско-финской системы демонстративов в языках-потомках усложнялось контактами между отдельными прибалтийско-финскими языками, приобретавшими общие ареальные черты. На основе наблюдавшихся изоглосс предлагается классификация прибалтийско-финских систем демонстративов на четыре группы: 1) западную (ливский, южноэстонский, северноэстонский), 2) центральную (водский, ижорский), 3) карельскую (олонецкий, севернолюдиковский) и 4) восточную (южнолюдиковский, вепсский). В-третьих, постпозитивный демонстратив «-то» и связанные с ним формы в севернорусских диалектах не являются результатом прибалтийско-финского субстрата, а скорее адстрата. Под взаимным влиянием вепсского демонстратива «se» и унаследованного из центрального диалекта среднерусского языка неизменяемого местоимения «-то» последнее постепенно приобрело формы склонения, а также начало чаще сочетаться с определённым референтом. Такая же прагматическая функция впоследствии распространилась и на людиковский и олонецкий языки.

Ключевые слова: указательные местоимения, прибалтийско-финские языки, севернорусские диалекты, субстрат, языковые контакты, ареальная типология, определённость, актуальное членение

บทคัดย่อ

งานวิจัยชิ้นนี้สืบวิเคราะห์การใช้งานและพัฒนาการของคำบ่งชี้ที่ตามหลังคำหลักของวลี กล่าวคือ คำบ่งชี้ตามหลัง ซึ่งเป็นลักษณะเฉพาะของภาษาฟินิกผิงตะวันออก และ ภาษารัสเซียถิ่นเหนือ งานวิจัยก่อนหน้านี้บางชิ้นจัดให้คำบ่งชี้ตามหลังเหล่านี้เป็นคำประเภทเดียวกับคำกำกับนามชี้เฉพาะ แต่งานวิจัยอื่น ๆ ในระยะหลังยังค้นพบการใช้งานในรูปแบบอื่น ๆ ที่เกี่ยวข้องกับโครงสร้างข้อความ และสัมพันธสารอีกด้วย เนื่องจากคำบ่งชี้ตามหลังมีใช้ลักษณะร่วมที่พบในกลุ่มภาษาสลาวิกตะวันออกทั้งหมด งานวิจัยจำนวนหนึ่งจึงเสนอว่า ปรากฏการณ์ดังกล่าวที่พบในภาษารัสเซียถิ่นเหนืออาจเป็นผลที่เกิดจากสัมผัสภาษากับประชากรภาษาตระกูลอูราล โดยเฉพาะกลุ่มฟินิก ผู้รับเอาภาษารัสเซียมาใช้เป็นภาษาที่สอง

จุดมุ่งหมายหลักของงานวิจัยชิ้นนี้ คือ การหาคำตอบสำหรับคำถามสามประเด็นต่อไปนี้

๑. คำบ่งชี้ตามหลังมีลักษณะการใช้งานเป็นคำไวยากรณ์รูปแบบใด
๒. พัฒนาการของคำบ่งชี้บอกอะไรแก่เราบ้างเกี่ยวกับประวัติศาสตร์ของภาษากลุ่มฟินิกและสลาวิก
๓. คำบ่งชี้ตามหลังในภาษารัสเซียถิ่นเหนือกำเนิดมาจากอิทธิพลภาษาฟินิกพื้นเดิมหรือไม่

ในการนี้ งานวิจัยชิ้นนี้ได้ตรวจสอบข้อมูลภาษาพูดที่ประกอบไปด้วยภาษาฟินิก ๑๓ ถิ่น และ ภาษารัสเซีย ๒ ถิ่น ซึ่งล้วนแล้วแต่มีการสัมผัสภาษาเรื่อยมาในช่วงสหัสวรรษหลังมานี้ รวมไปถึงจดหมายเปลือกไม้ เมืองนอฟโกรอดจากช่วงคริสต์ศตวรรษที่ ๑๑ ถึง ๑๕ การวิเคราะห์เชิงแบบลักษณ์ภาษาได้ค้นพบลักษณะ และการใช้งานรูปแบบต่าง ๆ ของคำบ่งชี้ตามหลังจากมุมมองการเรียงคำ การเชื่อมคำ วากยสัมพันธ์ และวัจนปฏิบัติศาสตร์ ซึ่งผลการวิเคราะห์ถูกนำมาบูรณาการกับข้อมูลเชิงภูมิศาสตร์ แสดงให้เห็นถึงความเกี่ยวพันระหว่างพิภพบริเวณภาษากับความคล้ายคลึงทางโครงสร้างภาษาระหว่างภาษาถิ่นต่าง ๆ

ผลลัพธ์ที่ได้จากงานวิจัยนี้แสดงเหตุผลสนับสนุนข้อเสนอดังต่อไปนี้ ประการแรก คำบ่งชี้ตามหลัง มีลักษณะการใช้งานเป็นคำไวยากรณ์เพื่อจัดเรียงโครงสร้างข้อความ ทั้งนี้ ลักษณะทางโครงสร้างข้อความ ในฐานะดัชนีกำหนดหัวเรื่องและจุดเน้นได้พัฒนาการใช้งานเพิ่มเติมเพื่อใช้กำกับตัวอ้างอิงชี้เฉพาะ รวมไปถึงการใช้แสดงการประเมินค่าข้อความตามวิจารณ์ญาณของผู้พูด โดยลักษณะดังกล่าวพบมากในภาษารัสเซียถิ่นเหนือและภาษาฟินิกผิงตะวันออกในบริเวณติดต่อกัน ประการที่สอง พัฒนาการของระบบ คำบ่งชี้จากภาษาฟินิกดั้งเดิมสู่ภาษาฟินิกปัจจุบันได้รับอิทธิพลจากสัมผัสภาษาระหว่างภาษาฟินิก สาขาต่าง ๆ ที่เกิดขึ้นในภายหลัง และพัฒนาไปในลักษณะร่วมที่พบในบริเวณใกล้เคียง จากลักษณะร่วมนี้ ระบบคำบ่งชี้ภาษากลุ่มฟินิกสามารถจำแนกได้เป็นสี่กลุ่ม คือ ๑. กลุ่มตะวันตก (ลีโวเนีย เอสโตเนียใต้ และ เอสโตเนียเหนือ) ๒. กลุ่มภาคกลาง (โวท และ อิงเกรีย) ๓. กลุ่มคาเรเลีย (โอโลเนตส์คาเรเลีย และ ลูตถิ่นเหนือ) และ ๔. กลุ่มตะวันออก (ลูตถิ่นใต้ และ เวปส์) ประการที่สาม คำบ่งชี้ตามหลัง -to และ รูปผันคำอื่น ๆ ที่พบในภาษารัสเซียถิ่นเหนือมิได้เกิดจากอิทธิพลภาษาฟินิกในฐานะภาษาพื้นเดิม หากแต่ในฐานะภาษาเท่าเทียม กล่าวคือ รูปคำไม่ผัน -to ที่สืบทอดมาจากภาษารัสเซียยุคกลางถิ่นภาคกลาง ได้พัฒนาคุณสมบัติการผันรูปคำได้ รวมไปถึงการใช้ร่วมกับตัวอ้างอิงชี้เฉพาะ ซึ่งทั้งหมดนี้ล้วนแล้วแต่เป็น พัฒนาการแบบสนับสนุนเกื้อกูลกันร่วมกับคำบ่งชี้ se ในภาษาเวปส์ การใช้งานเชิงวัจนปฏิบัติศาสตร์ ในลักษณะดังกล่าวยังได้แพร่หลายไปยังภาษาลูตและภาษาโอโลเนตส์คาเรเลียอีกด้วยในเวลาต่อมา

คำสำคัญ: คำบ่งชี้, ภาษาฟินิก, ภาษารัสเซียถิ่นเหนือ, อิทธิพลภาษาพื้นเดิม, สัมผัสภาษา, แบบลักษณ์ภาษาภูมิภาค, ความชี้เฉพาะ, โครงสร้างข้อความ

ACKNOWLEDGEMENTS

This piece of work is a mutual achievement of enumerable participating individuals who have contributed their physical, intellectual, mental, and financial support over the course of my doctoral studies, as well as in earlier phases of my life.

First and foremost, I own my deep gratitude to my two supervisors who have been teachers, mentors, colleagues, and good friends of mine, not only during my four years as a doctoral student but ever since I began my studies at the University of Helsinki in 2010. My first supervisor, Riho Grünthal, was a great support in both intellectual as well as bureaucratic parts of this project, by contributing his long experience in the field of Uralic linguistics and in the academic world. My second supervisor, Seppo Kittilä, has constantly offered good insights through the lenses of general linguistics, and has been a good reader who shared our adventure in exploring new aspects of the topic under focus of this book. Moreover, my thankfulness also goes to Max Wahlström for being a good colleague and friend, who shares similar topics of interest with me, particularly the topic of this dissertation, about which he has provided many useful thoughts, comments, and suggestions from a Slavic perspective, as my unnamed third supervisor.

The current work is a continuation of what I already began in my master's thesis. I also thank my previous supervisors, Janne Saarikivi and Jouko Lindstedt, for guiding me to this interesting research question, situated in between the scholarly traditions of Uralic and Slavic linguistics, and for helping me build a strong framework, the basis of which the present work developed from. They were also a great inspiration for me to keep exploring the historical-comparative aspect of these two language groups, which have been my major areas of studies since my bachelor's studies. Even after my master's project, they remained reliable experts to whom I was always able to turn to when a piece of advice was needed.

I also would like to express my respect and indebtedness towards the two external examiners of this doctoral dissertation, who sacrificed their valuable time to evaluate my work. Renate Pajusalu, who has a good understanding of the context of Uralic linguistics, particularly Finnic languages, has provided many insightful observations on issues, which I might have had overlooked or not paid enough attention to initially. Ekkehard König is a respected veteran in the European linguistics society, and has given many useful remarks, on the basis of which I was able to make the contents of this dissertation more accessible and readable to linguists outside the fields of Uralic and Slavic linguistics. Their comments, criticisms, and suggestions were an essential part, which helped me enormously in finalising this whole project into a complete book. As both are known as sympathetic persons, they also have often given me positive feedback, and encouraged me to keep on working on this topic, whenever we met in various conferences abroad.

Over the course of my dissertation project, I have benefited from many fruitful discussions with a number of brilliant scholars around the world, which has led me to the current results in the present work. Concerning the topic of the postposed demonstratives in Finnic and North Russian, I have exchanged many good ideas and remarks with Gerson Klumpp, Maria Vilkkuna, Ruprecht von Waldenfels, Motoki

Nomachi, Bernhard Wälchli, Maria Khachaturyan, Laura Becker, and the late Marja Leinonen, all of whom have been extremely supportive, and enthusiastically anticipating my research results. It was unfortunate that Marja Leinonen did not live long enough to see my final results, so I would like to dedicate this book as a memorial to her contribution to this research topic, as well as her interest towards my research.

I have learnt many essential methods which I subsequently applied in my research. It was a great opportunity to familiarise myself with areal linguistics under the guidance of the great linguists, Johanna Nichols, Ekkehard Wolff, and Eitan Grossman, during their frequent visits to Helsinki. I was lucky to have met Foong Ha Yap, who introduced me to the discourse-pragmatic domain of research, which crucially helped me solve some of the last puzzles of my research. It was also my great fortune to have a smart collaborator and a good friend, Pui Yiu Szeto, who showed me how interesting and effective quantitative methods can be for linguists. I also thank Kaius Sinnemäki for his consultation on good practices on the application of quantitative and statistical methods in language studies.

In my background, I became the linguist I am today thanks in part to the Helsinki linguistics family which has provided me training encompassing a broad expertise of the languages of the world. All teachers and friends from different subject units have contributed their bits to my mastery of languages across Eurasia and beyond. My home unit of Finno-Ugrian Languages and Cultures has always provided and maintained an inspiring learning and working atmosphere, thanks to the personnel: Riho Grünthal, Ulla-Maija Forsberg, Tapani Salminen, Petri Kallio, Janne Saarikivi, Rigina Ajanki, Florian Siegl, Arja Hamari, Svetlana Edygarova, Denis Kuzmin, Susanna Virtanen, Ilona Kivinen, Merja Salo, Santteri Junttila, Lotta Jalava, Magdolna Kovacs, Ildikó Vecsernyés, the late Auli Oksanen, and all of our Estonian visiting professors and lecturers from Tartu during 2010–2020; my doctoral fellows in the seminar of Finno-Ugrian Studies: Sampsa Holopainen, Niko Partanen, Timo Rantakaulio, Jaakko Raunamaa, Juha Kuokkala, Niklas Metsäranta, Alexander Pustyakov, Heidi Niva, Mari Saraheimo, Ulriikka Puura, Sofia Björklöf, and Santra Jantunen; and especially my classmates who have shared a joy of learning together since our first day at the university, the class of 2010: Markus Juutinen, Katri Hiovain, Vuokko Hangaslahti; and other fellow students of Finno-Ugrian studies: Jani Korhonen, Alexandra Kellner, Kimberly Mäkkäräinen, Jeongdo Kim, Zoli Görög, and Ryo Umeda.

I also feel very thankful to the unit of Slavic Languages and Cultures, where I studied my minor modules since my second year at the university. I learned many fascinating things about the world of Slavic languages and people among such an inspiring circle of great people: Jouko Lindstedt, Max Wahlström, Johanna Virkkula, Jouni Vaahtera, Päivi Saurio, Helena Lehečková, Miroslav Lehečka, Željko Jozić, and Dušica Božović. No lesser important are the units of Altaic Studies and Asian Studies, where I was able to remarkably broaden my cross-Eurasian perspective. I am glad to be part of a long scholarly tradition at the Helsinki School, following the steps of many great linguists in the past, and my teachers in the present day: Juha Janhunen, the late Volker Rybatzki, Okan Daher, Stefan Kuzay, Andrew Logie, and Uluhan Özalan. Last but not least, people in the unit of General Linguistics have been a group of amicable teachers and friends: Seppo Kittilä, Ekaterina Gruzdeva, Matti Miestamo, Kaius

Sinnemäki, Jack Rueter, Don Killian, Sonja Dahlgren, Ksenia Shagal, Erika Sandman, Sami Honkasalo, Olli Silvennoinen, Nicklas Bahrt, Stephan Schulz, Antti Laine, Jaakko Helke, Jouni Harjumäki, Heini Arjava, Alexandre Ryzhkov-Shukumine, Andrei Dumitrescu, Richard Kerbs, and many more. They all have provided me with the best environment possible for me to become a competent linguist.

Apart from my Helsinki linguistics family, I also thank all my language teachers around the world, initially my elementary and secondary schools in Bangkok, and later those who taught me in various summer/winter schools: Indo-European linguistics in Leiden, the Slovene language in Ljubljana, Cantonese in Hong Kong, Old Javanese in Heidelberg, and several Southeast Asian languages in Bangkok. In addition, my big thanks go to the native teachers of Belarusian, Mongolic, Turkic, and Tai-Kadai languages, who warmly welcomed me to their courses during my Erasmus exchange semester at INALCO, Paris, in the autumn semester 2015.

The whole doctoral dissertation project would not be possible without financial support from the Doctoral Programme in Language Studies at the University of Helsinki (HELSLANG), which has funded my dissertation project during the four years of my doctoral studies. Particular thanks go to my employer unit, the Department of Finnish, Finno-Ugrian and Scandinavian Studies, which has provided me with working facilities. In the bigger picture, the Faculty of Arts also generously supported my attendance in numerous international conferences, in which I learnt many essential things and good practices in the academic world, and, more importantly, got to meet many good people and discovered my future collaborators who share similar interests.

This book also would not have been complete without the hard work of my good colleague and friend, Don Killian, a great linguist who has shared many interesting thoughts about the world's languages, and, most importantly, contributed his time in carefully checking my writing and making this dissertation a readable English book. Needless to say, any remaining shortcomings are my own responsibility.

My studies in Helsinki would not have happened without help from sympathetic Finnish people during my upper secondary school's days. Saara-Liisa Hornela and her family hosted me at their cosy place in Pudasjärvi. Eeva Harju, the then principal of Pudasjärven lukio, saw my potential, and kindly offered me an opportunity to continue my upper secondary school after my AFS exchange year in 2008. The support from them, as well as my class teacher Pia Virranmäki, the guidance teacher Markku Ollila, and other subject teachers at Pudasjärven lukio was a great motivation for me to take on a challenge and complete the Finnish Matriculation Examination in 2010. This eventually enabled me to apply for a Finnish university, ending up at the University of Helsinki studying and writing this book.

Last but not least, words are not enough to say how grateful I am to my parents, Choomprin and Choowit Yurayong, and the entire family back in Bangkok, all who have always been supportive in whatever step I have taken in my life. Without their care and love, I would have never arrived to where I am as a person today.

ชิงดวง (แจ้ง) ยุระยง

Chingduang (Jæng) Yurayong

Helsinki, 20.10.2020

CONTENTS

Abstract	I
Tiivistelmä.....	II
Sammandrag.....	III
Резюме	IV
บทคัดย่อ	V
Acknowledgements	VI
Contents.....	IX
Abbreviations	XIII
1 INTRODUCTION	1
1.1 Research questions, context and goals.....	1
1.2 Data and methods.....	4
1.3 Structure of the work	6
2 FINNIC AND SLAVIC LANGUAGES AND PEOPLES OF NORTHWEST RUSSIA.....	8
2.1 Historical background of Northwest Russia	9
2.1.1 Ethnohistory	9
2.1.2 Sociopolitical situation.....	12
2.1.3 Literacy and language sociology.....	14
2.2 Taxonomy of Finnic and Slavic languages	17
2.2.1 Finnic as a western Uralic sub-branch	17
2.2.2 North Russian in a historical-comparative Slavic context	20
2.3 Finnic-Slavic interactions in Northwest Russia.....	25
2.3.1 Early Finnic-Slavic contacts as adstratum	25
2.3.2 Recent Finnic-Slavic contacts as substratum and superstratum.....	29
2.4 History of demonstratives in Finnic and Slavic languages	31
2.4.1 Demonstratives in Finnic languages	31
2.4.2 Demonstratives in Slavic languages.....	34

3	METHODOLOGICAL TRIANGULATION: LANGUAGE CHANGE, CONTACT AND AREA.....	38
3.1	Historical linguistics	38
3.1.1	Reconstruction as a method to study language change.....	38
3.1.2	Grammaticalisation as a tool to reconstruct morphosyntactic change.....	40
3.2	Contact linguistics	43
3.2.1	Stratification of languages as traces of language contacts.....	43
3.2.2	Contact-induced language change in grammatical systems	46
3.2.3	Identification of substrate language features	48
3.3	Areal linguistics.....	49
3.3.1	Linguistic areas as venues of large-scale language contact.....	49
3.3.2	Linguistic similarities due to areal convergence and diffusion	52
4	FUNCTIONAL DIMENSIONS OF DEMONSTRATIVES.....	55
4.1	Definitions and types of demonstratives	55
4.2	Grammaticalisation of demonstratives to definite articles	59
4.3	Demonstratives as a referential device	62
4.3.1	Deixis	62
4.3.2	Anaphora.....	66
4.3.3	Identifiability as a property of definite articles.....	68
4.4	Demonstratives in information structure	71
4.4.1	Definitions and behaviours of topic, focus, and contrast.....	72
4.4.2	Types of topics.....	76
4.4.3	Types of focus.....	78
4.4.4	Information-structural strategies with demonstratives	80
4.5	Demonstratives in evaluation	82
4.5.1	Intersubjectivity, stance and evaluation.....	82
4.5.2	Demonstratives as an indexical of stancetaking and evaluation....	85

5	MARKERS OF DEFINITENESS AND INFORMATION STRUCTURE IN URALIC AND SLAVIC LANGUAGES	90
5.1	Postposed demonstratives in Uralic languages	91
5.1.1	Demonstratives as markers of definiteness and information structure in Finnic languages	91
5.1.2	Postposed demonstratives in Saami languages	98
5.1.3	Definite nominal declension in Mordvinic languages.....	100
5.2	Demonstratives as markers of definiteness and information structure in Slavic languages	102
5.2.1	The development of adnominal demonstratives in Slavic languages..	102
5.2.2	Postposed demonstratives in Russian.....	105
5.3	Non-possessive uses of possessive suffixes in Uralic languages.....	111
5.3.1	Referential uses	112
5.3.2	Non-referential uses	115
5.4	Conclusions of the literature review	120
6	MICRO-AREAL INVESTIGATION OF FINNIC AND NORTH RUSSIAN.....	122
6.1	Constituent order and position	125
6.1.1	Finnic <i>*tämä/nämä(t)</i>	128
6.1.2	Finnic <i>*too/noo(t)</i>	130
6.1.3	Finnic compound demonstratives.....	131
6.1.4	Finnic <i>*se/ne(t)</i>	132
6.1.5	North Russian <i>ètot</i> and <i>tot</i>	134
6.1.6	Areal signals of constituent orders and positions.....	135
6.2	Host attachment and parts of speech.....	138
6.3	Syntactic functions.....	144
6.3.1	Clause-second versus other clausal positions.....	146
6.3.2	Clausal position effects in subject and direct object marking.....	148
6.3.3	Distribution of syntactic functions among preposed and postposed demonstratives	150

6.4	Pragmatic functions	152
6.4.1	Referential uses	153
6.4.2	Information-structural uses	163
6.4.3	Evaluative uses	173
6.5	Overview on the Old Novgorod data.....	180
6.6	Conclusions of the data analysis.....	186
7	DISCUSSION.....	188
7.1	Functional characteristics of postposed demonstratives.....	188
7.1.1	Parallels between demonstratives and definite articles.....	188
7.1.2	Demonstratives as topic and focus markers.....	192
7.1.3	Discourse-pragmatic uses of demonstratives.....	194
7.1.4	Areal-typological consideration on a grammatical category of postposed demonstratives.....	195
7.2	History of Finnic demonstratives in contact with Slavic.....	200
7.2.1	Grammatical changes in the Finnic demonstratives in comparison to North Russian	200
7.2.2	Classification and diversification of Finnic languages in the light of demonstratives.....	204
7.3	Do postposed demonstratives result from a Finnic substratum in North Russian dialects?	211
7.3.1	Evaluation of the Finnic substrate hypothesis	211
7.3.2	Understanding the emergence of postposed demonstratives in an areal context	213
7.4	Summary of the discussion.....	217
8	CONCLUSIONS	219
	References.....	223

ABBREVIATIONS

Languages

BCMS	Bosnian/Croatian/Montenegrin/Serbian
PIE	Proto-Indo-European

Corpora and databases

EDC	Estonian Dialect Corpus
RNC	Russian National Corpus
WALS	The World Atlas of Language Structures

Linguistic terms

1	1st person	ILL	illative
2	2nd person	IMP	imperative
3	3rd person	IMPF	imperfect(ive)
ABL	ablative	INDEF	indefinite
ACC	accusative	INES	inessive
ADES	adessive	INF	infinitive
ADJ	adjective	INSTR	instrumental
ADNOM	adnominal	INTJ	interjection
ALL	allative	LOC	locative
ALN	alienable	MASC	masculine
ANA	anaphoric	NARR	narrative
ANT	antecedent	NEG	negative
AOR	aorist	NEU	neutral
BEN	benefactive	OBL	oblique
CFP	clause-final particle	PASS	passive
CLF	classifier	PL	plural
CNG	connegative	POSS	possessive
COM	comitative	PRS	present tense
COND	conditional	PRF	perfect(ive)
CONJ	conjunction	PROL	prolative
COP	copula	PST	past tense
CVB	converb	PTCL	particle
DAT	dative	PTCP	participle
DEF	definite	PTV	partitive
DEM	demonstrative	P _x	possessive suffix
DEM.DIST	distal demonstrative	Q	question tag
DEM.MED	medial demonstrative	REFL	reflexive
DEM.PROX	proximal demonstrative	REL	relativiser
DU	dual	RESULT	resultative
ELAT	elative	SG	singular
ESS	essive	SPKR.ASYM	speaker-asymmetry
EVT	evidential	SPKR.SYM	speaker-symmetry
FEM	feminine	TERM	terminative
FOC	focus	TOP	topic
FREQ	frequentative	TRANSL	translative
GEN	genitive	VBLZ	verbaliser

1 INTRODUCTION

The present study examines the uses and development of “postposed demonstratives”, a term which stands for demonstratives that are used in a position following the head word of a phrase, e.g., Khmer *tʰŋai neh*, Malay *hari ini*, Thai *wan³³ ni⁵⁵* ‘today’ (literally ‘day/sun **this**'). This phenomenon has been previously reported from languages in the area of Northwest Russia such as in eastern Finnic (Larjavaara 1986), especially Veps (Kettunen 1943) and Lude (Pahomov 2011), as well as North Russian dialects (Kuz'mina & Nemčenko 1962; Trubinskij 1970; Leinonen 1998), for instance, Veps *leib se* and North Russian *hleb-ot* literally ‘bread **that**’¹. This constituent order pattern is atypical because it does not follow the canonical head-final syntax in Finnic and Slavic languages, making this phenomenon attractive for a number of scholars in the past as well as in the present.

1.1 RESEARCH QUESTIONS, CONTEXT AND GOALS

In earlier studies, the use of postposed demonstratives has often been compared with definite articles, for instance, in Scandinavian and Balkan languages (Kettunen 1943; Miklosich 1883; Kiparsky 1967; Vahros 1951). However, more recent studies (Kuz'mina & Nemčenko 1962; Trubinskij 1970; Leinonen 1998) have shifted the interpretation from the referential domain of marking definiteness towards the non-referential domain involving information-structural and discourse-pragmatic functions. In fact, the remarkable advances in theoretical frameworks of information structure and interactional linguistics during recent decades have shed light on actual uses of postposed demonstratives, which have extended to such contexts of use that are typical of topic, focus and discourse markers. For instance, such multifunctionality of postposed demonstratives is illustrated in the double use of demonstratives as in (1) and (2).

Central Veps

- (1) *nügüt silaaž om jo torvut se se?*
now 2SG.ADES be.3SG already **phonograph** DEM DEM
‘Have you already got the phonograph now?’ (Kettunen 1943: 402)

¹ Due to different conventions for transcription in Uralic and Slavic linguistics, postposed demonstratives will be written as separate words in Finnic examples (*leib se*), but being separated from a host word with hyphen in Russian examples (*hleb-ot*) throughout the whole work.

North Russian (Arkhangelsk)

- (2) *a eščě svěkor-ot-to eščě ženilsja na drugoj.*
 but still father_in_law-DEM-DEM still marry.PST.MASC on other.LOC
 ‘But still, the father-in-law got married again.’ (Russian National Corpus = RNC)

In both examples, the first postnominal demonstratives in *torvut se* and *svěkor-ot* add a definite reading to ‘phonograph’ and ‘father-in-law.’ Meanwhile, the second demonstratives have other non-referential functions such as marking question in (1) and clause topic in (2), or can generally be regarded as devices to gain the addressee’s attention. These examples show that the uses of postposed demonstratives in these languages involve more than just one functional dimension.

To participate in the ongoing discussion, the first goal of the current study is to describe the uses of postposed demonstratives in Finnic and North Russian by applying several perspectives such as referentiality, information structure, and evaluation (see the further description of methods in Section 1.2). This approach will provide understanding of the functions of postposed demonstratives from multiple areas of grammar, and not only from a single perspective of definiteness or topic marking, which was the case in a number of previous studies.

Apart from synchrony, the current study also investigates the emergence and development of postposed demonstratives in Finnic and North Russian. Especially for Finnic, the understanding of how demonstrative pronouns have developed into postposed markers involve various changes in the demonstrative system, as previously discussed for western Finnic in Pajusalu (1999) and eastern Finnic varieties in Larjavaara (1986). By tracing changes that have taken place from the Proto-Finnic to modern Finnic languages, the investigation can also provide useful information, which can be used for improving the subgrouping of Finnic languages that has been primarily based on evidence from historical phonology and morphology (Sammallahti 1977; Viitso 1985, 2000; Koponen 1991; Kallio 2014). Additionally, demonstrative systems can reveal a number of contacts between sub-branches and areal features, which are responsible for the formation of present-day “areal genetic units” (a term introduced by Helimski 2003) within the Finnic branch of the Uralic language family.

As for North Russian, an issue concerning the source of postposed demonstratives has been controversially discussed by scholars since the mid-20th century. On the one hand, scholars familiar with Uralic languages often consider postposed demonstratives as a Uralic substrate influence in North Russian dialects (Veenker 1967: 88–90; Kiparsky 1969: 25–26). On the other hand, this language feature is also considered by some scholars of Uralic languages as a Russian superstrate influence in heavily Russified eastern Finnic languages such as Veps (Larjavaara 1986: 321–323), while other scholars of Slavic languages have argued that it is ultimately a Slavic archaic feature retaining in North Russian dialects (Vahros 1951; Vaillant 1977: 265–266). Towards the end of the 20th century, recent studies have suggested that it is the product of a bidirectional contact-induced change (Leinonen 1998; Kasatkina 2007, 2008), which is ultimately due to an areal tendency in the linguistic area of the Russian North (Stadnik-Holzer 2006). In any case, few previous studies have taken an empirical

approach, and the language comparison often concerns a pair of a Uralic language and a North Russian dialect. Given this situation, the current study examines this matter by taking into consideration all Finnic varieties that have been more closely or remotely involved in contacts with North Russian dialects during the 2nd millennium (see description of the data in Section 1.2). This approach will shed light on the question whether it was a contact with the Finnic-speaking population that gave rise to the multifunctional postposed demonstratives in North Russian dialects.

Given the research contexts described above, the main goal of the current study is to answer the following questions:

1. How do postposed demonstratives function as grammatical markers?
2. What does the development of demonstratives tell us about the history of Finnic and Slavic languages?
3. Do postposed demonstratives result from a Finnic substratum in North Russian dialects?

In the present study, I will argue that postposed demonstratives in Northwest Russia are not definite articles, but they can be used to express various types of identifiability, which are characteristic of definite articles, as discussed in a number of previous studies of preposed demonstratives in Finnic such as in Estonian (Pajusalu 1997a), Finnish (A. Hakulinen 1985; Laury 1991, 1996, 1997; Vilkuna 1992; Dasinger 1995; Juvonen 2000), and in Votic (Agranat 2015), as well as Slavic languages such as in Upper Sorbian (Berger 1999; Breu 2002; Marti 2012), Polish (Bacz 1991; Mendoza 2005; Bartnik 2015), Czech (Cummins 1998), Slovene (Marušič & Žaucer 2006), Bosnian/Croatian/ Montenegrin/Serbian (Trenkić 2004), and Belarusian (Laryëna 2016). I will also argue that the use of postposed demonstratives does not show a strong tendency of becoming grammaticalised definite articles, which corresponds to a similar idea earlier proposed for spoken Estonian (Pajusalu 1997a: 172–173) and spoken Finnish (Juvonen 2000: 196–198), in contrast to the idea that preposed demonstratives in Finnish will eventually become definite articles (Karlsson 1975: 62), or have already become definite articles (Laury 1996: 178, 1997: 264).

From a historical-comparative perspective, I will show that unlike phonological and morphological evidence used for the subgrouping of Finnic languages (Sammallahti 1977; Viitso 1985, 2000, 2008; Koponen 1991; Kallio 2014), demonstratives are not genealogically stable grammatical elements. Instead, later contacts between sub-branches that took place after the diversification of Proto-Finnic play an important role in the development of the demonstrative system in individual languages and their subvarieties.

1.2 DATA AND METHODS

The language data used in the present study geographically is taken from the areas between Courland in North Latvia and Arkhangelsk Oblast in Northwest Russia, including thirteen Finnic and North Russian varieties from Vologda and Arkhangelsk Oblasts, as shown in Figure 1.

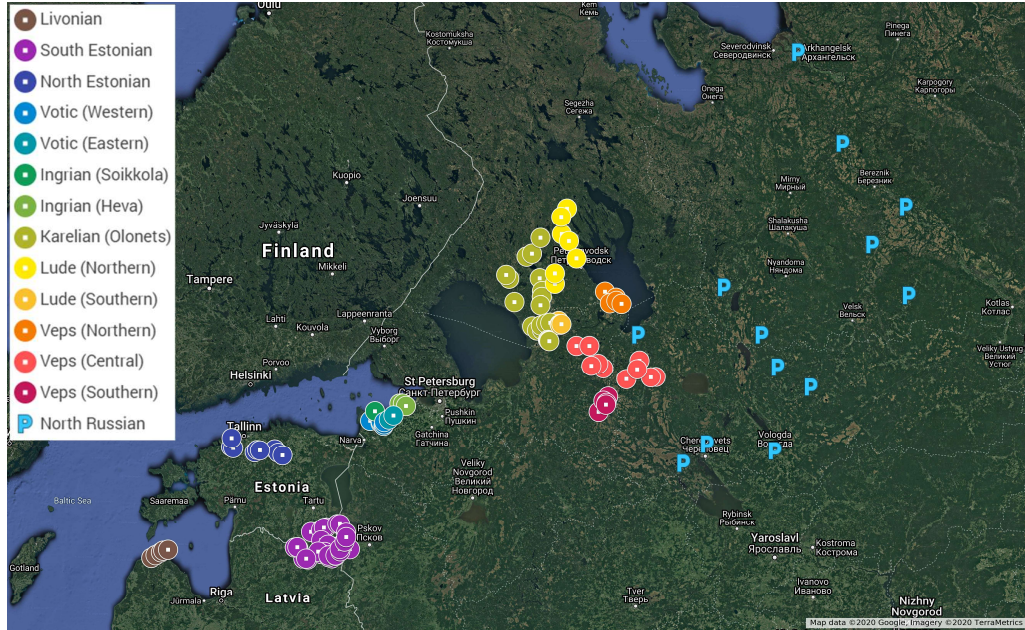


Figure 1 Geographical distribution of the Finnic and North Russian varieties under investigation

The reason for this language selection is the primary focus on geographical distribution and continuation of a historical Finnic-Slavic contact zone, to which Slavs arrived from the south. Considering that the speaking areas of modern Finnish and Karelian proper are results of later migration from the Finnic core zone of present-day eastern Estonia (see Saarikivi & Grünthal 2005; Janhunen 2014a; Frog & Saarikivi 2015), these Finnic varieties in the north do not belong to the east-west continuum illustrated in Map 1 and, therefore, are left out from the scope of the present study. Despite a number of Russian loanwords also found in eastern and southeastern Finnish varieties (see Ruoppila 1986), these borrowings date back further in the history and, thus, are also common to other Finnic branches. Moreover, since the Soviet era, the Russian language has been present in Finland mostly in its standardised form, so (North) Russian vernaculars no longer played an important role in the Finnic-Slavic contact that has taken place behind the Finland-Russia border during the most recent century.

The data is comprised of online language corpora and written sources, from which all the occurrences of demonstratives **tämä*, **se*, and **too* in Finnic, and *eto* and *to* in North Russian dialects as pronouns and proadjectives (excluding adjectival and adverbial demonstratives such as Estonian *selline* ‘such’ and *seal* ‘there’ or Russian

takoj and *tam* id.) are collected by the search engine tool within and thorough reading of the texts, respectively, and then classified according to their constituent orders and other morphosyntactic and semantic-pragmatic behaviours. The reason for choosing these particular corpora is an emphasis on non-standard varieties and spoken form of these languages. The data sampling concentrates on the varieties that by their location and ethnic history represent each dialectal group best. However, the corpora for each variety are not the most comparable by size and genre ideally, but this is what the currently available resources provide. Table 1 summarises the information on data of each Finnic and North Russian variety by source, size and genre.

Language	Variety	Source	Size	Genre
Livonian		Mägiste 2006	ca 25,000 words	Narrative
South Estonian		Estonian Dialect Corpus (EDC)	168,587 words	Narrative, dialogue
North Estonian			38,427 words	
Votic	Western	Kettunen & Posti 1932; The Archives of Estonian Dialects and Kindred Languages	ca 10,000 words	Narrative
	Eastern		ca 10,000 words	
Ingrian	Soikkola	Laanest 1966	ca 5,000 words	Narrative, dialogue
	Heva		ca 12,000 words	
Karelian	Olonets	Makarov & Rjagoev 1969	ca 35,000 words	Narrative
Lude	Northern	Ojansuu et al. 1934	ca 40,000 words	Narrative
	Southern	Pahomov 2011	ca 30,000 words	Narrative, dialogue
Veps	Northern	Kettunen & Siro 1935; <i>Korpus vepsskogo jazyka</i>	ca. 70,000 words	Narrative
	Central			
	Southern			
North Russian	Vologda	Kasatkina 1991; Russian National Corpus (RNC)	ca 300,000 words	Narrative
	Arkhangelsk			

Table 1. *Language corpora used in the present study*

As a remark, the South Estonian data is represented by Võro and Seto varieties whereas the representative varieties of North Estonian are spoken in Ambla, Keila and Kose. Throughout the entire work, the order of varieties presented both in the tables and charts is based on a geographical continuum from the southwest to the northeast as organised in Table 1. Beyond modern languages, data of an attested Slavic variety earlier spoken in the Russian North between the 11th and 15th centuries, Novgorod birch bark documents (Zaliznjak 2004), also falls under the investigation and discussion of the current study to provide a diachronic perspective.

The main approaches to the research questions of the current study are historical-comparative linguistics, study of language contacts and areal typology. This means that the aforementioned linguistic varieties in contact, both related (Finnic languages)

and unrelated (Finnic vs. Slavic), are placed into comparison both synchronically and diachronically. This comparison is based on identification of similarities and differences in the use of a grammatical element across the contact zone under investigation. In this case, it concerns the behaviour of postposed demonstratives in Finnic and North Russian speaking areas in Northwest Russia. This comparative analysis also places the micro-areal context of Finnic and North Russian contact zone onto a larger-scale theoretical framework of linguistic typology. Supplementing these approaches, an additional typological framework touches on several domains that serve as tools for analysing the uses of postposed demonstratives in Finnic and North Russian. Among others, concerned areas of linguistics are word order typology, word class, deixis, definiteness and referentiality, information structure and evaluation.

The data analysis is a combination of qualitative and quantitative approaches, which yields better understanding, concerning behaviour of postposed demonstratives and their geographical distribution on this micro-areal scale. The areas which can be quantified are morphology and syntax as well as referentiality and information structure, while the analysis of evaluative uses is limited to a qualitative level. The analysis also integrates achieved results with geographical data, which shows the correlation between the speaking areas and a degree of similarities between varieties.

1.3 STRUCTURE OF THE WORK

The present study is divided into eight chapters. Following the introductory remarks in **Chapter 1**, **Chapter 2** provides background information on the history and language sociology of Northwest Russia during the past two millennia as well as on the position of Finnic and North Russian in the Uralic and Slavic taxonomies, respectively, while also introducing the Finnic and Slavic demonstrative systems, which share some historical connections. **Chapter 3** introduces three primary fields and methods of linguistics to investigate development of language features from the perspectives of internal reconstruction, language contacts and areal tendency. **Chapter 4** discusses characteristics and properties of demonstratives from a typological perspective involving morphosyntax and semantics, grammaticalisation, referentiality, information structure and evaluation. Based on this understanding of demonstratives on the theoretical level, **Chapter 5** brings information from Finnic and North Russian as well as their cognate Uralic and Slavic languages into the discussion of how the typological properties and functions discussed in Chapter 4 have been understood and are concretely encoded in these languages. **Chapter 6** applies various criteria developed on the basis of previous theoretical (Chapter 4) and empirical studies (Chapter 5) to the analysis of the Finnic and North Russian as well as Old Novgorod data from multiple perspectives: constituent order and position, host attachment and parts of speech, syntactic functions and pragmatic functions. **Chapter 7** uses results from the data analysis in Chapter 6 to answer the three research questions addressed in Section 1.1. **Chapter 8** summarises the findings in the current study and suggests particular issues which remain open for further investigation in the future.

Throughout the entire work, unless the sources are otherwise indicated, language examples are own elicitations. The transliteration of Russian language examples in Cyrillic scripts principally follows the ISO/R 9 standard (with one deviation: Cyrillic *x* = Latin *h* pro *ch*). As for the Proto-Slavic reconstruction and ancient Slavic languages, the reduced vowels *ǫ* and *ǐ* are transcribed in their Cyrillic forms *ѳ* and *ѣ*, following the convention in Slavic linguistics, which emphasises their origin as vowels. Other noteworthy symbols are *ě*, standing for the Cyrillic *ѣ* (yat), the reflex of the Early-Proto-Slavic diphthong **ai* and later the Proto-Slavic long mid front vowel **ē*; and the nasal vowels *ę* and *ǫ̃*, used for transliterating the Cyrillic *ѣ* (small yus) and *ѣ* (big yus), respectively. Meanwhile, examples from other languages are presented as published in the original sources unless mistakes are observed and corrected.

2 FINNIC AND SLAVIC LANGUAGES AND PEOPLES OF NORTHWEST RUSSIA

Northwest Russia is ethnographically and linguistically a diverse area where two major speech communities, Uralic and Slavic-speaking populations, have been cohabiting over a millennium. The present study defines Northwest Russia as an area stretching from Novgorod in the southwest to the Ural Mountains in the east, illustrated as the red zone in Figure 2.



Figure 2 Northwest Russia as defined in the current study

Given this geographical context, this chapter discusses various aspects that provide a foregrounding understanding of how people were living and what languages have been spoken in Northwest Russia. Section 2.1 presents a historical background of the area from the ethnographic, sociopolitical and sociolinguistic perspectives. Section 2.2 introduces the taxonomic classification of Finnic and Slavic languages, while Section 2.3 discusses contact scenarios evolving between Finnic and Slavic-speaking populations, which have taken place during the past two millennia. For the context of postposed demonstratives and the question of Finnic substratum, the present study primarily focuses on contact scenarios that have taken place between Finnic and Slavic in the first half of the 2nd millennium and the later contacts in the second half of the 2nd millennium primarily in the east, while the west of the Finnic continuum is only briefly discussed. To connect the regional history to the linguistic phenomenon under investigation, Section 2.4 presents the history of demonstrative systems in Finnic and Slavic languages to illustrate similarities and possible historical connections of the systems, especially between Finnic and East Slavic.

2.1 HISTORICAL BACKGROUND OF NORTHWEST RUSSIA

The language situation and contacts in Northwest Russia can be approached from several aspects, such as the ethnic background of the speakers (Section 2.1.1), the political power competed for and distributed among different political units (Section 2.1.2), as well as the language competence and repertoire of the local inhabitants (Section 2.1.3).

2.1.1 ETHNOHISTORY

In a larger horizontal dimension, Eurasia is neatly divided into a northern and southern belt in terms of the ethnolinguistic setting and this division also applies to Northwest Russia. The former covers Tundra and Taiga zones, where the majority of Uralic languages, as well as extinct Paleo-European languages have been spoken, and the latter lies along the Steppe zone where Indo-European and Turkic languages, above all, have been extensively spoken (Nichols 1992; Janhunen 2014b; Saarikivi, in press). Considering the historical settlement, there is a remarkable amount of non-Indo-European toponyms in the northern belt mentioned in early written sources by both Scandinavian and Slavic chronicles as well as by foreign explorers (see, e.g., Aikio 2004; Saarikivi 2006: 26–39; Saarikivi & Lavento 2012).

As a territory in which the demographics and ethnicity have been continuously changing during the second millennium AD, Northwest Russia exhibits numerous traces of language and culture shifts of indigenous non-Slavic tribes to Slavic (Bjørnflaten 2006). In this respect, we can talk about “Slavicisation” of the erstwhile inhabitants rather than hybridisation of the Slavic community, because an emerging multiethnic and multilingual society has inherited the practices of the Slavic language and culture, even if it may have been modified or disrupted after the interaction with erstwhile non-Slavic inhabitants (Timberlake 2013). Previous studies of toponyms have shown that the Russian North has been resided by Uralic-speaking populations for a long time prior to the arrival of Slavic-speaking population (e.g., Ahlqvist 1998, 2006; Helimski 2006; Saarikivi 2004, 2007b). Most notably, several studies have identified lexical evidence in favour of contacts of Permian with Finnic and potentially also with other western Uralic languages, all which predate the Russification of the Russian North (Bartens 2000: 16–17; Saarikivi 2006: 33–38, 2018).

Ethnolinguistic data of the present-day Northwest Russia indicates that the only zone in which Uralic-speaking communities are missing from the northern Eurasian continuum is Arkhangelsk Oblast and the eastern part of Vologda Oblast (see Figure 3). Assuming that Uralic speakers used to be indigenous residents of the Russian North, this would imply that the speakers of vanished Uralic languages in the areas have shifted their language and identity to Slavic (e.g., Ahlqvist 1998, 2006; Mullonen 2007; Saarikivi 2007b).

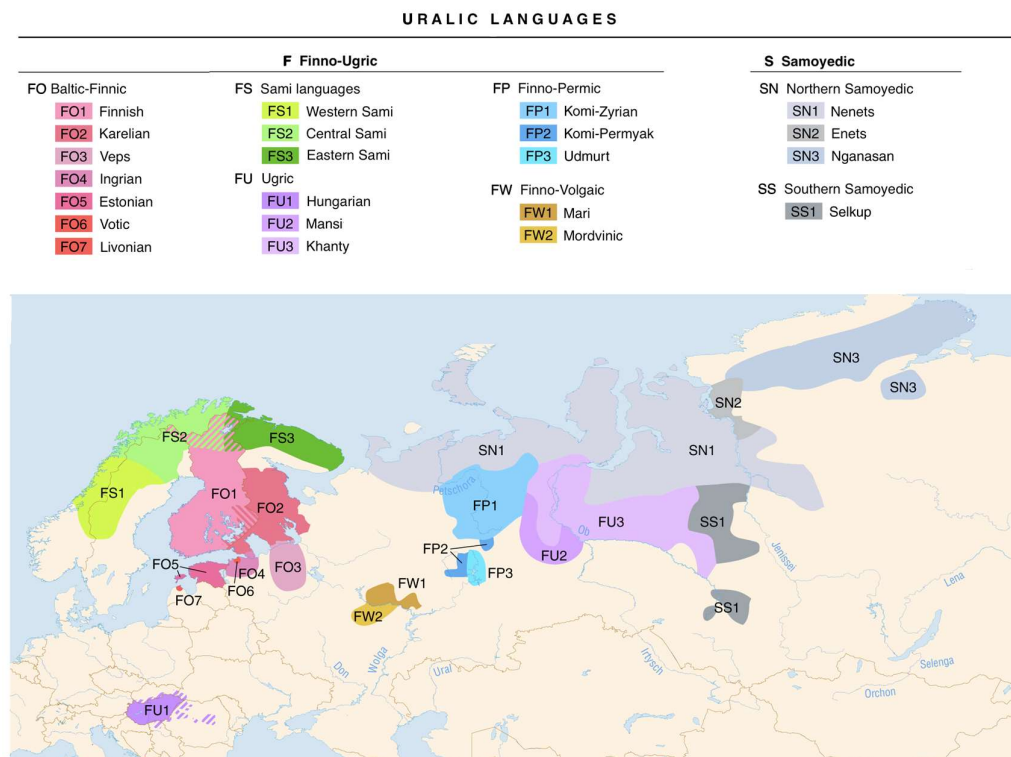


Figure 3 Linguistic map of the Uralic languages (en)²
(Wikimedia Commons: commons.wikimedia.org/w/index.php?curid=28408682)

Language shift of non-Slavic populations to Slavic, which has continued over the past millennium since the arrival of Slavs in Northwest Russia after the 5th century, is a crucial factor in the development of East Slavic languages (Lindstedt & Salmela 2020). This shift scenario is also supported by genetic data, as the Slavic-speaking population in Northwest Russia is genetically close to the Finnic-speaking neighbours (Khrunin et al. 2013; Kushniarevich et al. 2015: 3). In the same vein, the Slavic variety spoken in Northwest Russia is famous for its non-Slavic features transmitted from Uralic substrate languages. This particular topic of the Uralic substratum in North Russian dialects, therefore, is one of the essential topics in both Slavic and Uralic linguistics, as well as the studies of substratum in general.

Regarding the Uralic substratum in Russian, scholars have been focusing on many aspects from historical sources (Gippius 1997; Lind 2006) and toponyms (Mullonen 2002; Myznikov 2003; Saarikivi 2006; Rahkonen 2013; Kuzmin 2014) to language structure (Veenker 1967; Kiparsky 1969; Vostrikov 1990). In any case, the Uralic substratum in Northwest Russia is not a homogenous phenomenon in various areas because each North Russian variety illustrates distinct sets of characteristics that point to different Uralic substrate languages (Ojanen 1988; Sarhimaa 1992, 1995, 1999: 18; Myznikov 2003: 33).

² Note that the label “FO5 Estonian” in the map stands for two distinct languages: South vs. North Estonian.



Figure 4 Principalities of Kievan Rus' (1054–1132)
(Wikimedia Commons: commons.wikimedia.org/w/index.php?curid=12573422)

A description of the sociopolitical situation in Northwest Russia by Schenker (1995: 53–54) discusses the East Slavic “Primary Chronicle” (Old Rus’ *Načal’naja lětopis’*, a.k.a. “Tale of Bygone Years” *Pověst’ vremennykh lět*, or “Nestor’s Chronicle” *Nestorova lětopis’*), originally compiled in Kiev in 1113. This early historical source mentions a number of presumably Slavic-speaking tribes who

migrated from their hypothetical Slavic homeland in the headwaters of the Dnestr and Prut to the north of the Carpathians (see Barford 2001: 47–55). Several dominant groups among those tribes were Slovians (Old Rus' *Slověni*) who settled by Lake Ilmen around Novgorod, and Krivichians (Old Rus' *Kriviči*) who occupied the headwaters of Volga, Dvina and Dnieper with Smolensk functioning as their centre of political power. Compared to the West and South Slavs, the sociopolitical environment surrounding the East Slavs had fewer conflicts, smaller population density, and less stable communities. This calmer political climate was maintained, partially thanks to peaceful Baltic and Finnic neighbours to the north, while the Turkic neighbours in the east were mainly interested in trade rather than waging war. At the same time, the Vikings were living in adjacent areas in the west and their appearance among the East Slavs played an important role in the establishment of the first political unit of the East Slavs, Kievan Rus'. The only threat for the East Slavs was a great wave of migration and movement in the southern border of Kievan Rus', which did not, in any case, directly affect those tribes living in the north.

A dozen of local populations and principalities mentioned in the Primary Chronicle shows that the medieval Northwest Russia was politically not unified as a single organisation, but rather a territory where many political units, with Ancient Novgorod among others, were competing for power and domination of the Russian North, as illustrated in Figure 4. A noteworthy detail of Figure 4 is the labelling of an ethnic group “Chud” who resided in the northeastern territory of the Ancient Novgorod principality. The use of this ethnonym in Russian historical sources often makes association to Finnic tribes, with the exception of Livonian that has never been associated with this term (Grünthal 1997: 150). Its Early-Proto-Slavic etymology **tjūdji* ‘nation’ (~ Gothic *piuda*, German *deut-sch*) has received a new meaning in the Russian *Čud* ‘foreigner, stranger’ (see also Vasmer [1950–1958]1987: 4378), which has spread over the Russian North and Siberia as a term that Novgorodians, for instance, used to call ethnic groups of an unknown origin (Popov 1973: 69).

2.1.2 SOCIOPOLITICAL SITUATION

As a political centre in the northern part of Medieval Russia, Novgorod was historically significant in terms of language contacts between Finnic and Slavic-speaking populations. These early contacts have left traces as early Slavic loanwords in Finnic (discussed further in Section 2.3.1). In the studies of Finnic-Slavic contacts, it is also crucial to understand the political climate of the early-2nd millennium when the interaction between Finnic and Slavic tribes was the most early attested. The following description of the sociopolitical situation in Novgorod up to its fall and the adjacent regions is a summary from Schaeken (2012: 13–15).

The earliest archaeological evidence dates Novgorod as an early (medieval) urban centre back only to 910, while the first mention of Novgorod as a town dates back already to the oldest Russian chronicle (859/862 AD). A part of the chronicle mentions the arrival of Rurik, the founder of Rurik Dynasty that was in reign until 1598. In 947, Olga who was the wife of Novgorod's ruler, Igor, enlarged the territory of Novgorod

during the reign of Igor. In 988, the grandson of Igor, Vladimir I, succeeded the throne and named Ioakim as the first bishop of Novgorod in 989.

In the beginning of the 11th century, the son of Vladimir I, Jaroslav, established settlement in Novgorod's area for trading purpose and consequently named the place Jaroslav's Court after himself. In 1034, Pskov was also attached to the Novgorod principality. During this period, the Novgorodians wrote the "Novgorod Codex" that can be regarded as the oldest book found in Russia, predating the famous "Law Code of the Rus' (Old Rus' *Russkaja pravda*) in 1018, the oldest Slavic manuscript on parchment, "Ostromir's Gospel Book" (Old Rus' *Ostromirovo evangelie*) in 1056–1057 and "The Primary Chronicle" in 1113.

Later in the beginning of the 12th century, the first monastery was constructed in Novgorod. During this period of time, Novgorod also received merchants from Visby and Gotland, which also led to more contacts with Scandinavian-speaking populations. In 1136, the Novgorod state crowned Vsevolod Mstislavič of Pskov its ruler who turned the Novgorod principality into a republic. Towards the end of the 12th century, Novgorod governors constructed Peterhof (a.k.a. Latin *curia sancti Petri*), which functioned as a bureau for German merchants and Hansa Union's traders during the centuries that followed.

Between 1237–1242, the region of Kievan Rus' was invaded by the Mongol soldiers, but Novgorod remained outside the immediate area impacted by the invasion (see also Vlasto 1986: 303). During the same period of time, Novgorod's ruler, Alexandr Nevskij, won battles, first against the Swedes in 1240 around the area of Lake Neva, which may have also involved Finnic-speaking tribes, and then against the Germans in 1242 in Pskov. In the mid-13th century, Novgorodians started writing "The First Novgorod Chronicle" (Old Rus' *Novgorodskaja pervaja lětopisъ*) that reports the events taking place during the years 1016–1477. As of 1478, Novgorod was subsumed and attached to Moscow, which ended the period of Novgorod as an independent political unit and its history writing. Slightly later, in 1494, Ivan III also terminated the business in Hansa Union's bureau in Novgorod Peterhof.

The writing tradition of birch bark documents probably terminated at the end of 15th century, which is chronologically the latest date of the birch bark documents found on the archaeological sites in Novgorod and its neighbouring areas. In place of birch bark, the Novgorodians must have switched to paper at the latest in the 16th century. In 1570, Ivan IV of the Grand Duchy of Moscow (a.k.a. Ivan the Terrible) and his troops invaded and robbed Novgorod, driving the inhabitants out of the town. Eventually, between 1611 and 1617, Novgorod fell under the reign of the Swedish Kingdom, after which it has never gained back its status as an independent principality.

As is clear from the overview on history of Novgorod by Schaeken (2012), the Slavic Novgorod principality was not always the dominating polity in the medieval Russia. The power changed from one to another principality frequently during the medieval era. Therefore, Novgorod as a political unit experienced both rises in power as well as falling under other principalities' reign over periods of time, unsurprising given the political history. From the linguistic point of view, it is also notable that the ruling class of Northwest Russia between the late-1st and early-2nd millennium was

mostly Scandinavian-speaking, despite the principality being a multilingual society, which also comprises of a large group of Slavic-speaking and a smaller group of Finnic-speaking population (see Saarikivi 2007a). As a linguistic implication, the fall of Novgorod and its vernacular also seems to relate to the formation of the modern Russian dialect continuum, which began in the 16th century (discussed in Section 2.2.2). The sociolinguistic situation of the medieval Northwest Russia is discussed next.

2.1.3 LITERACY AND LANGUAGE SOCIOLOGY

Fortunately enough for the studies of language contact in Northwest Russia, modern-day scholars still have access to literary evidence of the Early Slavic variety, that is, the famous Novgorod birch bark documents. The Slavic variety written in these documents is relatively distinct from that of the Kievan Rus' (see Zaliznjak 2004; Nuorluoto 2006, 2007, and further discussion in Section 2.2.2). This entails that the surrounding language sociology, which also includes an early interaction with Finnic-speaking population, is also relevant for the present study and worth discussing with further details.

The majority of birch bark documents contain texts written down by scratching with pen or stylus, in the local Slavic variety and in the Cyrillic script. The age of these documents is estimated to be of the period between the 10th and 15th centuries. The discovery of this cultural treasure dates back to 26 June 1951, when a member of Russian archaeologist team, Nina Fëdorovna Akoelova, found the birch bark document №1 (Nosov et al. 1993: 49; Schaeken 2012: 84), as shown in Figure 5.

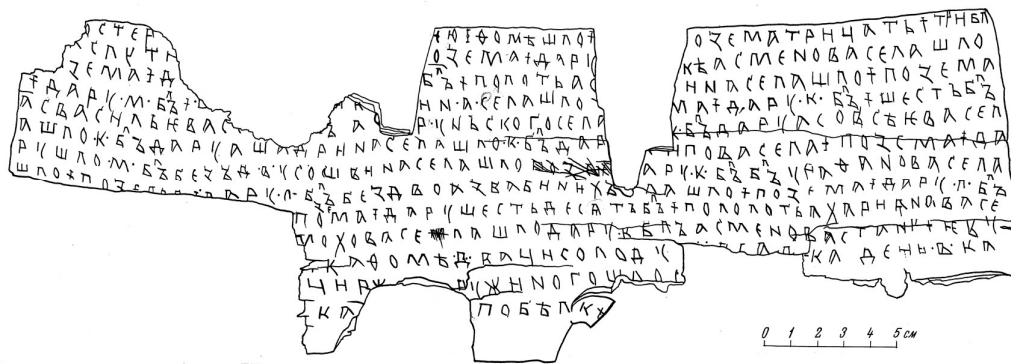


Figure 5 Novgorod birch bark document № 1 (ca 1380–1400)
(Russian Academy of Sciences: gramoty.ru/birchbark/document/show/novgorod/1/)

Thus began a new multidisciplinary field of archaeological and language studies named “berestology” (Russian *berestologija*, a term derived from *beresta* ‘birch bark’) (Schaeken 2012: 7–8).

The literacy of Ancient Novgorod inhabitants was relatively advanced and fascinating. As discussed in previous studies (Nosov et al. 1993: 49–55; Schaeken 2012: 29–30), birch bark was an affordable and easily accessible writing material for

ordinary people, in contrast to parchment, which was expensive. More importantly, the birch bark was usually about the size of half of A4 paper size, which facilitated portability and made written communication widely possible to the citizens of Novgorod and neighbouring areas. This is particularly relevant for historical linguistics, as written sources are most commonly carefully composed texts, while the Novgorod birch bark documents provide valuable and rarely accessible information on vernacular language use (cf. an issue with historical data in Labov 1994: 11, discussed in Section 3.1.1).

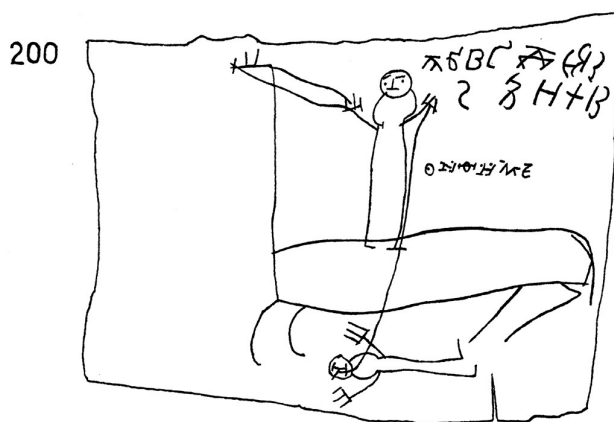




Figure 8 Novgorod birch bark document № 206 (ca 1240–1260)
(Russian Academy of Sciences:
gramoty.ru/birchbark/document/show/novgorod/206/)

Such birch bark documents are good evidence for an organised education at the community level in the medieval era (see also Nosov et al. 1993: 51).

From the sociolinguistic viewpoint, the birch bark documents illustrate both the low and high registers of a language. In other words, the use of language can be as formal as the Church Slavonic style or as informal as a “superregional” language use (Schaeken 2012: 40), which seemingly also allowed writing in other languages such as Greek, Latin or Finnic (see the Finnic document №292 in Section 2.3.1). This claim of diglossic language use is supported by the spreading area of birch bark writing tradition. For instance, Nosov et al. (1993: 49–50) identify relatively broad areas which include also territories of the present-day Ukraine, Belarus and even Moscow, as shown in Figure 9. This observation suggests that more birch bark documents are likely still lying under the muddy ground in the areas more to the south and east from Novgorod.

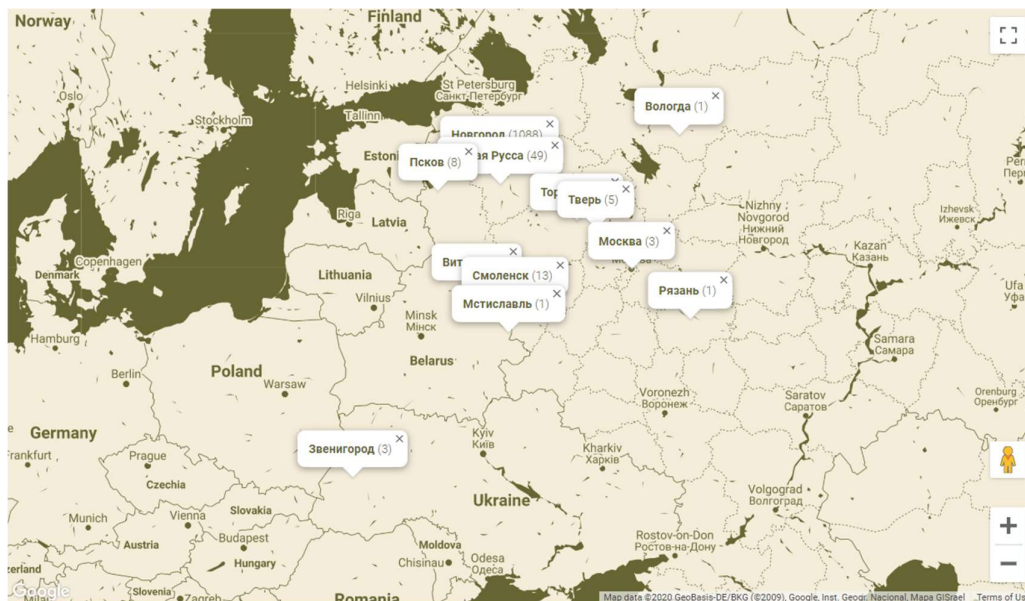


Figure 9 Locations where birch bark documents are found
(Russian Academy of Sciences: gramoty.ru/birchbark/maps/towns/)

In any case, Nosov et al. (1993: 50) also remark that in the end, birch bark is not an exotic writing material in Europe. In addition to Northwest Russia, birch bark inscriptions have also been found in Central and Western Europe, particularly in areas ranging from Germany to Sweden today. Evidence suggests that they were created by the Vikings, dating back to a century prior to the documents found in Novgorod. One significant difference between the birch bark writing tradition of the Slavs and Vikings, however, is that the Vikings primarily wrote with ink. In any case, the shift from birch bark to paper was likely taking place during the 15th century, when the production of birch bark documents also ends, according to the latest archaeological finding (discussed in Section 2.1.2).

Given the important role of Vikings in Northwest Russia, it is possible that the birch-bark writing tradition was introduced to Slavs by Scandinavian or some other Germanic-speaking peoples. Considering the etymology of the Slavic word for ‘alphabet, letter, book, document,’ the Early-Proto-Slavic reconstruction **bŭku* is proved as a borrowing from the Germanic **bōk-* id. (Vlasto 1986: 5–6; Pronk-Tiethoff 2013: 79–81). Moreover, the semantic field also extends to ‘beech’ in Slavic, and to ‘beech, oak’ or ‘birch bark’ in some Germanic languages. This etymological comparison shows a correlation of semantic fields and their extension in Slavic recipient-languages and Germanic donor-languages. Thus, words related to writing skill, such as ‘alphabet’, ‘letter’ and ‘book’, in these languages might have been originated from a word referring to wood or certain types of tree, which would suggest that the innovation of writing skill had ultimately spread from Germanic to Slavic peoples. This direction of spread might be partially related to the fact that many societies in early Russia were led by a Scandinavian elite, despite Slavic being the main language of (written) communication (discussed in Section 2.1.2).

2.2 TAXONOMY OF FINNIC AND SLAVIC LANGUAGES

The data used in the present study is comprised of two geographical peripheries of two language groups. On the one hand, Finnic languages are spoken on the western edge of the Uralic speaking areas (see Figure 3), while North Russian, on the other hand, was historically the easternmost extent of the Slavic-speaking areas.

2.2.1 FINNIC AS A WESTERN URALIC SUB-BRANCH

A subgrouping of the Uralic languages has been a topic of discussion over a century and various types of models have been proposed including a binary model (Janhunen 2009: 65), a bush model (K. Häkkinen 1984) and a comb model (Salminen 1999, 2002). In any case, the latest investigation of the Uralic language family by J. Häkkinen (2012), which takes into account both direct subbranching and later family-internal convergence, results in the tree structure illustrated in Figure 10.

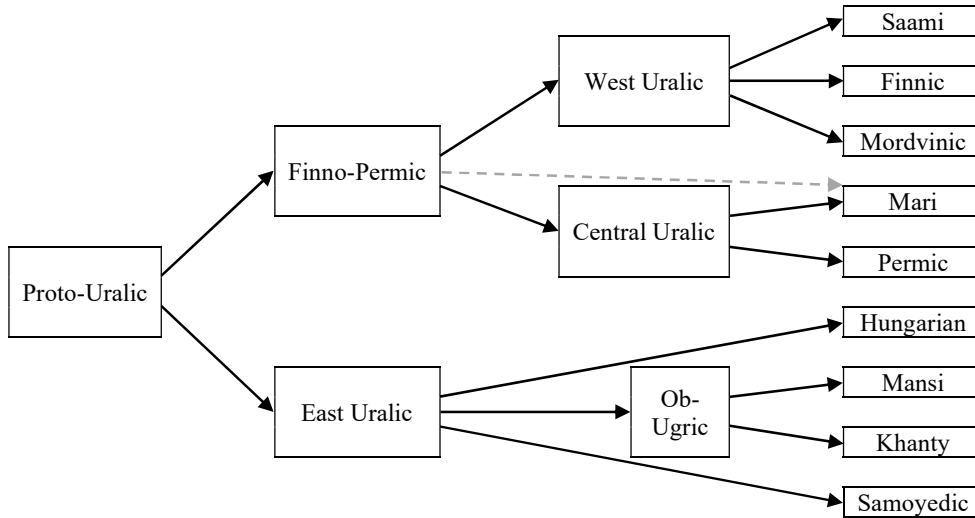


Figure 10 Subgrouping of the Uralic languages (J. Häkkinen 2012: 16)

Based on the tree in Figure 10, Finnic languages are classified as geographically western in the Uralic context, and are genealogically closer to Saami and Mordvinic languages. This idea is also reflected in earlier studies that regard Finnic and Saami languages as having derived from a common “Early-Proto-Finnic” ancestor (T. Itkonen 1998; Sammallahti 1999). This proposal has been recently questioned and criticised, however (see Saarikivi & Grünthal 2005: 124; Kallio 2007). In any case, J. Häkkinen’s model is also debatable and it is not the purpose of the current study to discuss the Uralic tree structure beyond the Proto-Finnic stage, so our interest is turning to the Finnic branch next.

In the similar principle as J. Häkkinen (2012), a tree structure that takes into account both sub-branch-specific innovations as well as later contacts across Finnic sub-branches is proposed by Kallio (2014) who largely applies the model from earlier studies by Sammallahti (1977), Koponen (1991) and Viitso (1985, 2000, 2008), illustrated in Figure 11. It is worth remarking that the idea of South Estonian being the first offshoot of Proto-Finnic goes back to Sammallahti’s proposal (1977: 133). Kallio’s criteria that result in the proposed intermediate stages are primarily based on historical phonology (see Viitso 1985, 2000, for morphological evidence). For instance, the first diversification stage from Proto-Finnic to Inland Finnic (South Estonian) is based on five regular sound correspondences.

Proto-Finnic		Coastal Finnic		Inland Finnic
* <i>kt</i> , * <i>pt</i>	>	* <i>ht</i>	~	<i>tt</i>
* <i>-pi</i> , * <i>-βi</i> [PRS.3SG]	>	* <i>-pi</i> , * <i>-βi</i>	~	Ø
* <i>c</i>	>	* <i>c</i>	~	* <i>cc</i> > <i>ts</i> , <i>ds</i>
* <i>kc</i> , * <i>pc</i>	>	* <i>ks</i> , * <i>ps</i>	~	* <i>cc</i> > <i>ts</i>
* <i>ck</i>	>	* <i>tk</i>	~	<i>tsk</i>

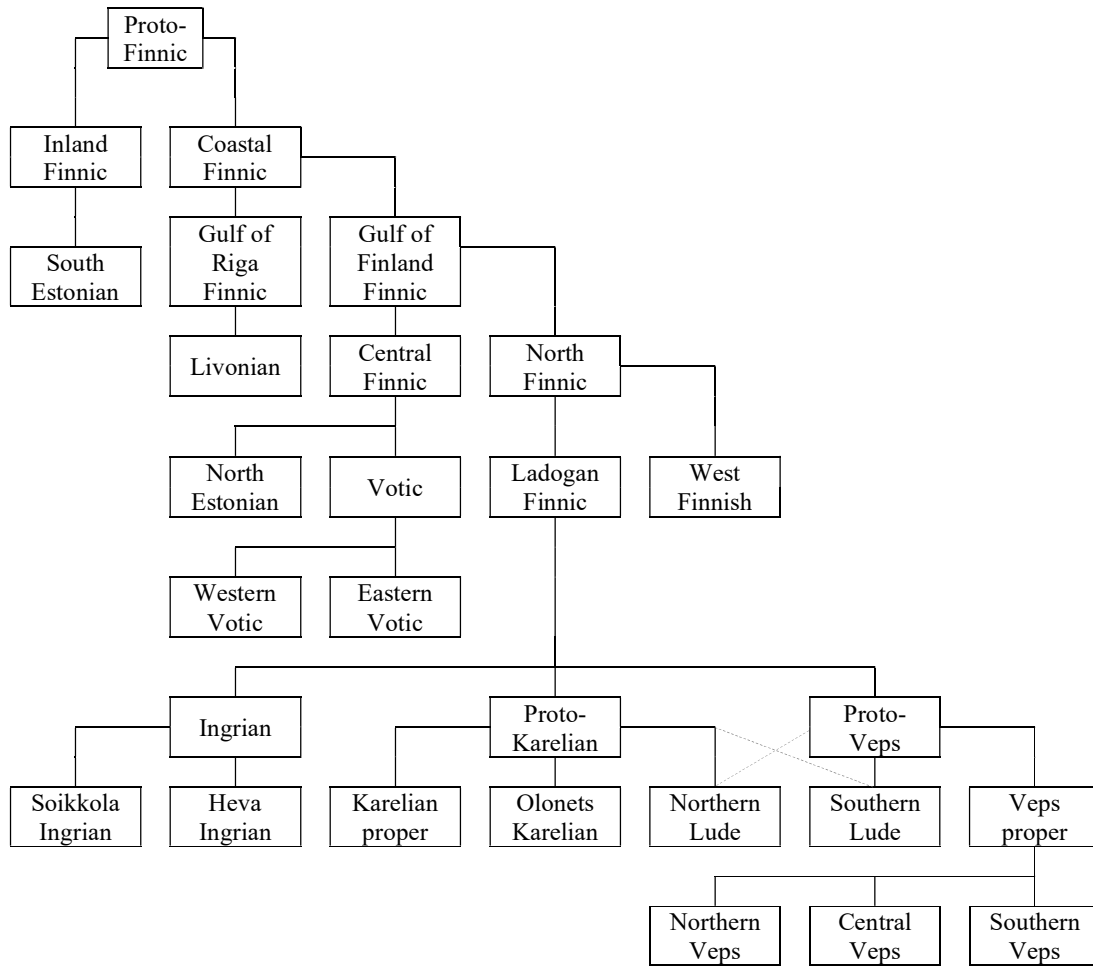


Figure 11 Subgrouping of the Finnic languages (adapted from Kallio 2014: 163)

Then, the branching of Coastal Finnic to Gulf of Riga Finnic (Livonian) is based on two sound correspondences, which are shared isoglosses with South Estonian.

Proto-Finnic		Gulf of Finland Finnic		Livonian & South Estonian
<i>*e-ä</i>	>	<i>*e-ä</i>	~	<i>*ä-ä</i>
<i>*ai</i>	>	<i>*ei-ä</i>	~	<i>*ai</i>

In the stage when Central Finnic languages (Votic and North Estonian) arguably diverged, four main sound changes and non-changes have been identified.

Proto-Finnic		North Finnic		Central Finnic
<i>*ë</i>	>	<i>*e</i>	~	<i>*ë (= ð)</i>
<i>*V_[+front]-o</i>	>	<i>*V_[+front]-ö</i>	~	<i>*V_[+front]-o</i>
<i>*-ksi</i> [COND]	>	<i>*-isi</i>	~	<i>*-ksi</i>
<i>*Rh</i>	>	<i>*Rh</i>	~	<i>*R</i>

Beyond this point, the divergence of Ladogan Finnic (Karelian, Veps, Lude and Ingrian) involves more recent changes, and complex contacts between Finnic sub-branches. The status of Lude in particular is worth remarking upon.

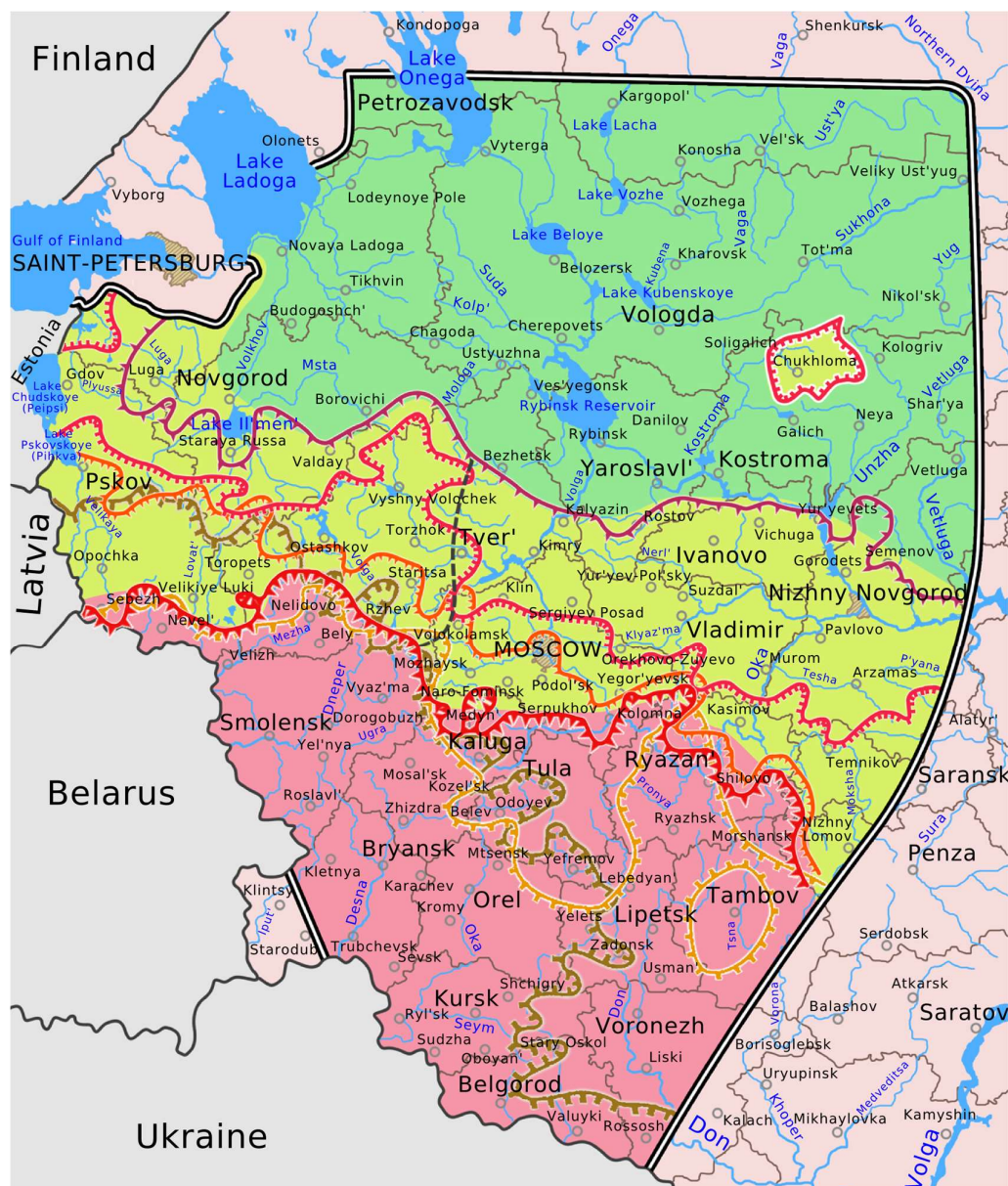
In a number of studies (Turunen 1946, 1947, 1950; Kettunen 1960; Kallio 2014), Lude is regarded as a mixed variety of Karelian and Veps. Given an areal-linguistic approach, the present study emphasises the equal contribution of all Finnic varieties, and treats Lude as independent varieties in the same way that Salminen (1998) does. Based on descriptions for an independent status of Lude by a number of scholars (Kujola 1910: 162–163; Setälä 1922: 730–731; Virtaranta 1972; Pahomov 2011, 2017), the current study maintains a working hypothesis that Northern Lude is a descendent of Proto-Karelian with Veps influence, whereas the southern variety is a Karelianised variety of Proto-Veps.

Following the proposed structure of the Finnic sub-branches presented above, the current study also takes data from the Finnic demonstrative systems into discussion, to supplement phonological evidence investigated in previous studies (discussed in Section 7.2.2).

2.2.2 NORTH RUSSIAN IN A HISTORICAL-COMPARATIVE SLAVIC CONTEXT

The speaking area of a Slavic variety called North Russian has changed over the course of time. This means that some specific regions that are reported to have been part of the North Russian dialect continuum a millennium ago might no longer be so today, and similarly, the dialect continuum might have expanded to new territories. As illustrated previously in Figure 4, the Medieval Novgorod was a political centre of the Russian North, and supposedly also the centre of the Old Novgorod Slavic language. As for the present-day situation, Novgorod falls out of the North Russian speaking area, seen in Figure 12. Additionally, the North Russian dialect continuum itself is not a homogenous area, and variation among North Russian varieties is expected, such as the distribution of *akanje* vs. *okanje* dialects (see also Avanesov 1958: 181).

Despite being classified as a dialect of Old Russian (i.e. Old East Slavic), some scholars have also proposed an unorthodox classification, in which Old Novgorod would not belong to the East Slavic sub-branch but rather form a separate “Northeast Slavic” (Ivanov 1990) or “North Slavic” sub-branch (Žuravlev 1994). For instance, the Slavic second and third palatalisations have, exclusively in Old Novgorod, taken place in different sound environments from the rest of Slavic languages due to a later date of monophthongisation of Proto-Slavic **ai* > **ē* > *ě* (yat), resulting in *kělъ* vs. Old Church Slavonic *cělъ* < Proto-Baltic **kailos* ‘whole’ (Vermeer 1986: 507–509). This isogloss reportedly still has some remnants in modern Northwestern Russian dialects (Zaliznjak 2004: 43). Another phenomenon of *cokanje*, meaning the merger of *c* [ts] and *č* [tɕ] to *c* [ts] such as *cъto/ceto* vs. Old Church Slavonic *čъto* ‘what’, is also unique for Old Novgorod, which is often regarded as an adstrate/substrate influence from Finnic (Vermeer 1986: 510; Vlasto 1986: 333–334).



Northern Russian
 Southern Russian

Central Russian

The territory of the primary formation
 The border between the West Central and East Central Russian

Isoglosses

- Voiced velar fricative /ɣ/
- Reduction of vowels in pretonic syllables (akanye, yakanye and others)
- Reduction of vowels in the second pretonic and post-tonic syllables
- Soft /tʲ/ in the 3rd pers. sg. and pl. present endings of the verb
- Dissimilative yakanye
- Prothetic /w~u/ before /o/ and /u/

Figure 12 Russian dialects (an English version of Kasatkin 1999: 96)
 (Wikimedia Commons: commons.wikimedia.org/w/index.php?curid=17591408)

Meanwhile, there are also several sound changes, in which Old Novgorod overlaps with West Slavic, e.g., the development *e > o* (*jokanje*) and the realisation of the Late-Proto-Slavic **o* and **ǔ* (see Mikkola 1894; Kalima 1956; Bjørnflaten 2006; Nuorluoto 2006, 2007; Frog & Saarikivi 2015: 76).

However, the current study follows conventional classifications, treating Old Novgorod as having branched off from Proto-East Slavic, as positioned in Figure 13.

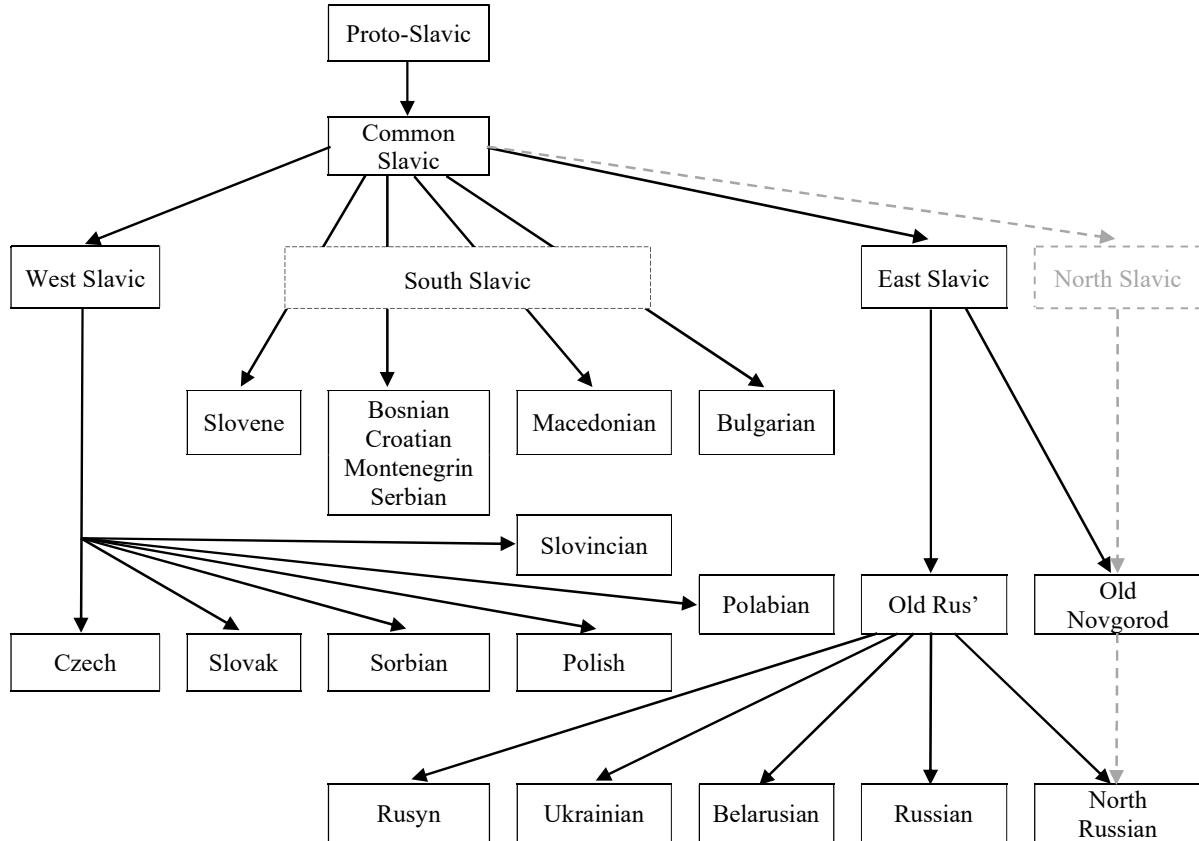


Figure 13 Subgrouping of modern Slavic languages

As a remark to the tree model in Figure 13, a recent study by Lindstedt and Salmela (2020) proposes the idea that the diversification of South Slavic sub-branches took place independently, meaning that no common intermediate proto-language, Proto-South-Slavic, can be reconstructed. Meanwhile, it is more certain that all West and East Slavic languages derive from common Proto-West-Slavic and Proto-East-Slavic, respectively. Alternatively, Slavic languages can also be classified as northern vs. southern groups, divided by the Hungarian speaking area. Namely, several isoglosses are shared exclusively by West and East Slavic, while South Slavic has taken a different direction of change, e.g., the reflexes of the Proto-Indo-European ending **-ens* > northern Slavic **-ě* vs. southern Slavic **-ę* (R. Kim 2019).

East Slavic languages in the early-2nd millennium can be roughly divided into two major dialectal groups based on the literary tradition: northern (Old Novgorod) vs.

southern (Old Rus') (see also Šahmatov 1915; Vlasto 1986: 303). At the same time, a dialectal comparison of phonology, morphology, and lexicon by Gorškova (1972: 136–142) points to even a more fine-grained classification of five distinct ancient dialects:

1. Northwestern I (Novgorod)
2. Northwestern II (Pskov)
3. Northeastern (Rostov-Suzdal')
4. Western (Smolensk-Polock)
5. Southern (Kiev, Černigovo-Severskij, Galič-Volyn)

Among these ancient East Slavic dialects, the western dialect later became Belarusian, while the southern dialect has continued its Kievan Rus' heritage as Ukrainian (Vlasto 1986: 336).

The modern Central Russian dialects go back to a variety of the ancient northeastern (Rostov-Suzdal') dialect spoken around the Moscow area in the transitional zone between northern and southern dialectal belts during the 13th–15th centuries (Vlasto 1986: 33–34, 323–324). This period of time corresponds to the emergence of Middle Russian and the formation of three East Slavic literary languages: Belarusian, Russian, and Ukrainian (Filkova 1973: 92–93; Saurio & Nuorluoto 2016: 16). Later in the early-16th century, an admixture with the Western (Smolensk) dialect, a direct source variety of literary Belarusian, gave rise to the vernacular, serving as the basis for literary (Moscow) Russian and has continued its life as Standard Russian today (Vlasto 1986: 336).

Vertically, modern North Russian dialects are sometimes thought to be direct descendants of Old Novgorod (cf. Figure 13). Due to continuous language shift to Russian, formerly non-Slavic speakers, particularly multilingual speakers with Finnic, Saami, and Permic background, have constantly been influencing the development of North Russian (see also Vlasto 1986: 301). However, the dialects in the north today might not have necessarily derived from Old Novgorod. Considering Moscow's invasions of Novgorod in the 15th and 16th centuries (discussed in Section 2.1.2), the Central Russian dialects must have shaped dialectal frontiers, influencing the local varieties of the north. For instance, the variety spoken in Novgorod today belongs to the Central Russian dialect group (shown in Figure 12), not the Northern.

The genealogical status of North Russian dialects and their relation to Central Russian is also crucial for explaining the emergence and development of multi-functional postposed demonstratives. As a starting point, the current study maintains the idea of modern North Russian dialects being descendants of a Central (Moscow) dialect of Middle Russian, which absorbed the Old Novgorod substratum after the fall of the Ancient Novgorod principality, in addition to the Uralic substratum (see further discussion in Section 7.3). This given scenario also entails a chronology that the diversification of modern North Russian dialects from Central Russian should have started in the 15th–16th centuries, after the assimilation of the erstwhile Old Novgorod speakers (see also Gorškova 1972: 144–146, 153–154; Vlasto 1986: 358, 363).

From an areal perspective, Russian, especially North Russian dialects, exhibit many features that distinguish them from the Slavic languages of the Standard Average European *Sprachbund* (a linguistic area proposed by Whorf [1939]1956, whose idea was later advanced by Haspelmath 2001). A number of previous studies have investigated potential Uralic substrate influences in the development of Russian, particularly in North Russian dialects (see also a similar comparison between Russian dialects in Vlasto 1986: 304–313), as shown in Table 2.

Feature		Uralic		Slavic			
		Mordvinic	Finnic	North Russian	Central Russian	Proto-Slavic	BCMS
1	Accentuation	initial stress	initial stress	initial stress	mobile stress	mobile stress	mobile stress
2	Unstressed <i>o</i>	<i>o</i>	<i>o</i>	<i>o/ʌ</i>	<i>ʌ/ə</i>	<i>a</i>	<i>o/a</i>
3	Diphthongisation of stressed <i>o</i> & <i>e</i>	<i>o</i> ' <i>e</i>	<i>oo/uo/oa</i> <i>ie/ia/æ</i>	<i>úo/óa</i> <i>ie/ia</i>	<i>ó</i> ' <i>é</i>	<i>á</i> <i>ǣ, æ</i>	<i>ó</i> <i>ě/jě/ije</i>
4	Palatal sibilants	<i>č, č', c</i> <i>š, š', s</i>	<i>c</i> <i>s</i>	(<i>č</i> >) <i>c</i> = <i>c</i> (<i>š</i> >) <i>s</i> = <i>s</i>	<i>č, c</i> <i>š, s</i>		
5	Grammatical gender	no gender		syncretism	3 genders		
6	Comparative form of noun	yes			no		
7	Clitic forms of the personal pronouns	no			yes		
8	Object case marking in negation	nominative	partitive	partitive-genitive		accusative nominative	accusative
9	Nominative object	nominative genitive ablative	nominative genitive partitive	nominative accusative			accusative genitive
10	Predicative possession	BE-possession					HAVE-possession
11	Possessive perfect construction	no	yes		no		
12	Postposed demonstrative (p)article	yes			less developed	no	

Table 2. *Potential Uralic substrate features in Russian*

Most previous studies primarily focus on what has changed in the development from the Proto-Slavic to modern Russian under the impact of Uralic-speaking language shifters, such as Features 1–8 (Veenker 1967; Vostrikov 1990; Saarikivi 2000). In contrast, other studies also show that certain archaic Slavic features have been lost in other Slavic languages such as in Bosnian/Croatian/Montenegrin/Serbian (BCMS), but have been retained in the Russian language due to contact with the neighbouring Uralic languages that use similar construction patterns, such as Features 9 (Timberlake

1974) and 10 (Grković-Major 2011; McAnallen 2011; Mazzitelli 2015; Yurayong 2019).

In addition to the two contrastive interpretations, innovation vs. retention, there are also cases in which the language speakers, still within a Uralic-Russian bilingual stage prior to the shift to Russian monolingualism, have developed and introduced a new contact-induced language feature to both the Russian and Uralic languages as a mutual reinforcement, such as Features 11 (Seržant 2012) and 12 (Leinonen 1998; Stadnik-Holzer 2006). As the current study primarily deals with Feature 12, previous studies suggest that the approach of contact and areal linguistics is needed for understanding the development of postposed demonstrative (p)articles in North Russian dialects. By considering all three aforementioned aspects, we can better understand the direction of development that the Russian language and its northern variety have taken in diverging from other Slavic sister languages.

2.3 FINNIC-SLAVIC INTERACTIONS IN NORTHWEST RUSSIA

The present study follows a view already proposed by a number of scholars in the 19th century (e.g., August Schleicher, Karl Brugmann, and later Jerzy Kuryłowicz, André Vaillant) that Slavic languages form a sub-branch under the Baltic branch of the Indo-European family (see Comrie & Corbett 1993: 62). Therefore, the earliest contacts between Finnic and Slavic-speaking populations can be said to have taken place prior to the branching of Proto-Slavic from Proto-Baltic (Viitso 1990a; Koivulehto 2006). Meanwhile, the early Slavic contacts refer to the stage between Common Slavic and Old East Slavic (Old Novgorod, Old Rus') whereas recent Slavic contact covers the periods from Middle to Modern Russian.

Taking into account the emergence of postposed demonstratives in Finnic and North Russian, the aforementioned historical contexts, in which language contact could have produced such a mutual language feature, are discussed in this section. As the scope of the present study is limited to the Slavic contacts as of the 5th century, the preceding Baltic contacts between the late-2nd millennium BC and the 5th century will not be discussed here (see, e.g., Thomsen 1893; Koivulehto 1983, 2006; Junttila 2016; Lang 2016, for a thorough discussion of Finnic-Baltic contacts). The focus of this section is to describe contact scenarios from different periods of time, while a detailed description of various contact scenarios (adstratum, superstratum and substratum) will be discussed later in Section 3.2.

2.3.1 EARLY FINNIC-SLAVIC CONTACTS AS ADSTRATUM

The Slavic influence on Finnic languages represents the third layer of the Balto-Slavic contacts with Finnic in chronology, following the periods of Pre-Baltic (an Indo-European *śatəm* dialect) and Proto-Baltic (see also Viitso 1990a: 140). Slavs migrated from their proposed homeland in the north of Carpathians towards the erstwhile non-Indo-European areas in the northeast. This migration route follows the Dnepr, leading

to the encounter with the Baltic tribes and then later the Finnic-speaking population who had earlier arrived from the east (see Timberlake 2013: 337). The first wave of migration and encounter with Finnic must have taken place at the earliest during the 5th century, prior to which historical sources provide no reliable information about Slavs (Birnbaum 1979: 5–6; Timberlake 2013; Lindstedt & Salmela 2020). The contact with Mordvinic and Permic-speaking communities would have happened later, after the expansion of Slavic-speaking population to the east and north, respectively.

After the Slavs established of multiethnic fortified settlements, in order to take control over the trade networks towards the east, they gradually began to Slavise the areas. Frog and Saarikivi (2015: 75–76) propose that one crucial factor was a disruption of silver trade in the mid-10th century, which reduced population movement and consequently could have reinforced stable settlements and the process of Slavisation. Along with trade, the introduction of Christianity by Slavs to the neighbouring Finnic-speaking population towards the end of the 1st millennium AD was another important event, which has established a stable Slavic contact ever since, especially in Karelia as the easternmost part of Finnic dialect continuum where Ladogan Finnic was spoken.

Even though not all Finnic groups were immediately baptised after the arrival of Christianity to their speaking areas, religious vocabularies still made their way into these languages, in which a lexical category of religious terms was still lacking. For instance, the Finnic population residing in Karelia in the east must have borrowed Slavic vocabularies related to Christianity already in the second half of the 1st millennium AD despite the archaeological evidence dating the widespread adoption of eastern Christian practices as late as the 13th century (A. Koivisto 2006). The same applies to Livonians in the west, whose language inherits Christian terminology from the eastern Christian tradition already long before adopting the western form of Christianity in the 13th century (Grünthal 2015b: 105, 132–133). Interestingly, the earliest phase of Finnic-Slavic contacts seems to correlate with the spread of the eastern Church from the east to the west.

The division of western and eastern Church also crucially divides the Finnic-speaking community into two different cultural areas, which have continued up till the present day. Catholicism and Protestantism were introduced later in the 16th century by Germanic-speaking peoples in the west, and the Eastern Orthodox Church was introduced by the Slavs in the east. Given different degrees of Slavisation among different Finnic-speaking populations (see T. Itkonen 1971; V. Koivisto 1990: 20, discussed further in Section 2.3.2), a similar variable distribution is also expected to be reflected also in the demonstrative system (see empirical evidence and further discussion in Chapters 6 and 7).

In the Russian North between the 11th–15th centuries, there were several principalities which took control over the others in turns (discussed in Section 2.1.2). Among the Slavic-speaking population, these economic centres were not a unified political unit neither a totally Russian-speaking community. In addition, the archaeological and genetic evidence shows that those centres also comprised many erstwhile residents with other ethnic backgrounds, above all, Uralic and Baltic-

speaking tribes (Tvauri 2007; Timberlake 2013: 337–338). Thus, the linguistic situation was accordingly very versatile. However, the diversity-favouring stage gradually started to vanish away after the invasion of Mongol troops in the 13th century, and later in the 15th century under the rise of the Moscow state, which took control over Slavic political units in the north and began the Pan-Slavic campaign (see further discussion in Section 2.3.2).

Archaeological evidence suggests that there were at least two migration waves of Slavs toward the north, dating between the 5th and 8th centuries (Birnbaum 1979; Sedov 1990; Kallio 2006; Kallio & Laakso 2020). Linguistically, the first Slavic wave can be connected to the Proto-Slavic stage, while the second one being connected to the Late-Proto-Slavic stage. The earliest contacts between Finnic and Slavic can be observed in Slavic loanwords, which still display the phonological shapes prior to several fundamental sound changes in East Slavic (Mikkola 1894; Setälä 1929; Kalima 1952, 1956; Ariste 1958; Kallio 2006; Saarikivi 2007b: 44, 2009: 114–116), as shown in the following three earliest Slavic loanwords in Finnic.

1. Old Russian *žbrdb* ‘bar’ ~ Finnish *hirsi* ‘beam’
 Early-Proto-Slavic **girdi-*
 > Proto-Slavic **žirdi-* → Early-Proto-Finnic **širti* > Late-Proto-Finnic **hirsi* > Finnish *hirsi*
 > Old Russian *žbrdb*
2. Old Church Slavonic *igo* (*ižesa* [PL]) ‘yoke’ ~ Finnish *ies* (*ikeet* [PL]) ‘id.’
 Early-Proto-Slavic **juga: jugesā*
 Proto-Slavic **jbgo: jbgesa ?* → Proto-Finnic **ikese-* > Finnish *ies: ikee-*
 > Late-Proto-Slavic **jigo: jižesa ?* → Proto-Finnic **ikese-* > Finnish *ies: ikee-*
 > Old Church Slavonic *igo: ižesa*
3. Old Russian *č̣melb* ‘bumblebee’ ~ Finnish *kimala(inen)* ‘id.’
 Early-Proto-Slavic **kimeli*
 > Proto-Slavic **č̣melb* → Proto-Finnic **kimala-* > Finnish *kimala-*
 > Late-Proto-Slavic **č̣melb*

(Kallio 2006: 157–159)

This is the primary linguistic evidence to support that the Slavic language in early contact with Finnic cannot be yet identified as a dialect of Russian, but rather some sort of Slavic variety in between East and West Slavic (discussed in Section 2.2.2).

On the Finnic side, some evidence of a medieval eastern Finnic language still survives up till today, namely birch bark document № 292 (Figure 14) (see Jelisejev 1961; Arcihovskij & Borkovskij 1963; Vermeer 1991; Laakso 1999, for interpretation attempts).



Figure 14 Novgorod birch bark document № 292 (ca 1240–1260)
(Russian Academy of Sciences:
gramoty.ru/birchbark/document/show/novgorod/292/)

As pointed out by Vermeer (1991: 334–340), some Finnic elements are also attested in two birch bark documents, for instance, the words *мол* (*moa*) < Proto-Finnic **maa* ‘land’ (line 1 in document № 56, Figure 15) and *касхи* (*kęski/ü*), *хапала* (*hapala*) < Proto-Finnic **käske-u* ‘order’, **hapa-* ‘sour’ (bottom left of document № 403, Figure 16).

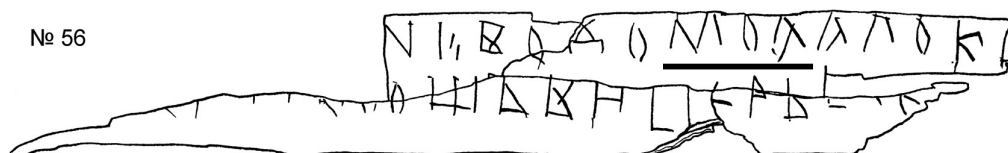


Figure 15 Novgorod birch bark document № 56 (ca 1240–1260)
(Russian Academy of Sciences:
gramoty.ru/birchbark/document/show/novgorod/56/)

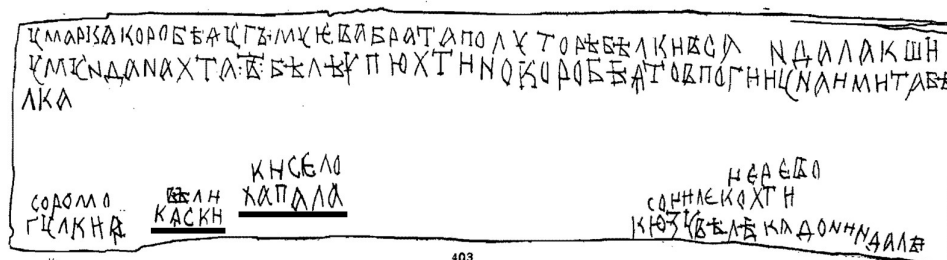


Figure 16 Novgorod birch bark document № 403 (ca 1360–1380)
(Russian Academy of Sciences:
gramoty.ru/birchbark/document/show/novgorod/403/)

Besides these documents, Finnic substrate elements concerning phonology, morphosyntax, lexicon and toponym are also observed in the Novgorod Slavic variety written on birch bark documents (Helimski 1986; Laakso 1999; Saarikivi 2006: 30). For instance, Saarikivi (2007a: 240–242, 2017) suggests that among a thousand of birch bark documents, there are at least 40 certain and 20 uncertain Finnic personal

names, which form a relatively large part of the name substance in the Novgorod birch bark corpus.

In this early phase of Finnic-Slavic contacts, there is no strong evidence that Slavic would have become a dominant language in Northwest Russia immediately after the arrival of Slavs. Instead, Slavicisation of the erstwhile Finnic-speaking areas would have taken place still in a slow pace from an initial adstrate scenario, in which both Finnic and Slavic speakers learned and spoke both languages, and consequently became multilingual, especially through intermarriage. This contact on the household level has given rise to a variety of Slavic with Finnic admixture that was later transmitted to the following generations of speakers and ultimately reached the community level (Timberlake 2013: 338).

2.3.2 RECENT FINNIC-SLAVIC CONTACTS AS SUBSTRATUM AND SUPERSTRATUM

This phase of Slavic contacts takes place in the stage after the transformation from the Old East Slavic (Old Novgorod, Old Rus') into Middle Russian during the 14–15th century (discussed in Section 2.2.2). Following the Mongol invasion and withdrawal, Novgorod was subsumed under the Grand Duchy of Moscow (discussed in Section 2.1.2). In any case, these historical events did not appear to have affected the multilingual environment in the eastern Circum-Baltic area in a significant way.

Maintaining the idea that modern North Russian dialects are possibly not direct descendant of Old Novgorod but rather of Central Middle Russian with the Old Novgorod and Uralic substratum (discussed in Section 2.2.2), the major area of recent Finnic-Slavic contact must have concentrated on the eastern part of Finnic dialect continuum, rather than the area adjacent to the historical centres Novgorod and Pskov, which hosted the earlier Finnic-Slavic contacts (discussed in Section 2.3.1). This view is supported by the fact that Russian varieties spoken in Novgorod and Pskov today are not classified in Russian dialectology as North, but rather Central Russian dialects (see Kasatkin 1999, illustrated in Figure 12).

Regarding various Uralic substrata in North Russian dialects (as discussed by Sarhimaa 1992, 1995, 1999; Myznikov 2003), there are at least four obvious types of Finnic substrata emerging in language contacts and shift from (i) Veps, (ii) Karelian, (iii) Votic, and (iv) South Estonian. Elsewhere, the recent influence from Russian such as loanwords, on the one hand, primarily concerns literary languages, such as the case of Estonian between the 18th–20th centuries (Mägiste 1962; Blokland 2005), considering that only 160 out of 2000 items have been identified as pre-18th-century Russian loanwords in Estonian dialects (Must 2000). Ultimately, the Russian loanwords may only be casual loans transmitted through other languages, such as from Russian through Latvian, or from Russian through Estonian to Livonian, a Finnic language which has never been in direct contact to Russian (Winkler 2002; Grünthal 2015b: 100, 138). Among these substrata, Veps and Karelian contact with Russian have continuously stayed the most intense since the Ladogan Finnic stage. Meanwhile, Votic and South Estonian contact with Russian can be traced back to the Ancient Novgorod state, but large-scale assimilation and language shift of Votic-speakers to

Russian must have started only in the 18th century (Ariste 1981: 77). Consequently, the degree of Russian interference in these Finnic speech communities is thought to be higher in the east and gradually decrease towards the west, meaning that Veps is a Finnic language that has undergone the heaviest Russian influence (see T. Itkonen 1971; V. Koivisto 1990: 20). The abovementioned contact scenarios are relevant for the diachronic investigation of demonstrative system, as they have taken place along the southern frontier of Finnic-speaking areas from the east (Veps) to the west (South Estonian), adjacent to the Slavic speaking areas (see Figure 1, Figure 3 and Figure 4).

Sarhimaa (1992, 1999: 43–50) describes that in the beginning of Slavic contact, most Uralic people preserved their languages alongside the acquisition of local Russian varieties, which created a stable multilingual practice and an interesting *Sprachbund* scenario between these languages in contact (see also Ariste 1981: 49). For this reason, Sarhimaa as well as Veenker (1967: 41) and Vostrikov (1990: 10–11) identify the early Finnic-Russian contact situation in Northwest Russia as an adstrate type (see Timberlake 2013, discussed in Section 2.3.1), which continued until the 18th century, at which point the language policy and practices of the Russian Empire changed in favour of Russian.

Thereafter, due to the stabilisation of the literary Russian language, the (even stronger) Russifying language policy of the Russian Empire, and the attempt to create a multiethnic nation state, the literary Russian language was promoted as an official state language. Given this dominant sociolinguistic position, Russian eventually became a socially more powerful language of the contact scenario (Kappeler 2001: 247–248). Following the revolution in the 1910s at the latest, the circumstance of Finnic-Slavic adstratum dramatically collapsed and Russian became a superstrate language to all minority languages of the Soviet Union (see discussion of superstratum and substratum in Section 3.2). This sociopolitical change also decreased the role of Turkic languages, most notably Chuvash and Tatar, as influential languages of traders in Central Russia.

Consequently, Russian interferences have been even more powerfully infiltrating into minority languages, not only to non-Slavic but also to Slavic varieties themselves such as Ukrainian and Belarusian as well as Russian regional varieties. Language shift to Russian occurred massively in the 1960s thanks to a rapid urbanisation, which had attracted people from suburban areas to come and work in bigger economic centres. Moreover, the socioeconomic situation inside the Soviet Union also forced many people to move and relocate themselves away from their home region to new areas where it was not possible to speak their native languages, only Russian (Lallukka 2001). This has turned even more minority languages into substrata of local Russian varieties, especially in several Uralic language communities investigated by Lallukka: Finnic (1990), Mari (2003) and Komi-Permyak (1995).

Sarhimaa (1999) also observes from her Karelian-Russian-speaking informants that in the transition period towards a Russian monolingual environment, language shift gave birth to many temporary Russian varieties, which illustrate obvious language features transmitted from regional substrate languages. In any case, these substrate features become subsequently extinct and levelled according to the norm of the Central (Standard) Russian, at latest in the following generation of speakers (see

also Vaahtera 2009 for a similar case of levelling in the phonology of Vologda variety). This shows that the language shift of the formerly Uralic speakers towards Russian proceeded in several steps and at different rates, which make various Uralic substrata distinct from one another according to individual local history (discussed above). Sometimes, the final phase preceding a complete language shift also causes significant changes, for instance, in the verb system in Finnic languages on the verge of extinction (Kehayov 2017; Jantunen 2019, forthcoming). This scenario is also expected to be the case for changes taking place in demonstrative systems (see further discussion in Section 7.2.2).

For the present study, the historical background discussed throughout this chapter provides useful information as a chronological parallel in dating the emergence of changes in demonstrative systems of Finnic languages in contact with Russian (see further discussion in Sections 2.4, 7.2 and 7.3).

2.4 HISTORY OF DEMONSTRATIVES IN FINNIC AND SLAVIC LANGUAGES

The linguistic phenomenon under investigation in the present study are demonstratives, the system of which has taken various development paths from Proto-Finnic and Proto-Slavic to modern Finnic and Slavic languages. Cross-linguistically, demonstratives can be organised differently in terms of proximity contrasts and morphosyntactic behaviours, even among genealogically related languages (see further discussion in Chapter 4). The following presentation will show that there is a tendency for both increasing and decreasing contrasts as well as semantic shifts of certain demonstrative series, while individual languages may also modify or build new series of demonstrative by applying additional deictic intensifiers or emphatic elements.

2.4.1 DEMONSTRATIVES IN FINNIC LANGUAGES

Across the Finnic languages, there is no single pattern of demonstrative system that would apply to all modern Finnic languages, as illustrated in Table 3 as pairs [singular] / [plural]. The following data is taken from literary languages (Laanest 1982: 196–199), while the actual spoken data of the present studies will be presented later in Chapter 6. Table 3 shows that the only etymological base that is still preserved in all Finnic languages is a series based on the roots *s-* in singular and *n-* in plural.

<div>West</div> <div></div> <div>East</div>	Language	Demonstratives			3rd person pronoun
		Proximal		Distal	
	Livonian	sie / ne			tāmā / nāmā (ta / nāā)
	South Estonian	sjoo / njooq	taa / naaq	tuu / nuuq	timā / nimāq (tiā / niāq)
	North Estonian	see / need			tema / nemad (ta / nad)
	Votic	kase / kane		se / ned	tāmā / nāmād / nāvād
	Ingrian	tāmā / nāmāt	se / net	too / noot	hän / he
	Finnish	tāmā / nāmā	se / ne	tuo / nuo	hän / he (se / ne)
	Karelian proper	tāmā / nāmā	še / né	tua / nua	hiän / he (še / ne)
	Olonets Karelian	tāmā / nāmāt	se / net	tua / nuat	hāi / hüö
neče / nenne					
Lude	tāmā / nāmād	se / ned	tuo / nuod	hāin / hüö (se / ned)	
neče / nehe					
Veps	nece / nene		se / ne	hän / hii	

Table 3. Demonstrative system in Finnic languages
(modified from the basis of Laanest 1982: 196–199)

Based on the evidence from modern Finnic languages, Larjavaara (1986: 69–75) identifies four etymological roots, three of which he includes in the following tripartite demonstrative system for Proto-Finnic (including my reconstruction for the plural forms).

	Semantics	Singular	Plural
1.	Proximal	* <i>tāmā</i>	* <i>nāmā(t)</i>
2.	Medial	* <i>se</i> < * <i>šej</i>	* <i>ne(t)</i>
3.	Distal	* <i>too</i>	* <i>noo(t)</i>

Notice that Kallio (2020) reconstructs the Early-Proto-Finnic singular medial series as **šej*. This is, on the one hand, based on the inflectional forms in modern Finnic languages, which have a long vowel *-ii-* in the stem, e.g., Finnish *se* [nominative], *siinä* [locative], *siihen* [illative], and *siitä* [elative]. On the other hand, this is compared to cognate forms in other Uralic branches, such as Proto-Mordvinic **šā* and Proto-Khanty **čī*. As some but not all Finnic languages also mark the plural forms with a Proto-Uralic plural suffix **-t* (shown in Table 3), the Proto-Finnic forms may or may not have the plural marker, the issue of which remains for further investigation in future studies.

Several remarks concerning significant changes from the Proto-Finnic to modern Finnic languages can be made. First, the Proto-Finnic proximal series **tämä/nämä(t)* has functionally shifted to mark 3rd persons in the western Finnic languages (Livonian, South Estonian, North Estonian and Votic). Synchronic 3rd person pronouns thus have two forms, full and shortened (in parentheses), the functional distinction based on marking emphasis or contrast (see Pool 1999; Kaiser & Hiietam 2003; Pajusalu 2005).

Second, the Proto-Finnic medial series **se/ne(t)* has been preserved across the modern Finnic languages. In some languages, however, it has taken over other spatial spheres, such as the reduction of the demonstrative system into a unipartite one in Livonian and North Estonian, or the semantic shift to a proximal demonstrative in South Estonian (see Pajusalu 1996a; Tomingas 2018). In other Finnic languages with the original tripartite distribution (geographically from Ingrian to Lude), **se/ne(t)* remain as medial³. Meanwhile, in some languages, **se/ne(t)* also become 3rd person pronouns in a spoken variety of Finnish, Karelian proper and Lude (see Saukkonen 1967; Laitinen 2002, 2005; Pahomov 2011; Priiki 2017). Due to the renewal of demonstrative paradigm into a bipartite system, the Proto-Finnic medial series **se/ne(t)* has become distal, while the proximal series **tämä/nämä(t)* has been replaced by newly emerged compound demonstratives *kase/kane* in Votic and *nece/nene* in Veps. In Veps, this new compound series is reportedly continuing to replace the synchronically distal demonstratives *se/ne* and consequently leading Veps towards a unipartite system (Grünthal 2015a: 277). In any case, the contrast is still occasionally observed in Veps as in (3).

Northern Veps

(3)	<i>ii</i>	<i>d'o</i>	<i>si-da</i>	<i>kät-t</i>	<i>an-nu,</i>
	NEG.3SG	already	that-PTV	hand-PTV	give-PTCP.PST
	<i>andj-i</i>	<i>ñeci-n</i>	<i>käde-n.</i>		
	give-PST.3SG	this-ACC	hand-ACC		

‘She did not give that hand but gave this hand.’ (Onegina & Zajceva 1996: 34–51)

³ Other studies in the framework of interactional linguistics and discourse analysis (e.g., Laury 1997; Seppänen 1998; Etelämäki 2006, 2009; Priiki 2017; Reile et al. 2019) may pay a closer attention to such delicate factors like inclusiveness between the speaker’s and addressee’s referential spheres and thereby classify the semantics of the Proto-Finnic medial demonstrative series **se/ne(t)* in the mentioned modern Finnic languages as “addressee-centred” (vs. “speaker-centred”) or alike rather than as “medial” (situating in between “proximal” and “distal”) in the traditional spatial-based classification (e.g., Larjavaara 1986, 1990). As the current historical-comparative study is primarily oriented towards a structuralist approach, I choose to use the simpler distance-based triad: “proximal”, “medial” and “distal”, for the sake of comparability with the equivalent demonstratives in Slavic languages, while reserving awareness of the aforementioned interactional-linguistic-oriented classification (see also a comparison between different approaches in Dixon 2003: 86–89).

As shown in (3), the two demonstratives *se* and *nece* can be used in parallel to create a contrastive reading between ‘that’ and ‘this’, even though the context of parallels is not frequently observed.

Third, the other Proto-Finnic medial series **taa/naa(t)* is exclusively observed in South Estonian (Pajusalu 2015) and marginally in some North Finnic varieties (Larjavaara 1986: 93–102). There has been a controversial claim by Larjavaara (1986: 73–75) that it could be a Baltic borrowing to Late-Proto-Finnic (cf. Latvian *tas/tā*, Lithuanian *tàs/taĩ/tà* and Old Prussian *stas/sta*), which has later partially merged with the Proto-Finnic proximal **tämä/nämä(t)*, resulting in the short back-vowel form of the proximal series *ta* from *tämä* in Livonian and from *tema* in North Estonian (see also Pajusalu 2015: 188; Tammekänd 2015).

Fourth, the Proto-Finnic tripartite system was reduced to a bipartite system (proximal vs. distal) in Votic and Veps. The new compound demonstratives derived from combinations of an emphatic element *ka-* (< a particle ‘well’) or *ne-* (< a particle *näged* ‘you see’) and the Proto-Finnic medial demonstrative **se/ne(t)*: Votic *kase/kane* and Veps *nece/hene* ‘this, that, it / these, those, they’ (see also Kettunen 1943: 403). The presence of the demonstrative series *nece/hene* in Olonets Karelian and Lude, the languages which mainly use the *nece*-series as medial or distal, is often regarded as result of borrowing from Veps (Larjavaara 1986: 154–155). This also has a parallel in Russian, in which the proximal demonstrative *ëtot* is a combination of a deictic intensifier *è-* and the Proto-Slavic medial demonstrative **tb(tb)* (see also Suhonen 1990: 95–96, and further discussion in Section 2.4.2). Furthermore, the Pinega variety of North Russian also has a new demonstrative series, having emerged from the combination of a particle *nu* ‘well’ and the Proto-Slavic medial demonstrative **tb(tb)*, resulting in the proximal demonstrative *nutot* ‘this’ (KTTGN 1970). Given these parallels, the formation of Finnic compound demonstratives may have eventually been motivated by the Russian model (see Blokland 2012 for a tendency of pronoun borrowings in Uralic languages, and further discussion in Section 7.2.2).

2.4.2 DEMONSTRATIVES IN SLAVIC LANGUAGES

Identical to Central (Standard) Russian, North Russian dialects have a bipartite demonstrative system with the proximal *ëtot* and distal series *tot*. However, a compound series *nutot*, (as mentioned in Section 2.4.1) is also observed in the Arkhangelsk variety of North Russian (see concrete data in Section 6.1.5). Similarly to the divergence among Finnic languages, Slavic demonstrative systems do not have a homogenous pattern either, as shown in Table 4 as triads [masculine] / [neuter] / [feminine] of the singular forms. Here, demonstrative systems of the extinct Slavic languages, Polabian and Slovincian, are also taken into account.

Language		Demonstratives			3rd person pronoun
		Proximal	→	Distal	
Old Church Slavonic		<i>sъ/se/si</i> <i>ovъ/ovo/ova</i>	<i>тъ/to/ta</i>	<i>онъ/ono/ona</i>	<i>онъ/ono/ona</i> <i>јъ/je/ja</i>
South Slavic	Slovene	<i>ta/to/ta</i>	<i>tisti/tisto/tista</i>	<i>oni/ono/ona</i>	<i>on/ono/ona</i>
	BCMS	<i>ovaj/ovo/ova</i>	<i>taj/to/ta</i>	<i>onaj/ono/ona</i>	<i>on/ono/ona</i>
	Macedonian	<i>ovoj/ova/ovaa</i>	<i>toj/toa/taa</i>	<i>onoj/ona/onaa</i>	<i>on/ono/ona</i> <i>toj/toa/taa</i>
	Bulgarian	<i>tozi/tova/tazi</i> <i>(toja/tuj/taja)</i>		<i>onzi/onova/onazi</i>	<i>toj/to/tja</i>
West Slavic	Polabian	<i>sъ/sü/so</i>		<i>tъ/tü/to</i>	<i>vъn/vâno/vâna</i>
	Slovincian	<i>tien/tuo/ta</i> <i>nien/nuo/na</i>			<i>vuon/vuono/vuona</i>
	Polish	<i>ten/to/ta</i>		<i>tamten/tamto/tamta</i>	<i>on/ono/ona</i>
	Lower Sorbian	<i>ten/to/ta</i>		<i>tamny/tamno/tamna</i> <i>wóny/wóne/wóna</i>	<i>wón/wóno/wóna</i>
	Upper Sorbian	<i>tón/to/ta</i>		<i>tamny/tamne/tamna</i> <i>wony/wone/wona</i>	<i>wón/wono/wona</i>
	Czech	<i>ten/to/ta</i> <i>tenhla/tohle/tahle</i>		<i>tamten/tamto/tamta</i> <i>onen/ono/ona</i>	<i>on/ono/ona</i>
	Slovak	<i>ten/to/tá</i> <i>tento/toto/tato</i>		<i>tamten/tamto/tamta</i> <i>henten/hento/henta</i>	<i>on/ono/ona</i>
East Slavic	Old Novgorod, Old Rus'	<i>(se)sъ/se/sja</i> <i>(sej/sije/sija)</i>		<i>(to)тъ/to/ta</i>	<i>онъ/ono/ona</i>
	Rusyn	<i>sys'/syse/sysja</i>		<i>tot/toto/tata</i>	<i>vin/ona/ono</i>
	Ukrainian	<i>ce/ce/cja</i>		<i>to/te/ta</i>	<i>vin/vono/vona</i>
	Belarusian	<i>gèty/gèta/gètaja</i>		<i>toj/toe/taja</i>	<i>ën/jano/jana</i>
	Russian	<i>ètot/èto/èta</i>		<i>tot/to/ta</i>	<i>on/ono/ona</i>

Table 4. Demonstrative system in Slavic languages

Based on Old Church Slavonic and modern Slavic languages, scholars usually reconstruct a tripartite system for Common Slavic, i.e. Late-Proto-Slavic (e.g., Leskien [1871]1955: 96–101; Babič et al. 2003: 122–123), which has a partial connection to the Proto-Indo-European (PIE) system (Kortlandt 1983; Derksen 2008). The system is organised as follows.

	Semantics	Form	Etymological remark
1.1	Proximal	* <i>sb</i>	< Proto-Balto-Slavic * <i>sis</i> < PIE * <i>ki-</i> ~ Hittite <i>kās</i> , Gothic <i>hina</i>
1.2		* <i>ovb</i>	< PIE * <i>h₂eu-o-</i> ~ Avestan <i>ava</i>
2.	Medial	* <i>tb</i>	< Proto-Balto-Slavic * <i>tos/tod/ta?</i> < PIE * <i>so/tod/seh₂</i> ~ Sanskrit <i>sá/tád/sá</i>
3.	Distal	* <i>onb</i>	< Proto-Balto-Slavic * <i>anos</i> < PIE * <i>h₂en-o-</i> ~ Greek <i>ān</i> , Latin <i>an</i>

Several changes that have taken place in later Slavic languages are worth mentioning.

First, the bipartite system, with a dichotomy proximal vs. distal, is predominant in West and East Slavic languages, while the tripartite system is still retained in most South Slavic languages except Standard Bulgarian. However, Lorentz (1903: 275–278) reports a unipartite system for Slovincian in which two demonstrative series *t-* and *n-* (~ Old Church Slavonic *nъně/nъnja* ‘now’) cover both proximal and distal.

Second, the Proto-Slavic proximal series **sb*, which was present in Old Church Slavonic, Old Novgorod and Old Rus’ (Vlasto 1986: 127–128; Dolgova & Maksimova 1996; Zaliznjak 2004: 125), has become obsolete in the majority of modern Slavic languages, while leaving some trace in the Standard Ukrainian and Rusyn systems as well as Polabian (Polański 1993: 812). Nevertheless, it still survives today in several fossilised adverbial phrases such as an adverb ‘now’ in Russian *sejčas*, BCMS *sad(a)*, Bulgarian/Macedonian *sega*.

Third, the other Proto-Slavic proximal series **ovb* is exclusively observed in South Slavic languages except Standard Bulgarian. It remains questionable whether this proximal series is simply a case of loss in other West and East Slavic languages, or a South Slavic innovation. In any case, Vaillant (1958: 380) regards this demonstrative series as secondary development, which replaces the original Proto-Slavic proximal series **sb* in South Slavic languages.

Fourth, the Proto-Slavic medial series **tb* is retained in all modern Slavic languages. In a number of languages (Bulgarian, Polish, Czech, Slovak, Slovene, Belarusian and Russian), it can receive additional deictic intensifiers, such as *tam-* ‘there’ in West Slavic languages, which help clarifying the spatial sphere contrast (Marušič & Žaucer 2012; Šimík 2016). Meanwhile, Vasmer ([1950–1958]1987: 4523) proposes that the East Slavic deictic intensifier (g)*è-* possibly traces back to the particles *èvo* and *èna* ‘here, well, see!’ (cf. BCMS *evo* id.), which also give variants *èvtot* and *èntot* ‘this’ in dialects (see the similar etymologies for Votic and Veps, given in Section 2.4.1), while double use of demonstratives is also found in the dialectal form *èstot* < **è-se-tot* (**se* < the Proto-Slavic proximal **sb*). As discussed in Section 2.4.1, the formation of modern Russian proximal demonstratives with a deictic intensifier *è-* and the Proto-Slavic medial demonstrative **tb*(*tb*), which emerged in spoken language no later than the 16th–17th centuries (Vlasto 1986: 129), seemingly provided a model for the Votic and Veps compound demonstratives (see further discussion in Section 7.2.2). In any case, the use of such deictic intensifiers in Polish, Czech, Slovak, Belarusian and Russian is obviously a secondary development motivated by the simplified paradigm, which only uses a demonstrative element from

the Proto-Slavic medial series **tǫ*. The intermediate stage after the neutralisation of proximity and before the application of deictic intensifiers to the demonstrative paradigm is seemingly portrayed in the Slovincian unipartite system.

Fifth, the nominative and accusative form of Proto-Slavic distal series **onǫ* has shifted to mark 3rd persons in most modern Slavic languages, but, for instance, not in Standard Bulgarian, which has instead shifted the Proto-Slavic medial series **tǫ* to 3rd person pronouns *tozi/tazi/tova* ‘he/it/she.’ In any case, the Proto-Slavic distal series **onǫ* has a suppletive paradigm, consisting of **onǫ* only in nominative (and inanimate masculine accusative), while the rest of paradigm is based on the old 3rd person pronouns **jǫ/je/ja*, which are scarcely attested in Old Church Slavonic, but later mainly survive as suffixes of long adjective forms *-ǫ-jǫ/-o-je/-a-ja* (see further discussion in Section 5.2.1). One important remark is that the use of **onǫ* as a distal demonstrative has seemingly become obsolete already during the Proto-East-Slavic period, resulting in the absence of the **onǫ* series from modern East Slavic demonstrative systems (cf. Dolgova & Maksimova 1996: 219).

The introduction of demonstratives has shown a parallel tendency that the original medial series **se/ne(t)* in Proto-Finnic and **tǫ* in Proto-Slavic are retained in all modern languages, although their semantics could have radically changed. This strong tendency of retention together with its dynamic of semantic change could be one factor that favours the functional extension of these erstwhile medial demonstratives to marking of definiteness, information structure and discourse, a phenomenon which will be discussed further in Sections 5.1 and 5.2.

3 METHODOLOGICAL TRIANGULATION: LANGUAGE CHANGE, CONTACT AND AREA

This chapter introduces the essential methodological components of the present study. The ultimate goal in the study of language change, contact, and area is to identify what is the motivation of a certain change or non-change in a language: internal development vs. contact influence. Given this goal, the present study combines methods from historical, contact, and areal linguistics, which are inseparable and complementary subfields in the studies of language change (as demonstrated in Campbell 1998). The advantage of this triangular approach is that languages in contact and contact-induced features can be understood in a vertical (language history), horizontal (language contact) as well as gradual scale dimension (areality). Section 3.1 introduces the methods, some issues, and key factors in the studies of language change, particularly in the area of morphosyntax. Section 3.2 classifies various types of language contact scenarios, as well as their characteristics and influences on language change. Section 3.3 extends the question of language change to a broader areal context, involving structural similarities and diffusion.

3.1 HISTORICAL LINGUISTICS

Since the 19th century when comparative linguistics was introduced by the Neogrammarian school as a scientific branch, the science of history is generally regarded as a field of research that describes general conditions of the existence of an object developing historically, as well as the nature and operations of the elements that remain constant throughout the changes (Paul 1880: 1). This understanding makes an analogy for languages as living organisms, which form themselves based on particular laws, evolve, and die off (Bopp 1827: 1). Based on this idea of life cycle of language, a diachronic approach can provide better understanding of how a language and its related phenomenon under investigation have become what they are today.

3.1.1 RECONSTRUCTION AS A METHOD TO STUDY LANGUAGE CHANGE

Contemporary studies of language change focus on three different stages of development: (i) origin, (ii) process, and (iii) results (Wiemer & Wälchli 2012: 9–14). In this scheme, the context of postposed demonstratives in Finnic and North Russian as examples of language change can be regarded as a phenomenon involving all the mentioned stages. In a similar principle, historical linguistics also aims to discover attestations of a specific linguistic phenomenon in different periods of time, which makes identification and chronologisation of changes possible (Meillet 1921: 19).

However, in reality, very few languages have been recorded long enough, e.g. Greek and Romance, in contrast to the majority of the world's languages. Thus, the investigation of language history, including that of Uralic languages, is often based on the reconstruction of missing information of the past (see Ross 1998; Spriggs 1998; Saarikivi & Lavento 2012). The ultimate goals of this method are to understand the history of individual languages, i.e. internal reconstruction, based on which instances of genealogical relatedness among languages can be identified, i.e. the historical-comparative method (Harrison 2003: 214). For instance, through these methods, Finnish and Hungarian have been identified as descendants of a common ancestor, the Proto-Uralic language.

Assuming that genealogically related languages evolve from a common ancestor language, i.e. proto-language, language change can be associated with two scenarios: (i) inheritance, and (ii) innovation. The former stands for retention of a language feature that has been continuously present, whereas the latter stands for an evolution from the proto-language stage. However, when comparing genealogically related languages, it is often difficult to determine which scenario is applicable for individual branches of a language family (Meillet 1921: 43). This is due to the uncertainty of how the proposed ancestor language eventually looked like, since a grammar of the attested language can easily be an admixture of two different language sources (Meillet 1928: 102; see also Tkačenko 1989, discussed in Section 3.2). For instance, modern Standard Finnish has an integrated grammatical system of both western as well as eastern varieties.

Innovations can either be an independent language-internal innovation or motivated by a language-external force, which analogously corresponds to the classic dichotomy drift vs. selection in Darwinist evolutionary biology (Darwin 1859, 1871). Drift refers to an internally-motivated change emerging from variation within the language system, which reorganises itself as 'compensation for something that was weakening within [a language]' (Sapir 1921: 170). Due to its inversive characteristics, drift often applies to languages spoken in isolated communities, such as Iceland, Faroe Islands and the islands in Oceania (e.g., Ross 1998: 155; Trudgill 2011: 6). Meanwhile, selection applies rather to majority of the world's languages because 'there is no evidence that any languages have developed in total isolation from other languages' (Thomason 2001: 8). Given the social nature of human beings, community, interaction, and communication generally lead to contact between languages, in the environment of which language resources from various sources are introduced, selected, and adopted to individual languages in contact (e.g., Dixon 1997; Kusters 2003; Trudgill 2011). Accordingly, language contact is crucially one of the most influential factors in language change because it can, for instance, guide the direction of change and selection of pre-existing structural variants in a language (see a variationist approach in Milroy & Gordon 2003, discussed in Section 3.1.2).

Given that studies in language history are diachronic by their nature, availability of data from different stages of a language is one of the main issues in historical linguistics. Qualitatively, what is available from the earlier periods might not necessarily represent the language entirely but only partly. In this vein, Labov (1994: 11) claims that many historical sources are only random collections or fragments of

text, which often show a bias towards language correctness and against vernaculars. Worse than that, the sociological background of the authors is poorly described, if available at all, and there is no method to prove what kind of features could be considered distinctive vernacular forms in regard to the normative language (see also Fox 1995: 2–3). These methodological challenges shape the Labovian view of historical linguistics as follows.

‘Historical linguistics can then be thought of as the art of making the best use of bad data. The art is a highly developed one, but there are some limitations of the data that cannot be compensated for.’

(Labov 1994: 11)

This is particularly true for Finnic languages, the written attestation of which dating back no further than the 13th century (see Novgorod birch bark documents with Finnic elements discussed in Section 2.3.1). This shortcoming of data in Finnic and Uralic languages in general highlights the importance of the application of historical-comparative methods and reconstruction.

3.1.2 GRAMMATICALISATION AS A TOOL TO RECONSTRUCT MORPHOSYNTACTIC CHANGE

Among different domains of language, scholars tend to agree that morphosyntax (not to mention pragmatics) is one of the most difficult areas to reconstruct, due to its internal complexity and abstractness (Harrison 2003: 221). Compared to phonology, morphosyntactic patterns such as word order can be easily replaced, leaving very little (if any) trace of its origin and earlier stage (Ravila 1966: 110; Fox 1995: 104–109, 190–194; Nichols 2003: 304–305). In any case, Campbell (1998: 251) is of the opinion that morphosyntactic reconstruction is not an impossible mission. For instance, despite the shift to SVO word order, Proto-Indo-European and Proto-Uralic SOV word order has left traces in the verb-final relative clause of several modern Germanic languages such as Dutch and German, while still being identifiable in the predominant head-final constituent order in noun and adposition phrase of Finnic and Saami languages.

In this regard, neighbouring languages play an important role in providing information on possible development tendency (see Campbell 1998: 244–246). This follows from the assumption that neighbouring languages, regardless of their original typological profile, tend to become structurally more similar to each other, i.e. converge, over the course of time, as is the case of languages in the Balkan and Circum-Baltic *Sprachbünde* (see also Romaine 1988: 80–81; Lindstedt 2000; Wiemer & Wälchli 2012, and further discussion in Section 3.3). Therefore, the reconstruction of morphosyntax often takes place after the genealogical relation between languages has already been established, primarily on the basis of regular sound correspondences (Harrison 2003: 225).

Another characteristic of morphosyntactic features is variation, because one function can be expressed by multiple morphosyntactic construction patterns which coexist in a language, without having to be dialectal features (Wilson & Henry 1998). This variationist approach operates on the idea that morphosyntactic changes emerge from the selection of variants (see also Nichols 2003, and earlier discussion in Section 3.1.1), and discusses how constructions with a shared function compete for the status of primary construction in a language (Milroy & Gordon 2003: 190). For instance, definite articles emerging in Slavic languages involve the selection of constituent position in which demonstratives are placed: postnominal position in Bulgarian, Macedonian and Southeast Serbian vs. prenominal position in other Slavic languages (see further discussion in Section 5.2.1). Often, the judgement is based on frequency of use, which implies that even a foreign construction can become domesticated and replace the erstwhile construction when the speech community has used the newly imported construction with a higher frequency than the erstwhile alternative (Pintzuk 2003; Heine & Kuteva 2005: 47).

The issue of selection also touches upon the contrast between notions of characteristic and marked features, which define what is, respectively, a canonical and non-canonical structure in a language system. The difference between characteristic and marked features in a language can be illustrated, for instance, with Talmy's (1985) example of a verb construction.

'Any language uses only one of these types for the verb in its most characteristic expression of Motion. Here, "characteristic" means that:

1. It is colloquial in style, rather than literary, stilted, etc.
2. It is frequent in occurrence in speech, rather than only occasional.
3. It is pervasive, rather than limited, that is, a wide range of semantic notions are expressed in this type.'

(Talmy 1985: 62)

This implies that the process of becoming characteristic takes place in gradual steps, the final stage of which is the achievement of characteristic expression. For instance, several studies have shown that the development of prenominal demonstratives to definite articles is an unfinished process, because of optionality in their usage (e.g., in Estonian by Pajusalu 1997a: 173; in Finnish by Juvonen 2000: 196; and in Votic by Agranat 2015: 51). This further implies that the use of this feature is still regarded as a marked construction, i.e. an unfinished process of grammaticalisation. The notion of characteristic feature is also a determining factor in contact-induced language change (Thomason 2001: 76–77, see further discussion in Section 3.2.2).

In the reconstruction of morphosyntax, grammaticalisation plays a significant role as a method, which helps in identifying and predicting morphosyntactic changes through "undoing grammaticalisation" and cross-linguistic information provided by the field of language typology (DeLancey 1994; Harrison 2003: 226; Heine 2003: 596, 598). The term "grammaticalisation", introduced by Meillet (1912: 132), refers to

development pathway in which a lexical element has semantically bleached or lost its original lexical meanings, being reanalysed and acquiring a new non-lexical assignment to express grammatical functions (see also Lehmann 1982; Hopper & Traugott 1993; Bybee et al. 1994). The change via grammaticalisation is more commonly thought to be unidirectional, as the lexical origin of a certain grammatical element is frequently, although not always, visible, whereas a lexical element generally does not illustrate any grammatical function beyond its part of speech (Hopper & Traugott 1993: 16–17; Heine 2003: 582–583).

However, there is also scepticism against grammaticalisation as a theory and its unidirectionality, because grammaticalisation is not a process *per se* but rather a by-product and ultimate outcome of semantic change and morphosyntactic reanalysis, which can also result in grammatical elements becoming less grammatical (Campbell 1998: 241–242, 2000; R. Janda 2000; Joseph 2000; Newmeyer 2000). For instance, French *derrière* has changed from the adverb ‘behind’ to a noun ‘buttocks’, while New Mexico Spanish *-mos* has changed from a 1st person verbal suffix to clitic *-nos* (R. Janda 1995). A similar degrammaticalisation process has given rise to the North Saami postposition *haga* ‘without’ and derivational morpheme *-naga* ‘stain, stained with,’ which originate from Proto-Uralic cases suffixes, abessive **-pta* and locative **-na*, respectively (Ylikoski 2016). Taking a cautious approach, the present study does not treat the notion of grammaticalisation as a process, but as a result of morphosyntactic reanalysis, functional extension, and semantic bleaching of the erstwhile meaning, which are observed in particular language elements such as demonstratives (see the three stages of language changes, origin, process, and result, discussed in Section 3.1.1).

In terms of historical reconstruction, undoing grammaticalisation is often conducted by “profiling” various domains of changes that have taken place in a certain morphosyntactic construction from a morphological, lexical, and semantic to syntactic profile (see a detailed description of the “profiling” method in Seržant 2015). Such a fine-grained structural analysis can capture changes and predict stages in the history of a language when certain changes started to occur. However, undoing grammaticalisation as a method has its own inadequacies in several aspects. Namely, it cannot provide information for establishing an absolute chronology, nor can it help in justifying subgrouping of cognate languages, as cross-linguistic similarities observed through this method may eventually be a result of massive diffusion (Harrison 2003: 220–221; Heine 2003: 598). As the present study investigates Finnic and Slavic varieties, the genealogical position of which is not controversial, the issue of genealogical relation is not generally relevant. Meanwhile, the subgrouping of Finnic varieties cannot be classified only on the basis of the reconstruction of demonstrative system, a morphosyntactic area of language. Therefore, the method of undoing grammaticalisation is applied to supplement an earlier subgrouping, which has been established on the basis of historical phonology that applied the conventional comparative method involving sound correspondences and changes (cf. Campbell 1998: 307–308, and see concrete application of this method in Section 7.2).

3.2 CONTACT LINGUISTICS

Every language has its own history. Such histories are more well known for a few languages, and less documented for the vast majority of the world's languages. As language changes take place in consecutive steps, each stage of change always leaves their trace and forms a "stratum" (literally 'layer'). Analogous to the use of tree rings to identify the climatological history in dendrochronology (e.g., Fritts 1976), each layer of language changes and non-changes can provide information on the space and time of when the language has been spoken. Most layers generally contain information on language contact, because a language that has developed in total isolation from other surrounding languages is rare (see Campbell 1998: 69; Thomason 2001: 8, as mentioned in Section 3.1.1). Such is not the case for Uralic and Indo-European languages. Therefore, language contact is methodologically considered an important clue for reconstructing language sociology of a certain speech community in specific periods of time.

3.2.1 STRATIFICATION OF LANGUAGES AS TRACES OF LANGUAGE CONTACTS

Language contact as an external input always affects the language repertoire of every speech community member and their languages in some way (Thomason 2001: 10). It can add more languages to their language repertoire, i.e. multilingual practice, or cause them to give up one language for another, i.e. language shift. The two aforementioned scenarios, multilingual practice and shift, are often related to the notions of substratum, superstratum and adstratum, which have been present and have played an important role in the development of languages in Northwest Russia, since the arrival of Slavs and their encounter with the Finnic-speaking population (discussed in Section 2.3).

The term "substratum" (literally 'lower layer') appeared for the first time in the linguistic scholarship in the investigation of substrate influences in Ibero-Romance languages by Ascoli (1870). As an antonym to substratum, von Wartburg in 1932 (1951: 155) introduces the term "superstratum" (literally 'upper layer'). This dichotomy does not concern only language (à Thomason 2001) but crucially also the sociolinguistic and sociocultural situation in a speech community. All of these contribute to the dialectal diversification of the Russian language (see also Gorškova 1972: 70). Accordingly, a superstrate language of the upper layer is spoken by elites who are in a number of cases newly arrived conquerors in the speech community, e.g. Russian, while a substrate language of the lower layer is more often a language of erstwhile inhabiting commoners such as Finnic and Saami languages in Northwest Russia (Veenker 1967: 8–11; Saarikivi 2006).

Applying these terms to language contact scenarios leads to the discussion of the possible source or motivation of language change after the language shift has taken place (e.g., Veenker 1967: 13–17). Substratum is used to describe the situation in which commoners adopt the elites' language, while superstratum refers to the opposite setting in which the elites adopt the commoners' language. Concretely, language

change in a substratum scenario is an influence from an abandoned substrate language to a target superstrate language, for instance, from the extinct Iberian substrate language to the superstrate Latin, which results in modern Ibero-Romance languages (Ascoli 1870), or from the Uralic substrate languages to the superstrate Russian (Veenker 1967; Vostrikov 1990). In contrast, superstratum refers to the motivating force coming from the opposite direction, i.e. influence from an abandoned superstrate language to a target substrate language such as the French superstrate vocabularies in English (Fischer 2003: 107).

From another dimension of contact, “adstratum” (literally ‘side layer’), a term introduced by Valkhoff (1932: 17, 22), is often considered as an early stage of contact preceding substratum or superstratum, which maintains a socially equal circumstance and stable multilingual environment to speakers of languages in contact. This indirectly also entails that the adstratum scenario typically emerges in a smaller speech community and vernacular languages, rather than a larger society with a dominant literary language and a greater pace of assimilation to the majority culture (Vostrikov 1990: 3–15; Myznikov 2003). In other words, enlargement of society is one of the factors that favour the weakening of adstratum, and faster linguistic (and also cultural) assimilation towards one particular speech community, i.e. “denationalising” (German *Denationalisierung*, Veenker 1967: 16). This is indeed the case in Northwest Russia discussed earlier in Section 2.3.

Another concept in the stratification of language is “perstratum” (literally ‘over layer’), which does not concern contact on the speech level, but rather on the literary level. In other words, perstrate language is an administratively, socio-politically, or religiously prestige literary language that gives model or influences the vernacular languages, such as Latin in European academia, German in Habsburg, Russian in the Soviet Union or Chinese in Korea and Japan. Therefore, Vostrikov (1990: 10–11) considers perstratum as an unnatural contact between written languages, unlike the other three types of contact discussed above, which are natural contact situation at the speech level. In any case, the perstratum scenario does not necessarily have to involve languages of different origins, but it can also concern (closely) related language, e.g. the influence of Classical Chinese on Chinese vernaculars, or Old Church Slavonic on the literary Bulgarian and Russian languages (see also Vlasto 1986: 346–347). This scenario resembles the case of literary Finnic minority languages and Standard (Moscow) Russian, that have been levelling language features in spoken Finnic and North Russian which deviate from the written norm (see further discussion in Section 7.1.1).

The ultimate goal of studies of language contact of the substrate type is to identify and understand the language structure of substrate languages. Individual cases of language contact and shift manifest different degrees of substratification, which Tkačenko (1989: 25) and Myznikov (2003) grade in two degrees: (i) a perfect substratum, and (ii) an imperfect substratum. A perfect substratum refers to a situation in which speakers of a substrate language have completely assimilated to the language and culture of elites, for instance, the Meryans who were mentioned in Russian chronicles and have left toponymic traces in Central Russia (Matveev 1996, 1998, 2001; Ahlqvist 1998, 2006). Meanwhile, an imperfect substratum is a more common

scenario widely observed in Northwest Russia, in which speakers of the Uralic substrate languages have not totally abandoned their native language yet, such as the Veps who are shifting to Russian but still maintain a Veps-Russian diglossia, despite the domain of use of Veps language becoming significantly restricted in recent decades (Puura et al. 2013: 23–29).

In the case of perfect substratum, supporting evidence for identifying substrate languages is often insufficient, and scholars need to rely on reconstruction (see a similar issue in historical linguistics, addressed in Section 3.1.1). Therefore, the degree of success in identifying a substratum varies from case to case, as in the following classification based on Veenker (1967, see also similar classifications in Saarikivi 2006; Rahkonen 2013).

X = substrate, Y = superstrate

1. **Extinct languages:** the genealogical identity of language X is uncertain and can be identified only by the external historical sources or on the basis of external reconstruction, and the direction of influence is uncertain whether $X \rightarrow Y$ or $Y \rightarrow X$, e.g., the Iberian substratum in Spain, the Etruscan substratum in the Italian Peninsula, or the Thracian and Illyrian substrata in the Balkans.

2. **Languages of the conquered nations:** language X has disappeared from the area conquered by speakers of language Y, but still survives elsewhere as a result of earlier expansion or migration of the speech community, e.g., the Celtic substratum in Great Britain and France or the Saami substratum in Southern Finland.

3. **Attested languages:** language X which is attested with a handful of written sources, e.g., the Latin substratum in Dalmatia.

4. **Abandoned languages:** the genealogical identity of language X is uncertain, but the language itself can be reconstructed on the basis of its attested or surviving cognate language(s) more distantly related, e.g., the Meryan substratum in Central Russia that can be reconstructed on the basis of other attested western Uralic languages (Finnic, Mordvinic and Saami).

5. **Fragment or tendency from the substrate languages:** individual features in language Y, which can be associated with the ones in language X and its cognate languages, e.g., lexical items with unknown etymologies in Germanic, Saami and Finnic languages, which point to earlier indigenous Paleo-European substrate languages.

(based on Veenker 1967: 14)

Fortunately, the current study is dealing with a context in which the substratification has not been completed, so information on the Finnic counterpart in contact with North Russian is still relatively well accessible.

3.2.2 CONTACT-INDUCED LANGUAGE CHANGE IN GRAMMATICAL SYSTEMS

All of the contact scenarios discussed in Section 3.2.1 can lead to various outcomes in the process of language change, which is termed “contact-induced change”. Despite her statement that contact-induced change is nearly unpredictable, Thomason (2001: Ch. 4) provides social and linguistic predictors as tools for understanding language changes due to contact. The social factors concern intensity of contact and degree of success in acquisition of a target language (L2). Contact intensity can be measured by its duration and the sociolinguistic relationship among speakers of the languages in contact. Meanwhile, given that language shift mostly occurs with adult speakers, the degree of success in L2 acquisition is expected to be significantly lower than in child language acquisition. Regarding the situation in Northwest Russia, the acquisition of Russian for the most part involved adult language learners.

Thomason (2001: 66–76) defines that imperfect L2 learning, often observed with adult L2 speakers, may cause a “shift-induced change”, which tend to transfer language structures from their native (L1) to target language (L2), such as the case of Karelian-Russian contact described by Sarhimaa (1999, discussed in Section 2.3.2). On this particular issue, Bakker and Matras (2013: 4–9) explain the mechanism in which adult L2 speakers with limited resources of the target L2 (or “truncated multilingualism” in Blommaert’s term 2010: 23–24) have no other solution than to replicate the structural model of their L1 while substituting lexical elements with those of the L2 (see also Weinreich [1953]1974). In contrast, imperfect learning is insignificant in the scenario of “borrowing”, which more often concerns vocabulary. These patterns are likely due to the borrowing scale in which vocabulary is typically borrowed more easily in a casual and less intense contact, before the borrowing of language structure that occurs in a more intense contact situation (see Tkačenko 1989; Thomason 2001: 70–71). This also raises the issue of the linguistic factor: the structural or typological similarity between contact languages (see also Stolz 1991; Thomason 2001: 76–77, discussed further in Section 3.3). For instance, it is reasonable to assume that with a higher possibility, Uralic as agglutinative languages could borrow and accommodate structural elements from the likewise agglutinative Turkic more smoothly than from the fusional Slavic languages (see also Hesselbäck 2001; Stadnik-Holzer 2006, for examples of a similar idea).

Given that substrate influence emerging from language shift usually causes a change on the language-structural level, Tkačenko (1989: 34, 73) presents an idea that two language systems are merged into one system whereby the lexical resources of the L2 are retained, but the grammatical structure is influenced or borrowed from the abandoned substrate L1. When discussing structural borrowing, we deal not only with form but function as well, as a language might not necessarily borrow the whole package, including form and function (or “metatypy” in Ross’ term 1996), from another language. Instead, the speakers may attempt to replicate L1’s “structural model” (in Weinreich’s term [1953]1974) or “conceptual templates” (in Heine and Kuteva’s term 2001: 410) by applying lexical resource of L2.

In terms of language acquisition, we might rather assume that L2 speakers with imperfect acquisition still use the grammar of their L1, while taking the lexical

resource from the target L2, which can be regarded as a “lexifier language” (A. Bartens 2013: 67–68). This mechanism of “polysemy copying” has been described in many accounts (Weinreich [1953]1974; Tkačenko 1989; Heine & Kuteva 2005; Wiemer & Wälchli 2012), the idea of which is portrayed in Table 5.

	Function α	Function β
Model language (native L1)	Form a	Form a
Replica language (target L2)	Form b \longrightarrow	Form b

Table 5. Polysemy copying from model to replica languages

In the mechanism shown in Table 5, the model language provides a pattern, in which Form a can encode both Functions α and β . By copying this polysemy pattern, the replica language applies and extends the usage of Form b , which initially encodes only Function α , also to Function β . As result, both model and replica languages can, in a similar way to each other, code Functions α and β with single forms a and b . For instance, the coding of expression ‘to like’ differs between the Tatar varieties spoken in Finland and Tatarstan as they correlate with different model languages, Finnish and Russian, respectively. In Finnish, the verbs *pitää* and *tykätä* ‘to like’ require an object in the elative case *-sta* (Form a), so Finland Tatar replicates this polysemy of locational case and extends the locational use (Function α) of the ablative case *-Tan* (Form b) to mark also the object of liking (Function β). Meanwhile, the construction of liking in Russian is formed by an experiencer in the dative case (Form a) with the intransitive verb *nravit’sja* ‘to like’, so Kazan Tatar likewise extends the locational use (Function α) of the dative case *-Ka* (Form b) to encode also the person who likes (Function β).

In an intense contact scenario, such polysemy copying may affect the entire grammatical system, the phenomenon of which is termed “resyntactification” (Romaine 1988: 80–81). Such a high degree of convergence can ultimately achieve “intertranslatability”, which means that clauses can often be translated morpheme-by-morpheme across the languages in contact. A good example of such advanced contact-induced development is, for instance, the Balkan *Sprachbund*, as stated in Kopitar (1829).

nur eine Sprachform herrscht, aber mit dreierley Sprachmaterie

‘only one structure is produced, but with three language materials [Albanian, Bulgarian, Romanian]’ (the author’s translation)

(Kopitar 1829: 86)

Apart from the Balkans, a number of similar examples of heavy resyntactification across languages in contact have been identified and empirically examined, for instance, in Amdo-Qinghai (Janhunen 2007), India (Masica 1976), Japan-Korea

(Yurayong & Szeto 2020), Mainland Southeast Asia (Enfield 2005) and Meso-America (Campbell et al. 1986). In the cases mentioned, a high degree of resyntactification has resulted in structural convergence among languages that together form linguistic areas (discussed further in Section 3.3).

3.2.3 IDENTIFICATION OF SUBSTRATE LANGUAGE FEATURES

Studies of substrata as a research method have provided satisfactory results that shed light on unattested languages and their speech communities. The method has been advanced particularly in the field of toponyms, in which a large number of Uralic place names have been identified in Northwest Russia, pointing to earlier settlements prior to Slavicisation (Matveev 1996, 1998, 2001; Mullonen 2002; Ahlqvist 1998, 2006; Saarikivi 2006; Rahkonen 2013; Kuzmin 2014). Given the functionality of this method, studies of substrata can also contribute to a diachronic investigation of language structures in the current study.

As languages in contact are assumed to share a large number of structural similarities, substrate languages in the proposed contact area, be they extinct or still healthy, are not an exception in this respect. In the cases where the substrate language or its cognate languages have been attested (see the classification in Section 3.2.1), we can apply Saarikivi's parameter (2000) for the identification of substrate language features, as shown in Table 6.

More likely	Marked feature in the shift-target language	→	Characteristic feature in the shift-target language	Less likely
	Common among the substrate language and its cognate languages		Common among the shift-target language and its cognate languages	
	Rare language-internal development		Expected language-internal development	
	Typologically rare among the world's languages		Typologically common among the world's languages	

Table 6. *Characteristics of substrate features (Saarikivi 2000: 398–399)*

The first criterion focuses on the status of the suspected substrate feature in the shift-target language, and whether it is a characteristic or marked construction in the grammatical system (see the discussion of characteristic feature in Section 3.1.1). Unless the feature has become well domesticated in the grammar of the shift-target language, the possibility of substrate influence can be considered as high (see also a similar idea for areal features in Muysken 2008: 8).

The second criterion concerns how common the candidate substrate feature is among the cognate languages of both sides. The less common it is in the shift-target language, the more likely it is a substrate feature and *vice versa* (see also Meillet 1928: 122–123). However, it is more difficult to apply this criterion when comparing cognate languages in contact (e.g., investigation of Uralic-internal substrata in Helimski 2003) or typologically similar languages such as Uralic and Turkic (e.g., Janhunen 1977;

Róna-Tas 1988), which can make the justification of a contact-induced feature less probable (Stolz 1991). In particular, such features that have similarly been lost or retained in cognate languages either within or outside the proposed contact area can make a contact explanation less convincing. This emphasises the point made by Thomason (2010) that the explanation of language change should primarily be based on internal reconstruction, before applying a contact explanation.

The third criterion examines how probable it is that the substrate feature is a language-internal development: that is, how stable and resistant towards change the behaviour of the language feature is in the contact situation. Meanwhile the fourth criterion concerns the question of how common or rare the proposed substrate feature is, cross-linguistically represented in a larger micro-areal, macro-areal, and cross-linguistic context. The latter two criteria are mostly based on typological surveys and, consequently, set a requirement for scholars to go beyond their specific contact scenario into a wider areal perspective, as well as look at language-universal questions (shown in Muysken 2008). For instance, the Eurasian linguistic area introduced by Trubetzkoy (1923) can have considerable significance for the studies of language contact in the Slavic sphere and Northwest Russia (e.g., Stadnik-Holzer 2004a, 2006).

This parameter will be used to evaluate the probability of the multifunctional postposed demonstratives in North Russian dialects as resulting from the Finnic substratum (see Section 7.3.1).

3.3 AREAL LINGUISTICS

Regarding the substrate type of contact-induced change, especially in the context of Finnic-Slavic contact in Northwest Russia, areal linguistics is an important supplementary aspect to language contact. Apart from the implication on how to reconstruct and subgroup a language family (Campbell 1998: 307–308), the areal-linguistic approach serves as a reminder that any proposed substrate feature can eventually be also an areal feature (discussed in connection to many proposed Uralic substrate features in Russian by Veenker 1967). Such an areal feature could have been spread from a more remote centre of innovation, or ultimately emerged as a mutual development without a clear source of borrowing (e.g., an example of Eurasian patterns that also have influenced Russian in Skalička 1933–1934). General conditions and effects of areality on languages participating in a particular contact zone will be discussed throughout this section.

3.3.1 LINGUISTIC AREAS AS VENUES OF LARGE-SCALE LANGUAGE CONTACT

The notion of a “linguistic area” goes back to the early 20th century when Trubetzkoy (1923) introduced the concept in Russian *jazykovej sojuz* ‘language union’ (German *Sprachbund*) to explain language contact, above all for a micro-area of the Balkans as well as even larger proposed macro-areas such as the Mediterranean and Ural-Altaic unions. This term refers to an area where multiple languages, which are not necessarily

related genealogically, are spoken, and the languages under this interaction structurally become closer to each other, i.e. converge (see also Campbell 1998: 299–300). In contrast to the historical-comparative method, such convergence in language structure, in any case, does not prove the genealogical tie between languages in the same way as regular sound correspondences do. Instead, the structural convergence can potentially be also result of analogy, i.e. an attempt to level asymmetries in the language system, which might coincidentally happen in a similar way across languages (Meillet 1921: 26–27, 36–37). Therefore, it is never enough to declare Uralic and languages of the Altaic type, for instance, as cognate only on the basis of their agglutinative morphosyntax (cf. the “Ural-Altaic hypothesis” in Castrén 1839, 1850).

As aid tools for the studies of human history, mutual features, i.e. isoglosses, emerging among the languages in a contact area can provide interesting and valuable pieces of information about the local history and sociolinguistic situation. Treating languages as dynamic populations (or living organism in Bopp’s 1827 term), Nichols (1992, 1998) discusses the relation between geographical factors, population movement and language diversity in the sense that the core area of a language family is stable and better maintains family-internal unity. At the same time, peripheral areas such as Northwest Russia tend to favour emergence of non-canonical structures, that is, structural divergence from the core zone through borrowing, stratification, or language mixing from genealogically unrelated languages in contact, which may typologically be distant such as in the case of Finnic-Slavic contact (discussed in Section 2.2). This idea correlates with the location of substrate areas, which are primarily observed in remote micro-areas, i.e. “contact superposition zone” (a term used in Koptjevskaja-Tamm & Wälchli 2001: 624–629), such as Arkhangelsk region in the Russian North (see Saarikivi 2007b).

Muysken (2008: 3–9) discusses six criteria which are crucial for defining a linguistic area. First, we assume a certain “geographic area”, in which a language contact hypothetically has taken place. Second, we identify “three or more languages”, which are spoken within the proposed area and potentially have had intense and long-lasting interactions. Third, we investigate “shared structural features”, which behave similarly or identically in all or the majority of languages in the proposed area. Fourth, we look at “contact situation” among languages from both the historical and sociolinguistic viewpoints in order to identify contact types (discussed in Section 3.1.2) and possible multilingual practices within the contact zone. Fifth, we consult cross-linguistic data whether the suspected mutual features, i.e. isoglosses, among the languages in the proposed area are probabilistically stable and resistant to contact influence, or unstable and prone to contact influence, in order to exclude the possibility of “similarity by chance” (see Janhunen 2014b, discussed in Section 3.3.2). Sixth, we try to identify whether convergent features are borrowed from a specific language in which the features can be considered as “inherited” from an ancestor language of the source language, or emerges through the mechanism of mutual reinforcement as a new feature among the languages in the proposed area (see Lindstedt 2000, discussed in Section 3.3.2). Together, these six criteria can provide sufficient information concerning the profile of a linguistic area in focus.

In the aforementioned principles, a number of linguistic areas have been proposed and investigated, especially the Balkan *Sprachbund* (e.g., Kopitar 1829; Sandfeld 1926; Trubetzkoy 1928; Lindstedt 2000), among others. Uralic languages have also been included in various linguistic areas. The largest are two macro-areas: Eurasia (e.g., Trubetzkoy 1923; Jakobson 1931; Skalička 1933–1934; Stadnik-Holzer 2004b), in which most Uralic languages except Finnic and Saami participate, and the later dichotomy of European Uralic vs. Rossic Uralic languages (Helimski 2003: 157–158). On a smaller scale, Haarmann (1977) and Helimski (2003) have proposed numerous micro-areas with Uralic members, as follows.

1. **Danube** (Carpathia): Hungarian, Slovak, Czech and German
2. **Baltic** (Hansa): Votic, Estonian, Livonian, Latvian, North Russian, Swedish, Danish and Low German
3. **Volga-Kama**: Permian, Chuvash, Tatar and Bashkir
4. **Volga-Oka**: Merya, Muroma, Meščera, Mari and Mordvin
5. **Core/Central Uralic**: Mordvin, Mari, Permian, Ugric
5. **Onogur**: Old Hungarian, Bolgar-Turkic, Ossetic and Avar
6. **Ostyakia** (Ob-Yenisei): Khanty, Selkup and Ket
7. **Ob-Ugra**: Hungarian, Mansi, Khanty
8. **Eastern Uralic**: Ugric and Samoyedic
9. **Upper Yenisei** (Yenisei Qyrghyz): Kamas, Mator, Khakas and Shor
10. **Yenisei**: Samoyedic, Yeniseic, Tuva and Tofalar
11. **Peripheral Uralic**: Finnic, Saami and Samoyedic.

Among these proposed micro-areas, the status of linguistic area has been investigated and supported by satisfactory empirical evidence mainly in the first three cases (Danube, Baltic and Volga-Kama), while the rest are still speculative. For instance, Baltic as a focused micro-area in the present study hosts Finnic, Baltic and Germanic languages as well as East Slavic (especially North Russian) that manifest a number of morphosyntactic isoglosses, on the basis of which Baltic can be considered as a proper linguistic area (Zeps 1962; Mathiassen 1985a, 1985b; Koptjevskaja-Tamm & Wälchli 2001; Klaas-Lang & Norvik 2014). Similarly strong empirical evidence has been presented for the Danube (Pilarský 2001) and Volga-Kama linguistic areas (Bereczki 2007). Regarding the current study, Danube, Baltic and Volga-Kama are such areas, in which the Uralic languages with demonstratives as a marker of definiteness are spoken, i.e. Hungarian, Finnic, Saami and Mordvinic (see further discussion in Section 5.1).

In each linguistic area, the intensity of contact may vary between the core and periphery areas, such as the case of the Baltic area, to which Finnic and North Russian belong. Based on this idea, Güldemann (2008) describes a contact zone in Northern Sub-Saharan Africa by classifying three types of concentric circles generally observed in a linguistic area: (i) areal hotbed, (ii) core circle, and (iii) periphery. In this

classification, density and number of similar language features that are shared between neighbouring languages are expected to be the highest in the areal hotbed and gradually decrease towards the periphery. Regarding the use of postposed demonstratives in Northwest Russia, this model can also be applied to clustering of properties and functions, i.e. isoglosses, observed in each variety, which will give a better picture of Northwest Russia as a linguistic area (see further discussion in Sections 7.2 and 7.3.2).

Concerning the consequences of contact-induced changes in a larger areal context, scholars still debate whether the convergence towards genealogically unrelated languages has made a language more or less complex. Contrasting the complexification view by Nichols (1998), McWhorter (1998, 2001, 2005) uses a development tendency from a synthetic towards an analytic morphosyntax (as observed in the Balkan *Sprachbund*, for instance) as an argument that language in contact does not become more but rather less complex. Namely, the analytic morphosyntax can level controversial and complex synthetic construction models that manifest variation across languages in contact (Lindstedt 2000). In any case, no ultimate answer has been proved because language complexity can be measured from different aspects such as epistemic, ontological and functional perspectives (Miestamo et al. 2008). Regardless of the disputable issue of language complexity, we can still certainly say that there is clearly a unified pattern and tendency of (non-)change between languages that are spoken in the adjacent areas (see further discussion of shared tendencies among Finnic varieties in Section 7.2.2).

3.3.2 LINGUISTIC SIMILARITIES DUE TO AREAL CONVERGENCE AND DIFFUSION

The picture of language change in contact situation is not always black and white, but it often takes place in a gradual scale across the contact zone, meaning that certain varieties may be affected by areal diffusion more than others. Based on this principle, areal linguistics as a discipline in linguistics operates on the comparison of language features that behave and/or look alike in several languages of the adjacent area. Aikhenvald and Dixon (2001: 1–4) discusses five factors that can explain similarities between languages: (i) universal properties vs. tendencies, (ii) chance resemblance, (iii) borrowing vs. diffusion, (iv) genealogical retention, and (v) parallel development vs. convergent development (see also Harrison 2003: 215).

First, a Greenbergian approach to language typology assumes that every language shares a number of comparable “universal properties” (Greenberg 1963) and eventually “universal grammar” in a formalist framework (Chomsky 1965). However, languages do not necessarily share a common tendency how they concretely encode particular meanings, but their speakers rather share common human cognition and physiology that can still produce cross-linguistic diversity in the grammatical encoding on the language-surface level (Evans & Levinson 2009). This principle is applicable to the question of contact-induced features, irrespective of whether languages in the proposed contact area share the use of grammatical features such as definite articles only in function, or also in form (see also the discussion on matter and

pattern borrowing, e.g., in Sakel 2007). For instance, several previous studies have compared two different grammatical elements: (i) the North Russian postposed demonstrative, and (ii) the possessive suffixes in eastern Uralic languages, and have identified a shared functional range between them (e.g., Leinonen 1998, discussed in Sections 5.2.2 and 5.3).

Second, similar-looking lexical and grammatical elements might be similar in several languages only by chance resemblance without any historical connection. In this respect, Janhunen (2014b: 318) discusses the degrees of structural similarities as tools for the identification of contact-induced features. The cases with no probability of areal contact can be either an “accidental similarity” that randomly occurs without any historical reason (no areal adjacency, genealogical relationship, or contact), or they can be a “secondary similarity” that can be successfully explained by the internal reconstruction method in any case. Meanwhile, a case with a higher probability of areal diffusion is “shared drift”, meaning that languages of unrelated families attempt to express the same grammatical function with the same pattern, but still use their own lexical resource. One such case is the North Russian postposed demonstratives and Uralic possessive suffixes mentioned above (see also Weinreich [1953]1974; Heine & Kuteva 2001, discussed in Section 3.2). Such degrees of convergence seem to be a minimal requirement for declaring a certain geographical area as a linguistic area. Nevertheless, a case with the most intense contact and the highest probability of being an instance of areal diffusion is a “borrowing of inflectional suffixes”, which means languages in contact copy the same construction in both form and function (whole-sale borrowing), such as the case of the Turkic borrowings in Mari (Hesselbäck 2001).

Third, the investigation should speculate whether there is a clear evidence of borrowing from a specific language or a result of “diffusion”, meaning that languages have converged without any identifiable source. In certain cases, such as in the Balkan *Sprachbund* discussed by Lindstedt (2000), there is not necessarily any trace of a source language as in Weinreich’s model ([1953]1974), but all the languages participating in contact can be simultaneously replica languages that have converged (see Table 5, discussed in Section 3.2.2). The driven force in such areal diffusion is a multilingual practice of people in contact, in which they express their ethnic and religious identity by maintaining their native language, but cannot resist strong influences from the structure of other languages in their daily repertoire (see the case of the Balkans in Lindstedt 2016, and the discussion of “resyntactification” and “intertranslatability” in Romaine 1988, discussed in Section 3.2.2). For conventional historical linguists with an emphasis on phonology, such an idea of regular morphosyntactic correspondences might not exist because it does not operate as systematically as regular sound correspondences, due to possible variation and instability in language contact (discussed in Section 3.1.2).

Fourth, “genealogical relationship” between the languages in contact can make the identification of contact-induced features difficult since cognate languages presumably tend to inherit the same grammatical categories and their forms from a common ancestor language, a challenge addressed by Stolz (1991) and Thomason (2001: 76–77). In the cross-family contact scenarios such as the Finnic-Slavic case, Curnow (2001: 422–425) discusses a problem, namely that applying a contact-based

explanation to the data might not suffice for discovering the whole story, without taking into consideration the force of internal development as well. Methodologically, a contact-based explanation should be applied only after internal reconstruction reaches its limit (Thomason 2010, as mentioned in Section 3.2.3). For instance, this has led several scholars (e.g., Veenker 1967; Kiparsky 1969) to consider the Russian intransitive predicative possession, BE-type, as a shift-induced change under Uralic contact, even though internal reconstruction shows that this is actually a case of retention in Russian by the reinforcement from Uralic (and Turkic) neighbouring languages which have the same structural model. Meanwhile, the majority of other Slavic languages have shifted to the transitive HAVE-type, following the tendency in Standard Average European *Sprachbund* (Grković-Major 2011; McAnallen 2011; Mazzitelli 2015; Yurayong 2019).

Fifth, genealogically related languages naturally share an inner dynamic, which results in a common direction of change, i.e. a “parallel development”. However, unrelated neighbouring languages which belong to the same contact zone may also adopt such a tendency, and follow the trend as a “convergent development”. For instance, the Balkan languages might not have been originally so similar in their morphosyntactic structure, but due to the long-lasting advanced multilingual practices of the inhabitants, speakers of all languages in contact have reached the point where they created a unified set of grammatical templates for all languages (see Romaine 1988, discussed in Section 3.2.2, and Lindstedt 2000, discussed above). In the Finnic-Slavic context discussed in the present study, the scenario of parallel development in the demonstrative system across the Finnic dialect continuum will be discussed in Section 7.2.2, while the convergent development between Finnic and North Russian will be discussed in Section 7.3.

These six factors will be taken into consideration when discussing the similarities between the uses of postposed demonstratives in Finnic and North Russian in Chapter 7, which evaluates whether postposed demonstratives in North Russian dialects should be regarded as a Finnic substrate feature or areal convergence in Northwest Russia.

4 FUNCTIONAL DIMENSIONS OF DEMONSTRATIVES

This chapter discusses the theoretical background and provides a typological perspective for the analysis of Finnic and North Russian postposed demonstratives. The discussion starts from the definition and identification of types (Section 4.1), as well as exploring scenarios of change from demonstratives to grammatical markers such as definite articles (Section 4.2). The investigation is then divided into three different aspects relevant for understanding the use of postposed demonstratives in Finnic and North Russian: referential domain (Section 4.3), information-structural domain (Section 4.4) and evaluative domain (Section 4.5).

4.1 DEFINITIONS AND TYPES OF DEMONSTRATIVES

Demonstratives are considered a universal grammatical category across the languages of the world, even though their form and function vary greatly in individual languages (Diessel 1999a: 1). The use of demonstratives touches on various grammatical domains such as syntax, semantics, and pragmatics.

In terms of syntax, demonstratives can either modify a noun or function independently by itself as a pronoun or as a locational adverb. Based on this division Anderson and Keenan (1985), and Diessel (1999a: Ch. 4; 1999b: 2–18) define three syntactic categories of demonstratives: (i) pronoun, (ii) determiner, and (iii) adverb, illustrated in the following examples (4) to (6).

Finnish (Finnic, Uralic)

- (4) *sa-i-n* *tänään* *kirjee-si.* Pronominal demonstrative
get-PST-1SG today letter-P_X2SG
lue-n *se-n* huomenna
read-1SG it-ACC tomorrow
‘I got your letter today. I will read it tomorrow.’

- (5) *tämä* *kirja* *on* *paksu.* Adnominal demonstrative
this book be.3SG thick
‘This book is thick.’

- (6) *mitä* *minä* *siellä* *tek-isi-n?* Adverbial demonstrative
what 1SG **there** do-COND-1SG
‘What would I do there?’

The pronominal demonstrative (4) is an independent noun-like demonstrative used to avoid a repetition of a previously mentioned noun, while the adnominal demonstrative (5) is a modifier of a head noun, and the demonstrative often shows agreement in case, number and/or gender with the latter. Meanwhile, the adverbial demonstrative (6) can express the location, time, or manner, in which the action takes place (see Etelämäki 2006). Related to the latter type, demonstratives can be used to express perspectives of an action or state-of-affairs such as horizontal or vertical, direction, manner, quality, degree, temporal aspects as well as conditional, causal and concessive (see König 2012, 2015, 2017; König & Umbach 2018). For instance, degree adjectivals are expressed with manner demonstratives in Finnish *näin/niin/noin* ‘like this/that, so’ and in English/German *so*, e.g., *niin hyvä* ‘so good/so gut’, while quality can be expressed with adjectival demonstratives like Finnish *tällainen/sellainen/tuollainen* ‘this/that kind of, such,’ or English *such* and German *solche*.

In addition to these, Diessel also discusses another category, an identificational demonstrative as in (7).

Russian (Slavic, Indo-European)

(7a)	<i>èto</i>	<i>moi</i>	<i>druz’ja.</i>	Identificational demonstrative
	this.SG.NEU	my.PL	friend.PL	

Serbian (Slavic, Indo-European)

(7b)	<i>ovo</i>	<i>su</i>	<i>moji</i>	<i>prijatelj.</i>
	this.SG.NEU	be.3PL	my.PL	friend.PL

German (Germanic, Indo-European)

(7c)	<i>das</i>	<i>sind</i>	<i>meine</i>	<i>Freunde.</i>
	this.SG.NEU	be.3PL	my.PL	friend.PL
	‘These are my friends.’			

The definition of this type of demonstrative is somewhat fluid, as it may also look like a pronominal demonstrative in an identificational predication, which may explain why it has also been labelled in many other ways, e.g. deictic predicator, predicative pronoun, existential demonstrative, deictic identifying pronoun. Its primary function is to serve as a clause predicate in place of verb, usually observed in a nonverbal clause such as in Russian (7a) where a copula is not used in the present tense, or a copular clause such as in Serbian (7b) and German (7c). The identificational demonstrative differs from the pronominal one in that no case/number/gender agreement between demonstrative and predicative is required as the demonstrative remains immutable. For example, the identificational demonstratives *èto*, *ovo* and *das* ‘this’ in (7) remain in singular neuter, regardless of the predicative *moi druž’ja*, *moji prijatelji* and *meine Freunde* ‘my friends’ as well as the predicates *su* and *sind* ‘are’, which are all in plural.

Regarding the adnominal type, Dryer (2013a) surveys variation in constituent order with demonstratives in the languages of the world, shown in Figure 17.

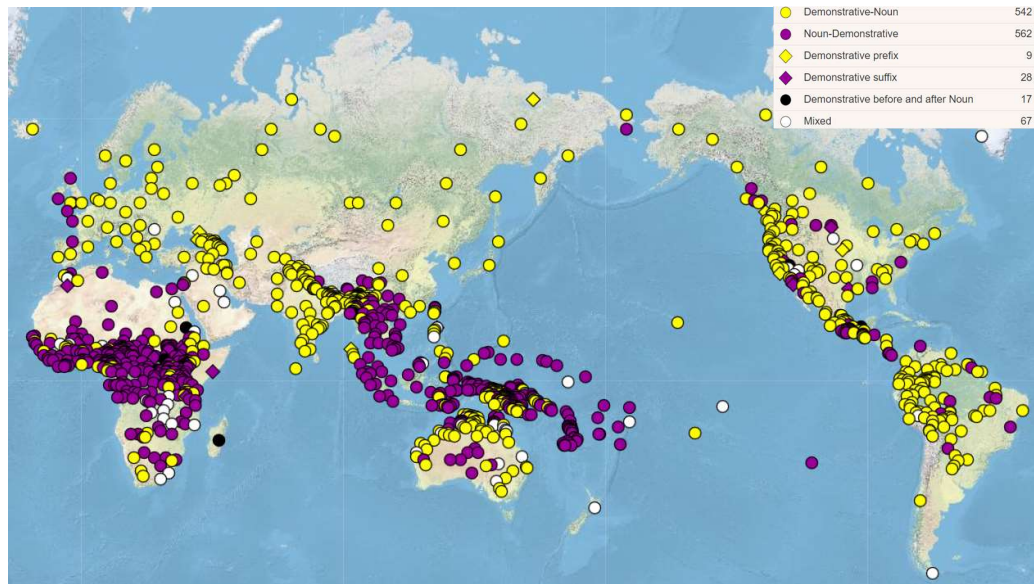


Figure 17 WALS Feature 88A: Order of Demonstrative and Noun (Dryer 2013a)

As can be seen in Figure 17, languages of Eurasia and the Americas largely prefer prenominal demonstratives, e.g., Dutch *dit boek*, Uyghur *bu kitap* ‘this book.’ In contrast, postnominal demonstratives are predominant in Africa, Southeast Asia, and Oceania, e.g., Malay *buku ini* [book this]. This speaks in favour of the fact that postposed demonstratives are not expected structural patterns in languages of Northwest Russia. Nevertheless, a free(r) constituent order within noun phrase, typical for a number of Indo-European languages, such as in Slavic (see Vaillant 1977: 266–267), receives no attention in Dryer (2013a).

Semantically, demonstratives often express a deictic contrast with regard to the distance between an entity in focus and the deictic centre (see the discussion of “origo” in Section 4.3.1). The expression of deictic contrast being a prototypical function of demonstratives, we find a great deal of variation in relevant deictic dimensions across the world’s languages. Concerning the adnominal type of demonstrative, a survey with simplified classification by Diessel (2013a) shows that some languages have no distance contrast, whereas others distinguish between two to five or even more values of distance, as shown in Figure 18. It is worth keeping in mind that distance is not always horizontally unidimensional, but a vertical dimension, such as different height levels ‘up’ and ‘down’, can also come into play.

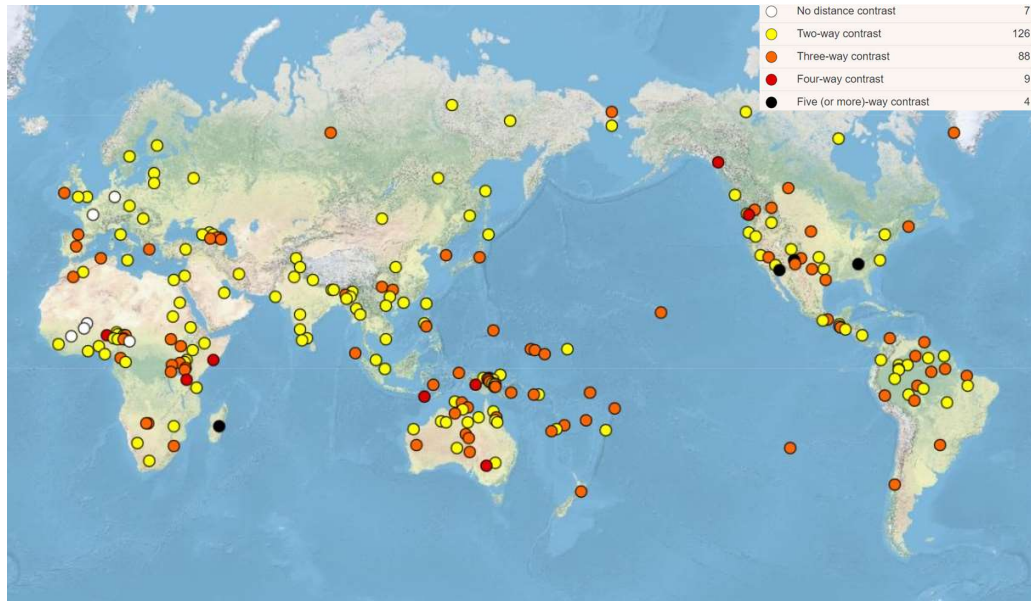


Figure 18 WALS Feature 41A: Distance Contrasts in [Adnominal] Demonstratives (Diessel 2013a)

In addition to distance contrasts, adnominal demonstratives also have additional functions, related to discourse pragmatics rather than spatial use, and are sometimes synchronically classified as definite article in language descriptions, e.g., English *the* and Egyptian Arabic *ʔil* (Halliday 1994: 180–186; C. Lyons 1999: 17–21).

In terms of pragmatics, the primary function of demonstratives is to draw a joint attention between the interlocutors and establish a spatial relation to a referred entity. At the same time, the demonstrative guides the addressee’s attention to objects or locations, often as medium for pointing gesture in the deictic sense as in (3) and (8).

- (8) *I asked you to bring me **that** cup, not **this** cup!*

Such referential uses are termed “exophoric” (Diessel 1999a: 93) or “situational” use (Himmelmann 1996, 1997). In addition to these functions, demonstratives play an important role in the organisation of discourse, as they keep track of the information flow as in (9).

- (9) *We had an organisation reform in our university in 2015.
This reform was driving everyone crazy!*

In Himmelmann’s term (1996: 224–229), (9) illustrates a propositional use, which is common in discourse when the speaker refers to the whole propositional content of the preceding statement. Such uses of demonstratives in a non-spatial sense are termed “endophoric” (Diessel 1999a: 93) or “tracking” use (Himmelmann 1996, 1997), which relates to concepts such as anaphora and definiteness (discussed in more details in Section 4.3).

4.2 GRAMMATICALISATION OF DEMONSTRATIVES TO DEFINITE ARTICLES

Diachronically, demonstratives play a role in terms of being a source for a number of grammaticalisation targets, particularly definite articles (e.g. Lehmann 1985; Hopper & Traugott 1993; Harris & Campbell 1995; Diessel 1999a; Heine & Kuteva 2002). Heine and Kuteva (2002: 106–116) mention eight possible targets for demonstratives:

1. Complementiser ‘that’
2. Clause connector ‘and, then’
3. Copula ‘that (is)’
4. Definite article ‘the’
5. Focus marker ‘so, just/only this’
6. 3rd person pronoun ‘he, she, it’
7. Relative pronoun ‘that’
8. Subordinator ‘when’

What is relevant for the purposes of the present study are definite articles and focus markers, both of which will be discussed in more detail in Section 4.3. At the same time, definite articles can originate from various lexical sources, above all, personal pronouns, demonstratives, numerals or even case suffixes. However, the current study primarily focuses on the grammaticalisation path from demonstrative to definite article, which is said to be taking place in some Finnic and Slavic languages (see the literature review in Chapter 5).

From the perspective of their morphosyntactic properties, Schroeder (2006) gives the following definition for definite articles:

‘[Definite articles are] a morphological device (free morpheme, enclitic morpheme, or affix) with the primary function to denote the definiteness of noun phrases with anaphoric and/or non-anaphoric reference.’

(Schroeder 2006: 553)

Meanwhile, a functional approach in Epstein (2001) defines functions of definite articles as follows:

‘... speakers select the definite article for a number of reasons: to distinguish (identify) discourse entities, certainly, but also to convey the prominence of a discourse entity, an entity’s status as a role function, or a shift in point of view.’

(Epstein 2001: 335)

Considering these aspects of definite articles, the present study adopts the following concise definition by Becker (2018):

‘An article [both definite and indefinite] is a marker that systematically occurs with a noun and whose primary function is to indicate the referential function of the noun.’

Becker (2018: 25)

This notion will become crucial when defining the Finnic and North Russian postposed demonstratives, which violate the systematic occurrence with noun (see further discussion in Sections 6.2 and 7.1.1).

Signs of grammaticalisation from demonstratives to definite articles can be observed in different areas of grammar. Diessel (1999a: 118) proposes that at least some of the following changes typically take place when demonstratives start developing to definite articles, as shown in Table 7.

Change	Area of grammar
1. They may have undergone a process of phonological reduction.	Phonology
2. They may have coalesced with other free forms.	
3. They are usually restricted to the distal or, less frequently, the proximal form.	Morphology
4. They may have lost their ability to inflect.	
5. Their occurrence is often restricted to a particular syntactic context.	Syntax
6. They are often obligatory to form a certain grammatical construction.	
7. Grammatical items that developed from demonstratives are no longer used to focus the addressee’s attention on entities in the outside world.	Pragmatics
8. They are deictically non-contrastive.	

Table 7. *Changes evolving in the grammaticalisation from demonstratives to definite articles (Diessel 1999a: 118)*

The previous diachronic descriptions of this grammaticalisation path (e.g., Hawkins 1978: 149–157; Diessel 1999b: 19–20, 25; C. Lyons 1999: 17–21; Givón 2001: 468–469) state that articles are often (but not always) a phonologically reduced or weakened form of the demonstrative that is originally a free morpheme but later loses stress and turns into a bound morpheme. Syntactically, a wide range of contexts of an adnominal use of demonstratives is narrowed to specific contexts where the use becomes obligatory (see also Greenberg 1978: 61). Often, but not always, the position of articles seems to be identical to the independent adnominal demonstrative in the earlier stage (see also C. Lyons 1999: 63–64). For instance, postnominal demonstratives in Scandinavian languages have turned into enclitic articles, as in Swedish *huset* ‘the house’ (< *hus* ‘house’ + **hit* [DEM]). Semantically, the grammaticalised demonstrative loses the ability to match a linguistic referent with some identifiable object on the basis of visibility or shared knowledge (Hawkins 1978:

154–157; Dixon 2003: 68–69). Therefore, the definite articles, as opposed to the demonstratives, tend to lose an ability to express a proximity contrast between a referred entity and the interlocutors, i.e. deixis, but has shifted to non-deictic functions, above all, tracking reference to its antecedent and establishing new referent, i.e. anaphora and cataphora (discussed further in Sections 4.3.2 and 4.3.3).

Across the world’s languages, articles are not universally attested as widely as demonstratives are, nor in the same way. As pointed out by C. Lyons (1999: 48–51), some languages might have both indefinite and definite articles, e.g., Germanic and Romance languages, while other languages might only have one type of articles, e.g., definite article in Irish and Ancient Greek vs. indefinite article in Turkish and Mam. In any case, most of the world’s languages do not have such a lexical category at all, including the Ural-Altaic type of languages spoken across Northern Eurasia (see Dryer 2013b). This variation of articles becomes clear when comparing different grammar traditions. As described by Krámský (1972: 13–29), articles were initially considered a grammatical category in the Stoic grammar (the 3rd century BC) and its follower traditions (e.g., Dionysius Thrax, the 1st century BC) under the category *árrhron*, including articles and relative pronouns. However, the article was later either excluded from the grammar (e.g., Remnius Palæmon 5–65 AD), merged with pronoun under the notion *pronomēn articulare* (e.g., Varron 116–27 AD), or even treated as attribute or determiner (see also the application of a term “definite determiner” in Juvonen 2000; Pajusalu 2001).

Although there are several types of articles as classified by Becker (2018), the current study only discusses properties and functions of definite articles. The geographical distribution of definite articles across the world’s languages has been investigated by Dryer (2013b), as shown in Figure 19.

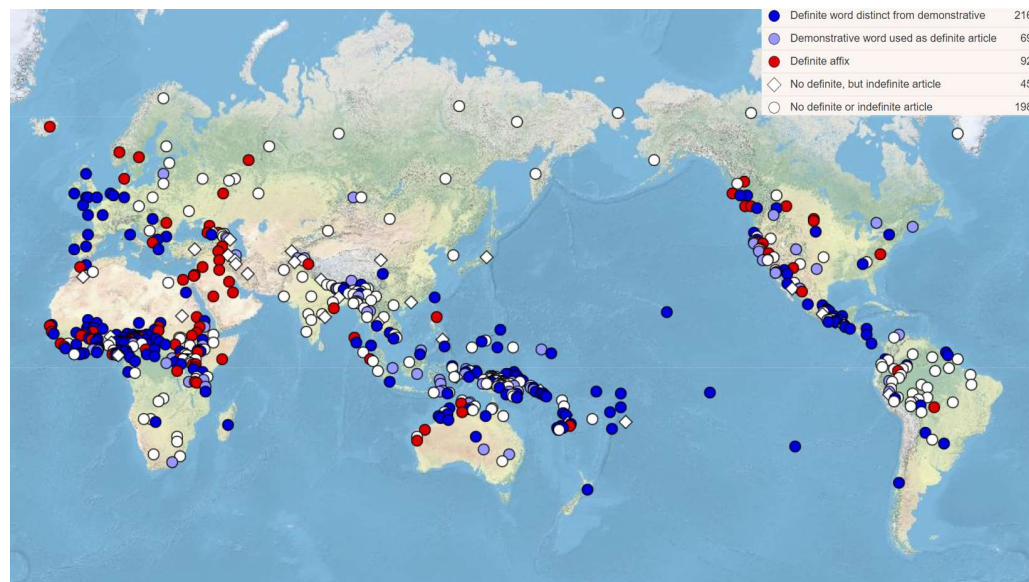


Figure 19 WALS Feature 37A: Definite Articles (Dryer 2013b)

From Figure 19, we see a tendency of not having definite or indefinite articles in Northern Eurasia. Nevertheless, Dryer (2013b) includes two red dots under the label “definite affix” for Mordvinic and Komi; however, the lexical source of the definite affix in Komi is not a demonstrative, but rather a possessive suffix (see the literature review of the phenomenon in Uralic languages in Chapter 5).

The most common view is that the primary context, in which demonstratives occur and which can lead to their grammaticalisation to definite articles, is through adnominal use (Greenberg 1978: 69; Diessel 1999a: 128). In any case, the current study presents another type of example from Northwest Russia, in which definiteness marking observed in postposed demonstratives does not necessarily emerge from the adnominal use, but more plausibly from pronominal use as phrase markers involving topicalisation and focalisation (see further discussion in Section 7.1).

4.3 DEMONSTRATIVES AS A REFERENTIAL DEVICE

Studies concerned with demonstratives may focus on the deictic use as well as other properties beyond spatial deixis such as anaphora, identifiability, and discourse organisation (see Dixon 2003). As for non-deictic aspects of demonstratives, some scholars also emphasise the relevance and role of demonstratives with definiteness, in topic or thematic structure (e.g., Givón 1983: 5; Curl 1999; Epstein 2001; Ihsane & Puskás 2001; Gernsbacher & Robertson 2002). In the scholarly history, definiteness has since the era of Ancient Greek been included as a part of an article’s characteristics. This section will discuss the notion and importance of deixis, anaphora and identifiability, which help to understand the use of demonstratives as markers of definiteness in the languages under investigation.

4.3.1 DEIXIS

A prototypical function of demonstratives is to indicate spatial deixis, that is, the concrete proximity relation between the speaker and an entity to which the demonstrative refers to on the basis of visibility (Hawkins 1978: 155). From the viewpoint of language evolution (e.g., Levinson 2004: 98; Diessel 2013b), such deictic characteristics are one of the most fundamental properties of human language, as gestures can be considered as one of the most primitive strategies in human’s communication. In terms of human language acquisition, a baby obviously starts to acquire the sense of deixis first by pointing, which it later replaces by speech after having acquired necessary vocabularies. Diachronically, this basic spatial-oriented function of demonstratives later gives rise to other secondary uses, in which the concrete space transforms into an abstract field in the discourse (Brugmann 1904: 7–8; Bühler 1934: 390; J. Lyons 1977: 671; Diessel 1999a: 109–113).

As a matter of terminologies, the concept of “deixis” by Bühler (1934) is parallel to “indexicality” by Peirce (cited in Bühler 1940, see also Silverstein 1976) and the selection of term varies according to scholarly traditions. Deixis is in a widely spread concept in linguistics, whereas indexicality is more frequently used in philosophy

(Nunberg 1993: 2). Levinson (2004), however, considers deixis as a narrower aspect of indexicality. As the current study does not focus on theory and the use of terminology, but rather how it is concretely realised in languages, the term “deixis” will be used henceforth.

In the discussion of deixis, the focus usually falls on the relation between entities in a shared deictic field, which minimally always comprises two major components: (i) the “origo” (also “ground zero”), and (ii) the object referred to. In Bühler’s sense (1934: 107), the origo is the deictic centre, from the perspective of which a narrator views the external world. This approach assumes that deixis is always subjective, and any objective attempt to describe the spatial orientation will fail because a certain statement can become false as soon as person, time or space of the origo changes. For instance, the statements (10) to (12) are true only if they are uttered by *me* (not *you*), on *Friday* (not on *Saturday*) and in *Moscow* (not in *P’yŏngyang*), respectively.

- (10) *She has called **me**.*
- (11) *Today is **Friday**.*
- (12) *Vladivostok is very far from **here**.*

In contrast, if statements (10–12) are uttered by *you* (not *me*), on *Saturday* (not on *Friday*) and in *P’yŏngyang* (not in *Moscow*), respectively, the truth value of the original statement immediately changes.

Furthermore, Diessel (1999a: 41) and Dixon (2003) also emphasise that visibility, elevation, geography, and movement play an influential role in the encoding of spatial deixis such as a height-based system in Lak (13), or a system which includes visibility distinctions in Muna (14).

Lak (Northeast Caucasian)

- | | | |
|------|----------------------|---|
| (13) | Demonstrative | Deictic value |
| | <i>aha</i> | close to speaker |
| | <i>hava</i> | farther from speaker, but on the same level |
| | <i>ho</i> | higher or lower than the speaker |
| | | (Khaidakov 1966: 12) |

Muna (Austronesian)

- | | | |
|------|----------------------|---|
| (14) | Demonstrative | Deictic value |
| | <i>aini</i> | near speaker |
| | <i>aitu</i> | near addressee |
| | <i>amaitu</i> | away from speaker and addressee, but nearby |
| | <i>awatu</i> | far away, lower than or level with point of speaking or orientation |
| | <i>atatu</i> | far away, higher than point of speaking or orientation |
| | <i>anagha</i> | not visible (may be audible), unspecified for time |
| | <i>awaghaitu</i> | not visible, was in view but no longer is |
| | | (van den Berg 1997: 199–201) |

As demonstratives are often regarded as a universal linguistic property (see Diessel 1999a: 1, mentioned in Section 4.1), deixis as the primary semantics of demonstratives is consequently also considered universal. Despite that, Levinson (2004: 112) presents a counter argument that not all languages necessarily express deictic categories in their grammar, since some languages might lack contrast in personal pronouns, motion verbs, tenses, or spatial demonstratives. Instead, it is the expression of deictic categories that is universal because even speakers of a language without the aforementioned contrastive person, temporal and spatial category will find a way to express deictic contrasts by another grammatical or lexical strategy. For example, Nimboran (a Papuan language spoken in Nimboran District, Jayapura Regency, Indonesia) has no spatial distinctions in demonstratives. It only has one demonstrative, which can be adverbial manner, locative, referential, or function as a modifier, making it very multifunctional (Don Killian, p.c.).

Previous studies on the functional dimensions of demonstratives provide various approaches to categorise the usage of demonstratives. The most basic classification by Levinson (1983) and Himmelmann (1996, 1997) operates on the contrast between two main uses. First, a usage of demonstratives to indicate reference to an entity in a speech situation is termed “deictic” or “exophoric” (à Levinson) or “situational” use (à Himmelmann). Second, a strategy to track elements in the ongoing discourse on a more abstract level is termed “non-deictic” or “endophoric” (à Levinson) or “tracking” use (à Himmelmann). In a similar vein to Levinson, Diessel (1999a: 93–109) extends Levinson’s classification into four categories by further splitting the endophoric use into three subfunctions: (i) anaphoric use, (ii) discourse deictic use, and (iii) recognitional use. Later, Levinson (2004) remaps the overall functional possibilities of demonstratives onto a more fine-grained scheme operating on the dichotomy of deictic vs. non-deictic uses, as shown in Figure 20.

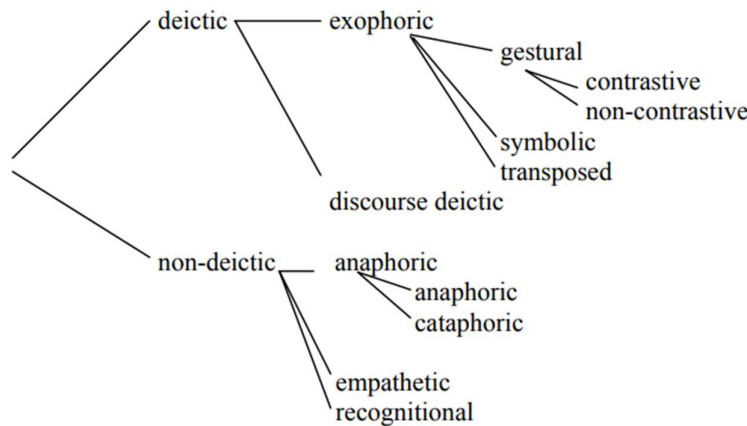


Figure 20 The distinct uses of demonstratives (Levinson 2004: 108)

Contrasting earlier descriptions by Himmelmann and Diessel, Levinson’s classification is more extensive as it goes further in distinguishing more uses under the exophoric and anaphoric uses (see concrete examples in connection to

identifiability in Section 4.3.3). However, a recent study by Talmy (2017) argues that the exophoric and endophoric uses of deictic indexicals cannot be always distinguished and there is often no need to do so. Therefore, he rather considers both functional categories present in Levinson's classification as a single system called "targeting" (see also a review of Talmy's approach in Huumo & Sivonen 2020).

From a cognitive perspective, Bühler (1934) views the origo as a specific kind of co-ordinate system as follows:

... das Koordinatensystem der „subjektiven Orientierung“, in welcher alle Verkehrspartner befangen sind und befangen bleiben.

'... the co-ordinate system of the subjective orientation, in which all the interlocutors are encompassed and become encompassed.' (the author's translation)

(Bühler 1934: 102–103)

This idea of the subjective orientation is also partially reflected in the theory of "Situation Semantics" (Barwise & Perry 1983), which propose three steps of processing in the interpretation of the utterance. First, the speaker maps a set of indices for both the interlocutors as well as the time and space of the utterance into intentions to foreground an utterance situation, i.e. an utterance is performed by whom, when, and where. Second, a speaker determines what other references (such as anaphora) are present in the context to establish a resource situation, i.e. what kinds of entities are present at the time of utterance. Third, a speaker places the propositional content into the real world to create a described situation, i.e. placing the right information to the utterance, e.g., *I wrote this book* is transferred into *Jæng wrote this book*. It is clear here that deixis often coexists with anaphora to the same referent, which can give a two-fold interpretation. Particularly for discourse deixis expressed by demonstratives, deictic interpretation is often mixed up with the anaphoric reading in interpretation (Nunberg 1993).

As discussed earlier, gesture and deixis, in general, are attentional, intentional and subjective (see also discussion on intersubjectivity in Section 4.5.1). Across the world's languages, there is a wide range of variation and complexity of deictic systems, in which the cultural and environmental setting plays a significant role (see also Dixon 2003; Sidnell & Enfield 2017). Assumedly, languages in the same linguistic area with a common cultural and natural living environment would also share similar deictic systems in the areal-typological sense, while the genealogically cognate languages might not necessarily do so (see concrete examples from the Finnic languages in Section 7.2.2).

As mentioned in the previous section, demonstratives without deictic values also exist in the world's languages which can also relate to the emergence of a definite article (Halliday 1994: 180–186; C. Lyons 1999: 17–21). In the grammaticalisation path from demonstrative to definite article, one of the major changes in semantics is the weakening or loss of deictic uses of demonstratives (C. Lyons 1977: 653–654; Hawkins 1978: 149–157; Anderson & Keenan 1985: 280), which is, to a certain

degree, observed in the case of postposed demonstratives in Finnic and North Russian (see further discussion in Section 6.4.1). Therefore, it is crucial to discuss also other non-deictic uses of demonstratives, especially anaphora, which is one of the main functions of definite articles.

4.3.2 ANAPHORA

As opposed to the concrete spatial orientation in deixis, anaphora relates to a reference strategy on the abstract level in the discourse. Huang (2000: 1–3) discusses the notion of anaphora, the term which goes back to Greek *anaphorá*, meaning ‘carrying back’ because anaphora refers to a relation between two linguistic elements. A target of the interpretation, i.e. “anaphor”, can be syntactically expressed by gap or empty category (15), pronoun (16), reflexives (17), and names (18). The anaphoric element (ANA) is to be identified by the other preceding element in the context or discourse, i.e. by its “antecedent” (ANT).

- (15) *Steve*_[ANT] comes from America and *Ø*_[ANA] likes baseball.
- (16) *Steve*_[ANT] comes America and *he*_[ANA] likes baseball.
- (17) *Steve*_[ANT1] said that *Stephanie*_[ANT2] hit *him*_[ANA1]/*herself*_[ANA2].
- (18) *Steve*_[ANT] thinks *Steve*_[ANA] is smart.

What can be referred to in the language might not necessarily only be a concrete object in the “real world”, but also fictional or non-existent entities, as long as they exist in the “universe of discourse” (Givón 2001: 438; Carlson 2004: 94–96). Two syntactic categories that tend to host anaphora are nominal expressions and verbs. As the present study only focuses on the use of demonstratives, the anaphora within verb phrase such as person and tense marking will not be discussed here.

As for discourse reference-tracking systems in general, a high degree of variation in morphosyntactic marking is observed in the world’s languages. This, again, returns to the question of universality as discussed earlier in the previous section in connection to deixis. Similarly to the claim by Levinson (2004: 112) about deictic categories, anaphora being expressed by grammatical markers is likewise not universal, but the expression of anaphora is common among all the human languages. Some languages may encode anaphora by a strategy such as gender/class system in Bantu languages, whereas the others may encode anaphora onto verbs such as the phenomenon of incorporation in Paleo-Siberian languages (e.g., Ainu and Chukotko-Lamchatkan) and objective conjugation in Uralic languages (see also discussion on Mordvinic and other eastern Uralic languages in Section 5.1.3).

When discussing the anaphoric use of demonstratives and definite articles, it is also important to justify different types of anaphor that demonstratives and definite articles can express. From the semantic perspective, Huang (2000a: 5–7) distinguishes four types of anaphora, in which pronouns and definite articles, especially, are among the primary encoding devices. This understanding of different anaphora types illustrates a connection between gesturing properties of demonstratives (deixis in

Section 4.3.1) and their extended uses as markers to encode identifiability (discussed in Section 4.3.3).

First, a “referential anaphor” refers to a unique fixed entity by establishing a co-reference to its antecedent in the same discourse, as in (19).

- (19) *Gorbachev*_[ANT] knew that *he*_[ANA] would be remembered as *the architect of perestroika*_[ANA].
(Huang 2000: 5)

In (19), a constituent ‘the architect of perestroika’ refers to its first mention ‘Gorbachev’ (see identifiability by anaphora in Section 4.3.3).

Second, if the anaphor lacks reference to a specific entity in the external world, the anaphor can establish a discourse-internal co-reference to its antecedent, i.e. “bound-variable anaphor”, which is ‘interpretable by the virtue of its dependency on some quantificational expression’ (Huang 2000a: 235). This type of anaphor can be expressed by similar strategies to (15–18) as well as by inference (20).

- (20) *Every doctoral student*_[ANT] thinks that *the supervisor*_[ANA] is intelligent. (Huang 2000: 6)

In (20), the bound-variable interpretation emerges from an inference that every doctoral student has a supervisor, i.e. the existence of ‘supervisor’ is affirmed by its dependency on the existence of ‘doctoral student’ (see a similar reading for a “situationally unique referent” in Section 4.3.3)

Third, “lazy anaphor” is a device for a repeated linguistic form, which does not refer to the same entity as the antecedent does, such as the repetition of ‘paycheque’ in (21).

- (21) *The man who gave his paycheque to his wife was wiser than the man who gave it to his mistress.* (Huang 2000: 7)

In (21), a pronoun of laziness *it* does not refer to same paycheque as mentioned formerly, as they are two different entities in the external world and, thus, have no coreferential relation, unlike the types of referential anaphor or bound-variable anaphora discussed above.

Fourth, “bridging cross-reference anaphora” also creates an association to some preceding expression in the same clause or discourse segment. However, it does not operate by direct reference to a certain antecedent (as in the first two types) but by the unambiguous semantic link between the antecedent and the bridging referent (also termed “anchoring set” and “trigger”, respectively, in Birner and Ward 1998), such as the inference in (22) that a concert is usually decorated by chandeliers.

- (22) *John walked into a concert hall*_[ANT]. *The chandeliers*_[ANA] were magnificent. (Huang 2000: 7)

Interesting and relevant to this study, Huang states that bridging cross-reference anaphor, particularly, is mostly encoded as a definite noun phrase (see similar views from the perspective of definiteness and identifiability in Section 4.3.3).

4.3.3 IDENTIFIABILITY AS A PROPERTY OF DEFINITE ARTICLES

Demonstratives and articles are closely tied to the notion of definiteness, such as the dichotomy of the indefinite and the definite article observed in the Standard Average European languages (Haspelmath 1998: 274). The current study adopts the definition by Sperber and Wilson (1986) that definiteness relates to the interlocutors' mutual identifiability and commitment to that knowledge. In discourse, an entity that is encoded as definite often refers to shared knowledge between the interlocutors (see also "Common Ground" in Krifka 2008, discussed in Section 4.4.1), while an indefinite expression is speaker-oriented information that is new or previously unknown to the addressee (Christophersen 1939: 28; Prince 1992).

Various studies have indicated what kind of grammatical elements are used for coding definite referents. As illustrated by Carlson (2004: 122–123), besides proper names and pronouns, a noun phrase can receive marking by a demonstrative (*this*, *that*), a definiteness marker (*the*), a possessive marker (*my*, *your*, etc.), a quantifier (*all*, *every*, etc.), a numeral classifier (*one piece of*), and a generic index when the context requires the expression to be definite (see Westerstål 1985; Ariel 1988, 1990; Prince 1992; Birner & Ward 1998). As the current study only concerns the definite use of demonstratives, the indefinite domain will be excluded from the remainder of this study.

The use of definite articles entails reference to an entity which is unique and identifiable in a given context (König 2018: 168). Abbott (2004: 125) also claims that the marking of uniqueness correlates with the contrast between definite and indefinite description, of which the definite one is usually unique referent. Concerning identifiability in the discourse, the speaker often evaluates whether the addressee could identify a unique referent that they are about to code in a definite form and utter. This emphasises the addressee-oriented relation of definiteness and identifiability (Andrews 2007: 148), while Etelämäki (2006: 25) also argues for the importance of speaker-oriented viewpoint, i.e. the speaker's mental proximity to information. Ultimately, both interlocutors make their contribution to interpretation (Gibbs 2017: 321). Nevertheless, these two different views show that the use of definite articles is based on the interlocutors' intersubjectivity how accessible a piece of information is to both interlocutors (see Etelämäki 2009, and further discussion in Section 4.5.1).

Definite articles are devices to identify a unique entity and can appear in various contexts in which they are identifiable. Such contexts have been discussed by Hawkins (1978: 106–130), C. Lyons (1999: 3–7), Becker (2018: 76–82) and König (2018: 169), among others. However, the mentioned previous studies often use different names for what are actually same concepts, so the discussion below will also take the difference in the choice of terminologies into account.

First, "situational" identifiability refers to knowledge that the addressee acquires only in the situation of utterance as in (23).

(23) *Pass me the salt.* (König 2018: 169)

This type of identifiability often contains a gesturing meaning that points to an entity in a visible range of the interlocutors, corresponding to what Diessel (1999a) refers to

as an “exophoric” use of demonstratives and what Becker (2018) classifies as a “deictic” use.

Otherwise, identifiability can also come from “general background knowledge” of the world that the interlocutors possess, as in (24) and (25).

(24) *Today **the sun** is shining.*

(25) ***The Pope** will come to Paris.* (König 2018: 169)

This kind of referent is usually a unique entity and unambiguous by itself, so Becker (2018) also names this context as an “absolutely unique referent” (see also “generic index” in Section 4.4.2).

On the discourse level, identifiability can emerge through “anaphora” (discussed in Section 4.3.2), i.e. the information that has been earlier mentioned or presented in the discourse, which Prince (1992) would classify as “discourse-old” as in (26).

(26) ***Somebody** stole my bike yesterday but they have already found **the thief**.* (König 2018: 169)

In (26), ‘the thief’ refers to its antecedent ‘somebody’, both of which stand for identical referents in this particular context and interpretation.

Similarly to anaphora, a term “bridging reference” (introduced by Clark 1975) also involves reference-tracking. However, in this case, the definite referent is not identical to its antecedent but creates some associative relationship to the antecedent, or an inferable relation (in Birner’s term 2006), such as part-whole or action-instrument (see also a detailed definition in Schwarz 2013: 536), for instance as in (27).

(27) *We laid out **the picnic**. **The coffee** was still warm.* (König 2018: 169)

This nature of association and inclusiveness between the unidentical referent and antecedent might be the motivation for a synonymic term “associative anaphora” by Hawkins (1978, see also C. Lyons 1999; Fraurud 2001; König 2018). C. Lyons (1999: 4–5) regards this type of identifiability as a functional combination of general background knowledge and discourse anaphora because in order to be able to connect unidentical entities, the speaker must have general background knowledge about the associative relationship, e.g., that people usually drink coffee when going picnic as in (27). The definiteness effect in possessive suffixes is also thought to have emerged or been inherited from associative anaphora (Fraurud 2001). As discussed in Section 4.3.2, Huang (2000a) states that bridging reference is one of the properties that has been developed under the grammaticalisation of demonstrative to definite article.

Another anaphora-like context is “larger situation” use discussed by Hawkins (1978: 115–123). Becker (2018: 80) rather names this as “situationally unique” referent because the referent is unique and identifiable through inference and an unambiguous link to the discourse situation, in which the speaker and the addressee are situated. Often, the coreferential link concerns a larger context of time or space

such as a construction, village, city, or country in which the discourse situation takes place as in (28).

- (28) A real estate agent is showing a house to customers. Once they enter the house, one customer asks:
*Where is **the kitchen**?*

Given a context in which the customer is in the house when uttering this question, it is unambiguous that ‘the kitchen’ in (28) refers to the one located in that house.

Becker (2018) and König (2018) also discuss two other uses that are not included in C. Lyons’ classification (1999). The first one is a “cataphoric” use, or “establishing” referent in Hawkins’ terminology (1978). Unlike anaphora, the speaker introduces a new referent completely new to the discourse which subsequently becomes identifiable through elaboration with more description afterwards as in (29).

- (29) ***The book** I bought yesterday is under discussion for the Nobel Prize.* (König 2018: 169)

Hawkins (1978: 150) and Himmelmann (1997: 93–101) argue that establishing reference, particularly in relative clauses, is one of the characteristic abilities of definite articles, something not typical for demonstratives.

Another context is “recognitional” use (also called *emploi mémoriel* in König 2018), in which definite articles do not create identifiability through the discourse but through personal memory with partial description as in (30).

- (30) *You remember **the restaurant** we went to recently. That is where I found a wallet.*
(König 2018: 169)

However, Becker (2018: 86–87) does not regard the recognitional use as a decisive context to identify definite articles as her typological database of articles in the world’s languages shows that some languages can have one type of article that is exclusively reserved for recognitional, but not for other referential uses.

As is clear from the discussion above, scholars have given different names for the same contexts of use. As a solution, the present study chooses to adopt Becker’s terms (2018), given that she has successfully used these terms to analyse articles in the world’s languages (cf. the language coverage consisting of only English in Hawkins 1978, and mostly European languages in König 2018).

Regarding the matter of grammaticalisation, a grammatical element can be considered as a grammaticalised definiteness marker when it develops a minimal capability to encode (i) anaphoric contexts, (ii) bridging contexts, (iii) situationally unique contexts, and (iv) establishing referent contexts (Becker 2018: 86), even though it may also encode (v) deictic, and (vi) recognitional referents (König 2018: 169). More importantly, the use of definite articles should be compulsory when expressing these reference relations, in order to claim the complete process of grammaticalisation (Greenberg 1978: 61). The present study will investigate whether the postposed demonstratives in Finnic and North Russian can fulfil this qualification of definite articles (see Section 6.4.1).

At the morphosyntactic level, there is a variable tendency in which syntactic function a definite expression is more likely to occur. Givón (2001: 473–474) proposes that such syntactic functions like subject, indirect object and comitative (‘with someone’) are more often definite due to their topichood and animacy in most cases; the definiteness status of direct object, in contrast, rather depends on context (see discussion on “differential object marking” in Comrie 1978; Aissen 2003; Dalrymple & Nikolaeva 2011). Besides these core arguments, adverbial functions such as location and time are often coded as definite; in natural discourse, these referents set a specific background for the conversation (see Birner & Ward 1998, and also “scene setting” in Chafe 1976, and “frame setting” in Krifka 2008, discussed in Section 4.4.1). This assumption will also be tested in the current study on whether it also applies to the use of postposed demonstratives in Finnic and North Russian (see Section 6.3).

As demonstratives and definite articles can also function in other non-referential uses (see Chesterman 1991: 142–148, 187), the next section will discuss how demonstratives can be used to modify information structure.

4.4 DEMONSTRATIVES IN INFORMATION STRUCTURE

More recently, a new functional and discourse analysis approach to demonstratives and definite articles has identified many other non-referential uses of demonstratives and definite articles for modifying information structure. Lambrecht (1994: 1–13) defines information structure as the interrelation between linguistic forms and the interlocutors’ state of mind: how they organise old and new pieces of information in discourse. Therefore, the information structure approach considers speech as communication activity. The similar view is shared by Dalrymple and Nikolaeva (2011) who give the following definition of information structure (see also a similar definition in Matić 2015):

‘We view exchange of information as the main function of language. Information structure is the level of sentence organisation which represents how the speaker structures the utterance in context in order to facilitate information exchange. Specifically, it indicates how the propositional content of an utterance fits the addressee’s perceived state of knowledge at the time of utterance. In human communication, new information is normally added to the already existing store of knowledge in the addressee’s mind. The distinction between familiar knowledge and the informational contribution of an utterance is manifested linguistically: propositions can receive different formal expression (are “packaged”) in accordance with what the speaker assumes to be old or new information for the addressee.’

(Dalrymple & Nikolaeva 2011: 45)

Given this nature, the term information structure is also called “information packaging” by some scholars (e.g., Chafe 1976: 28).

Apart from the knowledge-based description, information structure also relates to the dichotomy “old” vs. “new” information (Lambrecht 1994: 61), which can be regarded as synonymic to “theme” vs. “rheme” (Halliday 1994: 299–302). In discourse, a speaker makes an assumption whether the information they are going to utter is old or new in the discourse and to the addressee (see a similar process within “identifiability”, discussed in Section 4.3.3). Alternatively, information structure is also viewed as a codification of information updating. This view uses an analogy of file construction, in which file cards correspond to different referents in the discourse, and each file card bears record on the information about referents they represent, whether the card is new or old (see Karttunen 1969, 1976; Vallduví & Engdahl 1996; Erteschik-Shir 1997, 2007).

Regardless of the description manners, all the aforementioned descriptions commonly emphasise that the encoding of contents and referents in the information structure are also reflected in their grammatical marking at the language interface. Morphosyntactically, familiar knowledge is usually coded in speech as definite, while new information is coded as indefinite. In discourse, the speaker can manipulate this basic principle by information-structuring devices, a strategy to rearrange the information packaging. Significant components in information structure are the concepts of topic, focus, and contrast, which are discussed next.

4.4.1 DEFINITIONS AND BEHAVIOURS OF TOPIC, FOCUS, AND CONTRAST

The existence of topic in the discourse has been previously defined by a number of dichotomies: topic vs. comment (Andrews 2007), theme vs. rheme (Halliday 1994), given vs. new information (Dik 1989), presupposition vs. assertion (Lambrecht 1994), background vs. completive (Dalrymple & Nikolaeva 2011), or link vs. tail (Vallduví 1992). Importantly, this information-structural distinction also plays a crucial role in the morphosyntactic encoding in Uralic languages (Vilkuna 1998) and Russian (Grenoble 1998), in which demonstratives, among others, can function as marking devices for such purposes (see Chapter 5).

Despite the differences in naming conventions, the majority of the aforementioned frameworks point to the core function of topics, which is to define what a statement uttered in a certain moment is “about”, and relates a referent to its proposition (see Strawson 1964; Reinhart 1982; Dik 1989; Lambrecht 1994; Dalrymple & Nikolaeva 2011). Given that topics correspond to the core of an utterance, some scholars even claim that any utterance without topic is meaningless, while an utterance with multiple topics is fine (Givón 1983: 5–6; Erteschik-Shir 2007: 16–17). In any case, a topic need not be present in a clause, because the situation where a statement is uttered can already be considered as a given topic (Marty 1884).

Topics can also be regarded as devices which manage the interlocutors’ shared state of knowledge or “common ground” (a term introduced by Reinhart 1982), defined by Krifka (2008) as follows.

‘The topic constituent identifies the entity or set of entities under which the information expressed in the comment constituent should be stored in the C[ommon]G[round] content.’

(Krifka 2008: 265)

From this definition, we also see that a close relation between the topic and “comment” (see further discussion below). Due to this characteristic, topics are sometimes seen as “scene setting” (Chafe 1976). However, Lambrecht (1994: 118) points out a problem that the scene-setting function only applies to certain types of languages which often use clause-initial adverbial phrase, such as Chinese. Due to this language-specific characteristic, Krifka (2008: 269) does not treat such “frame setters” as topics but rather as information that instructs how the following expression should be interpreted, e.g., adverbial phrases like ‘health-wise’, ‘spiritually’ and ‘financially.’ In any case, this background-setting characteristic of topics also motivates Haiman’s idea (1978: 564) that topics are semantically equal to conditionals because ‘[they both] are givens which constitute the frame of reference with respect to which the main clause is either true (if a proposition), or felicitous (if not)’ (see concrete examples from Russian in Section 5.2.2).

As the typical place in the clause for topic and subject is similarly clause-initial position in non-verb-initial languages (Li & Thompson 1976; Keenan 1976), the behaviour of topics has often been discussed in parallel to subjects. Among others, Li and Thompson (1976: 461–466) give the following seven aspects where differences between topics and subjects can be found.

1. Definiteness
2. Selection relation with verb
3. Determination by verb
4. Verb agreement
5. Semantic role
6. Clause-initial position
7. Grammatical processes

First and foremost, a topic is often, but not necessarily always, definite (see also Givón 2001: 472–473) as the speaker evaluates that the topic entity is identifiable to the addressee, i.e. ‘an existing representation in memory’ (Gundel 1985, 1988), while a subject need not do so and can be indefinite. In terms of their relation to verb, a topic needs not have a selectional relation with any verb or predicative in the clause, i.e. be a verb argument, which also means that it is neither necessarily determined nor agrees with verb, while a subject is always a verb argument that is controlled by a verb. Semantically, a topic maintains a constant semantic role across clauses, as it provides a frame setting (in Krifka’s term 2008) for the main predication to take place (see also Chafe 1976). Syntactically, the default place for topic is clause-initial position (see also Keenan 1976), while a subject in languages with less strict word order, such as

Finnic and Slavic, can move (see, e.g., discussion of the free word order in Finnish by Vilkuna 1989). Last but not least, a topic does not involve grammatical processes such as reflexivisation and passivisation, which have a direct effect on subject.

Two relevant aspects of topics for the present study are definiteness and clause-initial position. Definiteness as a semantic characteristic of topics indicates that they are always identifiable, based on any of the criteria discussed in Section 4.3.3. Meanwhile, clause-initial position as a syntactic characteristic of topics, for instance, implies that topichood may also create other non-referential uses for the postposed demonstratives in the clause-initial topic constituents, such as evaluative uses (discussed in Section 4.5.2).

In contrast to topics that contain identifiable information, foci usually give “new information” (Halliday 1966; Selkirk 1984) that does not belong to the “presupposition” (Jackendoff 1972) with no “semantic antecedent,” and thus is not actively in the addressee’s attention (Rochemont 1986). While topics usually involve information packaging, foci rather create a scope of relevant alternatives from which an entity is selected (Jacobs 1983, 1988; Rooth 1985; König 1991: 46). A similar definition based the set of relevant alternatives is given by Krifka (2008) as follows.

‘Focus indicates the presence of alternatives that are relevant for the interpretation of linguistic expressions.’

(Krifka 2008: 247)

Based on the characteristics given above, foci are unpredictable and non-recoverable elements in the utterance (Lambrecht 1994: 207), but they capture the addressee’s attention for the new information (Erteschik-Shir 1986). This shows that the degree of identifiability within foci is considerably lower than that of topics, implying that foci are less frequently definite.

Foci also behave differently from topics in a clause. The position of a referent in focus in a clause varies, and depends on the constituent which receives focus marking (König 1991: 12), such as different stress patterns shown in (31).

(31a) *Surprisingly, George is **RUNNING** to Brooklyn.*

(31b) *Surprisingly, George is running to **BROOKLYN**.* (König 1991: 12)

Due to different places of stress, the foci in both examples clearly differ in their meaning in that the reason for being surprised focuses on the action ‘running’ in (31a), but on the destination ‘Brooklyn’ in (31b).

The discussion above shows that the presence of topic and focus also entails their complementary parts: “comment” and “presupposition”, respectively (or “completive” and “background” in Dalrymple & Nikolaeva 2011). Typically, both topic and focus constituents point to a single referent, but comment and presupposition can cover a larger unit of information, e.g., a complete verb phrase (see König 1991: 11; Andrews 2007: 148–151). In any case, this view is challenged by Lambrecht

(1994) who claims that focus can also be a lexical unit larger than a single phrase. Interestingly, Dalrymple and Nikolaeva (2011: 102–106) argue that the topic-comment structure is morphosyntactically an unmarked information structure (cf. Lambrecht 1994), which, at the same time, seems to entail that presupposition-focus structure should be, then, a marked construction. This might be true for such languages in which topics do not receive special morphological marking but foci do, such as in Germanic languages (see König 1991).

Similarly to the definition of focus discussed above, “contrast” also concerns the selection and exclusion of an entity from other members of the set, as defined by Neeleman et al. (2009) as given below.

‘Constituents that are contrastive are understood to belong to a contextually given set out of which they are selected to the exclusion of at least some other members of the set. Both topics and foci can be interpreted contrastively.’

(Neeleman et al. 2009: 17–18)

Moreover, contrast excludes an alternative, which is predicted or previously stated (Halliday 1966: 206), by presenting a “restricted universe” (Longacre 1996: 55). This exhaustive characteristic of contrast is frequently observed in the so-called *it*-cleft construction, as in (32).

(32) *It was **a HAT** that Mary picked for herself.* (Umbach 2004: 159)

In (32), ‘a hat’ presents ‘a subset of the set of contextually or situationally given elements for which the predicate phrase can potentially hold’ and ‘it is identified as the exhaustive subset of this set for which the predicate actually holds’ (Kiss 1998: 45). Compared to the referential domain, the exhaustive and selective characteristic of contrast resembles the use of situationally unique referent (see Section 4.3.3). In any case, the difference from focus is that contrast is not necessarily a category of grammar, but ‘the result of general cognitive processes referred to as conversational implicatures’ (Lambrecht 1994: 291).

Contrast can emerge from several factors, as discussed in Umbach (2004). In order to be contrastive, an entity should be similar or dissimilar from other members of the set, as in (33).

(33) *John bought **the beer**, and/but Mary bought **the port**.* (Umbach 2004: 162)

In (33), ‘the port’ is understood as belonging to the similar set with ‘the beer’, that is, a drink (similarity), while simultaneously creating a contrast to the other interpretation of ‘harbour’ (dissimilarity). Moreover, contrast also involves exclusion of an entity from the other possible alternatives in the set, such as in (34).

- (34) A: *Mary made the salad, and Anna made the hamburgers.*
 B: **RONALD** made the hamburgers. (Umbach 2004: 164)

The exhaustive characteristic of contrast uttered in B makes the statement stronger than in A, which gives the interpretation of B being “correction” of A.

As contrast can play a role in the interpretation of topics and foci, a number of scholars (e.g., Vallduví 1992; Vallduví & Vilkkuna 1998; Frey 2001; Neeleman et al. 2009; Neeleman & Vermeulen 2012) treat contrast as a category that involves the formation of information structure, as summarised in Table 8.

	Topic	Focus
	Aboutness topic [topic]	New information focus [focus]
Contrast	Contrastive topic [topic + contrast]	Contrastive focus [focus + contrast]

Table 8. *Information structural system (Neeleman et al. 2009: 15; Neeleman & Vermeulen 2012: 5)*

As a motivation for this parameter, Neeleman and Vermeulen (2012: 25) argue against the previous studies that do not treat contrast as a grammatical category, leading to the indifferent treatment between new information and contrastive focus (cf., e.g., Lambrecht 1994: 291; Krifka 2008). Based on their parameter, contrast as a grammatical category can be combined with topic and focus to create a new type of contrastive topic and focus. However, this semantic approach does not provide enough information on pragmatic uses of topics and foci, as well as ignoring many other characteristics of topic and focus such as definiteness and delimitation of a set of relevant alternatives discussed above. Therefore, the following sections will discuss different types of topic and focus, which have been proposed on the basis of cross-linguistic studies, by paying close attention to the pragmatics of topics and foci.

4.4.2 TYPES OF TOPICS

As discussed in the previous section, topics are often definite because they refer to an entity identifiable to the addressee as given information. Givón (1983) also emphasises the nature of “predictability” and “continuity” of topics. This entails that one of the most typical functions of topics is to track referents from earlier mention in the discourse, comparable to anaphoric use in the referential domain (discussed in Section 4.3.3). These functions of topic are named “expected topic” (Erteschik-Shir 1997: 18; Andrews 2007: 149), “given topic” (Dik 1989: 267), “continuing topic” (e.g., Pixabaj & England 2011; Jun 2015), or “theme” (Halliday 1994), meaning that the topic refers to a type of referent that the addressee, based on earlier mention, expects to hear as topic in the following utterance, as in (35).

- (35) *We built **a new house** over ten years ago. **The house** needs renovation now.*

In (35), ‘a new house’ is mentioned and still being active in the discourse, so it can be promoted to topic in the following expression. When occurring in the first utterance of discourse, a starting topic often refers to the first and the second person sphere, because the speaker and addressee are always participants of the discourse that need no background setting (Erteschik-Shir 1997: 18).

Another type of topic involves “switch of topic”. One may include this as a subtype of “contrastive topic”. However, the present study emphasises the aspect of pragmatic use and prefers to use a name that describes the function best: that is, “switch-topic” (Andrews 2007: 149). This type of topic can, on the one hand, mark switch of topic from a previous discourse context or, it can also reintroduce a previous referent that has been idle in the discourse for a period of time, which the addressee does not necessarily expect to hear (cf. expected topic discussed above), as in (36).

- (36) Bill mentioned about his brother’s wedding to Emily who is also coming. After their long conversation on other topics, Emily wants to make sure that she remembers the date of the wedding correctly, so she asks Bill:

*By the way, **the wedding** is next Saturday, right?*

In (36), ‘the wedding’ is reintroduced again but is still identifiable to Bill after not being mentioned for a long time in the discourse.

In addition to the expected topic and switch-topic, a “generic” index can often be a topic regardless of context, because they are unique entities in the real world (Givón 1983: 10). This type of topic resembles the absolutely unique referents in the referential domain (see Section 4.3.3) as it usually concerns encyclopaedic knowledge or significant historical events as in (37).

- (37) ***The Sun** is so large that a million Earths could fit inside it.*
(An article in Quora: [quora.com/Why-is-the-sun-so-powerful](https://www.quora.com/Why-is-the-sun-so-powerful))

In (37), ‘the Sun,’ as general knowledge that is identifiable to the interlocutors, can be a topic without any foregrounding context. Due to this characteristic, Erteschik-Shir (2007: 17) also gives a term “permanently available topic” for such generic indices.

Sometimes, a clause can also contain “multiple topics”, meaning that an utterance is about multiple entities, such as in (38).

- (38) A: *Did **you** see **Peter** yesterday?*
B: *Yes, **I** saw **him** at school.* (Erteschik-Shir 2007: 22)

Given that topics inform what a clause is about, both pronouns in B are the topics about which the answer is informing. This raises the question of “primary” vs. “secondary topic”, of which the former is pragmatically more salient than the latter (Dalrymple & Nikolaeva 2011: 53–57). Nikolaeva (2001: 26) defines the secondary topic as ‘an entity such that the utterance is construed to be ‘about’ the relationship between it and the primary topic.’ To explain this topic chain, Vallduví (1992) applies

a link-tail model, in which “link” refers to a stable topic that only presents the given aboutness without requiring additional information, while the “tail” is open for completing or modifying by additional information, i.e. comment (see also Erteschik-Shir 1997, 2007).

4.4.3 TYPES OF FOCUS

Apart from topic, an utterance is expected to have at least one focus that adds new information to the presupposition, and such obligatory focus of the clause is termed “(new) information focus” (Gundel & Fretheim 2004: 181–183; Neeleman & Vermeulen 2012), “non-contrastive focus” (Erteschik-Shir 1997: 11–13), or “semantic focus” (Gundel 1999). In (39), the answer in B simply gives new information to the question in A without creating any contrast to other possible alternatives.

- (39) A: *What did John read?*
 B: *He read **THE SELFISH GENE**.* (Neeleman & Vermeulen 2012: 6)

This is the only use of focus that does not involve the notion of contrast (see Table 8) because it ‘indicates covert questions suggested by the context’, such as *What happened? What was there? or What did she do?* (Krifka 2008: 251). Gundel (1999) states that the information focus is higher in the relevant hierarchy than contrastive focus, because each clause must have an information focus as an essential truth-conditional element in information processing. The contrastive focus, on the other hand, is optional, and its overt marking depends on the speaker’s intention (see also “Contrast-Hierarchy” in Molnár 2006: 211). Supporting evidence for this statement is the use of intonation, in which clausal stress primarily falls on the information focus (cf. Schmerling 1976; Gundel 1978; Selkirk 1984; Zacharski 1993; Vallduví & Vilkkuna 1998).

As discussed in Section 4.4.1, focus usually implies a set of alternatives. Most of the functions beyond the information focus relate to contrast. A similar issue with the case of topic is the category of “contrastive focus”, which does not satisfactorily describe the exact pragmatic uses observed in the actual language use. Instead of using such a broad umbrella term to cover a wide functional range of focus (as in Gundel 1999; Neeleman et al. 2009), the current study adopts the pragmatically more fine-grained classification by Krifka (2008), who treats contrastive focus as just one type among many uses, which is observed only in the context where the current utterance can really be contrasted.

Pragmatic uses of focus can be observed in various contexts, which involve the characteristic function of delimiting a set of relevant alternatives. First, foci can “correct” or “confirm” information previously mentioned, as in (40).

- (40) A: *Mary stole the cookie.*
 B: *(No,) **PETER** stole the cookie!*
 B': *Yes, **MARY** stole the cookie.* (Krifka 2008: 252)

In B, the answer differs from the context proposition in A so it gives a corrective interpretation, while the answer in B' gives a confirmative interpretation as the context proposition is identical to A.

Second, foci can “highlight parallels” which evoke the immediately surrounding contexts, as in (41).

(41a) *MARY stole **THE COOKIE** and PETER stole **THE CHOCOLATE**.*

(41b) *An **AMERICAN** farmer talked to a **CANADIAN** farmer, ... (Krifka 2008: 252)*

In (41), the foci create alternatives as they often do, but the alternatives ‘Peter’ and ‘the chocolate’ in (41a) and ‘a Canadian farmer’ in (41b), in this case, are also evoked in the surrounding contexts.

Third, foci can “delimit” the utterance to the constituent in focus, which gives an instruction to the addressee how an utterance should be interpreted, as in (42).

(42) *In **MY** opinion, JOHN stole the cookie. (Krifka 2008: 252)*

In (42), the frame setter part informs the addressee that the scope of the uttered statement is delimited to ‘my opinion.’

Beside these main pragmatic uses of focus, Krifka (2008: 257–259) also discusses other focus types, several of which are worth discussing, as they have been reported from Finnic and North Russian (see further discussion in Sections 5.1.1 and 5.2.2). First of all, Krifka divides the pure contrastive uses of focus into “closed” and “open focus”. This use involves the alternative set, which can be restricted to few alternatives provided, such as in (43).

(43) A: *What do you want to drink, tea or coffee?*

A': *What do you want to drink?*

B: *I want **TEA**. (Krifka 2008: 258–259)*

The question in A restricts a set of alternatives to ‘tea’ and ‘coffee’, while the question in A' leaves a possibility open to the addressee to choose any kind of drink. Due to the restrictive nature of focus in this context, Erteschik-Shir (1997: 12–13) also names a similar type of focus as “restrictive focus”.

Another noteworthy type of focus is the “exhaustive focus”, which is often observed in *it*-cleft construction as in (32). Equally worth discussing is the “scalar focus”, which highlights the least or greatest element in the context, and this use is often found with scalar particles like *even* and *at least* in (44).

(44) *Even/At least **WILD HORSES** wouldn't drag me. (Krifka 2008: 259)*

This type of construction establishes a scalar focus ‘even/at least’ to accompany a strong polarity item, ‘wild horses.’

As foci can appear more than once in a clause, a focus expression can be a combination of “complex foci” or “multiple foci”, which give different interpretations, as shown in (45).

(45a) *John only introduced **BILL** to **SUE**.*

(45b) *John only introduced **BILL** only to **SUE**.* (Krifka 2008: 258)

In (45a), the focus marker ‘only’ restricts a complex focus chain as ‘Bill’ being introduced ‘to Sue.’ In contrast, the two focus markers in (45b) establish a multiple foci construction, which highlights the introducing of someone ‘to Sue’ and that someone must only be ‘Bill.’

4.4.4 INFORMATION-STRUCTURAL STRATEGIES WITH DEMONSTRATIVES

Strategies for encoding information structure vary across the world’s languages, related to noun phrases such as the use of articles, demonstratives, or topic and focus markers (Givón 1983: 35), as well as morphosyntactic mechanisms such as case marking, agreement, word order, phrase structure position, and suprasegmental strategies like prosody (Dalrymple and Nikolaeva 2011: 67). For the context of postposed demonstratives in Northwest Russia, it is worth discussing a few strategies that have been previously reported in the literature, above all, those that concern noun phrase: articles, demonstratives, or topic and focus markers.

One issue concerning information-structuring markers is a question whether topics and foci always need to be explicitly marked. For this matter, it is worth borrowing Haspelmath’s idea (2019: 315) that ‘deviations from usual associations of role meanings and properties of referring expressions tend to be coded by longer grammatical forms.’ This can be applied to the marking of topics, which are not always marked due to their given saliency and predictability in the discourse. In contrast, foci are more often marked, due to their restrictiveness and exhaustiveness. Combining this idea with the typology of topics and foci discussed in Sections 4.4.2 and 4.4.3, non-contrastive use of topic and focus such as “expected topic” and “information focus” may not require any explicit marking, while other uses which involve contrast such as “switch-topic” and other types of focus more often do.

As discussed in Section 4.2, there have been a number of studies showing that in many of the world’s languages demonstratives are etymological sources for definite articles, topic marker, and focus markers (see Heine & Kuteva 2002: 109–112). Moreover, there are a number of grammatical descriptions that identify information-structural uses of demonstratives.

To begin, demonstratives are observed functioning as topic markers, for instance, in a number of Southeast Asian languages such as Green Hmong (Kunyt 1984: 121), Eastern Cham (Brunelle & Hân 2019: 544), Mon (Jenny 2009, 2014: 576), Thai (Warotamasikkhadit 1997; Singnoi 2004: 651–656; Iwasaki & Ingkaphirom 2005: 361), Hakha Chin (Barnes 1998), and some Karenic languages (Kato 2003: 635; Shee

2008: 135), all of which have adnominal demonstratives in a postnominal position (see also a cross-linguistic survey of topic markers in Wälchli, forthcoming). Their function can be seen, on the one hand, as a marker a boundary between topic and comment expressions, i.e. “information chunks” (in Matic’s term 2003), and, on the other hand, as a marker of dislocation constructions, so the typical place for this marker is in the clause-second position after the first constituent, such as in (46) to (48).

Hmong Njua (Hmongic, Hmong-Mien)

- (46) *crang*⁵⁵ *cra*¹¹ *nua*³⁵ *ku*³⁵ *yua*³⁵ *tua*⁴¹ *tao*²¹⁴ *ka*⁵⁵ *tua*⁴¹.
 [CLF knife DEM] 1SG buy come at market come
 ‘This knife, I bought it from the market.’ (Kunyt 1984: 121)

Mon (Monic, Austroasiatic)

- (47) *hmoŋ ʔələwiʔ kɔ̌h nùm kv kɾəð.cəh hnòk tao pùə.mə.lòn raʔ*.
 [king Alawi DEM] exist OBL glory big stay exceedingly FOC
 ‘This king Alawi was of great glory.’ (Jenny 2014: 576)

Pwo Karen (Karenic, Sino-Tibetan)

- (48) *ʔə- yeiN*⁵⁵ *nɔ̌*⁵⁵ *jə-* *lɪ*¹¹ *ʔe*⁵⁵.
 [his house DEM] 1SG go NEG
 ‘To his house, I did not go.’ (Kato 2003: 635)

In (46) to (48), the demonstratives separate the new information, i.e. comment, from the entities in left periphery ‘knife’, ‘king Alawi’ and ‘his house’, which are given information and, thus, topics of the clauses.

As for focus marking, demonstratives also appear as etymological sources for focus markers in a number of languages (Diessel 1999a: 148–149; Heine & Kuteva 2002: 111–112). Consider, for instance, the Ambulas and Mokilese examples (49) and (50) in which such focal use is observed on demonstratives.

Ambulas (Sepik, Papuan)

- (49) *véte dé wak a wan méné kaapuk yéménén*.
 and:see he said ah [DEM you] not you:went
 ‘He saw him and said, “Ah, so **you** did not go”.’ (Wilson 1980: 157)

Mokilese (Micronesian, Austronesian)

- (50) *ioar Wilson ma pwehng ih mehu*
 [DEM Wilson] REL told him that
 ‘It was **Wilson** who told him that.’ (Harrison 1976: 311)

In (49) and (50), the Ambelam distal demonstrative *wan* ‘that’ and the Mokilese demonstrative-based deictic marker *ioar* are used to emphasise the referents ‘you’ and ‘Wilson’ that follow. A factor that favours the functional extension of demonstratives to focus marking is likely due to their nature of gesturing proximity, which could highlight the restrictive and exhaustive characteristic of focus (as discussed earlier in Sections 4.4.1 and 4.4.3).

Beyond information structure, demonstratives can also be used for evaluative purposes such as stancetaking and expressing the speaker’s attitudes towards a state-of-affairs, discussed in the next section.

4.5 DEMONSTRATIVES IN EVALUATION

The functions of demonstratives discussed so far have shown that they are used for various communicative purposes, from creating identifiability to organising information structure. Beyond these aspects, however, demonstratives may also serve as devices for establishing intersubjectivity, as well as for coding evaluation and stancetaking, reported from in studies of individual languages in recent decades, as well as a theoretical framework discussed by Du Bois (1980). Therefore, it is also worth discussing and understanding how demonstratives behave in the evaluative domain. Previous observations from languages with rich demonstrative systems and robust evaluative uses of demonstratives can help in exploring more functional dimensions of postposed demonstratives, and provide good tools for the analysis of Finnic and North Russian data later in Section 6.4.3.

4.5.1 INTERSUBJECTIVITY, STANCE AND EVALUATION

Language as a communicative tool between social actors contains the notion of “intersubjectivity”, which can be understood as ‘the relation between one actor’s subjectivity and another’s’ (Du Bois 2007: 140). This entails that intersubjectivity is based on the interaction between individual social actors’ subjectivities on how they perceive sociocultural values.

This social act gives rise to “stance”, which emerges through an individual’s evaluation by assertion or inference. Importantly, stance can often be expressed by demonstratives in a number of languages (see further discussion in Section 4.5.2). In Du Bois’ framework of a stance triangle (2007), stance can be defined as follows.

‘Stance is a public act by a social actor, achieved dialogically through overt communicative means, of simultaneously evaluating objects, positioning subjects (self and others), and aligning with other subjects, with respect to any salient dimension of the sociocultural field.’

(Du Bois 2007: 163)

According to Du Bois, stancetaking involves three participant components: (i) the stancetaker, (ii) the act of positioning or evaluating, and (iii) the object of stance, as illustrated in (51).

	Speaker	Stance subject	Positions/Evaluates	Stance object	
(51)	JAMIE;	<i>I</i>	<i>like</i>	<i>this song.</i>	(Du Bois 2007: 153)

In (51), a stancetaker, Jamie, performs the stance action by positioning himself as ‘liking this song’, while also simultaneously evaluating that ‘this song’ is good in his opinion. In a dialogue with multiple participants, the stance action can also represent their individual subjectivities, which may match or collide, such as in (52).

	Speaker	Stance subject	Positions/Evaluates	Stance object	
(52)	SAM;	<i>I</i>	<i>don't like</i>	<i>those.</i>	
	ANGELA;	<i>I</i>	<i>don't</i>	<i>either.</i>	
	CINDY;	<i>I</i>	<i>like</i>	<i>it, though.</i>	(based on Du Bois 2007: 159)

In (52), Sam’s and Angela’s subjectivities are “aligned”, while Cindy’s subjectivity is “disaligned” with the negative pole ‘don’t like’ shared by Sam and Angela. By adding this process of “alignment” to the three components of stancetaking, Du Bois proposes the following cartography in Figure 21, which describes the mechanism of stancetaking.

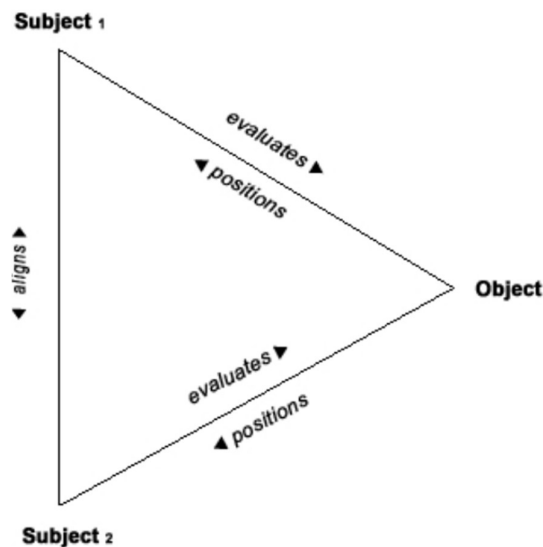


Figure 21 Du Bois' stance triangle (2007: 163)

Previous studies have shown that stance can be the outcome of several sociocognitive processes. The most central notion in terms of stancetaking action is “evaluation”, used for three communicative purposes described by Hunston and

Thompson (2000). First, evaluation can express “the speaker’s opinion” and reflect the value system of that person and the community they belong to. This is often the case of a communal or cultural-specific ideology, which instructs whether a certain matter should be regarded, e.g., as good or bad, acceptable or unacceptable, true or untrue. Cognitively, the speaker’s evaluation can emerge from their affect (emotion), judgement (ethics, norms) or appreciation (aesthetics, social value), according to the “Appraisal framework” (see, e.g., White 2015). Second, evaluation can construct and “maintain interactional relations” between the speaker and addressee through manipulation, hedging, or politeness, which can be associated with the act of face-saving (see also Fraser 1990). Third, evaluation can “organise the discourse” by creating solidarity within a mental sphere in which the speaker and addressee assumedly share common attitudes, values, or reactions to a specific state-of-affairs.

Stancetaking and evaluation are closely related to epistemic states of the interlocutors under the notion of “evidentiality”, which deals with information source and the speaker’s access to it. In other words, the focus is on how the speaker has acquired information before transmitting it to the addressee, i.e. by personal experience or observation through senses, such as visual, sensory, inference, assumption or reported information (see, e.g., Aikhenvald 2004; Tournadre & LaPolla 2014). Applying this framework to the evaluation mechanism, we can say that the speaker’s opinion and value presented to the addressee can evolve from either the speaker’s personal experience or observation, which the speaker can use for maintaining interactional relations and organising the discourse.

In addition to the epistemic aspect, stancetaking and evaluation also involve the interlocutors’ attentional states. An empirical observation from Andoke, a South American language, by Landaburu (2007), followed by a theoretical account by Evans et al. (2018a, 2018b), introduce a framework of “engagement”, which stands for ‘a grammatical system for encoding the relative accessibility of an entity or state of affairs to the speaker and addressee’ (Evans et al. 2018b: 142). In the concrete discourse, the speaker simultaneously takes a stance and evaluates the addressee’s state of knowledge, whether the information uttered is only in the speaker’s attention, or shares joint attention for both sides (compared to identifiability discussed in Section 4.3.3). A symmetry evolving in the interlocutors’ states of knowledge and attention can also be utilised as a means for the speaker to take or give epistemic authority to the addressee (Evans et al. 2018a: 118). Previous theoretical studies have emphasised that this aspect of the interlocutors’ attention is also crucial for understanding contexts in which the use of definite articles, for instance, is required (see, e.g., Clark & Marshall 1981; Epstein 1997).

The aforementioned functions related to evaluation can be coded with various lexical and grammatical elements, among which the use of demonstratives is also observed in a number of languages, especially those spoken in Southeast Asia and Papunesia (see Section 4.5.2) as well as in Finnic and North Russian (see Section 6.4.3). This matter of grammatical encoding will be discussed with concrete examples in the next section.

4.5.2 DEMONSTRATIVES AS AN INDEXICAL OF STANCETAKING AND EVALUATION

Various linguistic elements, including demonstratives, are used to express a speaker's stancetaking and evaluation. Hunston and Thompson (2000: 14) discuss a number of lexical items that establish stancetaking in the discourse, such as various kinds of adjectives (*splendid, terrible*, etc.), adverbs (*unfortunately, interestingly*, etc.), nouns (*failure, triumph*, etc.) and verbs (*doubt, win, lose*, etc.). What is more interesting is the case when stance is expressed by grammar, i.e. a marker of stance. Such marker can have their root in noun or verb phrases which develop into fixed expressions of stance such as adverbs (*for sure, really, sort of*, etc.), conjunctions (*while, though, since*, etc.), or longer verb phrases (*all being well, whatever that means, I did wonder if...*, etc.), for instance (see Labov 1972; Stubbs 1986; Biber & Finegan 1989). Moreover, markers of stance can emerge from pre-existing grammatical elements, which have extended their original semantics to pragmatic uses such as demonstratives, among others.

Most importantly, the pragmaticisation of demonstratives to devices for stancetaking and evaluation has been reported from Papuan languages (Kratochvíl 2011; Schapper & San Roque 2011), Burmese (Simpson 2008), Thai (Iwasaki & Dechapratumwan, forthcoming), Vietnamese (Lê 2002; Adachi 2016, forthcoming), as well as several Austronesian languages (Cleary-Kemp 2007: 336–337) such as varieties of Malay (Ernanda & Yap, forthcoming) and Tagalog (Nagaya 2011), or even in Finnish (Yurayong & Kittilä, forthcoming). Here, proximity distinctions across demonstratives play an important role in the development of versatile evaluative uses in the reported languages.

A good example can be taken from the description concerning the use of demonstratives as clause-final particles (CFP) in Vietnamese by Adachi (2016, forthcoming), in which several demonstratives are associated with specific evaluative uses, illustrated in Table 9.

Deixis	Proximal (independent)	Medial (independent & adnominal)	Distal (independent & adnominal)	Medial (adnominal)
Form	<i>đây</i>	<i>đấy</i>	<i>kia / cơ</i> (reduced form)	<i>ấy / ý</i> (reduced form)
Clause-final use	Speaker-centred	Addressee-oriented	Counter-expectation	Shared knowledge

Table 9. *Clause-final uses of Vietnamese demonstratives (adapted from Adachi, forthcoming)*

The first contrast is found between the speaker-centred and addressee-oriented demonstratives, which indicate whether a statement is based on the speaker's personal evaluation (53), or new information that should be relevant to the addressee (54). Otherwise, the demonstratives can also express whether a statement is assumed as a counter-expectation to the addressee (55), or as shared knowledge between the interlocutors (56).

Vietnamese (Vietic, Austroasiatic)

- (53) An employee complains how difficult the assignment that she is working on is.

bài này khó đây.
 assignment DEM.PROX.ADNOM difficult DEM.PROX.CFP
 ‘This assignment is difficult **as far as I can see.**’ (Lê 2002: 60)

- (54) The mother is telling a news to her daughter about her friend being transferred to another class, which the mother assumes that her daughter might not know about this.

à, Hương, nó bị chuyển lớp đấy.
 INTJ Hương she PASS transfer_to class DEM.MED
 ‘Ah, Hương, she was transferred to another class **which you may not know about, but may be interested to know.**’ (Adachi, forthcoming)

- (55) The mother is telling the father about their mutual friend’s son going to New York and how expensive his parents had to pay for the flight tickets.

cái kia đắt lắm.
 CLF DEM.DIST.ADNOM expensive very
những hai nghìn mấy cơ.
 some two thousand several DEM.DIST.CFP
 ‘(In fact,) that ticket was very expensive. It cost more than 2,000 (US dollars), **more expensive than you would expect.**’ (Adachi, forthcoming)

- (56) The mother is talking to the father about popular Vietnamese souvenirs among Japanese tourists. As they have many Japanese friends in common, she assumes that he will understand what she means.

người Nhật thích ăn phở,
 people Japanese like eat rice_noodle
phở ăn liền ý.
 rice_noodle eat instant DEM.MED
 ‘The Japanese like rice noodle, **to be exact**, instant rice noodle, **you know.**’
 (Adachi, forthcoming)

The proximal demonstrative *đây* in (53) expresses the employee’s stance emerging from her direct experience, and the evaluation of the difficulty of the assignment. Meanwhile, the medial demonstrative *đấy* in (54) expresses the mother’s inference that the news should interest and be highly relevant to her daughter. In (55), the distal demonstrative *cơ* expresses the mother’s evaluation that the information she is telling may not be familiar, and should come as a surprise to the father. In König’s term (1991: 181), the expression of counter-expectation creates “inconsistency” between the truth and previous assumption (see also a similar notion of “mirativity” in DeLancey 1997, 2001). As for (56), the other adnominal medial demonstrative *ý*

expresses the mother's evaluation that the information she is telling should also be familiar and shared with the father. A similar functional distribution of demonstratives as discourse particles are also reported from languages in the neighbouring areas of Southeast Asia and Papuanesia, especially in Thai (Iwasaki & Dechapratumwan, forthcoming), Tagalog (Nagaya 2011), Abui (Kratochvíl 2011) and several other Papuan languages (Schapper & San Roque 2011).

In addition to clause-final uses, demonstratives in other clausal positions can also perform stancetaking when they are syntactically used as “place holders” or “lexical fillers”, and a similar evaluative use of demonstrative is also observed with the postposed demonstratives in Finnic and North Russian dialects (see Section 6.4.3). A typological investigation by Podlesskaya (2010) shows that pronouns, particularly demonstratives, are typical lexical items that have further developed into place holders. Hayashi and Yoon (2006: 490) define place holders as ‘a referential expression that is used as a substitute for a specific lexical item that has momentarily eluded the speaker’, such as in (57).

Korean (Koreanic)

- (57) *ne ce-ke cwu-lkka? chicukheyikh?*
 you DEM.DIST-CLF give-shall cheesecake

‘You, would you like to have that (thing)? Cheesecake?’ (Hayashi & Yoon 2006: 492)

In (57), the speaker was momentarily lost in her own mind on what she was going to give the addressee. While searching for the referent from her memory, the speaker temporarily fills the required syntactic slot with a place holder ‘that (thing)’ before elaborating afterwards that it was ‘cheesecake’ that she was going to offer to the addressee. According to Hayashi and Yoon, the place holder use of demonstratives can be paralleled to the recognitional use in the referential domain (discussed in Section 4.3.3), in which the identifiability is similarly arrived to through personal memory.

In any case, demonstratives in the place-holder position can also perform other evaluative uses such as expressing “avoidance” as in (58) and “hesitation” as in (59).

Lao (Tai, Tai-Kadai)

- (58) The speaker is checking whether the addressee's mother is still in a good health.

me:³³ caw⁵³ naŋ³⁵ bɔ:³³ ʔan¹³-nan⁵³ ju:³³ ti:⁵³?
 mother 2SG still NEG CLF-DEM.MED stay PTCL

‘Your mother is still not **that (thing)** [getting better/cured from sickness], right?

(Enfield 2003: 108)

In (58), the speaker may feel it impolite to mention that the addressee's mother is sick, so she avoids such an utterance by using the medial demonstrative filler *ʔan¹³-nan⁵³* ‘that (thing).’

Japanese (Japonic)

- (59) H is teasing K, who once unthinkingly took off her short pants in front of a crowd after running in a 100-meter race.

H: *iya konkai ano hashi-tte-mo,*
 well this_time DEM.DIST run-CVB-even
 ‘Well, this time, *ano* [=um], even if you run in a race,’
ano gooru shi-ta ato ni,
 DEM.DIST goal do-PST after LOC
 ‘*ano* [=um], after you enter the goal,’
nanka ano tanpan o nuga-nai y(h)oon(h)i.
 like DEM.DIST short_pants 0 take_off-NEG.IMP INTJ
 ‘like, *ano* [=um], do not take off your short pants.’

K: *nan desu ka sore wa?*
 what COP Q DEM.MED TOP
 ‘What on earth are you talking about?’ (Hayashi & Yoon 2006: 508)

In (59), while hesitating what to say next, the speaker uses the distal demonstrative *ano* for extra time for thinking, and to keep reminding the addressee about her embarrassing memory, so that she would pay attention and react. Similar uses of demonstratives as place holders, as described above, are also reported in Estonian (Keevallik 2010) and Finnish (Etelämäki 2006).

A similar hedging effect is also observed in manner demonstratives such as German *so*, when occurring in a non-gradable expression (Umbach & Ebert 2009). For instance, the speaker uses the demonstrative *so* in (60) as a hedging device to widen the tolerance range of ‘three o’clock,’ as she does not memorise the precise moment when Marie came.

German (Germanic, Indo-European)

- (60) *Marie kam so um drei.*
 Marie come.PST.3SG DEM about three
 ‘Marie arrived at about three o’clock.’ (Umbach & Ebert 2009: 162)

The examples from the languages presented above imply that the deictic nature of demonstratives could have motivated their pragmaticisation to devices for stancetaking and evaluation. These examples can be related to the classification of evaluation by Hunston and Thompson (2000) as well as to evidentiality (Aikhenvald 2004) and engagement (Evans et al. 2018a, 2018b). First, we see that the proximal demonstrative puts the focus on the speaker’s stance and personal experience, thereby giving them epistemic authority as the attention is not shared with the addressee (53). Meanwhile, medial demonstratives can maintain interactional relations between the interlocutors by expressing the speaker’s consideration of the addressee’s stance and attention, i.e. affiliation and inclusiveness, through thoughtfulness and politeness,

which hand epistemic authority to the addressee (54, 56, 58). At the same time, the distal demonstrative organises the discourse by establishing joint attention and stance based on the speaker’s evaluation (55), or by expressing the speaker’s hesitation and allusion that invite assertion from the addressee (57, 59).

Apart from demonstratives, languages across Eurasia can also use other grammatical elements to convey stancetaking and evaluation, such as possessive suffixes in Khalkha Mongol (Brosig et al. 2018), Malay (Englebretson 2003, 2007; Yap 2011), and Uralic languages (see Section 5.3).

Khalkha (Mongolic)

- (61) *en-čín’ yaa-san xünd we!*
 this-Px.2SG do_what-PTCP.PRF heavy Q
 ‘How heavy this is!’ (Brosig et al. 2018: 83–84)

In (61), the use of the 2nd person possessive suffix *-čín’* familiarises the addressee with a disagreeable ongoing situation, often used for dealing out blame (‘just to let you know how hard I have to lift this thing’) or for insistently asking for assistance (‘could you possibly help me?’). This shows that the functional resemblance of possessive suffixes and demonstratives is not only related to referential uses (discussed in Section 4.3.3) but also to evaluative uses (see also Section 5.4).

This section has illustrated that evaluative uses of demonstratives obviously became very robust in languages with rich demonstrative systems, seen in the languages mentioned above. At the same time, it is also interesting to see whether a similar pragmaticisation of demonstratives would also be observed in Finnic and North Russian, which show less complex demonstrative systems (see the discussion in Sections 5.1.1 and 5.2.2).

5 MARKERS OF DEFINITENESS AND INFORMATION STRUCTURE IN URALIC AND SLAVIC LANGUAGES

This chapter presents state of the art of research on postposed demonstratives in Uralic and Slavic languages. The discussion will show how the theories and mechanisms described in Chapter 4 are concretely realised in these two language groups.

Uralic languages prototypically do not possess articles, although certain contemporary languages have developed definiteness marking strategies on nouns, such as articles in Hungarian and definite declension in Mordvinic. Since articles are a less common category among the Uralic languages, we have to look at other grammatical elements that serve a comparable function to articles, such as demonstratives and possessive suffixes. Synchronically, the possessive suffix is a common marker of definiteness and topicality among the majority of Uralic languages, while postposed demonstratives are only observed in the western Uralic languages, Finnic, Saami, and Mordvinic. Nevertheless, the use of Finnic postposed demonstratives has not been as extensively investigated as the non-possessive uses of possessive suffixes in other Uralic languages. Therefore, this chapter will show that semantic and pragmatic tools used to analyse possessive suffixes in previous studies are also applicable to the analysis of postposed demonstratives in the current study.

Apart from demonstratives and possessive suffixes, Uralic languages may also use other morphological and syntactic strategies to encode a meaning expressed by the opposition between definite and indefinite articles, or between topics and comments. In the case of Finnish and Russian, for instance, E. Itkonen (1966: 258), L. Hakulinen (1968: 509–512), Wexler (1976) and Chesterman (1991: 113–116, 183–184) also consider the following strategies:

1. **Subject case marking:** nominative vs. partitive
2. **Object case marking:** accusative vs. partitive
3. **Conjugational number of predicate:** singular vs. plural
4. **Voice:** active vs. passive
5. **Word order:** unmarked/canonical vs. marked/non-canonical.

These strategies are also applicable to the majority of Uralic languages, which have been previously discussed elsewhere (see general overview on information structure in Skribnik 2001; Sosa 2017, differential object marking in Klumpp 2012; Virtanen 2013, and word order variation in Vilkuna 1989, 1998).

In any case, this chapter will delimit the scope to demonstratives (Sections 5.1 and 5.2) and possessive suffixes (Section 5.3), as well as highlight the relevance of possessive suffixes to the studies of postposed demonstratives (Section 5.4).

5.1 POSTPOSED DEMONSTRATIVES IN URALIC LANGUAGES

As mentioned in Section 4.1, the typical place for demonstratives in the head-final Uralic languages is a prenominal position. Meanwhile, the use of demonstratives in postnominal position is likely an areal feature, observed among the languages in Northwest Russia. This includes Ladogan Finnic languages (Karelian, Lude, and Veps), as well as eastern Saami languages (Kildin and Ter Saami). In addition, the discussion in this section also includes Mordvinic languages, in which a definite declension has also developed from a postposed demonstrative. Hungarian also uses a definite article *á(z)* that has developed from demonstratives (see, e.g., Egedi 2014), but its position in a noun phrase is prenominal. Given also that the speaking area of modern Hungarian itself represents another type of Slavic contact, mainly with the Slavic languages in neighbouring areas (Slovak, Slovene, BCMS, Rusyn and Ukrainian) but without a direct connection to Northwest Russia, Hungarian is excluded from the discussion in the present study.

5.1.1 DEMONSTRATIVES AS MARKERS OF DEFINITENESS AND INFORMATION STRUCTURE IN FINNIC LANGUAGES

An adnominal use of demonstratives in Finnic languages has often been compared with the use of definite article in other European languages (Kettunen 1943: 130–131; E. Itkonen 1966: 257; A. Hakulinen 1985: 342; Nordlund et al. 2013). Diachronically, E. Itkonen claims that such a phenomenon should have emerged in Finnish through the literary language, in which Mikael Agricola introduced in the 16th century a construction with prenominal indefinite *yksi* ‘one’ in (62) and definite article *se* ‘that’ (63) and (64) as a replication of Swedish (and potentially also other Western European languages).

Old Finnish

- (62) *ydhen* *oijkea-n* *aija-n* Indefinite ‘one’
 one-GEN right-GEN time-GEN
 ‘at a right time’ (E. Itkonen 1966: 257)

- (63) *sille* *wimeijs-se* *domio* *peijwe-lle* Definite demonstrative
 DEM.MED.ALL last-GEN judgement day-ALL
 ‘to the last Judgement Day’ (E. Itkonen 1966: 257)

- (64) *sihen* *cunniallis-ehen* *hää* *waatte-hen*
 DEM.MED.ILL last-ILL wedding dress-ILL
 ‘in(to) the honourable wedding dress’ (E. Itkonen 1966: 257)

In the spoken language, however, L. Hakulinen (1968: 509–512) also points out that there is a sign of language-internally motivated article formation by means of lexical elements like *se* ‘that’ and other (attributive) pronouns *tämä* ‘this’, *tuo* ‘that’, *eräs* ‘one’, *joku* ‘some’, *muu* ‘other’, such as in (65) and (66).

Spoken Finnish

- (65) *on-ko se isäntä koto-na?*
 be.3SG-Q DEM host home-ESS
 ‘Is the landlord home?’ (L. Hakulinen 1968: 509)

- (66) *maa-sta se pieni-kin ponnista-a.*
 land-ELAT DEM small-PTCL push-3SG
 ‘Even the little one will take off from the ground.’ (L. Hakulinen 1968: 509)

Later researches (e.g., Laury 1997: 250–264; A. Hakulinen et al. 2004: §1418) have also confirmed this tendency in spoken Finnish.

Towards the end of 20th century, an interactional linguistic approach brought a new perspective, which does not regard the use of Finnic prenominal demonstratives as definite articles proper. Instead, demonstratives rather share functions that are characteristic of definite articles, as has been discussed for South Estonian *tuu* (Pajusalu 1998, 2006), North Estonian *see* (Pajusalu 1997a, 1997b, 2001; Sahkai 2016), Finnish *se* (A. Hakulinen 1985; Chesterman 1991; Vilkuna 1992; Dasinger 1995; Sundbäck 1995; Juvonen 2000; Etelämäki 2006), and Votic *se* (Agranat 2015), and as a shared tendency among Standard Estonian *see*, Finnish *se* and Russian *этот* (Pajusalu 1996b; Hint et al. 2017). Most studies do not deny that demonstratives still primarily function in a deictic use, typical for demonstratives, but expected to weaken when they grammaticalise to definite articles (discussed in Section 4.2). More importantly, the use of demonstratives is still optional in most contexts (cf. a criterion of obligatoriness in Greenberg 1978), especially absolutely unique, recognitional, bridging and situationally unique referents, in which the use of grammaticalised definite articles would be obligatory (see the minimal requirement of definite articles in Becker 2018; König 2018, discussed in Section 4.3.3).

Although proposed demonstratives in Finnish are claimed to become definite articles in the future (Karlsson 1985: 62), or have already become definite articles (Laury 1991, 1996, 1997), language use of individual speakers show significant variation, pointing to the article-like uses of prenominal demonstratives *se* (and also *tämä*) being sociolinguistically motivated (Juvonen 2000: 195–197). In the case of North Estonian, on the other hand, article-like uses might not be on the rise but rather on the decline, potentially due to a decreasing degree of German influence during the latest century (Pajusalu 1997a: 173). In any case, language contact might not necessarily be fully responsible for the development of demonstratives towards marking definiteness in Estonian and Finnish. The phenomenon and tendency could have also existed in the language prior to the contact with Germanic languages (Nordlund et al. 2013), and we can also talk about a language-internal “forced

grammaticalisation” (from the original German term *Zwangsgrammatikalisierung*, introduced by Nau 1995) that has given rise to such development (Metslang 2011).

However, little has been discussed about functions of demonstratives in the postnominal position, which are mostly observed in easternmost Finnic languages (Karelian, Lude and Veps). Meanwhile, the postnominal use as adverbs has been discussed for Estonian *siin* and *seal* (Pajusalu 1997a), for Finnish *tässä* (Etelämäki 2006) and from a comparative Estonian-Finnish perspective (Reile et al. 2019). A description of Veps by Kettunen (1943: 130–131, 399–403) defines four contexts of use for postposed demonstratives *se/ne* ‘that/those.’

1. Definite article
2. Being used in the place of possessive suffixes which are becoming obsolete
3. Clause adverbial
4. Question particle

As also being proposed by other scholars, namely that Finnic demonstratives can take functions equivalent to definite article, in Veps the postposed demonstratives can do the same, as shown in (67), where the demonstrative is marking a bridging referent ‘spirit’ as belonging to ‘corps.’

Southern Veps

- (67) *hibja-n kop-ha pano-bad, a henk se käulo-b*
 corpse tomb put-3PL but spirit DEM move-3SG
 ‘A corpse is put into a tomb, but the spirit is moving around.’ (Kettunen 1943: 400)

Due to the use of possessive suffixes becoming obsolete in Veps, the replacement by postposed demonstratives is also observed, as in (68) in which ‘ass’ refers to that of the writer of the message himself.

Central Veps

- (68) *kirjuta-b kirjaiže-n perske-le se ka jä-b siloi ...*
 write-3SG letter-ACC ass-ALL DEM so remain-3SG then
 ‘He writes a letter (and puts) to his own ass, so it (the information) then remains (unknown to the officers).’ (Kettunen 1943: 401)

Often when occurring in a phrase expressing time, manner, or location, postposed demonstratives can mark adverbial phrases as a scene setter (in Chafe’s term 1976), or as a frame setter (in Krifka’s term 2008), such as ‘spring’ in (69).

Central Veps

- (69) *kevade-l se ištu-i-n*
spring-ADES DEM sit-PST-1SG
 ‘During the spring time, all I did was sitting.’ (Kettunen 1943: 401)

Apart from using a rising intonation in a Russian-like manner, Veps also uses demonstratives in the clause-final position to indicate a question, as in (70).

Central Veps [identical to (1)]

- (70) *nügüt silaaž om jo torvut se se?*
 now 2SG.ADES be.3SG already **phonograph DEM DEM**
 ‘Have you already got the phonograph now?’ (Kettunen 1943: 402)

In (70), a clause-final demonstrative appears to indicate an interrogative mood, while the preceding demonstrative expresses definiteness of the noun phrase ‘phonograph.’ Interestingly, the functional parallel to the aforementioned uses in Veps is also observed with North Russian *-to*, which points to a contact-induced development (see further discussion in Section 5.2.2).

Nevertheless, the concept of “definite article” later receives less support and recent descriptions tend to classify postposed demonstratives as “emphatic focus particles” (Grünthal 2015a: 279–281, see also a similar case of North Russian in Section 5.2.2). Meanwhile, the use of a postposed demonstrative *se* is also reported from Southern Lude, which Pahomov (2011: 54) regards as a modal particle with similar discourse functions to Finnish *-han* (see also a thorough description of Finnish *-han* in Duvallon & Peltola 2012; Duvallon 2014; Liefländer-Koistinen 1989). In any case, discussion in Chapter 6 will show that we still cannot ultimately ignore the notion of definite article as the data also illustrate the use of postposed demonstratives as definiteness marker besides other pragmatic uses that are more typical for discourse markers.

Concerning the areal distribution among the Finnic languages, a thorough investigation of (eastern) Finnic demonstratives by Larjavaara (1986: 307–310), and a micro-areal investigation by Yurayong (2020), divide the use of postposed demonstratives into three types, with different degrees of morphosyntactic restrictions. The first type is observed in the Finnic languages of the outer ring, including all Finnic languages except Karelian, Lude and Veps. In these outer ring languages, the Proto-Finnic medial demonstratives **se/ne(t)* occur only once in a clause-second position, i.e. Jakob Wackernagel’s law (1892), as a kind of a marker to separate a topic from comment, or as a marker of left dislocation constructions (see the case of Finnish in Vilkkuna 1989: 145–147; Priiki 2015, 2017, Estonian in Amon 2015, and the discussion in Section 4.4.4 with similar examples from Southeast Asian languages). Morphologically, they decline in the same form as the head noun, or remain indeclinable after other nominal parts of speech (pronoun and name). Ojansuu (1922: 82), Kettunen (1943: 399–400), E. Itkonen (1966: 257), and L. Hakulinen (1999: 45) report occurrences in Eastern Finnish in (71) and (72), as well as in the literary language of Agricola in (73).

Eastern Finnish

- (71) *poika se laula-a.*
 boy DEM sing-3SG
 ‘The boy is singing.’ (Kettunen 1943: 399)

- (72) *poja-t ne marsi-t.*
 boy-PL DEM.PL march-3PL
 ‘The boys are marching.’ (Ojansuu 1922: 82)

Old Finnish

- (73) *Herra se antoi, Herra se ott-i.*
 Lord DEM give-3SG.PST lord DEM take-3SG.PST
 ‘The Lord gave and took.’ (L. Hakulinen 1999: 45)

There is also a suggestion by Kiuru (1990) that this topical use of postposed demonstrative *se* alongside the postposed 3rd person pronoun *hän* in Old Finnish bible translations by Agricola was motivated by the early modern Swedish original text. In Swedish, namely, in the 3rd person pronouns *han* ‘he’ and *hon* ‘she’ are used in the clause-second position as resumptive pronouns of (seemingly) definite referents, such as parallel examples in (74) and (75).

Early Modern Swedish

- (74a) *Herren han weet menniskiernes tankar.*
 lord.DEF 3SG know.PRS man.PL.DEF.GEN thought.PL

Old Finnish

- (74b) *HERRA hen tietä ninen wijsasten aijatoxet.*
 lord 3SG know.3SG DEM.PL.GEN man.PL.GEN thought.PL
 ‘The Lord knows those people’s thoughts.’ (Kiuru 1990: 289)

Early Modern Swedish

- (75a) *men kerleken han förbättrar.*
 but love.DEF 3SG ameliorate.PRS

Old Finnish

- (75b) *mutta rackaus se ylesrakenda.*
 but love DEM ameliorate.3SG
 ‘But love will ameliorate.’ (Kiuru 1990: 289–290)

This observation suggests that in Agricola’s dialect, the demonstrative *se* was already used as a 3rd person pronoun (mentioned in Section 2.4.1), which makes the alternation in this topical use possible. Nevertheless, the current study will not discuss the development of the Proto-Finnic 3rd person pronouns **hän/he* to discourse

markers, instead limiting the scope to that of demonstratives (see a detailed description of morphosyntactic and pragmatic functions of the enclitic particle *-hAn* in Finnish by Laitinen 2002, 2005; Yurayong & Kittilä, forthcoming).

The second type is observed in Karelian, in which the Proto-Finnic medial demonstratives **se/ne(t)* can also appear with non-nominal parts of speech, such as a verb in (76) or interrogative pronoun in (77). This extends the context of use, making the demonstratives less article-like. However, the morphosyntactic and pragmatic behaviour of the Karelian postposed demonstratives does not deviate from the outer ring Finnic type discussed above.

Karelian proper

- (76) *en še mie ole kaunis, ...*
 NEG.1SG DEM 1SG be.CNG beautiful
 ‘I am not beautiful.’ (Konkka 1963: 120)

- (77) *ka midä še miä roa-n?*
 well what DEM 1SG do-1SG
 ‘Well, what do I do?’ (Konkka 1963: 135)

In (76), the demonstrative follows a negative verb *en* ‘(I am) not’ while follows an interrogative pronoun *midä* ‘what’ in (77).

The third type is observed in Veps in which the postposed demonstratives have a free host attachment, meaning that the demonstratives can follow any part of speech (see also the description in M. Zajceva 1981: 170, 298–299; Grünthal 2015a: 279–281). Morphologically, the demonstratives agree with the head noun in number but not in case, and always remain in the singular form *se* after other parts of speech. Consequently, Veps speakers only use two forms, *se* for singular and *ne* for plural, e.g., *kodižespei se* ‘from the house’, *praznikeižed ne* ‘the feasts’ and *om se* ‘(it) is.’ Syntactically, there is no restriction on which position in the clause the phrase with demonstrative can occur. Additionally, they can occur multiple times in the clause, as shown in (78).

Central Veps

- (78) *aka-l se mužik se Piteri-š*
 wife-ADES DEM husband DEM Saint_Petersburg-INES
 ‘The husband of the lady is in Saint Petersburg.’ (Kettunen 1943: 166)

As a side remark, the demonstrative *se* can also be used to form indefinite pronouns in Veps, e.g., *ken-se* [who-DEM] ‘someone’, *mi-se* [what-DEM] ‘something’, which is considered as a polysemy copying of the Russian model, cf. Russian *kto-to* ‘someone’, *čto-to* ‘something’ (see also N. Zajceva 1995: 104–105; Karjalainen 2019: 73–77). The same pattern is also observed in Estonian with the interrogative pronouns *kes*

‘who’ and *mis* ‘what’, as in *kes-se* ‘someone’ and *mis-se* ‘something.’ In any case, this type of postposed demonstrative will not be discussed any further in the present study.

Compared to the four contexts of use identified by Kettunen (1943), the more recent studies by Larjavaara (1986) and Yurayong (2020) define three pragmatic functions for all the three types of Finnic postposed demonstratives presented above. Commonly, Finnic postposed demonstratives can encode either referential uses such as anaphor in (79), non-referential uses such as topicalisation or focalisation in (80), and discourse-pragmatic uses such as expressing surprise in (81) (see also Yurayong & Kittilä, forthcoming, for a discussion on the extended dimension of pragmatic functions).

Central Veps

- (79) *lu-d ne kaik hānou d’ät-tud oli-ba.*
 bone-PL DEM.PL all 3SG.ADES leave-PTCP.PASS be.PST-3PL
mina lu-d ne kogo-ho tačl-i-n.
 1SG **bone-PL** **DEM.PL** pile-ILL throw-PST-1SG
 ‘All the bones of his were left. I was throwing those bones to a pile’.
 (Zajceva & Mullonen 1969: 173–176)

- (80) *lapse-d ne voike-ta-s, sõ-da paki-ta-s,*
child-PL **DEM.PL** cry-PASS-PST food-PART ask-PASS-PST
a leiba-t ei-le. a staruh se pagiže-b: ...
 and bread-PART NEG.3SG-be.CNG and **old_man** **DEM** say-3SG
 ‘The children cried, craved for food, there is no bread. The old man says: ...’
 (Zajceva & Mullonen 1969: 187–188)

Finnish

- (81) *sinä se osaa-t tämä-n.*
 2SG **DEM** can-2SG this-ACC
 ‘(Wow), you can do this. (You really can.)’ (Yurayong & Kittilä, forthcoming)

Based on the contexts of use and functions preliminarily provided by the previous studies above, the current study is capable of diving deeper into pragmatics, and identifying even more fine-grained domains of use of Finnic postposed demonstratives (see the data analysis in Chapter 6).

Regarding contact between Finnic sub-branches, Larjavaara (1986: 2–9, 307–323) states that Karelian and Veps lost contact with other Finnic languages to their west after the formation of Proto-Ladogan-Finnic (see Figure 11 in Section 2.2.1). Consequently, the primary contact switched to Russian, which has remarkably separated Karelian and Veps from their western sister languages in the whole language system. Particularly with regard to the use of postposed demonstratives, Russian is the most relevant contact language to investigate. Given that Larjavaara considers the postposed demonstrative as a borrowing from Russian to Karelian and Veps, this also

implies a chronology that the postposed demonstrative could hardly be reconstructed to Proto-Ladogan-Finnic, when Karelian and Veps separated themselves from East Finnish, in which no productive use of the same construction is observed.

Nevertheless, Larjavaara does not pay attention to the fact that this language feature is not common in Central and Southern Russian dialects, nor in fact in the majority of Slavic languages. Thus, this is likely not a unidirectional contact-induced phenomenon, but rather a bidirectional reinforcement between North Russian dialects and adjacent Finnic languages (see Yurayong, 2020, and further discussion in Sections 5.2.2 and 7.3.2). Given that Kettunen also associates the use of Veps postposed demonstratives to possessive suffixes, this makes a good parallel with the uses of possessive suffixes in other Uralic languages in the east (see the discussion in Section 5.3).

5.1.2 POSTPOSED DEMONSTRATIVES IN SAAMI LANGUAGES

The use of demonstratives as articles is usually not considered in descriptions of Saami languages without intense Slavic contact (cf. Bergsland 1950; Nickel & Sammallahti 2011; Wilbur 2014; Feist 2015; Ylikoski 2020). However, E. Itkonen (1966: 257) does mention a North Saami variety spoken in Finnmark, where the prenominal demonstrative *dât* ‘this’ is used in an article-like function, presumably copied from Norwegian. As for demonstratives in the postnominal position, the observation of such phenomena is often given by scholars who have themselves observed from speakers, especially in eastern Saami languages of the Kola Peninsula (see, e.g., T. Itkonen 1958: 583–585). Kildin Saami, for instance, uses the proximal demonstrative *tedt* ‘this’ after the head noun in an article-like manner (Kert 1971: 224–225). This is most likely a polysemy copying from the use of the postposed demonstrative *-to* in North Russian dialects. The (North) Russian *-to* has also been borrowed into Kildin Saami, and used in the same function as in Russian (Rießler, in press).

Reports from language fieldworkers imply that the tendency of demonstratives becoming a postposed article has spread further to the northern part of Northwest Russia as an areal feature (see also Bartens 1999: 83). One characteristic that eastern Saami languages share with Indo-European languages in contact is an absence of possessive suffixes, the use of which has become obsolete, especially in Kildin and Ter Saami, partially due to apocope of word-final vowels that also took place in Finnic languages in the south. The obsolescence of possessive suffixes entails that no other grammatical element is competing with demonstratives for the postnominal position (see the similar case in Finnic discussed in Section 5.1.1). In any case, the frequency of use in the mentioned Saami languages varies enormously between speakers (see also a similar sociolinguistic variation in Finnish in Juvonen 2000: 195–197, discussed in Section 5.1.1). Some informants abundantly use postposed demonstratives, while others do not use them at all. The general assumption for this variation is relative to the degree of competence and exposure to the Russian language among individual Saami speakers (Michael Rießler & Rogier Blokland, p.c.).

In Kildin Saami, concretely, the initial consonant in the proximal demonstrative *tedt* may also assimilate into *d* according to the phonological environment where the

final syllable of head word ends with a vowel or voiced consonant (Kert 1971: 224). This obviously indicates the loss of accent and cliticisation of demonstratives (see Rießler, in press, and earlier discussion in Section 4.2). For instance, the language archive of Kildin Saami (Rießler et al. 2005+) contains examples for the use of the postposed demonstrative in singular nominative in (82) and accusative form in (83).

Kildin Saami

- (82) *kas't tedt gāz-dedt el'ke? gāz aps!*
 wherefrom DEM.PROX gas-DEM.PROX son gas smell.3SG
 'Where does this gas (smell) come from, my son? The gas does smell!'
 (Rießler et al. 2005+, sjd20070421kudckullj.eaf)

- (83) *pēnnga kuennt sūrmass-tenn.*
 dog carry.3SG ring.ACC-DEM.PROX.ACC
 'The dog carries the ring.' (Rießler et al. 2005+, sjd19750000asf-piennemoajnas.eaf)

Example (83) shows that the Kildin postposed demonstratives also morphologically agree with the head noun. The use observed from this Kildin Saami material seems to be related to definiteness marking. Ultimately, this might be due to change in constituent order, in which postnominal demonstratives have simply become more frequent in eastern Saami languages, as a similar tendency is widely observed also in Akkala, Skolt and Ter Saami (see also T. Itkonen 1958: 583–585), resembling the ongoing constituent order shift in Finnic languages spoken in Russia (see Ojanen 1985: 230–235).

In addition to the referential use above, the information-structural use of a proximal demonstrative *dat* in a clause-second position as topic marker, similar to the Finnish examples (70), (71), (72), and (74), is also reported from North Saami by Fernandez-Vest (2009: 49) as well as eastern Saami languages such as Akkala, Skolt and Ter Saami (Markus Juutinen, p.c.), shown in (84) and (85).

North Saami

- (84) *moai Ándi-in dat oinn-ii-me su Guovdageainnu-s.*
 1DU Ánde-COM DEM.PROX see-PST-1DU 3SG.ACC Guovdageaidnu-LOC
 'It was I and Ánde who saw him/her in Guovdageaidnu.' (Fernandez-Vest 2009: 49)

Skolt Saami

- (85) *ton tōt kuāhtlas-n'ouč.*
 2SG DEM.PROX pretender
 'You (are) such a pretender!' (T. Itkonen 1958: 583)

Clause-second demonstratives in North Saami appear to be indeclinable as in (84), but the plural forms are also attested in eastern Saami languages, e.g., Skolt Saami *oummu tōk* ... [human.PL DEM.PL] 'people, they ...' (T. Itkonen 1958: 583). Furthermore, the

clause-second demonstrative *dat* in North Saami can also be substituted by the enclitic particle *-han* borrowed from Finnish *-hAn*, as in (86), or a Saami 3rd person pronoun *-son* as, in (87) (see Toivonen [1955]1987: 98).

North Saami

- (86) *áhčči-han nu muiŋal-ii.*
 father-HAN so tell-PST.3SG
 ‘Father told (us) so.’ (Toivonen [1955]1987: 98)

- (87) *mii-son go in gávnna?*
 how-3SG when NEG.1SG find.CNG
 ‘What if I cannot find (it)?’ (Toivonen [1955] 1987: 98)

This topical use of demonstratives in the sense of Wackernagel’s law may be common across Saami languages (see the similar tendency across the Finnic languages discussed in Sections 5.1.1 and 7.2).

5.1.3 DEFINITE NOMINAL DECLENSION IN MORDVINIC LANGUAGES

When it comes to the question of the postposed demonstrative as a marker of definiteness, the most representative Uralic branch is Mordvinic. As opposed to Hungarian, which only uses prenominal indefinite and definite articles, Mordvinic languages are the only Uralic branch to have a complete definite nominal declension, which can even be considered as the most representative language group of an areal phenomenon of the postposed demonstrative, observed in Northwest Russia (see R. Bartens 1996, 1999: 83–88).

Morphologically, the Mordvinic definite declension comprises of four types of elements depending on number and case, as the Erzya example in Table 10 shows. In the nominative form, a medial demonstrative *še* is attached to noun as *-ś*. R. Bartens claims that there is no trace of a separate accusative form of *še* since Proto-Mordvinic. For this reason, other case suffixes in singular are formed by a combination of genitive marker *-ń* and a proximal demonstrative *íe*, which gives the suffix base *-ńí-*. As for locational cases, the suffix base is a distal pronoun *e-* and its derivational stem *ez-*, e.g., Moksha *eza* ‘there to’, *esa* ‘there’ ja *esta* ‘there from.’ The last group of suffixes is found in plural forms, whose marker is the medial demonstrative *ńe* attached to the plural suffix *-ť*, yielding the suffix base *-ńíe-*. Case suffixes are attached after this base, e.g., Erzya *veľe-ť-ńe-ste* [village-PL.NOM-PL.DEF-ELAT] ‘from the villages.’ However, several locational cases can be alternatively constructed by a genitive form and postposition, e.g., *ejste* for elative, *ejse* for inessive, *ezga* for prolative and *ejška* for comparative. Raun (1988: 101–102) also concludes that apart from the definite nominative *-ś* or *-ć* (after *ń*), genitive *-ť* and allative *-ťi* in Moksha, the rest of the cases are considered as postpositional suffixes in Moksha grammars (see also grammars by native speakers, Moksha by Aljamkin et al. 2000, and Erzya by Cygankin et al. 2000).

Case	Singular	Plural
Nominative	-ś	-í-ñe
Genitive-Accusative	-ñ-t	-í-ñe-ñ
Dative-Illative-Allative	-ñ-t-eñ	-í-ñe-ñeñ
Ablative	-dE-ñ-t	-í-ñe-dē
Inessive	-sE-ñ-t	-í-ñe-se
Elative	-stE-ñ-t	-í-ñe-ste
Prolative	-va/Ka-ñ-t	-í-ñe-va
Translative	GEN + <i>ladso</i> ‘in a manner’	GEN + <i>ladso</i> ‘in a manner’
Comparative	-ška-ñ-t	-í-ñe-ška
Abessive	-(v)tEmE-ñ-t	-í-ñe-víeme

Table 10. *Erzya definite declension*

In the language system, the choice between pronominal demonstrative *śe* and suffixed demonstrative *-ś* in the definite declension is based on an addressee-related pragmatic function, that is, the former bears a restrictive and contrastive reading, while the latter functions as a marker of definiteness (R. Bartens 1999: 83–84). As for the definite declension, it can also encode establishing a referent (in Becker’s term 2018) which is discourse-new, to be elaborated on in a later statement. This type of cataphoric use is characteristic of definite articles in the criteria by Hawkins (1978: 150) and Himmelmann’ (1997: 93–101). Generally, a definite subject tends to get definiteness marked directly on the noun, while a definite direct object is more frequently marked through the object conjugation of a verb, rather than the definite declension on noun (cf. Finnic languages in which the definiteness of a direct object is controlled by differences in case marking, accusative vs. partitive, mentioned in the beginning of Chapter 5).

From a diachronic viewpoint, Mordvinic languages tend to maintain a relational system of definite direct object marking. That is, if there is a definite nominal declension, there will also be an objective conjugation, with both having simultaneously developed side-by-side (Keresztes 1999). Keresztes identifies the demonstrative *śe* as a source of the suffix *-ś* in the objective conjugation, and uses this scenario to explain a parallel in Hungarian, Mansi, Khanty, and Samoyedic, in which the equivalent suffix is etymologically a set of person possessive suffixes (see also a similar description of Nenets definite verbal conjugation in Sebestyén 1976).

Keresztes also describes the co-occurrence of possessive suffixes and objective conjugation, which shows a similar tendency among Mordvinic, Mansi, Khanty, and Samoyedic languages, in which a nominal declension has extended to verbal conjugation. Keresztes even suggests that the use of 3rd person possessive suffix to express definiteness in verbs might have already occurred in Proto-Uralic, still preserved in modern Mordvinic and possibly also in Permic languages. To a certain degree, this hypothesis is possible because several branches of Uralic languages, including the geographically removed Hungarian, still use such objective conjugation forms today. The exception is Finnic and Saami, which do not have the objective conjugation. Whether this is due to Germanic contact, or a misinterpretation by Keresztes about the archaism of this construction, is subject to further investigation.

5.2 DEMONSTRATIVES AS MARKERS OF DEFINITENESS AND INFORMATION STRUCTURE IN SLAVIC LANGUAGES

Grammaticalised definite articles are not observed in all Slavic languages, although the Balkan type of enclitic definite articles in Bulgarian, Macedonian and in the majority of Torlak dialects (transitional between Bulgarian and Macedonian, on the one hand, and Serbian, on the other) is well-known in the literature on definiteness. In any case, a variety of studies have tried to identify the grammaticalisation of demonstratives across Slavic languages (Section 5.2.1), particularly Russian postposed demonstrative *-to* (Section 5.2.2).

5.2.1 THE DEVELOPMENT OF ADNOMINAL DEMONSTRATIVES IN SLAVIC LANGUAGES

A question of adnominal demonstratives as definite articles in Slavic languages has been disputed for over two centuries. Kurz (1937–1938) gives a comprehensive summary of the state of knowledge achieved until the Second World War. Kurz mentions two remarkable names, Josef Dobrovský and Jernej Kopitar, who conducted initial work on this research question. In their early language descriptions, Dobrovský (1792: 68) and Kopitar (1808: 214–215) simply regard the article-like use of preposed demonstratives in Czech and Slovene, respectively, as a replication of Germanic syntactic model. This idea persists today as well (see, e.g., Miklosich 1883: 115–130; Vaillant 1977: 266; Trovesi 2004; Marušič & Žaucer 2006).

Later, Kopitar (1857: 230–243, 320–323) familiarised himself with the Balkan *Sprachbund* and observed that postposed demonstratives are used in languages from multiple Indo-European branches: Albanian, Bulgarian, and Romanian. As this feature is not attested in Latin, he proposed an idea that the Romanian postposed article could have been a substrate feature acquired from the Paleo-Balkan languages such as Illyrian, Thracian or Gothic (also Dacian in Gebaura & Gebauer 1874: 43–56). To this newly proposed idea, Dobrovský (1818: 24–25; 1822: 608–611) also brought Old Church Slavonic to the discussion, as he began to recognise the use of article-like element in long (i.e. definite) forms of adjectives, to which the anaphoric pronoun **jb* is attached, e.g., *novъ/novo/nova* [INDEF] vs. *novъ-jb/novo-je/nova-ja* [DEF] ‘new.’ A similar observation is made by Meillet (1921: 37–38) who proposed that the definite paradigm of adjectives should have already emerged in Late-Proto-Slavic, as he observed such occurrences in many recensions of Church Slavonic. Of course, the paradigm is not preserved in all modern Slavic languages. Most languages have, at least, however, lost the indefinite paradigm. Consequently, only the definite paradigm survives, which explains why the adjectives are a formally separate class in Slavic.

A generalisation of the use of articles in the entire Slavic language branch by Miklosich (1883: 115–130) and Vaillant (1977: 265–266) suggests that articles might have already been present in the Proto-Slavic stage. Miklosich argues that Baltic languages, such as Old Lithuanian, also show a similar construction, supporting the idea of articles in Slavic being an old feature. He even considers an extreme scenario

in which the use of articles could be reconstructed to a higher level of intermediate proto-language, namely Proto-Germano-Balto-Slavic. However, the classification of intermediate proto-languages based strictly on articles is not sufficient to verify such a wide hypothesis on genealogical relationship, and we know from a Germanic case that the definite article is a recent innovation which has secondarily developed from a demonstrative. Moreover, the presence of such a tendency in Baltic and Slavic languages does not necessarily imply any Proto-Baltic inheritance, but rather a “parallel development” (see the definition in Section 3.3.2), which would have taken place after the branching of Proto-Slavic from Proto-Baltic.

Still in the same study, Miklosich lists occurrences of articles in different Slavic languages. In any case, he only considers a postnominal relative pronoun *jb-že* ‘who’ (and its whole declensional paradigm) as the only *per se* article in Old Church Slavonic, while never regarding the adnominal demonstrative *tb* as a proper article. He suggests in conclusion that the use of articles is likely adopted through language contact. For instance, articles in Slovincian and Sorbian are a replication of German model, as well as Bulgarian articles as a structural copy of Illyrian (see a similar idea of copying in Dobrovský 1792; Kopitar 1808 and later studies mentioned above). In any case, the most crucial thing in his work, directly related to the current study, is that Miklosich was the first scholar to publish his finding that North Russian dialects also use postposed articles, in a similar way to postposed definite articles in Bulgarian. Nevertheless, articles in Slavic languages had not been studied in a systematic way until Leskien (1879: 163, 522–523) who discussed article-like uses of all three series of demonstratives in the Balkan Slavic languages: *-s/v-* (proximal), *-t-* (medial), and *-n-* (distal).

A speculative claim by Miletich (1887: 305–331) argues for a connection of another *s-* article in Bulgarian dialects to Indo-European masculine nominative suffix **-s*, which had been reduced and dropped out already in the stage of Proto-Slavic, e.g., Proto-Slavic **vīlku-Ø* ~ Proto-Baltic **wilkó-s* ~ Sanskrit *vṛka-s* < PIE **ulḱʷ-o-s* ‘wolf’ [NOM.SG.MASC]’ (Derksen 2008: 536–537). However, this attempt is far from convincing, and I shall reject this idea of archaism, as the Proto-Indo-European case ending suffix *-s* must have been dropped at the latest in Late-Proto-Slavic, so the article *-s-* in Bulgarian dialects must have secondarily developed from the proximal demonstrative **sb*.

Furthermore, Miletich claims that after the loss of case inflection, Bulgarian dialects develop an analytical article phrase, where a short genitive form of the 3rd person pronoun (*e*)*go* ‘him, it’ is placed after the head noun, e.g., *dete go* ‘the child (lit. child of its).’ This is an interesting parallel with Uralic and Turkic languages, among others, which use possessive suffixes to express definiteness (discussed in Section 5.3).

Miletich also points out such fossilised forms in Old Church Slavonic as *gradosb* ‘the town’ and *dbnbtb* ‘the day’, which indicate phonological reduction and cliticisation of the demonstratives toward the head noun. Vaillant (1977: 265–266) also discusses similar contexts, in which a demonstrative is cliticised to its head noun, the final reduced vowels *ɔ/b* of which becoming vocalised, e.g., Old Church Slavonic *dbnbsb* (cf. Bulgarian/Czech *dnes* and BCMS *danas* ‘today’). A similar mechanism lies behind the development of enclitic articles in eastern South Slavic languages. The

grammaticalisation of the Old Church Slavonic demonstrative *t-* to an enclitic definite article was becoming consistently more productive in Middle Bulgarian texts from the 13th century, e.g., *zlbjio**t** rab**b*** ‘the bad servant’ (Svane 1961, 1962, and see a discussion on the absolute chronology in Lindstedt 2014).

In any case, the status of adnominal demonstratives as articles is ultimately difficult to justify, because Old Church Slavonic texts are usually written without space, and words are put together word after word. To prove that a phonological reduction had already taken place by that time, the eternal question of how word prosody works in Old Church Slavonic plays an important role, for which there is unfortunately no longer any native speaker for us to consult with. Another issue discussed by Vaillant (1977: 259–260) as well as Vahros (1951) is contact influence from Greek in Old Church Slavonic syntax, especially in the use of clitic words. Nevertheless, Slavic languages tend to replicate the Greek model only in the use of clause-second particles, i.e. Wackernagel’s clitic.

Based on his quantitative analysis of all demonstratives in a corpus of Old Church Slavonic, Kurz (1939–1946: 286–287) states that postposed demonstratives are not articles, and their position in noun phrase is interchangeable, later illustrated by Večerka (1993: 82–85) with an example ‘this good man.’

- | | |
|--|----------------|
| 1. <i>sb dobrŭjb možb</i> | [DEM good man] |
| 2. <i>sb možb dobrŭjb</i> | [DEM man good] |
| 3. <i>dobrŭjb sb možb</i> | [good DEM man] |
| 4. <i>možb sb dobrŭjb</i> | [man DEM good] |
| 5. <i>dobrŭjb možb sb</i> | [good man DEM] |
| 6. <i>možb dobrŭjb sb</i> | [man good DEM] |

The set of possibilities presented above can be divided into three main types of demonstratives: preposed (1–2), interposed (3–4), and postposed (5–6). In a similar quantitative manner to Kurz, Večerka reports the frequency of these patterns in Old Church Slavonic with the ratio 3:2:6:2:2:4. A striking observation here is that the use of interposed demonstratives in pattern 3, as in modern Balkan Slavic languages like Standard Bulgarian *dobrij**ŭ**t m**ă**ž* [ADJ.DEF + NOUN], is the most frequent. Despite obvious influence from Greek syntax, Večerka attempts to describe interposed demonstratives as semantically unstressed components, usually found in noun phrases that contain more words. Meanwhile, pattern 1 resembles a frequent construction in West Slavic languages which are in contact with Germanic. In the end, this simply seems to be a question of syntactic variation and stylistics, by which Kurz describes the postnominal use of demonstratives as “expressive reading”. Functionally, all these demonstratives still carry a prototypical anaphoric use, that is, referring to a previously mentioned referent, but no discussion of other referential uses (discussed in Section 4.3) has been given for Old Church Slavonic.

The uses of preposed demonstratives in modern Slavic languages have been discussed rather extensively more recently. For instance, the use of German definite

articles *der/das/die* [MASC/NEU/FEM] is said to have introduced a similar model of prenominal demonstratives in West Slavic, e.g., Czech *ten/to/ta* and Upper Sorbian *tón/to/ta* (identically corresponding to German), as well as in Baltic, such as in Old Prussian *stas* (Vaillant 1977: 266; Nau 1995: 122). Despite there being various approaches and criteria to identify the emergence of definite articles, the general line of thinking still remains the same, and does not consider such adnominal demonstratives as articles proper, but rather as demonstratives with extended referential uses, e.g., in Upper Sorbian (Berger 1999; Breu 2002; Marti 2012), Polish (Bacz 1991; Mendoza 2005; Bartnik 2015), Czech (Cummins 1998), Slovene (Marušič & Žaucer 2006), BCMS (Trenkić 2004) and Belarusian (Laryěna 2016).

As for modern South Slavic languages with postposed demonstratives as definite articles, Lindstedt (2009: 81–82, 2014) regards the use of such enclitic articles in Balkan languages not only as a syntactic device to form a definite noun phrase in phrase-second position, but also to mark the status of a noun phrase in a thematic structure of a clause such as topic (see the connection between topicality and definiteness in Givón 2001: 472–473, discussed in Section 4.4.1). This implies that the uses of postposed demonstratives in Slavic languages cannot be understood only on the basis of referential uses related to definiteness and specificity (cf. Mladenova 2007). Instead, other non-referential uses involving information structure (particularly topicality) must also be considered (see a similar line of thinking on the uses of postposed demonstratives in Finnic in Section 5.1.1, and in Russian in Section 5.2.2).

5.2.2 POSTPOSED DEMONSTRATIVES IN RUSSIAN

Following the mention of article-like uses of postposed demonstratives *-to* in North Russian by Miklosich (1883), Vahros (1951) also discusses a similar use in the Old South Russian chronicle. Kiparsky (1967: 148–152) even observes such uses across the entire declensional paradigm in an autobiography of Avvakum (1620/21–1682), an Orthodox priest from the area of present-day Nizhny Novgorod. This written source has arguably imitated Avvakum's actual speech, which contains a productive use of demonstratives after head noun, as reconstructed by Kiparsky in Table 11. This article-like use of the North Russian postposed demonstratives has also gained attention from scholars working on definiteness marking in a typological framework (Krámský 1972: 187; C. Lyons 1999: 48). From a functional viewpoint, however, Mendoza (2011) does not regard these demonstratives in Avvakum's autobiography as articles but rather as “articloids”, i.e. definite articles *statu nascendi* (in Voge's term 1958).

Regardless of their controversial status as definite articles, Kiparsky (1969: 25–26) compares the North Russian paradigm with the Mordvinic definite declension (discussed in Section 5.1.3), proposing that it is unlikely that the North Russian postposed demonstratives would have been imported from South Slavic (cf. Vahros 1951 who believes in the ultimate possibility of Balkan influence coming through the Church Slavonic literature to Russian), but rather that this is a Uralic substrate feature. Kiparsky uses the geographical argument that Avvakum is known to have resided in areas only a hundred kilometres away from Mordvinic speaking regions, and this could have influenced his Russian vernacular.

Case			Nominative	Accusative	Genitive	Dative	Locative	Instrumental
Singular	Noun	MASC	<i>-o-tъ</i> <i>-e-tъ</i>	↔	<i>-a-togo</i> <i>-a-tovo</i>	<i>-u-tomu</i>	<i>-ě-tomъ</i> <i>-i-tomъ</i>	<i>-omъ-těmъ</i> <i>-emъ-těmъ</i>
		NEU	<i>-o-to</i> <i>-e-to</i>					
		FEM	<i>-a-ta</i> <i>-ta</i>	<i>-u-tu</i> <i>-Ø-tu</i>	<i>-y-toj</i> <i>-i-toj</i>	<i>-ě-toj</i> <i>-i-toj</i>		<i>-oju-toju</i> <i>-iju-toju</i>
	Adjective	MASC	<i>-oe-tъ</i>	↔	<i>-ago-togo</i>	<i>-omu-tomu</i>	<i>-omъ-tomъ</i> *	<i>-imъ-těmъ</i>
		NEU	<i>-oe-to</i>					
		FEM	<i>-aja-ta</i>	<i>-uju-tu</i>	<i>-oj-toj</i> <i>-aj-toj</i>		<i>-oj-toju</i> * <i>-ej-toju</i> *	
Plural	Noun	MASC	<i>-y-te</i> <i>-i-te</i>	<i>-y-tě</i> <i>-i-tě</i> *	<i>-ovъ-těhъ</i> <i>-ej-těhъ</i>	<i>-amъ-těmъ</i> <i>-jamъ-těmъ</i>	<i>-ahъ-těhъ</i> <i>-jahъ-těhъ</i> *	<i>-ami-těmi</i> <i>-jami-těmi</i>
		NEU	<i>-a-tě</i> / <i>-a-te</i> <i>-ja-tě</i> / <i>-ja-te</i> *					
		FEM	<i>-i-te</i>	<i>-y-tě</i> <i>-i-te</i>	<i>-ъ-těhъ</i> <i>-ej-těhъ</i>			
	Adjective	MASC	<i>-yje-te</i> <i>-ije-te</i> *	<i>-yje-tě</i> * <i>-ije-tě</i> *	<i>-yhъ-těhъ</i> <i>-ihъ-těhъ</i> *	<i>-ymъ-těmъ</i> <i>-imъ-těmъ</i> *	<i>-yhъ-těhъ</i> * <i>-ihъ-těhъ</i>	<i>-ymi-těmi</i> * <i>-imi-těmi</i> *
		NEU	<i>-yje-tě</i> * <i>-ije-tě</i> *					
		FEM	<i>-yje-te</i> * <i>-ije-tě</i> *	<i>-yja-te</i> <i>-ija-te</i> *				

Table 11. Paradigm of definite declension in Avvakum's Russian (* = the author's supplementary reconstruction)

Another fact that might support the Uralic substrate hypothesis is the distribution of postposed demonstrative *-to*, which is only observed in Russian, predominantly in the northern variety, with less frequent occurrences in central and southern varieties. There is no trace of the postposed demonstrative *to* in any other East Slavic language without intense Uralic contact, i.e. Belarusian, Rusyn, and Ukrainian (Vlasto 1967: 305, 312; Kasatkina 2007: 109; information confirmed by data in the parallel corpus of Slavic languages *Parasol*, Ruprecht von Waldenfels, p.c.).

Sharing the same thought, Angere (1956: 223), Serebrennikov (1956: 54), Veenker (1967: 88–90), and Kusmenko (2008: 118–119) even present a speculative idea that the use of possessive suffixes in Uralic languages could have motivated the functional extension of postposed demonstrative in North Russian, considering that the non-possessive pragmatic uses of the possessive suffixes have developed in many Uralic branches such as Mari, Permic, Mansi, Khanty, and Samoyedic (see also Tauli 1966: 148; R. Bartens 2000: 122–123). This idea of a parallel to the Uralic possessive suffixes frees scholars from the restricted notion of definiteness, and opens a new territory in the discussion on the functions of the North Russian postposed demonstratives (see a similar trend change in Uralic linguistics in Section 5.1).

Discourse-pragmatic uses of the postposed demonstrative *-to* in Central (Standard) Russian has been discussed in many accounts. Functionally, Russian *-to* can be regarded as a discourse marker that contributes to information structure through topicalisation and delimitation of a set of alternatives (i.e. focus and contrast), or expresses the interlocutors' state of knowledge such as familiarity or surprise (as previously discussed in Rathmayr 1985; Bonnot 1986, 1987, 1990, 1991; Zybatow 1990; Grenoble 1998: 199–200; McCoy 2001, 2003; Bolden 2008). Diachronically, Gvozdanović (2019: 126) argues that *-to* could have originated from the use of the demonstrative *to(tь)* as a connective particle to separate main clause from conditional (or temporal) subordinate clause in a formula: X, *to* Y '(if/when) X, then Y', as attested in Old Rus' (88) and Old Novgorod (89).

Old Rus'

- (88) *ažb ubьetь mužb muža, to mьstiti bratu brata ...*
 if beat.3SG man man.ACC DEM avenge.INF brother.DAT brother.ACC
 'If one freeman kills another, (then) a brother may avenge his brother ...'
 (*Russkaja Pravda*, in Vlasto 1986: 237)

Old Novgorod

- (89) *prisъli kuny. ože li ne prisъleši to ti vъ poly.*
 send.IMP.2SG kuna.PL if PTCL NEG send.2SG DEM PTCL to half.DU
 'Send the money. If you do not send [the money], it [the loan] will become half
 [at 50% interest rate].' (document № 915 1050–1075, Zaliznjak 2004: 243)

In (88) and (89), the connective particles *to* separate the acts of freeman killing others and not paying the debt as conditional subordinates from the main clauses providing a parallel of vengeance between brothers, and a consequence of delayed payment, respectively (see also Kopotev 2006 for a historical and typological analysis of a similar construction with the manner demonstrative *tak* 'so': X *tak* X '(if/when) X, so X'). Typologically, such clause connectors are also one of the more common grammaticalisation targets of demonstratives mentioned in Heine and Kuteva (2002: 107–108, discussed in Section 4.2).

Gvozdanović's postulation above also corresponds to the proposal by Haiman (1978, discussed in Section 4.4.1), namely that topics are conditionals. Interestingly, this construction dates already back to the very beginning of the East Slavic literary tradition, as it was already attested for instance in (88), which originates from the law text *Russkaja Pravda* written in the 11th century (Vlasto 1986: 237), as well as in Novgorod birch bark documents written between the 11th–15th centuries, seen in (89) (Zaliznjak 2004: 192). Ultimately, the development of *to* from a connective particle to a postposed demonstrative might have related to the narrowing of its use, from marking a subordinate clause to the separating of individual shorter phrases, i.e. information chunks (à Matić 2003), which can also be regarded as a strategy to maintain speech rhythm, i.e. "rhythmic-syntactic barriers" (see also Zaliznjak 2008: 47–51).

In a parallel to Russian, the use of the demonstrative *to* in clause-second position as an information-structuring device in spoken language is also reported from Czech (Šimík 2009) and Polish (Tabakowska 1989; Huszcza 2000; Cegłowski & Tajsner 2006; Rutkowski 2006). For example, the demonstrative *to* may be used to mark the topicalisation of clause-initial elements, such as those seen in (90) and (91).

Czech

- (90) *v kolik hodin jste mu ho to*
 [in how_many hour.GEN.PL be.2PL 3SG.MASC.DAT 3SG.NEU.ACC DEM]
tam tehdy chtěli dát?
 there then want.PST.PL give.INF
 ‘At what time did you want to give it to him there then?’ (Šimík 2009)

Polish

- (91) *Waterloo to wydaje się być zwycięstwo.*
 [Waterloo DEM] seem.3SG REFL be.INF victory
 ‘Waterloo seems to be a victory.’ (Rutkowski 2006: 171)

Syntactically, Wackernagel’s law is applied to the use of *to* in Czech in (90), as it is the final constituent of a clitic chain, following the other clitics such as the auxiliary verb *jste* ‘to be’, and personal pronouns *mu* ‘to him’ and *ho* ‘it.’ In contrast, Polish *to* is not a Wackernagel’s clitic, but rather a marker of dislocation constructions, as it does not form a single clitic chain with the reflexive pronoun, i.e. ***się to* in (91), but rather exclusively marks the topicalisation of *Waterloo*. In terms of stancetaking, Šimík (2009) also makes a remark that the Czech demonstrative *to* can also express the speaker’s surprise towards the mentioned state-of-affairs, i.e. mirativity (see evaluative uses of demonstratives discussed in Sections 4.5.2, and 5.1.1 for Finnic, as well as observation from the Finnic and North Russian data in Section 6.4.3).

In addition to referential uses (definiteness and specificity), other aspects of information structure and evaluation have also made their way into a discussion of the North Russian postposed demonstratives (see Kuz’mína & Nemčenko 1962: 4; Serebrennikov 1963: 129–133; Avanesov & Orlova 1965: 197; Leinonen 1998; Kasatkina 2007, 2008; Ahlborn 2018). This has resulted in scholars of the late-20th century and early-21st century no longer calling the North Russian *-to* an article but rather an “emphatic focus particle” (see the same solution for the Veps *-se/-ne* by Grünthal 2015a, discussed in Section 5.1.1), or as a “pseudo-article” (in Kasatkina’s term 2007, 2008). Again, the data discussed in Section 6.4.1 will show that the article-like use of the North Russian *-to* cannot be totally ignored.

A quantitative investigation by Trubinskij (1970) of the Pinega variety spoken in Arkhangelsk Oblast gives a good overview on variation in both form and function of the postposed demonstrative *-to*. In terms of form, despite being observed in the entire Russian dialect continuum (with some exception with dialects spoken near the borders of Belarus and Ukraine), an inflectable postposed demonstrative, e.g., *dom-ot* ‘the house’, is geographically restricted to North Russian and is observed already in

dialectological materials at the end of the 19th century (Kuz'mina 1993: 184–187; Mendoza 2011: 249). The form of demonstrative does not always match its head word in case, but often has an identical vowel to the preceding syllable, i.e. the last syllable of head word, the phenomenon which Trubinskij (1970: 66) and Meščerskij (1972: 250) name as “harmonisation”, e.g., *na beregu-tu* ‘on the shore’, *bez soli-ti* ‘without the salt’ and *na reke-te* ‘on the river.’

From the functional point of view, Trubinskij (1970) provides six contexts of use where the postposed demonstratives are observed in the Pinega variety (see also a similar description in Leinonen 1998: 77–84). Interestingly, many functions resemble those observed in Veps by Kettunen (1943: 399–403, discussed in Section 5.1.1).

1. Emphatic marker
2. Anaphoric marker
3. Contrastive marker
4. Question particle
5. Insertive marker
6. An equivalent to the Standard Russian adverb *daže* ‘even, though’

The first two contexts are referential uses, in which a postposed demonstrative can emphasise and refocus an accessible new or old referent, i.e. a topic in (92), or a previously mentioned referent in (93). Leinonen (1998) observes a pattern for the emphatic use in which the demonstrative follows the phrase (*vot*) *etot* NOUN ‘(so) this NOUN’ as in (92).

North Russian

- (92) ... *i v Ševerodvinss vyzvūli vot èta doci-ta Ol'ga-ta ...*
 and to Ševerodvinss call.PST.PL so this daughter.ACC-DEM Ol'ga-DEM
 ‘... and to Ševerodvinss, the daughter Ol'ga was called, ...’ (Leinonen 1998: 78)

- (93) *mužik da ženka žili. mužik-ot i pomēr.*
 husband and wife live.PST.PL husband-DEM also die.PST.MASC
 ‘There were a husband and wife. The husband has died.’
ženka-ta žale mužika-ta i placě, i placě vsě ...
 wife-DEM regret.3SG husband.ACC-DEM and cry.3SG and cry.3SG all
 ‘The wife mourns the husband (‘s passing away) and keeps on crying.’ (Leinonen 1998: 78)

The other four contexts, however, do not relate to referentiality, but rather non-referential uses with other discourse functions (see also Leinonen 1998; Kasatkina 2007, 2008). In such non-referential uses, the postposed demonstratives can mark a contrast between a pair of entities as in (94), or emphasise a question as in (95).

North Russian

- (94) *odna-to defka išo, defčěnka byla,*
 one-DEM girl still girl be.PST.FEM
a druga-ta už žěn'čina byla.
 and other-DEM already woman be.PST.FEM
 ‘One was still a girl, the other was already a lady.’ (Leinonen 1998: 79)
- (95) *skolb djon-ta?*
 how_many day.GEN.PL-DEM
 ‘How many days?’ (Trubinskij 1970: 56)

The use as a question particle is potentially related to a characteristic of North Russian, which lacks the dynamic intonation common to other Russian dialects, forcing a compensation through some other morphological marking, in this case the postposed demonstrative *-to* (Serebrennikov 1963: 131; Stadnik-Holzer 2006: 350; Kasatkina 2007: 106). Two other contexts of use concern the insertion of a noun which follows a pronoun of the same referent as in (96), or an “additive” use (in König’s term 1991), equivalent to the Standard Russian adverb *daže* ‘even, though,’ as in (97) (see also a similar use of “scalar focus” in Krifka 2008, discussed in Section 4.4.3).

North Russian

- (96) *u nej u sverkovi-tb*
 at 3SG.FEM.GEN at Sverkova.GEN-DEM
 ‘at her place, at Sverkova’s’ (Trubinskij 1970: 56)
- (97) *dy ja adnago-tb nja vynjančila*
 and 1SG one.GEN-DEM NEG stitch.PST.FEM
 ‘And I have not stitched even a single thing.’ (Trubinskij 1970: 56)

An important characteristic of the latter four contexts of use is that *-to* can also follow non-nominal parts of speech, such as verbs (primarily infinitives) and adverbs, making it less article-like (Meščerskij 1972: 249–251). Strikingly enough, the contexts in which postposed demonstratives are used in North Russian are very similar (if not identical) to the case of Veps *se/ne* (discussed in Section 5.1.1), in addition to similar functions of Komi possessive suffixes identified by Leinonen (1998, 2006). This leads to further discussion concerning the comparability to the use of Uralic possessive suffixes (discussed in Section 5.3), and, ultimately, the emergence of multifunctional postposed demonstratives in both Finnic and North Russian.

Retaking the Uralic substrate hypothesis in the light of new interpretation of the data, the discussion concerning emergence of the inflectable North Russian *-to* has been extended to the framework of areal linguistics. Among others, Leinonen (2002) highlights that the postposed demonstrative and likewise a “possessive-definite suffix”

are a widespread areal feature, observed not only in Russian and Uralic but also in Turkic languages of the area, which share common functions as marker of emphasis, discourse prominence and definiteness. Leinonen pays special attention to the contact zone of Veps, Komi, and North Russian as a potential “epicentre” of this micro-area (see also a similar idea in Stadnik-Holzer 2006: 352). Supporting the argument are historiography and toponymy, which illustrate early contact between Finnic and Komi speech communities (see Saarikivi 2006: 33–38, 2018, discussed in Section 2.1).

Stadnik-Holzer (2006: 352–353) argues that this is unlikely an internal development in Russian. Instead, the construction was rather introduced to Russian from the neighbouring Uralic languages which robustly use grammatical suffixes as postposed elements. A similar idea was earlier proposed also for borrowing the adnominal possessive construction of possessive suffixes from Turkic to Balkan Slavic, e.g., Turkish *baba-m* [father-Px.1SG] and Bulgarian *bašta mi* [father 1SG.DAT] ‘my father’ (Stadnik-Holzer 2004b: 11; Yurayong 2016: 147). A statistical study by Tommola (1998) also speaks in favour of an areal effect being more significant than a genealogical inheritance regarding the emergence of articles across European languages. The question of the origin of the North Russian postposed demonstratives will be discussed in the light of the data of the present study in Section 7.3.2.

5.3 NON-POSSESSIVE USES OF POSSESSIVE SUFFIXES IN URALIC LANGUAGES

Possessive suffixes share a number of functional properties with demonstratives regarding the marking of definiteness, identifiability as well as stancetaking (as already discussed briefly in Sections 4.3 and 4.5). In the same vein, many scholars state that the Uralic possessive suffixes can also express non-possessive functions such as definiteness marking and therefore regard them as equivalent to definite articles (Collinder 1960: 203; Tauli 1966: 148; R. Bartens 2000: 122–123; Fraurud 2001; Künnap 2004; Schroeder 2006; Gerland 2014). This phenomenon is particularly common among central and eastern Uralic branches (Collinder 1960: 203; Abondolo 1998: 22), for instance, in Mari (Collinder 1957: 251; Alhoniemi 1988: 90; Kangasmaa-Minn 1998: 229), Permian (Collinder 1957: 276, 301; Rédei 1988: 118; Csúcs 1988: 137, 1998: 285; Hausenberg: 1998: 313), Mansi and Khanty (Collinder 1957: 322, 349; Honti 1988: 168, 194), and Samoyedic (Castrén 1854: 207; Collinder 1957: 426, 459, 494; Sebestyén 1976; Helimski 1998: 496; Siegl 2013: 371–377).

Interestingly, Castrén (1854: 207) already observes the use of the 3rd person possessive suffix as a marker of definiteness in Samoyedic languages as early as the mid-19th century, during his pioneering fieldwork. Given its wide distribution, Janhunen (1982: 32), Künnap (2004), Stachowski (2010), and Gerland (2014) propose that the use of 3rd person possessive suffixes as definiteness markers could even be traced to Proto-Uralic. However, these attempts do not provide satisfactory empirical evidence, relying instead on shallow generalisations in descriptive grammars of each Uralic language. Furthermore, they do not take any further account of individual languages and other dimensions of functional extension beyond the marking of

definiteness. Given this problem, several studies have investigated functional parallels between Komi possessive suffixes and the North Russian *-to*, showing that their use extends beyond definiteness, to the marking of information structure (see Batalova 1975: 203; Leinonen 2002, 2009: 315; Blokland 2012: 5).

A number of previous studies have reported non-possessive uses of possessive suffixes beyond definiteness, for instance, in Mordvinic (Feoktistov 1963), Mari (Kokla 1963; Bereczki 1990), Permian (Serebrennikov 1963; Kel'makov 1996; Leinonen 2006), Mansi and Khanty (Nikolaeva 2003; G. Janda 2015), and Samoyedic (Tereščenko 1979; Hajdú 1982; Kuznecova et al. 1980; A. Kim 1987; Wagner-Nagy et al. 2002; Siegl 2013; Zayzon 2015). Based on theoretical grounds discussed in Chapter 4 and empirical evidence from these Uralic languages, two domains of use can be classified: (i) referential and (ii) non-referential uses, summarised in Table 12.

Functions		Geographical distribution	Forms	Host attachment
Referential	Deictic, Recognitional	Mordvinic, Mari, Permian, Mansi, Khanty Samoyedic	3rd person	Noun
	Anaphoric, Situationally	Mari, Permian, Mansi, Khanty, Samoyedic	2nd / 3rd person	
Non-referential	Topical	Mari, Permian, Mansi, Khanty, Samoyedic	any person	Noun or Free (Mari & Permian)
	Contrastive	Mari, Permian	3rd person	
	Evaluative	Mordvinic, Mari, Permian, Mansi, Khanty, Samoyedic	1st / 2nd person	

Table 12. *The uses of possessive suffixes in non-possessive functions*

Table 12 illustrates observations from each Uralic branch by showing the functional range of different person-marking suffixes, host attachment possibilities, and the distribution across the Uralic languages. This comparison also shows that there is a specific functional distribution between different personal pronouns, and different types of usage can break the host attachment boundary. Next, each function will be elaborated on with concrete examples.

5.3.1 REFERENTIAL USES

Resembling the use of demonstratives, many Uralic languages employ 3rd person possessive suffixes for referential uses. Several functions are observed of the classification of referential uses of definite articles by Becker (2018) and König (2018) as discussed in Section 4.3.3.

First, possessive suffixes can carry a “deictic” function that contains gesturing expression such as ‘the grass that is on the field (not elsewhere)’ in (98) and ‘the ground in the swamp (not elsewhere)’ are present in (99).

Udmurt (Permic)

- (98) *guždor vyl-yn turyn-ez čeber.*
 field on-INES grass-Px.3SG beautiful
 ‘The grass on the field is beautiful.’ (Alatyrev 1970: 85)

Forest Enets (Samoyedic)

- (99) *mod’iñ lota ne-on d’odi-ŋa-i morgan-di-ñ*
 1DU swamp on-PROL go-FREQ-1DU cloudberry-BEN-Px.ACC.PL.1DU
pe-ŋa-i morgan koo-b’ d’a-đa ŋul’ naldi-ŋa
 search-FREQ-1DU cloudberry find-SG.1SG ground-Px.3SG very be_red-FREQ.3SG
 ‘We (two) walked along the swamp. We were looking for cloudberry. I found cloudberry.
 The ground was very red.’ (Siegl 2013: 374)

Additionally, Komi can use different person suffixes to create a spatial contrast in a cognitive setting (see E. Itkonen 1996: 258; Nikolaeva 2003; Künnap 2004; Leinonen 2006). In such strategy, the 3rd person suffix would be addressee-distant as in (100), whereas the 2nd person is suffix addressee-close as in (101).

Komi (Permic)

- (100) *vöra-s lymjy-s syl-öma ñin*
 forest-Px.3SG snow-Px.3SG melt-PST.3SG PTCL
 ‘In the forest, the snow has already melted (I am telling you).’
- (101) *vöra-d lymjy-d syl-öma ñin*
 forest-Px.2SG snow-Px.2SG melt-PST.3SG PTCL
 ‘In the forest, the snow has already melted (you already heard that).’ (R. Bartens 2000: 122)

In (100), the referents are entities that are psychologically remote to the addressee, while the counterparts in (101) are psychologically close to the addressee (see also evaluative uses of demonstratives discussed in Section 4.5.2). Nevertheless, Leinonen (2006: 101) also shows that this strategy can even express a concrete spatial contrast, as in (102).

Komi (Permic)

- (102) *Palad’ poč-yd-lön matynžyk kerka-ys, da setć-ö,*
 Palad’ aunty-Px.2SG-GEN closer house-3SG and it-ILL
 ‘Aunty Palad’ has a house closer here, and she ...’
- köñkō mun-is. mi ord-o oz ñin lok. aćis*
 probably go-PST.3SG 1PL place-ILL NEG PTCL come.CNG self
 ‘...probably went there. To our place, she no longer comes.’

mövyšt-is ‘med Palad’ *pöč-ys* *oz* *neršy* *daj.*
 think-PST.3SG then Palad’ aunty-Px.3SG NEG insult.CNG PTCL
 ‘He thought to himself: may they not insult Aunty Palad’.
 (Kozlova 2002: 7, in Leinonen 2006: 101)

In (102), the 2nd person suffix indicates that ‘aunty Palad’ is, in general, living close to the addressee, whereas at the time of utterance ‘aunty Palad’ was not around in the addressee’s sphere, and was therefore referred to with the 3rd person suffix.

Second, possessive suffixes can express “recognitional” referents that arise from shared knowledge between the interlocutors, as in (103) and (104).

Forest Enets (Samoyedic)

(103) *tu-ďa* *čiki* *ibľeigo-n* *čuo.*
 fire-Px.3SG this little-PROL burn.3SG
 ‘The fire, it burned a little bit.’ (Siegl 2013: 374)

Nganasan (Samoyedic)

(104) *sürü-ďü* *donü-po.*
 snow-Px.3SG melted-3SG.PST
 ‘The snow has melted.’ (Tereščenko 1979: 95)

In (103) and (104), ‘fire’ and ‘snow’ are natural phenomena that both the interlocutors can identify without necessarily being mentioned previously in the discourse.

Third, possessive suffixes are used to mark “anaphora”, which refer to their antecedent in the discourse as in (105) and (106), in which the referents ‘park’ and ‘boat’ are mentioned again.

Komi (Permic)

(105) *važša* *bazarnej* *ploščad* *vyľ-e* *śod-isny* *park.*
 old market square top-ILL make-PST.3SG park
 ‘At the old market square, they have made a park.’
interesne *kučëm* *loę* *park-ys.*
 interesting which be.FUT.3SG park-Px.3SG
 ‘Interesting, what will become of the park?’ (Serebrennikov 1963: 129)

Selkup (Samoyedic)

(106) *qoltyt* *qanyq-qyn* *anty* *tot-ta.* *anty-ty* *lapy-kə:l* *ε:ŋa.*
 river bank-on boat stand-3SG boat-Px.3SG oar-without be.3SG
 ‘A boat stands on the riverbank. The boat doesn’t have an oar.’ (Kuznecova et al. 1980: 187)

In addition to the 3rd person suffix, anaphoric use of the 2nd person suffix is also observed. As in the case of the deictic use discussed above, the 2nd person suffix adds

a privative and intimate connotation towards the referent, such ‘the old woman,’ and ‘the girl’, towards which the speaker feels sympathy in (107) and (108).

Nenets (Samoyedic)

- (107) *puxacyako yilyewi°. puxacyako-nt° syidya səwa nyu-da.*
 old_woman live.NARR.3SG old_woman-GEN.Px.2SG two nice son-Px.3SG
 ‘There is an old woman. The old woman has two nice sons.’ (Nikolaeva 2003: 137)

Nganasan (Samoyedic)

- (108) *baarbə-ðuŋ hon-ti kobtua. kobtua-rə četuami ħeəniŋku*
 master-GEN have-AOR.3SG girl.ACC girl-Px.2SG very pretty
 ‘Their master has a daughter. The girl is very pretty.’ (Tereščenko 1979: 95)

Possessive suffixes can also mark “situationally unique” referents, which have not been previously mentioned but are unambiguous from the context as in (109), where it is unambiguous that ‘four’ refers to the group of ‘65 households’ mentioned in the previous phrase.

Komi (Permic)

- (109) *tajö grezdja-s-as völä-ma 65 ovmös,*
 this village-INES-3SG be-PST.PL 65 household
 ‘In these villages, there were 65 households, ...’
ñol’-ys na pövst-ys tyrtöm. ...
 four-Px.3SG piece they-ELAT empty
 ‘... four of them are empty, ...’ (Bondarenko et al. 2000: 52, in Leinonen 2006: 105)

As for “bridging” referents, possessive suffixes with their possessive use by default can establish such an association on the basis of inference from a part-whole or action-instrument relation (see Fraurud 2001, discussed in Section 4.3.3), so no further detailed discussion is needed here.

5.3.2 NON-REFERENTIAL USES

Beyond referential uses, other discourse-pragmatic uses are observed in connection to possessive suffixes in Uralic languages (compared to the non-referential uses of demonstratives in Finnic and Saami, discussed in Sections 5.1.1 and 5.1.2). First, possessive suffixes can function as “topic markers”. On the one hand, they anchor one entity to another previously mentioned in the discourse similarly to the anaphoric use. On the other hand, they also organise information structure which specific entity is to be raised as topic. Apart from keeping track of an active topic, this encoding often involves the switching of topic from one clause to another, as in (110) to (112).

Mari

- (110) *vaštarešy-žy* *officer* *šinč-a* ... *ruki* *vverx!* *kyčkyr-em*,
in_front_of-Px.3SG officer sit-3SG [hands] [up] shout-1SG
‘An officer was sitting in front (of her). ‘Hands up’, I shouted.’
vintovk-em *vikt-em* ... *sade* *officer-em* *šul-yš*.
gun-Px.1SG aim-1SG this officer-Px.1SG melt-PST.3SG
‘I aimed my gun. The officer got frightened (lit. melted).’ (Tužarov 1987: 68)

Khanty

- (111) *amp* *elti* *pälta-məm* *xojat* *xönt-əs*.
dog from frighten-PTCP man run-PST.3SG
‘The man who was frightened of a dog ran away.’
amp *xőtalt-əs* *lūw* *elti* *pälta-məm* *xojat-əl* *xönt-əs*.
dog get_up-PST.3SG 3SG from frighten-PTCP man-Px.3SG run-PST.3SG
‘The dog got up. Frightened of it, the man ran away.’ (Nikolaeva 2003: 138)

Nenets (Samoyedic)

- (112) *møny°* *syany°* *po-h* *tyuku°* *to-h* *xew-xøna* *xanye-d°m*.
1SG many year-GEN this lake-GEN near-LOC fish-1SG
‘I have gone fishing for many years in this lake.’
tyiki° *to-xøna-nyi* *xalya-da* *ŋoka* *ŋæsyati*.
this lake-LOC-Px.1SG fish-Px.3SG many be.3SG
‘There are a lot of fish in the lake.’ (Nikolaeva 2003: 139)

Given that the possessive markers are used to organise information structure, it has also been reported that this function is observed in clauses with a (reversed) marked word order, i.e. dislocation (R. Bartens 2000: 122; Leinonen 2006: 102–104).

Second, the use of possessive suffixes can create a “contrast” between parallels in the discourse, (i.e. “highlighting parallels” in Krifka’s term (2008), discussed in Section 4.4.3) and shown in (113) to (116)

Mari

- (113) *vüt-šö* *jog-a,* *ser-že* *kode-š*.
water-Px.3SG run.3SG bank-Px.3SG remain-3SG
‘The water is running (and/while) the riverbanks remain.’ (Bereczki 1990: 43)

Komi (Permic)

- (114) *karandaš* *me* *ñeb-i,* *a* *ručka-se* *eg* *ñeb*.
pencil 1SG buy-PST.1SG but pen-Px.ACC.3SG NEG.1SG buy.CNG
‘I bought a pencil, but as for a pen, I did not buy it.’ (Nikolaeva 2003: 140)

Udmurt (Permic)

- (115) *uli-zy* *vyl-i-zy* *kyk* *brat-jos*,
 live-PST-3SG be-PST-3SG two brother-PL
 ‘There were two brothers, ...’
pokči-ez *kuaner*, *byžym-ez* *užyr*.
 younger_brother-Px.3SG poor older_brother-Px.3SG rich
 ‘... the younger one was poor, the older one was rich.’ (Serebrennikov 1963: 133)

Khanty

- (116) *ullə-ŋən* *kat* *niŋ* *sar* *ur* *mor* *ur* *kutna*.
 be-3DU two woman thick forest tight forest in
 ‘There are two women in a thick forest.’
i *niŋ-əl* *nur* *atsam*, *i* *niŋ-əl* *nur* *nomsəŋ*.
 one woman-Px.3SG very stupid one woman-Px.3SG very clever
 ‘One woman is very stupid, while another woman is very clever.’ (Nikolaeva 2003: 134)

The parallels in contrast can be both marked as in (113), (115), and (116), but sometimes only one entity of the parallel is marked, as in (114).

Third, possessive suffixes can have an “evaluative” use that determines how a referent is psychologically related to the interlocutors on the basis of affectedness (discussed in Section 4.5.1). It points to the fact that this function can be expressed only in the dimension of the speaker and addressee, that is, by only the 1st or 2nd person possessive suffixes (see Kel’makov 1996), illustrated in (117) to (120).

Komi (Permic)

- (117) *oj*, *ylyn* *əd* *siję* *Kijev-yd*, *musa* *nylej*.
 oh far be.3SG it Kiev-Px.2SG dear girl
 ‘Oh, Kiev is far, dear girl.’ (Serebrennikov 1963: 135)

Udmurt (Permic)

- (118) *ton* *čeber* *apaj-e!*
 2SG beautiful sister-Px.1SG
 ‘You, my beautiful sister!’ (Nikolaeva 2003: 135)

Khanty

- (119) *wanta* *tām* *mašinaj-en* *jowra* *mān-əs*.
 see.IMPR.2SG this car-Px.2SG awry go- PST.3SG
 ‘Look, that car went awry.’ (Nikolaeva 2003: 136)

Selkup (Samoyedic)

- (120) *tü-lj:* *čə:tyŋylj:*!
 fire-P_x.2DU make.IMPR.2DU
 ‘Make (you two) a fire!’ (Kuznecova et al. 1980: 188)

In (117), (119), and (120), 2nd person suffixes express that the referents are important for the addressee, while 1st person suffix in (118) shows the speaker’s affectedness towards his ‘sister’ (see also Halm 2018 for similar examples in Hungarian).

Note that the possibility of host attachment of the possessive suffixes as markers of topic, contrast, and evaluation in Mari and Permic languages is not restricted to nouns, and they can occur with any part of speech. For instance, possessive suffixes are observed after pronouns in (121), adjectives in (122) and (125), finite verbs in (123), and non-finite verbs and adverbs in (124).

Mari

- (121) *tyj-že* *kuze* *ila-š* *tüŋal-at?*
 2SG-P_x.3SG how live-INF begin-2SG
 ‘And how are you going to live?’ (Tužarov 1987: 67)
- (122) *šöŋy-št-vlak* *pört* *wokten* *šinča-t,* *rwezy-št-vlak* *mody-t.*
 elderly-P_x.3PL-PL house near sit-3PL young-P_x.3PL-PL play-3PL
 ‘The elderly people sit by the house, (whilst) the young people play’ (Nikolaeva 2003: 141)
- (123) *a* *tyj* *kuze* *jörat-et-še?*
 and 2SG how love-2SG-P_x.3SG
 ‘And how do you love?’ (Bereczki 1990: 43)

Komi (Permic)

- (124) *me* *öd* *og* *kuž* *börd-ny-sö.*
 1SG you know NEG.1SG can.CNG cry-INF-P_x.3SG
 ‘I, you know, cannot cry.’
- da* *kydzi* *ber-se* *menym* *tat-eš* *mun-ne?*
 and how back-P_x.3SG 1SG.DAT here-ELAT go-INF
 ‘And how should I go back from here?’ (Leinonen 1998: 86)

Udmurt (Permic)

- (125) *buskel-jos-len* *badžym-ez* *pi-zy* *armi-yš* *bert-yz* *ini.*
 neighbour-PL-GEN elder-P_x.3SG son-3PL army-ELAT return-PST.3SG already
 ‘The elder son of the neighbours has already returned from army service.’
 (Kel’makov 2001: 179)

This functional characteristic implies that the Mari and Permic possessive suffixes have developed from nominal suffixes to clitics which can also attach to other parts of speech than noun. This makes a good parallel to the use of postposed demonstratives in eastern Finnic and North Russian (discussed in Sections 5.1.1 and 5.2.2).

Regarding the function of emphasis and contrast, Nikolaeva (2003) speculates about the possibility of contact-induced change, as this function is not common in other Uralic branches. It is common in neighbouring Turkic languages in the Volga basin, however, namely Tatar and Chuvash (see also Serebrennikov 1963: 133). However, I will also add that both referential as well as non-referential uses of possessive suffixes are not limited only to the Volga basin, but are also observed in other Turkic, Mongolic, and Tungusic languages across Northern Eurasia (see also Grønbech [1936/1979]1997: 92; Johanson 1998: 51; Fraurud 2001; Nikolaeva & Tolskaya 2001; Leinonen 2002). Ultimately, a contact-induced change shared with Uralic, especially with Samoyedic, is definitely possible (Stachowski 1998; Pakendorf 2007a, 2007b: 242–270).

Particularly, these languages of the Altaic type exhibit the use of possessive suffixes in the following contexts: (i) topicalisation or marking of unique referents such as ‘the snowdrops’ in (126), (ii) nominalisation of adjectives such as ‘the youngest (brother)’ in (127), and (iii) expressing the interlocutor(s)’s close psychological relation to referents such as ‘your Sergei (who is possibly the addressee’s relative or close person)’ in (128).

Yakut (Lena Turkic, Turkic)

- (126) *ńurguhun-a ol da buol-lar erde taxs-ar*
 snow- Px.3SG that also become-COND early go_out-PTCP.PRS
 ‘The snowdrops, nevertheless, come out early.’ (Pakendorf 2007b: 242)

Khalkha (Mongolic)

- (127) *tom ah žolooč, dund ah edijn zasagč, бага nj barilgačny.*
 big brother drive middle brother economist small Px.3SG labourer
 ‘the eldest brother is driver, the second brother is economist, **the youngest (one)** is labourer.’
 (course material, Pratique orale du mongol 1, INALCO, autumn semester 2015)

Udege (Orochic, Tungusic)

- (128) *si Sergej-ni: ə-s’ə mamasa-la?*
 2SG Sergei-Px.ALN.1/2SG NEG-PRF wife-VBLZ
 ‘Has not your Sergei got married?’ (Nikolaeva & Tolskaya 2001: 139)

This similarity observed in Turkic, Mongolic, and Tungusic languages shows that non-possessive uses of possessive suffixes in Uralic languages are not unique, from an areal-typological perspective.

5.4 CONCLUSIONS OF THE LITERATURE REVIEW

As illustrated in this chapter, both Finnic and North Russian postposed demonstratives as well as central and eastern Uralic possessive suffixes can similarly serve not only referential but also in non-referential uses. Based on observations in previous studies discussed above, both grammatical elements share the usage as anaphoric and topic markers. Given that the use of possessive suffixes has become remarkably less productive or even obsolete in most spoken Finnic varieties, it is not surprising that another grammatical element, a demonstrative in this case, would functionally take over the slot where the possessive suffix had erstwhile occupied.

Figure 22 illustrates the distribution of these two grammatical elements used for similar pragmatic purposes among Uralic languages in Russia. A contact-based explanation suggests that the contact with Russian plays an important role in the development of postposed demonstratives, particularly for the eastern Finnic languages.

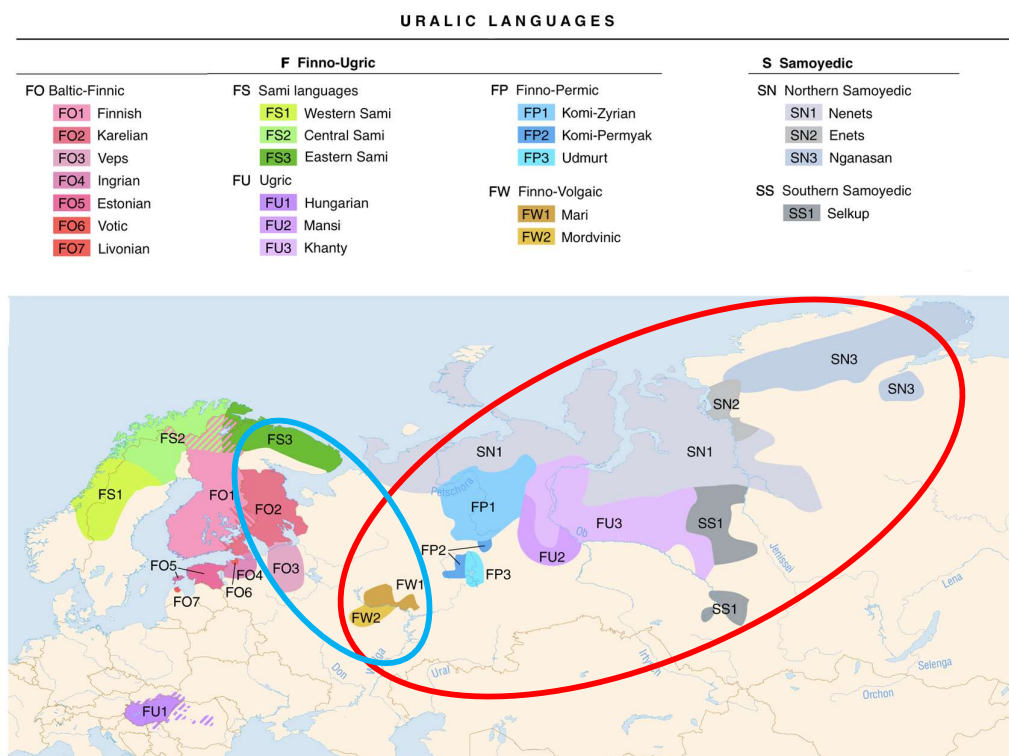


Figure 22 Marking of definiteness and information structure in the Uralic languages (a modification on Figure 3)
blue circle = postposed demonstrative, red circle = possessive suffix

Preliminary observation of this functional parallel confirms and suggests that the investigation of Finnic and North Russian postposed demonstratives and their uses should apply similar criteria as that of possessive suffixes found in other Uralic languages (in the same spirit as Leinonen 1998, 2006). As this chapter has shown,

postposed demonstratives have not been studied as extensively as possessive suffixes in non-possessive functions. Therefore, this approach can help defining both referential and non-referential uses of postposed demonstratives with more fine-grained categories. The data analysis in Chapter 6 will add further concrete evidence to the discussion.

6 MICRO-AREAL INVESTIGATION OF FINNIC AND NORTH RUSSIAN

This chapter investigates the primary data of the present study, giving an overview of the functions of postposed demonstratives in Finnic and North Russian from different grammatical perspectives. Sources of the data are text collections and online corpora, as described in Table 1. As stated in Section 1.2, the current study only uses data of spoken language registers, which assumedly exhibit pragmatic and discourse uses of postposed demonstratives better than that of the written language registers. Additionally, the use of postposed demonstratives as adverbs, as in Estonian *see maja seal* ‘the house there’ and Finnish *tämä talo tässä* ‘the house here’ (discussed in Pajusalu 1997a; Etelämäki 2006; Reile et al. 2019), is not included in the current study, which only pays attention to postposed demonstrative pronouns and proadjectives. The aim is to identify in a quantitative manner whether there is an areal signal that accounts for a particular phenomenon in this micro-area.

Due to the fact that dialects spoken in the surroundings of historical urban centres of Novgorod and Pskov today as well as those in the Saint Petersburg area do not belong to the North but rather Central Russian dialect continuum (see the classification of Russian dialects in Kasatkin 1999: 96, shown in Figure 12), the diffusion of modern Finnic and North Russian dialects are more likely to have taken place after Moscow’s invasions of Novgorod during the 15th–16th centuries. Therefore, this more recent contact, which likely affected the demonstrative system, must have been concentrated in Vologda Oblast rather than Novgorod (see earlier discussion in Section 2.3.2).

As discussed in van Gijn and Wahlström (forthcoming), various measures have been employed in the studies of dialectal diversity and language contact, such as “elevation” involving mountain ranges (Nichols 1992, 1997; Bickel & Nichols 2006; van Gijn 2014; van Gijn & Muysken, in press), “bodies of water” such as river networks, particularly influential in Amazonia (Hornborg 2005; Eriksen 2011), “traveling time”, which looks at the types and speeds of transportation within specific speaking areas (Gooskens 2005), and “ecological circumstance” such as the degree of risk to have natural disasters (Nettle 1999). The current study conservatively employs a basic measurement of a straight-line “geographical proximity” between the varieties under investigation. The reason for this choice is that cartographic data is easily accessible and able to be organised well for a quantitative analysis of this particular contact zone. Given the dialectological information above, Vologda Russian as the westernmost North Russian dialect today is selected as the nexus, from which the distance of each variety as a single data point is calculated as shown in Table 13. Distances are based on average of minimal and maximal extend of speaking areas by using coordinators (x = longitude, y = latitude) acquired from Google Maps tools. The distance values are calculated in the Pythagorean theorem $x_{average}^2 + y_{average}^2 = distance^2$. The results of the centralised speaking locations of each Finnic and North Russian variety in Table 13 is transferred to Figure 23.

Language	Variety	West x_{min}	East x_{max}	$x_{average}$	South y_{min}	North y_{max}	$y_{average}$	Distance from Vologda
Livonian		22.10736	22.58818	22.34777	57.62027	57.74244	57.681355	960 km
South Estonian		26.20057	27.77222	26.986395	57.59418	58.13252	57.86335	698 km
North Estonian		24.38439	25.8513	25.117845	59.1507	59.40401	59.277355	754 km
Votic	Western	28.34904	28.75136	28.5502	59.582	59.65441	59.618205	556 km
	Eastern	28.88292	29.01633	28.949625	59.67448	59.72888	59.70168	532 km
Ingrian	Soikkola	28.48579	28.48579	28.48579	59.78919	59.78919	59.78919	547 km
	Heva	29.18358	29.36949	29.276535	59.86463	59.89912	59.881875	512 km
Karelian	Olonets	32.23022	33.48588	32.85805	60.7794	62.17141	61.475405	349 km
Lude	Northern	33.61694	34.23417	33.925555	61.55941	62.56158	62.060495	339 km
	Southern	33.73074	33.80237	33.766555	61.00199	61.05568	61.028835	284 km
Veps	Northern	35.04835	35.51798	35.283165	61.28755	61.45086	61.369205	235 km
	Central	34.23242	36.49177	35.362095	60.24753	60.71418	60.480855	181 km
	Southern	34.86347	35.13555	34.99951	59.78206	59.99781	59.889935	193 km
North Russian	Vologda	36.00001	40.92166	38.460835	59.03583	60.86759	59.95171	0 km
	Arkhangelsk	38.44537	43.72107	41.08322	60.86665	64.54725	62.70695	338 km

Table 13. *Distance from the areal nexus, Vologda (in kilometre)*

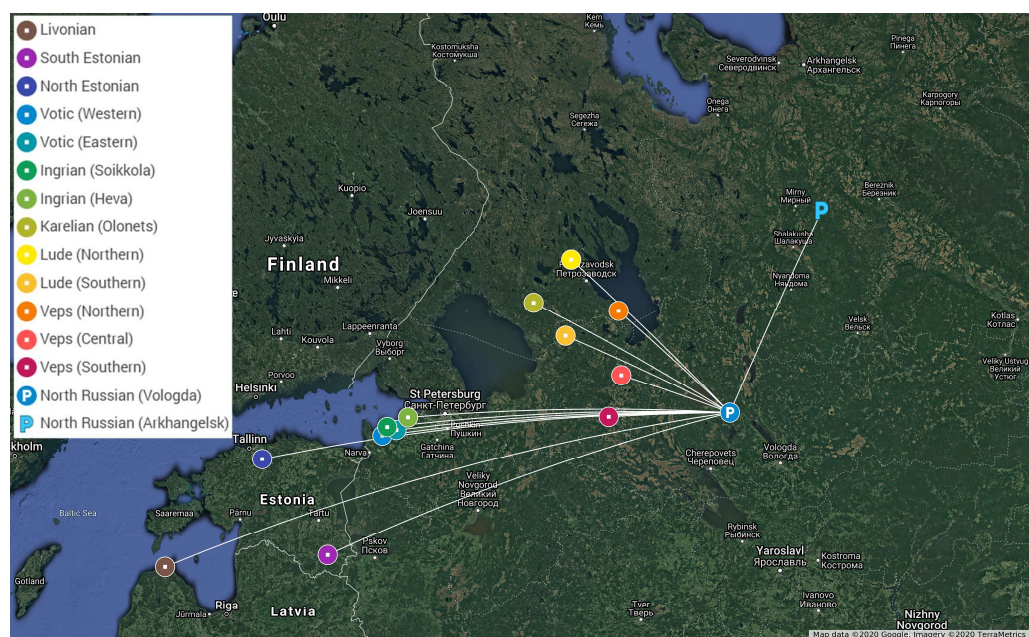


Figure 23 Centralised speaking areas of the Finnic and North Russian varieties under investigation

As mentioned in Section 1.2, Finnish and Karelian proper are not taken into account in the present study, due to their speaking areas which lie outside the Finnic-Russian contact zone under investigation.

The data analysis applies a functional-typological approach using various categories from syntax, morphology, semantics, and pragmatics, motivated by previous studies of postposed demonstratives and possessive suffixes in Uralic languages and North Russian dialects (e.g., Kettunen 1943; Trubinskij 1970; Leinonen 1998, 2006; Nikolaeva 2003, among others). Table 14 illustrates the criteria and their categories used for analysing the uses of postposed demonstratives in Finnic and North Russian varieties.

Criteria	Categories
Constituent order	Preposed / Postposed / Independent
Constituent position	Clause-Initial / Clause-medial / Clause-final
Parts of speech	Nominal: Noun / Pronoun / Name / Adjective / Numeral Others: Verb / Adverb / Adposition
Agreement with head word	Case / Number / Gender / Phonological harmonisation
Syntactic functions	Subject / Direct object / Indirect object / Verb / Adverbial / Determiner / Predicative
Referentiality	Deictic / Recognitional / Absolutely unique / Anaphoric / Bridging / Situationally unique / Establishing
Information structure	Continuing topic / Switch-topic / Beginning topic / Highlighting parallels / Afterthought
Stancetaking and evaluation	Subjective evaluations: Affection / Surprise & admiration Intersubjective evaluations: Shared knowledge / Counter-expectation / Hesitation / Warning & reminder

Table 14. *Typological criteria used for the data analysis*

Profiling the uses of demonstratives in each perspective (in Seržant’s terms (2015)) can bring better understanding of the functional dimensions and development of the demonstratives in question, as previously discussed from a theoretical perspective in Chapter 4.

“Constituent order” and “position” give a picture of how non-canonical the use of postposed demonstrative can be in certain languages in terms of word order typology, and how freely postposed demonstratives can move around the clause (Sections 6.1). “Parts of speech” illustrate how demonstratives can extend the possible range of what type of host they attach to, ranging from a prototypical use attached to a nominal to other parts of speech such as verb or adposition (Section 6.2). “Agreement with head word” sheds light on how demonstratives can develop into a bound morpheme of a lexical unit by such process like cliticisation and suffixation (Section 6.2). “Syntactic functions” help understanding the tendency of various verb arguments that are either more or less likely to co-occur with demonstrative such as subjects, direct and indirect objects, and adverbials (Section 6.3).

In terms of pragmatics, three criteria are used for data analysis. First, “referentiality” shows how a demonstrative can extend its prototypical deictic function towards other uses of identifiability typical for definite articles, including anaphoric reference, bridging reference, situationally or universally unique reference, recognitional reference, and establishing reference (Section 6.4.1). Second, “information-structural uses” provide another perspective beyond definiteness marking, by illustrating the information flow on whether the host word followed by demonstrative keeps track of the continuation of a previously active topic, switches to another topic, introduces a primary topic, creates a contrast between several entities, or expresses the speaker’s afterthought (Section 6.4.2). Lastly, “evaluative uses” explore other functions beyond conventional uses of definite articles, topic markers, and focus markers, such as different ways of stancetaking, and intersubjectively evaluating the state of knowledge between interlocutors (Section 6.4.3). Each of these categories will be discussed further, with concrete examples from Finnic and North Russian varieties under discussion.

Apart from the evaluative uses, which can be open to alternative interpretations, other morphosyntactic and pragmatic properties discussed in this chapter will also be quantified and integrated with the geographical data presented above (Table 13 and Figure 23). Where possible, the Fisher’s exact test (the χ^2 algorithm from Fisher 1922) will also be performed to check statistical significance and dependence of certain properties of demonstratives.

A diachronic comparison is also made with two demonstrative series observed in the Novgorod birch bark documents (Zaliznjak 2004): 1) *sb* and 2) *tb* as well as the clausal particle *ti* that potentially derived from a demonstrative (Section 6.5). Despite a Finnic text fragment also being attested in the Novgorod birch bark corpus (shown in Section 2.3.1), there is not a single occurrence of a demonstrative, so our current knowledge of how demonstratives behave in that particular Finnic dialect is still lacking for the time being. The analysis throughout this chapter will provide general patterns among the languages examined as well as some exploration on what kind of syntactic and pragmatic contexts or conditions the use of postposed demonstratives would be expected (Section 6.6).

6.1 CONSTITUENT ORDER AND POSITION

Among the types of demonstratives discussed in Section 4.1, the current study primarily focuses on dependent demonstratives that can be seen as accompanying another host word. Therefore, independent demonstratives without a possible head word before or after are excluded from the analysis of grammatical (Sections 6.2 and 6.3) and pragmatic uses (Sections 6.4). As some constructions, particularly left dislocation (discussed in Amon 2015 for Estonian; in Vilkuna 1989: 145–147; Priiki 2015, 2017 for Finnish), can often put the interpretation of morphosyntactic dependence on the borderline, I also treat such demonstratives that follow the clause-initial element as Wackernagel’s clitics (such as in (71), (72), (73), (75b) and (81)) as dependent demonstratives. This follows from the working hypothesis in the present

study that this type of syntactic context is the preceding stage prior to the extension to other syntactic positions, as will be illustrated in Section 6.1.6, as well as various pragmatic uses discussed in Section 6.4 (see also a development path proposed in Section 7.2.1).

As for the dependent demonstratives, classification of constituent order is based on the position of a demonstrative in relation to its host word. In a head-preceding construction, the demonstrative is labelled as “preposed” in (129a), (130a), (131a), and (132a). Meanwhile, a demonstrative in a head-following position is labelled as “postposed” in (129b), (130b), (131b), and (132b) (see also Dryer 2013a, illustrated in Figure 17). At least one of the two possible orders in the following examples is observed in the actual data, while the other may be constructed examples to illustrate the two possibilities.

South Estonian

- (129) a. *n'eoq mõtsa-q*
 b. *mõtsa-q n'eoq*
 DEM.PROX.PL forest-PL DEM.PROX.PL
 ‘these/the forests’

Heva Ingrian

- (130) a. *ne karhu-t*
 b. *karhu-t ne*
 DEM.MED.PL bear-PL DEM.MED.PL
 ‘those/the bears’

Northern Lude

- (131) a. *se dragun*
 b. *dragun se*
 DEM.MED dragon DEM.MED
 ‘that/the dragon’

North Russian (Vologda)

- (132) a. *to seno*
 b. *seno-to*
 DEM.DIST.NEU hay-DEM.DIST.NEU
 ‘that/the hay’

The constituent order of demonstratives in a phrase has a varying distribution across the Finnic varieties. Moreover, each set of demonstratives also shows different tendencies of constituent order. Among the Finnic demonstratives discussed in Section 2.4.1 (Table 3), the most divergent series is the Proto-Finnic medial **se/ne(t)*, while the other Proto-Finnic series (proximal **tämä/nämä(t)*, distal **too/noo(t)*, and compound demonstratives) follow a more unified pattern across all Finnic varieties.

Chart 1 shows the proportion of all demonstrative series observed in the data. The fourth series **taa/naa(t)* is exclusively observed in South Estonian and it has become synchronically confused with the short forms of the 3rd person pronouns *ta/nad* (a diachronically proximal series) borrowed from North Estonian (see Larjavaara 1986: 310–311; Pajusalu 2015: 188, discussed in Section 2.4.1). For this reason, the South Estonian series *taa/naaq* is not taken into the comparison, as it is not directly relevant for the Finnic-Russian contact.

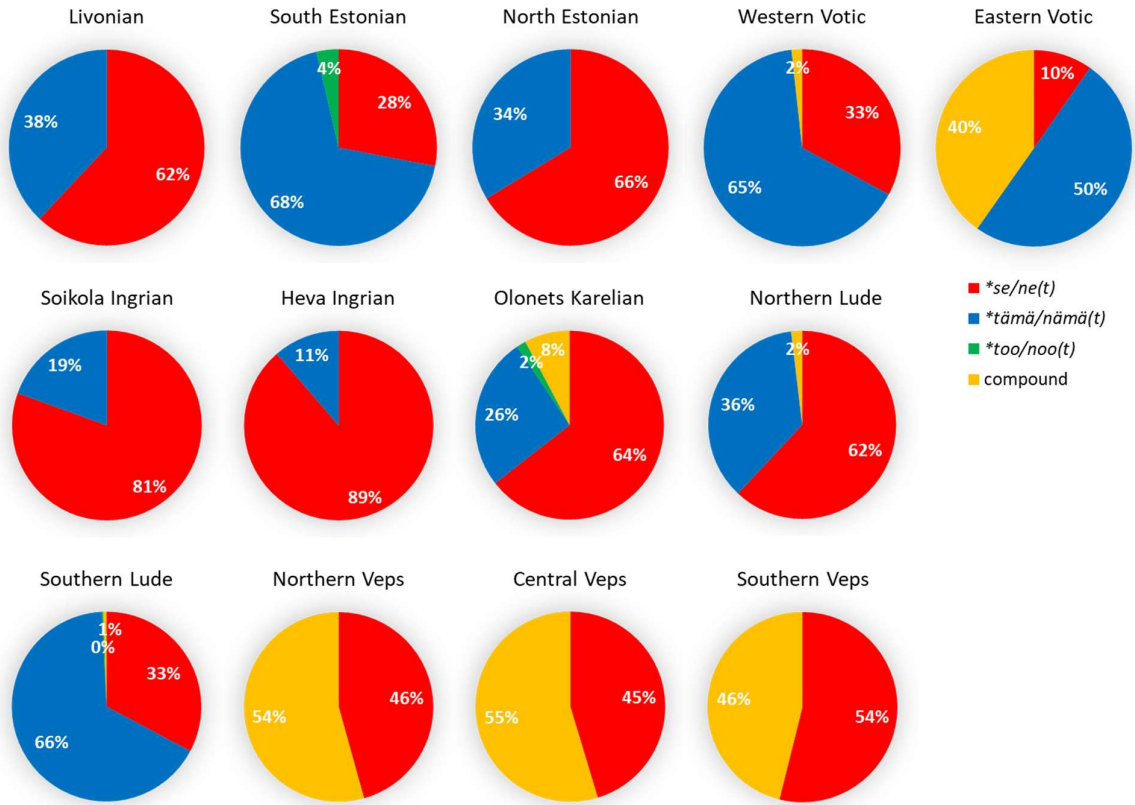


Chart 1 Proportion of the different series of demonstratives observed in the Finnic data

At this point, we can observe obvious micro-areal patterns across the Finnic varieties. For instance, Livonian resembles North Estonian, possibly due to recent intense contact between fishermen towards the end of the 19th century (see Ariste 1981: 78). As for Lude, the northern variety clearly behaves similarly to Olonets Karelian, while the southern variety behaves similarly to Veps (see also Pahomov 2011: Ch. 3, discussed in Section 2.2.1), given that the use of *tämä/nämä* in Southern Lude is comparable to the compound demonstratives *nece/ñene* in Veps. Parallels between the different varieties will be discussed further from a historical-comparative perspective in Section 7.2.2.

Beginning from less complex cases, the data analysis begins by initially showing the occurrences of the Proto-Finnic series **tämä/nämä*, **too/noo(t)*, and compound demonstratives, before discussing a more complex issues with the series **se/ne(t)*, and a comparison with the North Russian demonstratives *ètot* and *tot*. To avoid the confusion of proximity contrast, which can vary between the proto-language and modern varieties, such as the case of **se/ne(t)* in Finnic, the demonstrative series will be henceforth labelled only according to their stems as in Chart 1.

6.1.1 FINNIC **tāmā/nāmā(t)*

As discussed in Section 2.4.1, the Finnic series **tāmā/nāmā* has been lost in Veps, so the geographical distribution of this series ends in the Lude-speaking area, shown in Table 15 for the singular and Table 16 for the plural forms.

Demonstrative	Language	Variety	Preposed		Postposed		Independent		Total
<i>*tāmā</i>	Livonian		12	1.68%	1	0.14%	703	98.18%	716
	South Estonian		13	0.70%	-		1837	99.30%	1850
	North Estonian		2	0.38%	-		527	99.62%	529
	Votic	Western	2	0.64%	-		312	99.36%	314
		Eastern	1	0.38%	-		263	99.62%	264
	Ingrian	Soikkola	21	48.84%	-		22	51.16%	43
		Heva	12	37.50%	-		20	62.50%	32
	Karelian	Olonets	34	29.06%	9	7.69%	74	63.25%	117
	Lude	Northern	253	73.76%	10	2.92%	80	23.32%	343
		Southern	291	49.91%	32	5.49%	260	44.60%	583
	Veps		-		-		-		-

Table 15. *Frequencies of the Finnic demonstrative *tāmā*

Demonstrative	Language	Variety	Preposed		Postposed		Independent		Total
<i>*nāmā(t)</i>	Livonian		21	23.08%	-		70	76.92%	91
	South Estonian		2	0.36%	-		560	99.64%	562
	North Estonian		1	0.54%	-		185	99.46%	186
	Votic	Western	2	2.00%	-		98	98.00%	100
		Eastern	3	4.55%	-		63	95.45%	66
	Ingrian	Soikkola	6	60.00%	-		4	40.00%	10
		Heva	-		-		1	100%	1
	Karelian	Olonets	6	42.86%	2	14.29%	6	42.86%	14
	Lude	Northern	9	56.25%	-		7	43.75%	16
		Southern	169	66.27%	18	7.06%	68	26.67%	255
	Veps		-		-		-		-

Table 16. *Frequencies of the Finnic demonstrative *nāmā(t)*

The only occurrence of the Finnic demonstrative **tāmā* after a noun in Livonian is the case of a synchronic 3rd person pronoun *ta* ‘he, she,’ functioning as a resumptive pronoun that marks a left dislocation of the topic ‘grandmother’ in (133).

Livonian

- (133) *izā jemā ta vól dūoņđnikā rištīng.*
 father.GEN mother 3SG be.PST.3SG Dundaga.GEN people
 ‘Grandmother was a person from Dundaga.’ (Mägiste 2006: 2)

However, this example is not an attributive use of a demonstrative to modify a head noun, but rather an information-structural use to mark a clause topic (see further discussion in Section 6.4.2). As for Finnic varieties in the east, the occurrences of postposed **tämä/nämä(t)* are largely attributive, in which demonstratives agree with head nouns in number and case, e.g., Olonets Karelian *kolhōzas täs* [kolkhoz.INES DEM.PROX.INES] ‘in this kolkhoz’ (Makarov & Rjagoev 1969: 194), and Southern Lude *ruaduud nämiiid* [work.PL.PTV DEM.PROX.PL.PTV] ‘these works’ (Pahomov 2011: 82).

On a dialect continuum scale, Chart 2 shows a clear difference in distribution across the Finnic varieties.

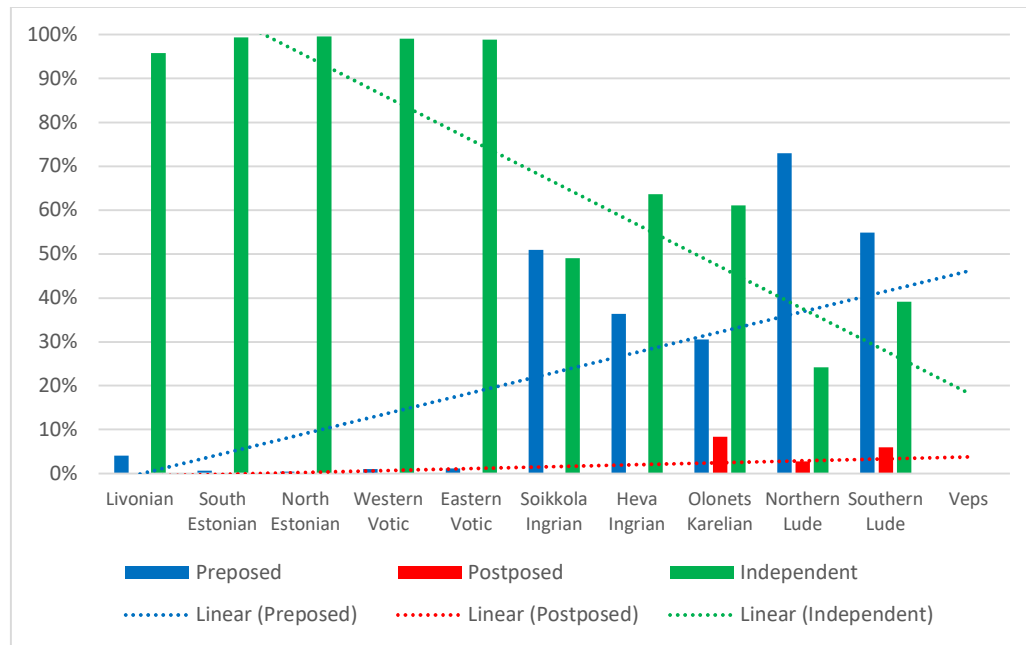


Chart 2 Constituent orders of the Finnic demonstratives **tämä/nämä(t)* across modern Finnic varieties

Chart 2 illustrates that preposed attributive use of demonstratives **tämä/nämä* is primarily observed between the speaking areas of Ingrian and Lude. The southwestern Finnic languages (Livonian, South Estonian, North Estonian and Votic), on the other hand, mostly use demonstratives **tämä/nämä(t)* as independent pronouns. Functionally, they have replaced the Proto-Finnic 3rd person pronouns **hän/he* ‘he, she/they,’ shown in (134).

North Estonian

- (134) *ta metsa-s oli.*
 3SG forest-INES be.PST.3SG
 ‘She was in the forest.’ (EDC)

This semantic change within the Finnic demonstrative systems results in a contrast between the predominant attributive use in the east and a pronominal use in the west, as the trendlines in Chart 2 (Linear) indicate.

From a typological viewpoint, this micro-areal feature is, however, a relatively rare case of change, as 3rd person pronouns more commonly derive from medial or distal demonstratives, as in the case of the Proto-Slavic distal **onъ/ono/ona* in most modern Slavic languages (cf. Greenberg 1978: 61; Givón 2001: 226; Diessel 1999a: 161; Bhat 2013). Nevertheless, the Bulgarian 3rd person pronouns *toj/to/ta* (and Macedonian *toj/toa/taa*) synchronically share the same roots as proximal demonstratives *tozi/tova/tazi*, or *toja/tuj/taja* in short forms, despite their etymologies ultimately going back to the Proto-Slavic medial series **tъ/to/ta*, as discussed in Section 2.4.2.

Despite this semantic shift from the Proto-Finnic proximal demonstrative to 3rd person pronouns in Finnic varieties in the west, there are still several cases of attributive use observed in the data. First, the occurrences of preposed demonstratives in South Estonian are possibly due to the confusion of the short forms of the Finnic demonstratives **tämä/nämä(t)* and South Estonian *taa/naaq* as *ta/na* (discussed in Section 2.4.1, and earlier in this chapter), e.g., *tast Tal'linast* [DEM.PROX/MED.ELAT Tallinn.ELAT] ‘from this Tallinn’ (EDC). A few cases of pronominal use in North Estonian are sporadic occurrences however, e.g., North Estonian *talle karjatselle* [DEM.PROX.ALL cattle.ALL] ‘to these cattle’ (EDC). In addition to Estonian, Livonian and Votic speakers still used the Finnic demonstratives **tämä/nämä* in the non-personal function in several time and manner adverbial expressions, e.g., Livonian *täm õödõg* [DEM.PROX.GEN evening] ‘this evening’ (Mägiste 2006: 160), and (both western and eastern) Votic *tü-nä vō-na* [DEM.PROX-ESS year-ESS] ‘this year’ (Kettunen & Posti 1932: 2, 130).

6.1.2 FINNIC **too/noo(t)*

The data of the present study do not contain many occurrences of the Finnic demonstratives **too/noo(t)*, which occur only in South Estonian (129 cases), Olonets Karelian (8 cases), Northern Lude (1 case), and Southern Lude (3 cases). Furthermore, the plural forms **noo(t)* are only observed in South Estonian. This possibly relates to the disappearance and transformation of the Proto-Finnic distal demonstratives into a marker denoting human referent (see Pajusalu 2006). On the other hand, the low frequency of **too/noo(t)* might simply be due to genres of text in the data, which are more monologue than dialogue.

Due to the lack of empirical evidence, it is difficult to give a firm generalisation. Although the South Estonian data contains over a hundred of occurrences of demonstratives *tuu/nuu(q)* ‘that/those’, none of them occur after a head noun, and likewise with the Olonets Karelian and Lude data.

6.1.3 FINNIC COMPOUND DEMONSTRATIVES

Compound demonstratives in Finnic varieties cannot be compared as such, because they derived from different combinations of source elements. In any case, the data still shows interesting insights about tendencies with word order, shown in Table 17 for singular forms, and Table 18 for plural forms.

Demonstrative	Language	Variety	Preposed		Postposed		Independent		Total
-	Livonian		-		-		-		-
	South Estonian		-		-		-		-
	North Estonian		-		-		-		-
<i>kase</i>	Votic	Western	8	80.00%	-		2	20.00%	10
		Eastern	102	64.97%	3	1.91%	52	33.12%	157
-	Ingrian	Soikkola	-		-		-		-
		Heva	-		-		-		-
<i>netše</i>	Karelian	Olonets	12	41.38%	3	10.34%	14	48.28%	29
	Lude	Northern	6	50.00%	-		6	50.00%	12
		Southern	-		-		7	100%	7
<i>nece</i>	Veps	Northern	164	79.23%	15	7.25%	28	13.53%	207
		Central	365	74.34%	28	5.70%	98	19.96%	491
		Southern	78	80.41%	8	8.25%	11	11.34%	97

Table 17. *Frequencies of singular compound demonstratives*

Demonstrative	Language	Variety	Preposed		Postposed		Independent		Total
-	Livonian		-		-		-		-
	South Estonian		-		-		-		-
	North Estonian		-		-		-		-
<i>kane</i>	Votic	Western	1	100%	-		-		1
		Eastern	82	75.93%	2	1.85%	24	22.22%	108
-	Ingrian	Soikkola	-		-		-		-
		Heva	-		-		-		-
<i>ñene</i>	Karelian	Olonets	7	70.00%	1	10.00%	2	20.00%	10
	Lude	Northern	5	83.33%	-		1	16.67%	6
		Southern	-		-		-		-
	Veps	Northern	31	81.58%	1	2.63%	6	15.79%	38
		Central	107	82.31%	4	3.08%	19	14.62%	130
		Southern	10	66.67%	3	20.00%	2	13.33%	15

Table 18. *Frequencies of plural compound demonstratives*

Chart 3 shows the distribution of compound demonstratives across Finnic varieties.

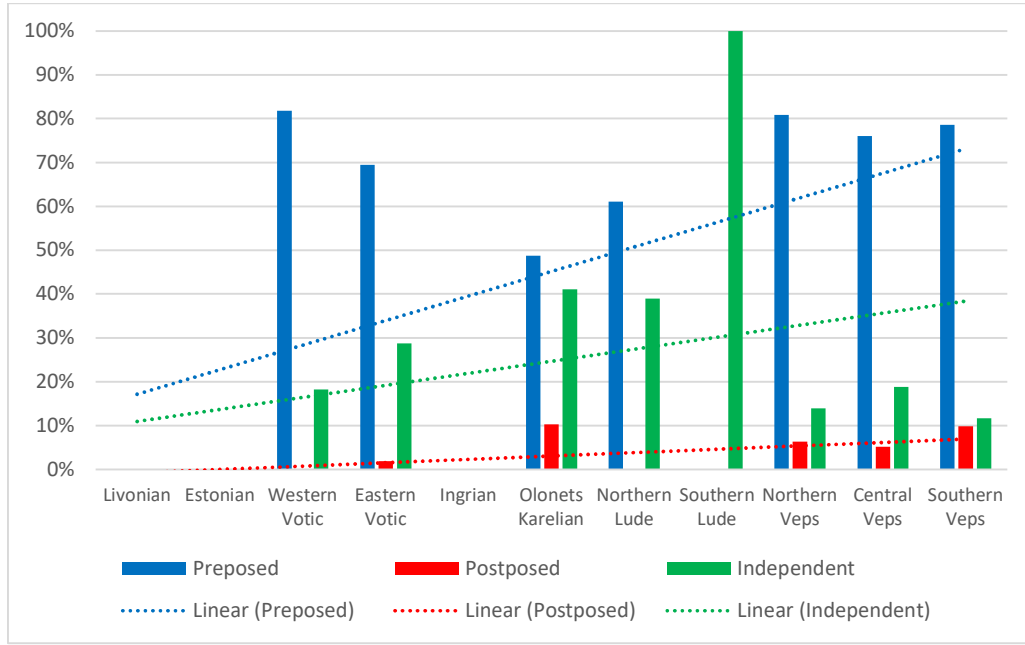


Chart 3 Constituent orders of compound demonstratives across modern Finnic varieties

The trendlines in Chart 3 show that postposed demonstratives are observed more frequently towards the eastern end of the Finnic dialect continuum, while Votic varieties always use compound demonstratives before a head noun, aside from a handful of postposed compound demonstratives, such as Eastern Votic *lehmä-d kannee* [cow-PL DEM.PL] ‘these cows.’ This areal trend becomes even clearer when we discuss the uses of the Finnic series **se/ne(t)*.

6.1.4 FINNIC **se/ne(t)*

Unlike the other demonstrative series, the Finnic demonstratives **se/ne(t)* are well preserved in all modern varieties. At the same time, however, they behave distinctly across the different varieties. As discussed earlier in Section 2.4.1, **se/ne(t)* have become proximal demonstratives in South Estonian, neutral in Livonian and North Estonian, remained as medial demonstratives in varieties with a tripartite system in the east, and changed into distal in the renewed bipartite system of Votic and Veps (see Table 3). In terms of frequency, this is the most used demonstrative series in the majority of Finnic varieties (see Chart 1). Table 19 and Table 20 illustrate the frequencies of **se* and **ne(t)*, respectively, in different constituent orders.

Demonstrative	Language	Variety	Preposed		Postposed		Independent		Total
*se	Livonian		455	46.57%	10	1.02%	512	52.41%	977
	South Estonian		278	40.88%	6	0.88%	396	58.24%	680
	North Estonian		376	35.04%	13	1.21%	684	63.75%	1073
	Votic	Western	89	46.84%	13	6.84%	88	46.32%	190
		Eastern	14	29.79%	4	8.51%	29	61.70%	47
	Ingrian	Soikkola	74	45.68%	4	2.47%	84	51.85%	162
		Heva	71	33.02%	12	5.58%	132	61.40%	215
	Karelian	Olonets	63	23.86%	51	19.32%	150	56.82%	264
	Lude	Northern	217	42.88%	48	9.49%	241	47.63%	506
		Southern	68	18.94%	131	36.49%	160	44.57%	359
	Veps	Northern	14	7.11%	113	57.36%	70	35.53%	197
		Central	40	8.28%	317	65.63%	126	26.09%	483
		Southern	7	5.93%	84	71.19%	27	22.88%	118

Table 19. *Frequencies of the Finnic demonstrative *se*

Demonstrative	Language	Variety	Preposed		Postposed		Independent		Total
*ne(t)	Livonian		74	21.64%	6	1.75%	262	76.61%	342
	South Estonian		92	29.87%	6	1.95%	210	68.18%	308
	North Estonian		107	32.03%	5	1.50%	222	66.47%	334
	Votic	Western	10	55.56%	-	-	8	44.44%	18
		Eastern	6	35.29%	-	-	11	64.71%	17
	Ingrian	Soikkola	30	52.63%	1	1.75%	26	45.62%	57
		Heva	16	38.10%	4	9.52%	22	52.38%	42
	Karelian	Olonets	17	28.81%	8	13.56%	34	57.63%	59
	Lude	Northern	50	45.46%	9	8.18%	51	46.36%	110
		Southern	22	40.74%	6	11.11%	26	48.15%	54
	Veps	Northern	2	22.22%	6	66.67%	1	11.11%	9
		Central	1	3.13%	31	96.87%	-	-	32
		Southern	2	15.38%	11	84.62%	-	-	13

Table 20. *Frequencies of the Finnic demonstrative *ne(t)*

Transforming these data into a graphic illustration with the geographical distribution, we can notice from the trendline in Chart 4 that the uses of postposed demonstratives are observed increasingly more towards the east.

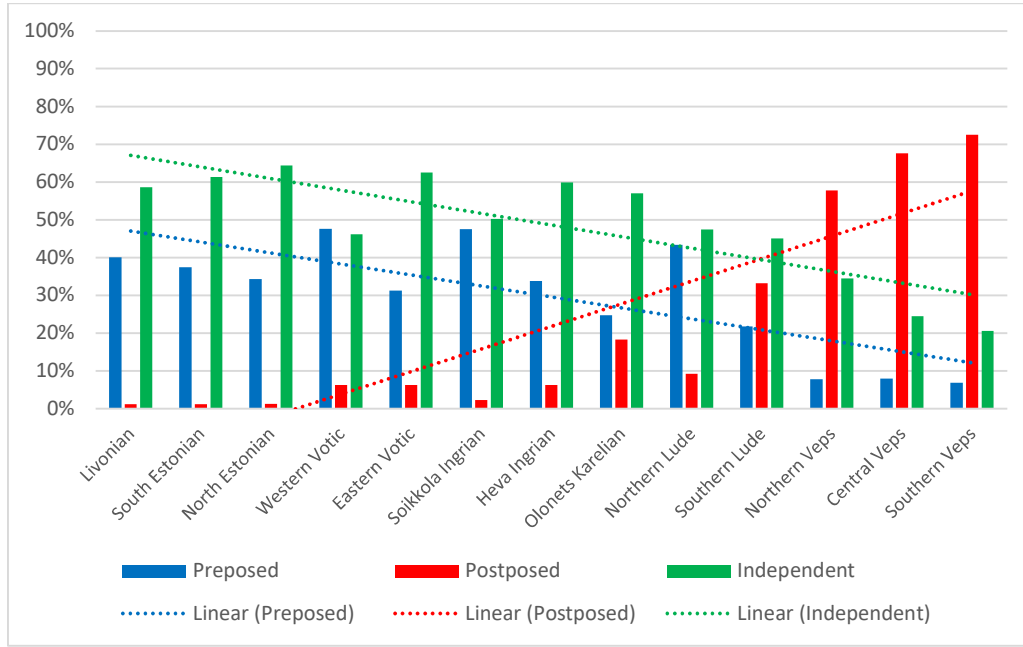


Chart 4 Constituent orders of the Finnic demonstratives **se/ne(t)* across modern Finnic varieties

There is a direct correlation with the use of preposed demonstratives decreasing as one moves towards the east. This tendency obviously entails that something is happening in the eastern end of the Finnic dialect continuum (discussed further in Section 6.1.6). From this point onwards, the investigation will focus on the use of the Finnic demonstratives **se/ne(t)*, the main object of the present research.

6.1.5 NORTH RUSSIAN *ètot* and *tot*

The Russian language has a bipartite demonstrative system, consisting of the proximal *ètot* and the distal *tot* (discussed in Section 2.4.2). Regarding constituent order, the data shows a distinct contrast between the two series of demonstratives. The *ètot* series is predominantly used preposed, seen in Table 21.

Demonstrative	Region	Preposed		Postposed		Independent		Total
<i>ètot</i>	Vologda	91	37.29%	8	3.28%	145	59.43%	244
	Arkhangelsk	306	47.59%	30	4.67%	307	47.74%	643
	Total	397	44.76%	38	4.28%	452	50.96%	887

Table 21. Frequencies of the North Russian demonstrative *ètot*

On the other hand, the *tot* series has a strong tendency for being postposed, seen in Table 22.

Demonstrative	Region	Preposed		Postposed		Independent		Total
<i>tot</i>	Vologda	17	5.31%	268	83.75%	35	10.94%	320
	Arkhangelsk	56	8.08%	526	75.90%	111	16.02%	693
	Total	73	7.21%	794	78.38%	146	14.41%	1013

Table 22. *Frequencies of the North Russian demonstrative tot*

In addition to these basic demonstratives, compound demonstratives mentioned in Section 2.4.2 are also observed in three cases in the Arkhangelsk variety: independent *nuto* ‘that’, as well as preposed *nuti nizki-ti* ‘those lower ones’ and *nuti varežki-ti* ‘those mittens’ (RNC). These forms are used alongside *tot* in a non-proximal sense with a lower frequency, and do not appear to create a separate additional series of demonstratives at this stage. This type of demonstrative, however, is not attested in the Vologda variety.

6.1.6 AREAL SIGNALS OF CONSTITUENT ORDERS AND POSITIONS

Previous sections already provide a hint of the areal effect that influences the distribution of constituent orders within certain series of demonstratives across the Finnic varieties. This section tests whether areality is a trigger, by also taking the North Russian *tot* into comparison with the Finnic **se/ne(t)*. Table 23 describes the proportion of preposed and postposed demonstratives occurring in the data.

Language	Variety	Preposed + Postposed demonstratives	Tendency of preposed demonstrative		Tendency of postposed demonstrative		Distance from Vologda
	Livonian	545	529	97.06%	16	2.94%	960 km
	South Estonian	382	370	96.86%	12	3.14%	698 km
	North Estonian	501	483	96.41%	18	3.59%	754 km
Votic	Western	112	99	88.39%	13	11.61%	556 km
	Eastern	24	20	83.33%	4	16.67%	532 km
Ingrian	Soikkola	109	104	95.41%	5	4.59%	547 km
	Heva	103	87	84.47%	16	15.53%	512 km
Karelian	Olonets	139	80	57.55%	59	42.45%	349 km
Lude	Northern	324	267	82.41%	57	17.59%	339 km
	Southern	227	90	39.65%	137	60.35%	284 km
Veps	Northern	135	16	11.85%	119	88.15%	235 km
	Central	389	41	10.54%	348	89.46%	181 km
	Southern	104	9	8.65%	95	91.35%	193 km
North Russian	Vologda	285	17	5.96%	268	94.04%	0 km
	Arkhangelsk	582	56	9.62%	526	90.38%	338 km

Table 23. *Tendencies of preposed vs. postposed demonstratives*

Chart 5 visualises the correlation between the tendency of postposed demonstratives and the distance from Vologda, the areal nexus.

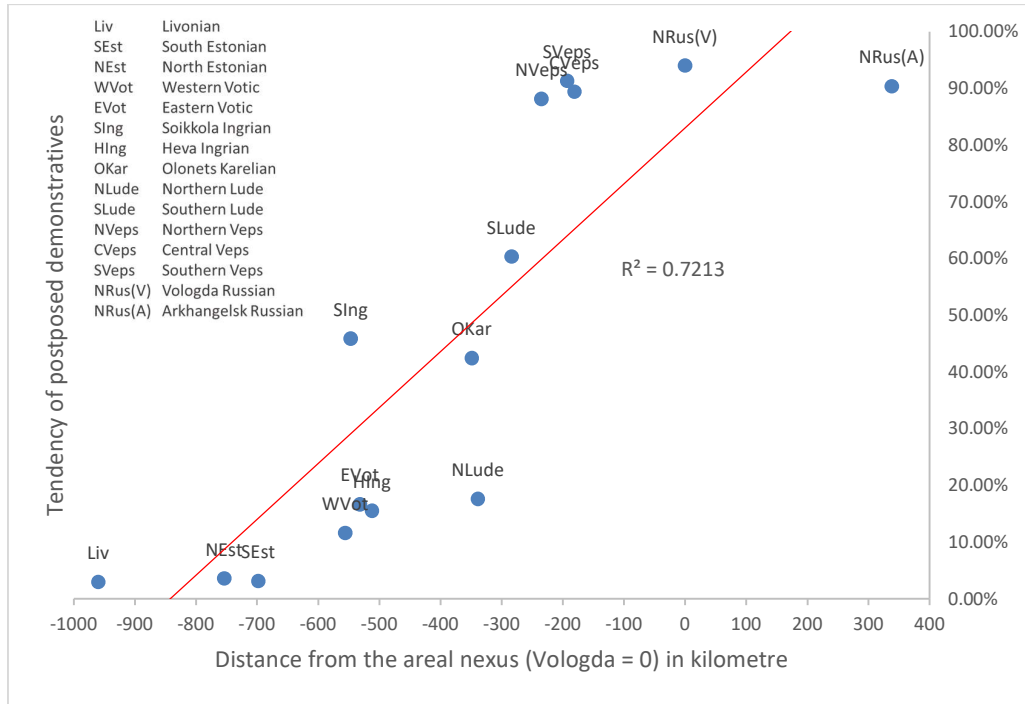


Chart 5 Tendency of postposed over preposed demonstratives

An exceptionally high value of $R^2 = 0.7213$ is seen in Chart 5. This means that the correlation covers approximately 72% of the data, confirming that the geographical location of each variety is a crucial factor in explaining the distribution of postposed demonstratives in the data.

Another factor to be investigated from an areal perspective is the position in the clause where the phrase accompanied by the postposed demonstrative can be found. The classification focuses on three contexts: clause-second position after the first stressed constituent of the clause in (135), clause-medial position in (136), and clause-final in (137).

Northern Veps

- (135) *no, Maša se kodi-š näg-i, kut baba kiita-b.*
 but **Maša** DEM home-INES see-PST.3SG how grandmother cook-3SG
 'Well, Maša saw at home how her grandmother was cooking.' (Onegina & Zajceva 1996: 31–34)

- (136) *nu ka hö miska-d ne söd'-he, söd'-he, söd'-he, ...*
 but so 3PL **dish-PL** DEM.PL eat.PST-3PL eat.PST-3PL eat.PST-3PL
 'Well, they kept eating dishes (of meat).' (Onegina & Zajceva 1996: 34–51)

- (137) *potom dö hõ läm-ha ma-ha lähto-ba, süguze-l se.*
 then already 3PL warm-ILL land-ILL leave-3PL **autumn-ADES** **DEM**
 ‘Then, they (the geese) will already be leaving for a warm place in the autumn.’
 (Onegina & Zajceva 1996: 34–51)

The frequencies of each position across the varieties are summarised in Table 24 and Chart 6.

Language	Variety	Postposed demonstratives	Tendency of clause-second position		Tendency of clause-medial positions		Tendency of clause-final positions		Distance from Vologda
Livonian		16	16	100%	0	0.00%	0	0.00%	960 km
South Estonian		12	10	83.33%	2	16.67%	0	0.00%	698 km
North Estonian		18	16	88.89%	1	5.56%	1	5.56%	754 km
Votic	Western	13	13	100%	0	0.00%	0	0.00%	556 km
	Eastern	4	2	50.00%	0	0.00%	2	50.00%	532 km
Ingrian	Soikkola	5	5	100%	0	0.00%	0	0.00%	547 km
	Heva	16	15	93.75%	0	0.00%	1	6.25%	512 km
Karelian		59	50	84.75%	6	10.17%	3	5.08%	349 km
Lude	Northern	57	36	63.16%	14	24.56%	7	12.28%	339 km
	Southern	137	50	36.50%	72	52.55%	15	10.95%	284 km
Veps	Northern	119	72	60.50%	31	26.05%	16	13.45%	235 km
	Central	348	190	54.60%	114	32.76%	44	12.64%	181 km
	Southern	95	52	54.74%	19	20.00%	24	25.26%	193 km
North Russian	Vologda	268	111	41.42%	89	33.21%	68	25.37%	0 km
	Arkhangelsk	526	213	40.49%	222	42.21%	91	17.30%	338 km

Table 24. *Frequencies of postposed demonstratives in different clausal positions*

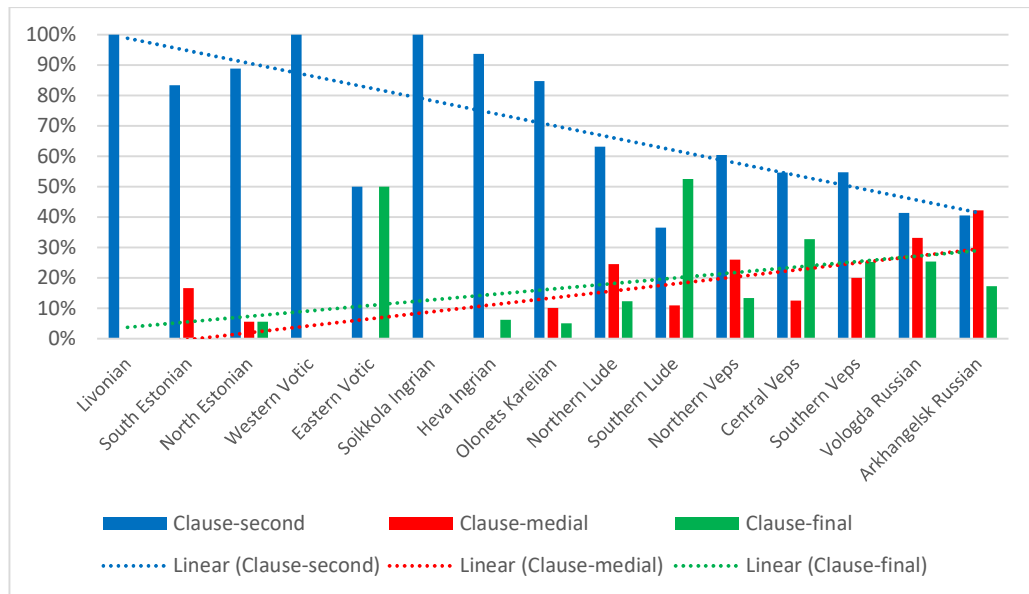


Chart 6 *Frequencies of postposed demonstratives in different clausal positions*

The trendline in Chart 6 shows that the frequency of using postposed demonstratives in other positions than clause-second gradually increases as one moves towards the east. Chart 7 visualises the correlation between the tendency of demonstratives in non-clause-second positions and the distance from Vologda, the areal nexus.

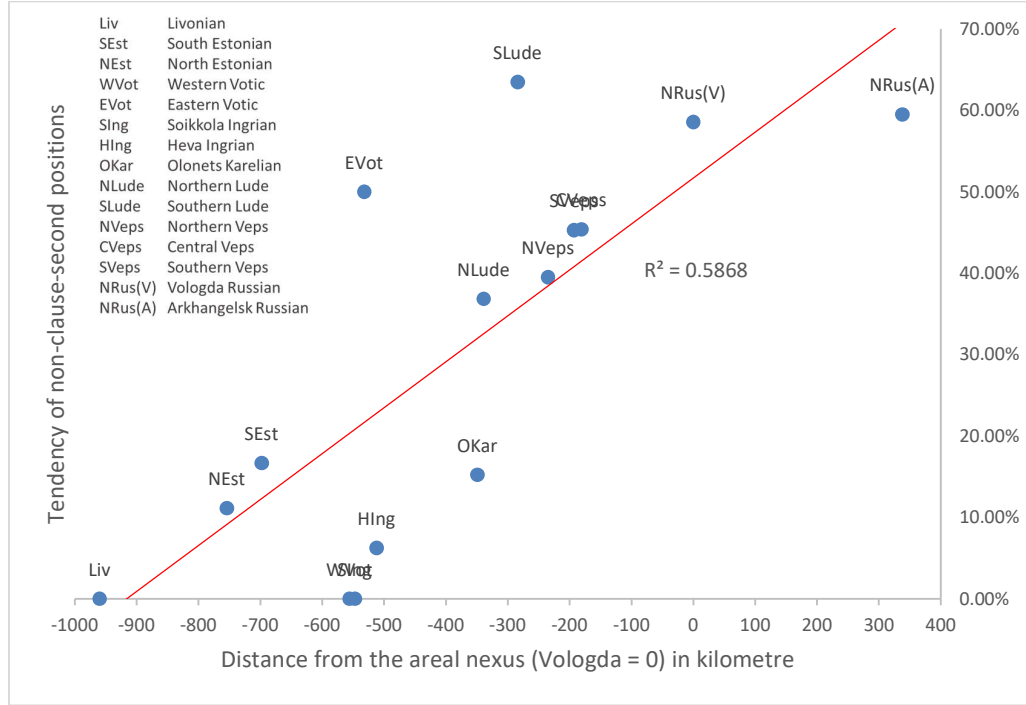


Chart 7 Tendency of non-clause-second over clause-second positions

A relatively high value of $R^2 = 0.5868$ is seen in Chart 7. This means that the correlation covers approximately 59% of the data, once again speaking in favour of the geographical location of each variety being an important factor.

As demonstrated in this section, the correlations based on the distance from Vologda clearly point towards language contact, and a stronger degree of diffusion with North Russian on the eastern end of the Finnic dialect continuum (discussed further in Section 7.3.2).

6.2 HOST ATTACHMENT AND PARTS OF SPEECH

Another important factor for the investigation of the uses of postposed demonstratives in Finnic and North Russian is the part of speech of the host word to which the demonstrative attaches. Prototypically, the demonstrative functions as determiner in a noun phrase, as seen in examples (129) to (132). In the current study, noun phrases are classified according to their morphological structure, namely case, number and gender. Note that gender is a category only used for Slavic languages, not Uralic.

In Votic, Ingrian, Olonets Karelian, and Lude, postposed demonstratives agree in case (138) and number (139) with the head noun. However, in Veps, postposed demonstratives agree with their head noun only in number, not case, as in (140).

Eastern Votic

- (138) *antõ-vad näväd poigaa sennee Petter-ii, juhd zavod-aa.*
 give.PST-3PL 3PL **boy.ACC** **DEM.ACC** Saint_Petersburg-ILL leather factory-ILL
 ‘They gave the boy away to Saint Petersburg, to a leather factory.’ (Kettunen & Posti 1932: 132)

Olonets Karelian

- (139) *üksikai poija-t net kod'i-h to-i.*
 one-by-one **boy- PL** **DEM.PL** home-ILL bring-PST.3SG
 ‘One-by-one, he brought the boys (teddy-bears) home.’ (Makarov & Rjagoev 1969: 141)

Central Veps

- (140) *nu, a vell-i-le ne tariž oli naida.*
 but and **brother-PL-ALL** **DEM.PL** need be.PST.3SG get_married.INF
 ‘Well, and the brothers also need to get married.’ (Onegina & Zajceva 1996: 182–184)

Several examples, particularly from Southern Lude and Veps, show a conflict in number agreement between the head noun and the postposed demonstrative, seen in (141) and (142).

Southern Lude

- (141) *vot samvaruo-d, samvara-t se ol'-d'i-he mednuo-d.*
 well samovar-PL.PTV **samovar-PL** **DEM.SG** be-PST-3PL brass-PL.PTV
 ‘Well (as for) the samovars, the samovars are made from brass.’ (Pahomov 2011: 230)

Central Veps

- (142) *priha-d se vāhān jü-i-ba.*
guy-PL **DEM.SG** little drink-PST-3PL
 ‘The guys drank [wine] a little bit.’ (Gerd et al. 2002: 70–72)

Truncation of number and case agreement in the eastern end of the dialect continuum, namely Southern Lude and Veps, suggests that internal agreement in the noun phrase has weakened, as demonstratives move to after the head noun. This might also explain why the use has extended from nominal to non-nominal parts of speech. With non-nominal hosts such as pronoun, verb and adverb, postposed demonstratives remain in their basic singular form *se*, as in (143) and (144).

Olonets Karelian

- (143) *nu häi se o'li vai 60 ümbäri kond-ie-n tappa-nut.*
 but 3SG DEM be.PST.3SG only 60 around bear-PL-GEN kill-PTCP
 ‘Well, he had killed only about sixty bears.’ (Makarov & Rjagoev 1969: 140–141)

Central Veps

- (144) *ihastji-ñ se ka surmha-zesei, surmha-zesei ihastji-ñ.*
 rejoice-1SG DEM so death-TERM death-TERM rejoice-1SG
 ‘I will, indeed, rejoice until the end of my life.’ (Onegina & Zajceva 1996: 133–138)

The North Russian data, in contrast, shows a higher degree of variation in agreement. As mentioned in Section 5.2.2, postposed demonstratives can agree in case, number, and gender with their head noun as in (145), or they can remain in the nominative singular neuter form *to* as in (146), which they also do with non-nominal hosts, such as with the adverb *teper* ‘now’ in (147). Moreover, we also encounter the phenomenon of a type of phonological agreement, a.k.a. “harmonisation” (Trubinskij 1970), in which demonstratives copy the vowel of the immediately preceding syllable as in (148), regardless of whether there would be a conflict in agreement with the head noun.

	With agreement		No agreement
(145)	<i>koljõsa-te</i> mill.NEU.PL-DEM.PL ‘those/the mills’	(146)	<i>voda-to</i> water.FEM.NOM-DEM.NEU.NOM ‘that/the water’
	Non-nominal host		Harmonisation
(147)	<i>teper-to</i> now-DEM.NEU.NOM ‘now’	(148)	<i>malo mjaša-ta</i> little meat.NEU.GEN-DEM.FEM.NOM ‘little of meat’

These four contexts of agreement found in the data also correspond to earlier descriptions by Kuz'mina (1993: 186) and Mendoza (2011: 249).

As mentioned previously, Finnic varieties of the east as well as North Russian also use postposed demonstratives after other non-nominal parts of speech. Apart from nominal categories such as nouns, pronouns, proper names, adjectives, and numerals, demonstratives can also follow verbs, adverbs and adpositions. All these are possible in Central and Southern Veps, as well as North Russian, seen in Table 25. Moreover, this does not seem to be affected by the number of syllables, as postposed demonstratives can attach to short as well as long words.

Part of speech	Central Veps	North Russian (Arkhangelsk)
Noun DEM	<i>vezi se</i> ‘water + DEM’	<i>voda-ta</i> ‘water-DEM.FEM’
Pronoun DEM	<i>mina se</i> ‘1SG + DEM’	<i>ja-to</i> ‘1SG-DEM’
Name DEM	<i>Miša se</i> ‘Miša (Mikhail) + DEM’	<i>Miša-to</i> ‘Miša (Mikhail) -DEM’
Adjective DEM	<i>lām’ se</i> ‘warm + DEM’	<i>teplo-to</i> ‘warm-DEM’
Numeral DEM	<i>kakś se</i> ‘two + DEM’	<i>dva-ti</i> ‘two-DEM.PL’
Verb DEM	<i>elān se</i> ‘live.1SG.PRS + DEM’	<i>živete-to</i> ‘live.2PL.PRS-DEM’
Adverb DEM	<i>nügde se</i> ‘now + DEM’	<i>teper’-to</i> ‘now-DEM’
Adposition DEM	<i>kuzon au se</i> ‘spruce.GEN + under + DEM’	<i>vperēd-to ljudej</i> ‘ahead-DEM + people’

Table 25. *Postposed demonstratives attaching to different parts of speech*

A noteworthy point here is the contrast with preposed demonstratives, which never modify any other part of speech than nominal (noun, pronoun, name, adjective, and numeral).

Table 26 shows the occurrences of different parts of speech which precede a demonstrative of the same phrase unit.

Language	Variety	Part of speech								Total
		Nominal					Other			
		Noun	Pronoun	Name	Adjective	Numeral	Verb	Adverb	Adposition	
Livonian		14	-	-	2	-	-	-	-	16
South Estonian		8	1	1	2	-	-	-	-	12
North Estonian		17	-	-	1	-	-	-	-	18
Votic	Western	13	-	-	-	-	-	-	-	13
	Eastern	4	-	-	-	-	-	-	-	4
Ingrian	Soikkola	5	-	-	-	-	-	-	-	5
	Heva	15	-	1	-	-	-	-	-	16
Karelian	Olonets	53	4	1	1	-	-	-	-	59
Lude	Northern	50	-	5	-	-	2	-	-	57
	Southern	58	21	18	8	4	11	17	-	137
Veps	Northern	95	6	12	1	1	3	1	-	119
	Central	269	7	10	6	4	34	16	2	348
	Southern	73	6	1	1	1	9	3	1	95
North Russian	Vologda	147	25	12	20	1	43	20	-	268
	Arkhangelsk	315	22	15	30	9	76	58	1	526

Table 26. *Frequencies of different parts of speech which precede a postposed demonstrative*

Table 27 shows the contrast between tendencies of nominal and non-nominal hosts across varieties.

Language	Variety	Postposed demonstratives	Tendency of nominal hosts		Tendency of non-nominal hosts		Distance from Vologda
Livonian		16	16	100%	0	0.00%	960 km
South Estonian		12	12	100%	0	0.00%	698 km
North Estonian		18	18	100%	0	0.00%	754 km
Votic	Western	13	13	100%	0	0.00%	556 km
	Eastern	4	4	100%	0	0.00%	532 km
Ingrian	Soikkola	5	5	100%	0	0.00%	547 km
	Heva	17	17	100%	0	0.00%	512 km
Karelian		59	59	100%	0	0.00%	349 km
Lude	Northern	57	55	96.49%	2	3.51%	339 km
	Southern	137	109	79.56%	28	20.44%	284 km
Veps	Northern	119	115	96.64%	4	3.36%	235 km
	Central	348	296	85.06%	52	14.94%	181 km
	Southern	95	82	86.32%	13	13.68%	193 km
North Russian	Vologda	268	205	76.49%	63	23.51%	0 km
	Arkhangelsk	526	391	74.33%	135	25.67%	338 km

Table 27. *Tendency of nominal vs. non-nominal hosts*

From Table 27, it is interesting to see that the easternmost Finnic languages, Veps and Lude, are as productive as North Russian in using postposed demonstratives with non-nominal parts of speech. Both Northern Lude and Northern Veps show a low frequency, however, deviating from their cognate varieties further to the south. The case of Northern Lude might be due to the genealogical connection to their Proto-Karelian root, which could have maintained the same tendency of not using demonstratives with non-nominal hosts across different Karelian varieties. At the same time, a low frequency in Northern Veps might simply be due to the difference in text genres, even though areal adjacency and contact with Northern Lude along with Karelian proper could also be used as explanation.

Chart 8 transforms the numeric data in Table 27 into a scale of geographical distribution. A high value of $R^2 = 0.6843$ is seen in Chart 8. This means that this correlation covers approximately 68% of the data, confirming that the geographical location of each variety also affects the frequency of use of postposed demonstratives with non-nominal hosts (similar to the two other cases discussed in Section 6.1.6).

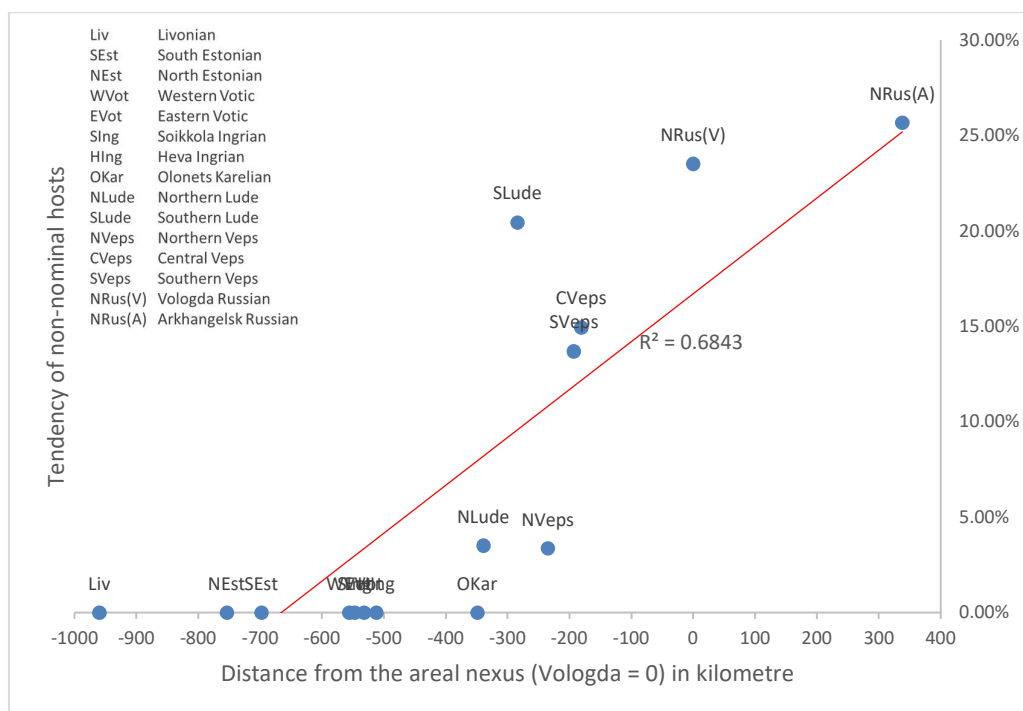


Chart 8 Tendency of non-nominal hosts

In many examples, it is not clear whether a postposed demonstrative only marks its preceding constituent or the entire preceding phrase. This is observed in Lude, Veps, and North Russian, i.e. varieties which can use postposed demonstratives with non-nominal parts of speech, seen in Table 26 and Table 27. As it is often ambiguous, and difficult to judge whether the range of marking covers the whole phrase, the rest of this section will not discuss this issue quantitatively, but rather present several examples which illustrate the phenomenon hand (see also Hayashi & Yoon 2006 for the discussion of this interpretation problem).

First of all, postposed demonstratives might not follow only an adjacent noun, but might in fact follow several preceding nouns, in a comitative construction consisting of two people such as ‘mom and dad’ in (149) and ‘we with our granny’ in (150).

Central Veps

- (149) *no, tat da mam se koud-i-he,*
 well **dad** **and** **mom** **DEM** die-PST-3PL
a gä-i hii-le kuume härgä-d.
 but remain-PST 3PL-ADES three ox-PTV
 ‘Well, mom and dad passed away, but three of their oxen remain.’
 (Onegina & Zajceva 1996: 178–182)

North Russian (Vologda)

- (150) *u nas byla korova telënok i desjat' ovec,*
 at 1PL.GEN be.PST.FEM cow calf and ten sheep
my s babusei-to vot kogda žili.
 1PL with granny.INSTR-DEM so when live.PST.PL
 'We had a cow, a calf, and ten sheep when we were living with granny.' (RNC)

Meanwhile, postposed demonstratives can also highlight not just the preceding verb, but the whole preceding verb phrase, as in (151) and (152).

Central Veps

- (151) *hebo-u aja-nu se, ka heboine tanha-u.*
 horse-ADES ride-PTCP.PST DEM so horse yard-ADES
 'I came riding a horse, the horse (stood) in the yard.' (Zajceva & Mullonen 1969: 104–107)

North Russian (Vologda)

- (152) *ak kak mama tvūrit'-to vet' ja ne znaju.*
 but how mother make-DEM after_all 1SG NEG know-1SG
 'But how (my) mother made (it), I do not know.' (Kasatkina 1991: 96)

This observation hints that postposed demonstratives in Southern Lude, Veps, and North Russian might have extended their constituent marking function into more of a phrase marker. From a certain perspective, it has started to behave somewhat similarly to topicalising demonstratives in some Southeast Asian languages, which generally highlight the entire chunk of information uttered before the demonstrative marker (discussed in Section 4.4.4). This functional parallel with other languages is discussed further in Section 7.1.4.

6.3 SYNTACTIC FUNCTIONS

The use of postposed demonstratives as modifiers of head words can occur in various syntactic functions in Finnic and North Russian. This section investigates the relation between the positions where the phrases containing demonstrative are observed in a clause (clause-second vs. other positions), and the syntactic functions in which the observed phrases are employed (subject, direct object, indirect object, verb, adverbial, determiner, and predicative). The reason for separating clause-second position from other clause positions (Sections 6.3.1 and 6.3.2) is that the marking of syntactic functions, along with their distribution, may be partly affected by clause position, due to some effects like subjecthood (discussed in Givón 2001: 473–474).

Andrews (2007: 152–193) classifies three main groups of syntactic functions: (i) nuclear functions, (ii) oblique functions, and (iii) external functions. First, the nuclear function consists of a “subject” in (153), a “direct object” in (154), and an “indirect object” in (155), which semantically and structurally belong to intransitive, transitive, and ditransitive verbs.

Southern Veps

- (153) *mamš se nagra-b, dei basi-b: ...*
 old_lady DEM laugh-3SG and speak-3SG
 ‘The old lady laughs and speaks: ...’ (Zajceva & Mullonen 1969: 240–241)

- (154) *oi, nagrhude-n se voi-b otta.*
 oi turnip-ACC DEM can-3SG take-INF
 ‘Oi, the turnip can be taken already.’ (Zajceva & Mullonen 1969: 228–232)

Olonet Karelian

- (155) *hää meil’e se tavari-a and-u.*
 3SG 1PL.ALL DEM goods-PTV give-PST.3SG
 ‘He gave some goods to us.’ (Makarov & Rjagoev 1969: 25)

Second, oblique functions are an addition to the nuclear function presented above. This group consists of adverbials expressing semantic roles such as “instrument” in (156), “manner” and “location” in (157), and “time” in (158).

North Russian (Vologda)

- (156) *sejčas nikto igrat’-to ne umeet na garmoške-to.*
 now no_one play-DEM NEG know.3SG on accordion-DEM
 ‘Now, no one knows how to play the accordion.’ (RNC)

- (157) *a potom on otošël i my-to: ur-ra-a,*
 but then 3SG leave-PST.MASC and 1PL-DEM hurray
da begom-to do školy-to.
 and run.INSTR-DEM till school.GEN-DEM
 ‘But then, he left and we (cried): hurray, and we ran to school.’ (RNC)

- (158) *a ja-to konešno dak ne vidala snacjala-to.*
 but 1SG-DEM definitely so NEG see.PST.FEM initially-DEM
 ‘But I definitely did not see (it) at first.’ (Kasatkina 1991: 91)

In addition to the categories mentioned above, I also add to the oblique function two additional categories, not included in Andrew’s classification. The reason for this addition is that postposed demonstratives can also be observed after a “determiner” of a phrase in a genitive construction ‘Y is of X’ in (159) and (161), and “predicative” of a copula phrase ‘Y is X’ in (160) and (162).

Southern Lude

- (159) *i tämä žüda-n se-n luod, oi!*
 and this heck-GEN DEM-GEN islet oh
 ‘And this damn islet, oh!’ (Pahomov 2011: 156)

- (160) *Šürd’a-lpiä häin om ninkuuhe se da.*
 Syrjä-ABL 3SG be.3SG such DEM yes
 ‘He is such a person from Syrjä.’ (Pahomov 2011: 198)

North Russian (Arkhangelsk)

- (161) *vot v podpol dveri-to*
 well in underground door.GEN-DEM
 ‘Well, on the underside of the door.’ (RNC)

- (162) *vot èto byla Nikifiriha-to.*
 well this be.PST.FEM Nikifiriha-DEM
 ‘Well, it was that Nikifiriha.’ (RNC)

In the predicative function, postposed demonstratives normally can accompany predicate adjectives (160) and nouns (162), but no instance of a postposed demonstrative in a predicate locative construction is observed in the data (see also the typology of predicative in Payne 1997: 114–123).

Following the report of syntactic function assignment in each variety (Sections 6.3.1), statistical tests will be conducted to show whether there is an effect of clausal position influencing subject and direct object marking (Section 6.3.2), and the correlation between differences in the constituent orders of demonstrative (preposed vs. postposed) and distribution of different syntactic functions (Section 6.3.3).

6.3.1 CLAUSE-SECOND VERSUS OTHER CLAUSAL POSITIONS

In all clausal positions shown in (135) to (137) (see Section 6.1.6), the subject is the syntactic function that most often co-occurs with demonstratives, seen in Table 28 for clause-second position (i.e. Wackernagel’s law), and Table 29 for other clausal positions (clause-medial and clause-final).

Language	Variety	Syntactic functions							Total
		Subject	Direct object	Indirect object	Verb	Adverbial	Determiner	Predicative	
	Livonian	14	2	-	-	-	-	-	16
	South Estonian	9	1	-	-	-	-	-	10
	North Estonian	17	-	-	-	-	-	-	17
Votic	Western	12	1	-	-	-	-	-	13
	Eastern	2	-	-	-	-	-	-	2
Ingrian	Soikkola	4	-	-	-	1	-	-	5
	Heva	12	3	-	-	-	-	-	15
Karelian	Olonets	35	9	2	-	4	-	-	50
Lude	Northern	25	6	-	1	2	-	-	34
	Southern	22	12	1	-	13	-	2	50
Veps	Northern	55	4	1	2	10	-	-	72
	Central	115	24	2	8	41	-	-	190
	Southern	32	10	1	1	8	-	-	52
North Russian	Vologda	44	18	7	11	31	-	-	111
	Arkhangelsk	96	30	3	10	72	-	2	213

Table 28. *Syntactic functions co-occurring with postposed demonstratives in clause-second position*

Language	Variety	Syntactic functions							Total
		Subject	Direct object	Indirect object	Verb	Adverbial	Determiner	Predicative	
	Livonian	-	-	-	-	-	-	-	-
	South Estonian	2	-	-	-	-	-	-	2
	North Estonian	1	-	-	-	-	-	-	1
Votic	Western	-	-	-	-	-	-	-	-
	Eastern	-	2	-	-	-	-	-	2
Ingrian	Soikkola	-	-	-	-	-	-	-	-
	Heva	-	1	-	-	-	-	-	1
Karelian	Olonets	2	5	-	-	2	-	-	9
Lude	Northern	8	8	-	1	6	-	-	23
	Southern	28	9	-	11	32	1	6	87
Veps	Northern	15	18	2	1	11	-	-	47
	Central	32	41	1	26	55	1	2	158
	Southern	8	7	2	8	17	-	1	43
North Russian	Vologda	33	27	5	32	54	-	6	157
	Arkhangelsk	57	69	4	66	106	2	9	313

Table 29. *Syntactic functions co-occurring with postposed demonstratives in non-clause-second positions*

Table 28 and Table 29 show that there is an obvious west-east division across the dialect continuum. In the Finnic varieties in the west, from Livonian to Ingrian, generally, postposed demonstratives occur in core functions (e.g., subject or direct object primarily in clause-second position), with some sporadic occurrence in adverbial phrases. At the same time, in eastern Finnic varieties and North Russian, postposed demonstratives occur in constituents in all syntactic functions, and in any clause position in the clause.

6.3.2 CLAUSAL POSITION EFFECTS IN SUBJECT AND DIRECT OBJECT MARKING

In terms of information structure, clause-initial position is the default place for subject (see, e.g., Keenan 1976). On the one hand, the subject is generally prominent enough that it needs no overt morphological marking. On the other hand, a direct object is more prone to be marked in general, especially when occurring in the topic position (see, e.g., Onishi 2001 for a theoretical discussion, and Ceglowski & Tajsner 2006 for the case of a topicalised object marking with the demonstrative *to* in Polish).

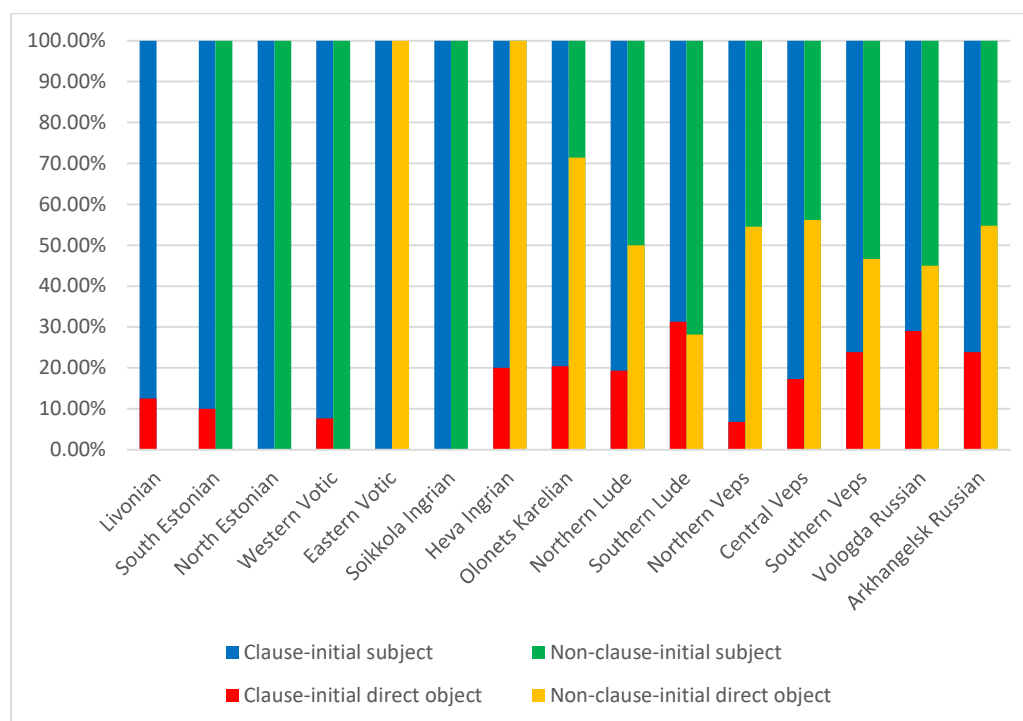


Chart 9 Co-occurrences of demonstrative with subject and direct object in clause-initial and other positions

Based on the proportion illustrated in Chart 9, the Fisher's exact test is conducted to see whether the difference between demonstratives that mark subjects and objects in clause-second and other positions is statistically significant. The hypothesis to be tested is whether marking of subjecthood is dependent on clausal position; that is,

subjects would still co-occur with demonstratives more frequently than direct objects in clause-initial position, as shown in Table 30.

Language	Variety	Clause-initial position		Other clause positions		Statistical significance Fisher's exact (greater)
		Subject	Direct object	Subject	Direct object	
Livonian		14	2	-	-	$p = 1$ (not significant)
South Estonian		9	1	2	-	$p = 1.0$ (not significant)
North Estonian		17	-	1	-	$p = 1$ (not significant)
Votic	Western	12	1	-	-	$p = 1$ (not significant)
	Eastern	2	-	-	2	$p = 0.17$ (not significant)
Ingrian	Soikkola	4	-	-	-	$p = 1$ (not significant)
	Heva	12	3	-	1	$p = 0.25$ (not significant)
Karelian	Olonets	35	9	2	5	$p = 0.013$ (significant)
Lude	Northern	25	6	8	8	$p = 0.034$ (significant)
	Southern	22	12	28	9	$p = 0.90$ (not significant)
Veps	Northern	55	4	15	18	$p = 0.00000054$ (significant)
	Central	115	24	32	41	$p = 0.000000096$ (significant)
	Southern	32	10	8	7	$p = 0.093$ (not significant)
North Russian	Vologda	44	18	33	27	$p = 0.050$ (significant)
	Arkhangelsk	96	30	57	69	$p = 0.00000037$ (significant)

Table 30. *Does subject co-occur with demonstrative more than direct object in clause-initial position?*

The degrees of significance as tested and shown in Table 30 suggest that in the clause-initial position, a postposed demonstrative co-occurs with a subject more often than a direct object in Finnic varieties in the east and North Russian. However, due to insufficient data, the dependency test cannot provide a reliable result for such a tendency in Finnic varieties in the west. Nevertheless, the numbers of occurrences in Table 30 alone already indicate that clause-initial position is almost certainly the place for subject, regardless of the presence of postposed demonstrative, confirming the tendency of definite subject marking discussed in connection to Mordvinic definite declension in Section 5.1.3 (see also Givón & Li 1976). In any case, the reason for a subject co-occurring with a postposed demonstrative more often than a direct object co-occurring with a postposed demonstrative in clause-initial position may not be related to a subjecthood effect, but rather to a factor involving information structure (cf. Onishi 2001; Cegłowski & Tajsner 2006, discussed above).

To summarise the findings thus far, in western and central Finnic varieties, subjects and direct objects in clause-initial position tend to co-occur with postposed demonstratives more often than in eastern Finnic and North Russian, possibly due to the main pragmatic use of topicalisation (discussed in Section 6.4.2). Meanwhile, in eastern Finnic and North Russian, referents can freely co-occur with postposed demonstratives, regardless of their clausal position. Therefore, it is possible to claim

that the use of postposed demonstrative around the areal nexus (Vologda) has moved beyond the clausal position effect, likely related to their other evaluative uses beyond referentiality and information structure (discussed further in Section 6.4.3).

6.3.3 DISTRIBUTION OF SYNTACTIC FUNCTIONS AMONG PREPOSED AND POSTPOSED DEMONSTRATIVES

Another interesting factor to test statistically is the distribution of clausal constituents between nuclear vs. other syntactic functions co-occurring with preposed demonstratives (Table 31), in comparison to postposed demonstratives (Table 32). As some Finnic varieties still prefer preposed demonstratives, the test is limited to those Finnic varieties and North Russian, where the frequency of postposed demonstratives outnumbers that of preposed ones (see Table 19, Table 20, and Chart 4).

Language	Variety	Syntactic functions							Total
		Nuclear			Other				
		Subject	Direct object	Indirect object	Verb	Adverbial	Determiner	Predicative	
Lude	Southern	22	13	-	-	50	5	-	90
Veps	Northern	4	3	-	-	8	-	1	16
	Central	7	5	1	-	27	1	-	41
	Southern	1	3	-	-	5	-	-	9
North Russian	Vologda	-	-	-	-	17	-	-	17
	Arkhangelsk	7	2	-	-	46	-	1	56

Table 31. *Distribution of syntactic functions among co-occurrences with preposed demonstratives*

Language	Variety	Syntactic functions							Total
		Nuclear			Other				
		Subject	Direct object	Indirect object	Verb	Adverbial	Determiner	Predicative	
Lude	Southern	50	21	1	11	45	1	8	137
Veps	Northern	70	22	3	3	21	-	-	119
	Central	147	65	3	34	96	1	2	348
	Southern	40	17	3	9	25	-	1	95
North Russian	Vologda	77	45	12	43	85	-	6	268
	Arkhangelsk	153	99	7	76	178	2	11	526

Table 32. *Distribution of syntactic functions among co-occurrences with postposed demonstratives*

Chart 10 clearly shows that among the varieties that use postposed demonstratives more than the preposed demonstratives, the division of labour is very similar. Namely, there is a common tendency that postposed demonstratives are more likely found with nuclear syntactic functions, whereas preposed demonstratives more often mark other functions.

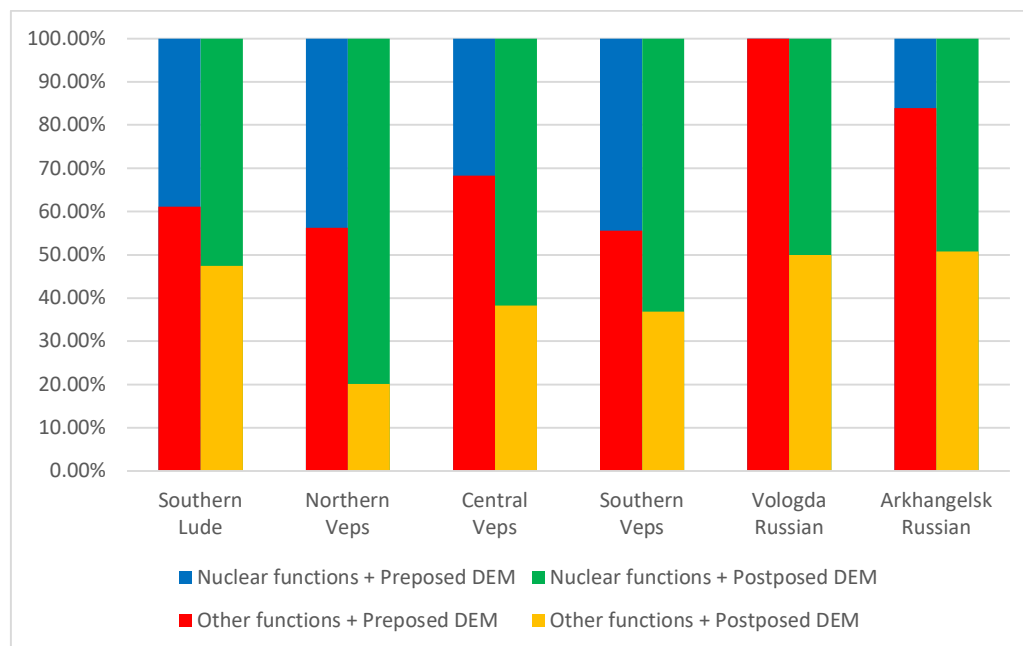


Chart 10 Co-occurrences of nuclear functions and other functions with preposed and postposed demonstratives

The Fisher's exact test is conducted in Table 33, in order to test the validity of the correlation.

Language	Variety	Preposed demonstratives		Postposed demonstratives		Statistical significance Fisher's exact 2-tail
		Nuclear functions	Other functions	Nuclear functions	Other functions	
Lude	Southern	35	55	72	65	$p = 0.057$ (not significant)
Veps	Northern	7	9	95	24	$p = 0.0037$ (significant)
	Central	13	28	215	133	$p = 0.00033$ (significant)
	Southern	4	5	60	35	$p = 0.30$ (not significant)
North Russian	Vologda	-	17	134	134	$p = 0.000015$ (significant)
	Arkhangelsk	9	47	259	267	$p = 0.00000098$ (significant)

Table 33. Statistical test on the correlation between syntactic function types and constituent orders

The values of some varieties in the Fisher's exact test are $p < .01$, suggesting that the correlation between constituent order and assignment of syntactic function is highly significant in these varieties. The correlation is not statistically significant in Southern Lude and Southern Veps, but this might be due to insufficient data, or the genre of the language data.

6.4 PRAGMATIC FUNCTIONS

From a pragmatic perspective, the use of postposed demonstratives can add various meanings to their host words. As introduced in Section 1.1, the working hypothesis of the current study is that the functions of postposed demonstrative can be compared with those of markers of definiteness, topic, focus, and discourse (discussed in Chapter 4).

Interestingly, the Veps and North Russian examples (163) and (164) below even show stacking of two postposed demonstratives, which seemingly carry different functions.

Central Veps [identical to (1) and (70)]

- (163) *nügüt silaaž om jo torvut se se?*
 now 2SG.ADES be.3SG already **phonograph** **DEM** **DEM**
 'Have you already got the phonograph now?' (Kettunen 1943: 402)

North Russian (Arkhangelsk) [identical to (2)]

- (164) *a eščě svėkor-ot-to eščě ženilsja na drugoj.*
 but still **father_in_law-DEM-DEM** still marry.PST.MASC on other.LOC
 'But still, the father-in-law got married again.' (RNC)

These examples are obvious evidence showing that postposed demonstratives may carry more than one pragmatic dimension. Hypothetically, the former demonstratives *se* and *-ot* have a referential function, whereas the latter demonstratives carry some other non-referential meaning, such as an interrogative mood in (163) and the speaker's evaluation in (164). An extreme case is seen in (165), which contains a multitude of postposed demonstratives in a single clause.

North Russian (Vologda)

- (165) *wot ta-to wot ščo-to to-to to-to goworila.*
 well **DEM.FEM-DEM** well what- **DEM** **DEM-DEM** **DEM-DEM** speak.PST.FEM
 'Well, she, uh, talked about something this and that.' (Kasatkina 1991: 100)

Functionally, the first instance, *ta-to*, marks a topic 'as for her, she ...'. The second use is as an indefinite pronoun *ščo-to* 'something' comparable to Standard Russian *čto-to*. The third and fourth uses might have something to do with referentiality and

evaluation. This, again, is an indication that the use of postposed demonstrative is decidedly multifunctional in North Russian, thus it is interesting to see whether comparable uses are also observed in Finnic varieties of the neighbouring areas.

Given the challenges in determining some of the meanings in the examples above, it would appear crucial to investigate the different uses of postposed demonstratives from multiple perspectives. Section 6.4.1 discusses postposed demonstrative usage in relation to different reference relations, while Section 6.4.2 focuses on different uses from an information-structural perspective of information structure. As there could potentially be other pragmatic uses which do not relate to referentiality or information structure, Section 6.4.3 introduces evaluative uses, which need yet another type of approach, which takes into account factors like stancetaking and epistemic modality.

6.4.1 REFERENTIAL USES

Discussion of definite articles in previous studies usually touches on the notion of referentiality (discussed in Section 4.3). Adopting the classification by Becker (2018) and König (2018), reference relations which postposed demonstratives (as potential definite articles) can express are classified into seven categories: (i) deictic, (ii) recognitional, (iii) absolutely unique, (iv) anaphoric, (v) bridging, (vi) situationally unique, and (vii) establishing (described in Section 4.4.3). Referential uses primarily concern noun phrases, so phrases in which demonstratives follow non-nominal parts of speech (i.e. verb, adverb, and adposition) are excluded from the analysis in this section.

The focus of this section is mainly to identify article-like uses of postposed demonstratives, not their obligatoriness (cf. Greenberg 1978: 61), so a comparison of similar contexts with and without postposed demonstratives will not be discussed here. Moreover, this section takes such stringent criteria, paying attention only to the semantics and referential relations *per se*, without consideration of information-structural effects, discussed later in Section 6.4.2.

Resuming the discussion of Section 4.3.3, any grammatical marker of definiteness should be able to express, at the very least, anaphoric referents, bridging referents, situationally unique referents, and establishing referents, in order to be called a definite article (see Becker 2018: 87; König 2018: 169). In particular, a demonstrative by itself can mark deixis and anaphora by default (see Hawkins 1978: 149–157; C. Lyons 1999: 331–332; Dixon 2003), so an extension to bridging referents, situationally unique referents, and establishing referents, for example, is decisive for determining whether a demonstrative has developed into a definiteness marker (see also Himmelmann 1997: 93–101; Huang 2000, discussed in Sections 4.3.2 and 4.3.3). The following examples can shed light on whether the Finnic and North Russian postposed demonstratives are functionally comparable to definite articles.

A deictic referent is frequently new information identifiable in the situation of utterance through its visibility, shared knowledge, or a restrictive relation in regard to previously mentioned referents (Hawkins 1978: 155; Dixon 2003). Deictic referents still show a gesturing expression, distinguishing an entity from others of its kind, for instance, spatially, as in (166) and (167).

Northern Veps

(166) A wedding cortege was moving on the road along the river shore. Suddenly the bride screamed.

no, siga druška-d ne i kuulište-t'he.
 well there friend-PL DEM.PL also hear-PST.3PL
sanu-tas: mii nece kidošta-b se?
 say-3PL what this scream-3SG DEM

‘Well, those friends over there also heard her. They say: Who is this screaming?’
 (Onegina & Zajceva 1996: 34–51)

North Russian (Vologda)

(167) *my vot tam na senokose na rec'ke byli,*
 1PL well there on hay_time.LOC on river.LOC be.PST.PL
u nas po ozeru-to vot, tut tako nissko mesto.
 at 1PL.GEN around lake.LOC-DEM well here so low place

‘We were there by the river during the hay time. Our place around the lake was low land.’
 (Kasatkina 1991: 91)

Based on these examples above, we can say that postposed demonstratives indicate a concrete location of entities in the real world: ‘the friends over there in the wedding cortege’ in (166), and ‘the lake around our place’ in (167).

A *recognitional referent* usually concerns information that has not been mentioned in the discourse but is nonetheless identifiable through common or shared knowledge or experience between the interlocutors, as in (168) to (170).

Olonets Karelian

(168) *šit kümminen vuvve-n südäme-h tul'-i tämä revol'ucii meil'e,*
 then ten year-GEN heart-ILL come-PST.3SG this revolution 1PL.ALL
rošia-h ... konzu kerenskoi-n pravi'eľstvu rodi-ih, cuari še
 Russia-ILL when Kerensky-GEN government.ACC be_born-PASS.PST Tsar DEM
l'ükä-ttih, meidü čotai-ttih bolševieko-i-kse.
 overthrow-PASS.PST 1PL.PTV regard-PASS.PST Bolshevik-PL-TRANSL

‘Then for ten years came this revolution to us, to Russia ... When the Kerensky government emerged, the Tsar was overthrown and we were regarded as Bolsheviks.’

(Makarov & Rjagoev 1969: 164)

Southern Veps

(169) A: *miš-ak lambas se minu-n?*
 where-PTCL sheep DEM 1SG-GEN
 B: *lambaz, san, om, a busiš-t' ii-le.*
 sheep say be.3SG but lamb-PTV NEG-be.CNG

- A: *mihe-k hān kado-o?*
 where-PTCL 3SG disappear-3SG
 A: ‘Where are my sheep?’
 B: ‘There are, say, sheep, but there is no lamb.’
 A: ‘Where did they disappear?’ (Zajceva & Mullenon 1969: 228–232)

North Russian (Vologda)

- (170) An old lady was sitting and pondering how she would live at the old age. She talked about her daughter-in-law whom she was not pleased about.

... *skazet: “mamka ne nado”*.
 say.3SG mommy NEG necessary
a bog znaet, snohi-ti kakie.
 but God know.3SG daughter_in_law.PL-DEM.PL how.PL
 ‘... she [the daughter-in-law] will say: “Mommy, please do not [interfere my affairs]”.
 But God knows how daughters-in-law are.’ (RNC)

In (168), both interlocutors shared an experience witnessing the Russian Revolution, so it was obvious which of the Russian Tsars they were talking about. In (169), the speaker was certain that the addressee had seen his sheep before and knew exactly which sheep he was asking about. In (170), the speaker assumes that the addressee shares the same thought with her, that her daughters-in-law are annoying and unpleasant, typical for daughters-in-law in the interlocutor’s community.

In relation to information structure, recognitional referents can be considered identifiable without a foregrounding context, due to their strong status as common-ground knowledge shared between the interlocutors (Krifka 2008, discussed in Section 4.4). Moreover, as recognitional referents usually involve shared knowledge that the speaker presents, the postposed demonstratives in (168) to (170) above can also be substituted by first person possessive suffixes, e.g., ‘our Tsar’, ‘my sheep,’ and ‘our daughters-in-law’ (see further discussion of “shared knowledge” in Section 6.4.3).

An absolutely unique referent stands for such an entity that there is no contrast with another referent. There are cases which can be considered being such referents, e.g., a language name ‘the Russian language’ in (171), the religious concept ‘God’ in (172), a specific date or historical event such as ‘the Victory Day’ in (173), and a natural phenomenon ‘the Sun’ in (174).

Southern Lude

- (171) *ka “kläps”, kapkan meil, vena-ks se “kapkan” om.*
 so trap[Lude] trap[Rus] 1PL.ADES Russian-TRANSL DEM trap[Rus] be.3SG
 ‘So *kläps* is a trap metal at our place. In the Russian language they say *kapkan*.’
 (Pahomov 2011: 206)

Central Veps

- (172) *a jumou se pagasta-s teda-b, kudam min moli-še.*
 but **God** **DEM** church-INES know-3SG how who pray-3SG
 ‘But God in the church knows the matters that people pray about.’
 (Onegina & Zajceva 1996: 130–132)

North Russian (Vologda)

- (173) *v den'-ot pobedy my šli v akkurat iz domu*
 in day-DEM victory.GEN 1PL come.PST.PL to precision from house.GEN
i vstretili počtal'ona (utrom rano na val'gu
 and meet.PST.PL postman.GEN morning.INSTR early to Val'ga
šel-ot počtal'on).
 come.PST.MASC-DEM postman
 ‘On Victory Day, we had just come out of our house and met a postman (every morning the postman came early to Val'ga).’ (RNC)
- (174) *už i solnyško-to saditsja ...*
 already also sun-DEM set.3SG
 ‘Already now, the Sun is also setting ...’ (RNC)

In (171) to (174), the referents accompanied by postposed demonstratives are unambiguous by themselves, without earlier mention or a foregrounding context. From an information-structural perspective, this unambiguity and uniqueness may be due to the fact that these are permanently available topics (see Erteschik-Shir 2007: 17, discussed in Section 4.4.2).

In any case, the use of postposed demonstratives in the context of the absolutely unique referents above may not in the end be purely referential, but rather due to topicality, as the majority of examples for absolutely unique referents show an occurrence of demonstratives in clause-second position, where the use of demonstratives can alternatively be interpreted as a marker of topicality (see the approach taken in Section 6.4.2). Given that there are also cases in which language names as in (171), ‘God’ as in (172), and ‘the Sun’ as in (174) do not co-occur with demonstratives, the data implies that the use of postposed demonstratives with such referents is rather motivated by topicality effect, rather than being a true referential use.

An anaphoric referent, in contrast to recognitional and unique referents, refers back to an antecedent, i.e. an entity previously mentioned in the discourse, as in (175) to (178).

Livonian

- (175) *izā izā vōļ Dāvid, nim Dāvid.*
 father.GEN father be.PST.3SG Dāvid name Dāvid

izā izā se vól Pizānikā rištīng.
father.GEN father DEM be.PST.3SG Piza.GEN person
 ‘My grandfather was Dāvid, his name was Dāvid. My grandfather was from Piza.’
 (Mägiste 2006: 4)

Western Votic

(176) *võtt-i tämä naisiuzii. sa-i ene-le üvää*
 take-PST.3SG 3SG marry.PST.3SG get-PST.3SG self-ALL good.ACC
ilozaa naizõõ. a sene-l naizõ-ll(õ) õli varjo.
 joyful.ACC woman.ACC but that-ADES woman-ADES be.PST.3SG **mirror**
varjo se õli mokoma että se pajatt-i.
mirror DEM be.PST.3SG such that it talk-PST.3SG
 ‘He (the emperor) took (a woman) and married (her). He got himself a good joyful woman.
 But that woman had a mirror. The mirror was such a thing that it could talk.’
 (Kettunen & Posti 1932: 58)

Olonet Karelian

(177) *ol'i sie kowme vel'Pes-tü. nu vel'Pekse-t net ero-ttih.*
 be.PST.3SG there three **brother-PTV** but **brother-PL DEM.PL** separate-PST.3SG
 ‘There were three brothers. But the brothers got separated.’ (Makarov & Rjagojev 1969: 177)

North Russian (Arkhangelsk)

(178) *ja zapomnila u nas tatja-to prišël s germanskoj*
 1SG remember.PST.FEM at 1PL dad-DEM come. PST.MASC from **German.GEN**
vojny eščë vojna-ta vot byla germanska, ne eta-ta ...
war.GEN still war-DEM well be.PST.FEM German NEG this- DEM
perva vojna, ak on prišël s vojny-ta.
 first war so 3SG come. PST.MASC from **war.GEN**
 ‘I remembered that our father came back from the German war, still, the war was the one against
 Germany, not this first war, so he came back from the war.’ (RNC)

Omitting postposed demonstratives in these contexts does not cause ungrammaticality, but it may interrupt the identifiability for the addressee (see a similar thought on the optionality of Finnish prenominal demonstratives in Larjavaara 1990: 146; Chesterman 1991: 102–103; Juvonen 2000: 193). The anaphoric use can be a short-distance reference, as in (175) to (178), but a long-distance reference that reactivates an entity mentioned earlier (over a 50–200 word gap) is also observed (see further discussion on topic continuing and switching in Section 6.4.2).

A bridging referent usually establishes an associative relation to antecedent, which is not identical to the referent as in the anaphoric cases (175) to (178), but inferable through inalienable relations, as in (179) to (182).

North Estonian

- (179) *ja siis pärast, vihku kirju-ttud ... aga etteütlemese*
 and then after instalment write-PST.PTCP but prompt.GEN
kiri se ol'i ikka tahvle pia-l ja ...
 note DEM be.PST.3SG still table.GEN head-ADES and
 'And then after, an instalment agreement was written ... but the note of prompt payment was still on the table and ...' (EDC)

Southern Lude

- (180) *"mechiine om siga" da. "kävel'o-b" da. a vett,*
 forest_angel be.3SG there yes walk-3SG yes but well
d'äl'gi-i-d se ni ken ei näh-nu. san-daze om, aa.
 footprint-PL-PTV DEM NEG who NEG.3SG see-PST.PTCP say-PASS be.3SG ah
 'The forest angel is there yes. He roams, but well, no one saw the/his footprints, they only say that he exists.' (Pahomov 2011: 212)

Central Veps

- (181) *süčas kol-i, ka hibj se, näge-d, zdrogai-b*
 now die-PST.3SG so corpse DEM see-2SG shiver-3SG
 'Now, he (your friend) has died, so the/his corpse, you see, was shivering.'
 (Zajceva & Mullonen 1969: 91–93)

North Russian (Arkhangelsk)

- (182) *vyryli, tudy odin-to zalez, grob-to stal*
 dig_up.PST.PL here one-DEM climb_up.PST.MASC coffin-DEM become.PST.MASC
otkryvat', nu vot, kryšku-to svernul,
 open but well lid.ACC-DEM turn.PST.MASC
 'They dug up, one (of them) is climbing up here, the coffin started to open and then he turned the (coffin's) lid, ...' (RNC)

As discussed in Sections 4.3.2 and 4.3.3, inalienable relations can emerge between components of a process 'instalment payment' > 'note of prompt payment' in (179), producer/production and product 'walking' > 'footprint' in (180), event and participant 'dying' > 'corpse' in (181), or construction and its inseparable part 'coffin' > 'lid' in (182).

A *situationally unique referent* is similar to bridging referent in that it creates an associative anaphoric relation (Hawkins 1978, discussed in Section 4.3.3). However, it need not be an inseparable entity from its antecedent. Instead, the context is clear enough to indicate that the referent is related to an antecedent in the given context, despite being its first mention in the discourse, as in (183) to (185).

Northern Lude

- (183) *knagin ende veñtša-d vie-dau küliüi i šiid*
bride before wedding-PTV bring-PASS sauna.ILL and then
hälläi hībd' se kai maido-л pes-tau.
 3SG.ADES **body** **DEM** all milk-ADES wash-PASS
 ‘A bride, before a wedding, is brought to the sauna, and then all the/her body is washed by milk.’ (Ojansuu et al. 1934: 261)

Northern Veps

- (184) *no, ajaškat'-he randa-d se vert' möto dorog se, ...*
 well drive.PST-3PL bank-PTV DEM along with **road** **DEM**
 ‘Well, they [the wedding cortege] were moving on the road along the riverbank.’
 (Onegina & Zajceva 1996: 34–51)

North Russian (Vologda)

- (185) *i vot ètih ... kozlovyh kranov, tam v Ponizov'e-to*
 and well this.PL.GEN gantry PL.GEN crane PL.GEN there in Ponizov'e.LOC-DEM
u nas-to este na ploščadke-to dak vot.
 at 1PL.GEN-DEM be.PRS.3SG at **platform.LOC-DEM** so well
 ‘And well, these ... gantry cranes, there in Ponizov'e we have them at the assembly site, right.’
 (RNC)

In (183), the referent *hibd'* ‘body’ is unpossessed, but it refers to the body of the *knagin* ‘bride.’ In (184), the speaker does not talk about just any old road, but rather a specific road in the village that goes along the riverbank. In (185), the speaker indeed refers to gantry cranes located at the assembly site in the town of Ponizov'e, not those of any other town.

In a way, both bridging and situationally unique referents in other Uralic languages can be marked by possessive suffixes as the semantics are closely related to possession (see Fraurud 2001, discussed in Section 4.3.3), and omitting the possessive suffixes may sometimes cause ambiguity in interpretation. In any case, most Finnic varieties no longer use possessive suffixes productively, so it might be the case that postposed demonstratives have inherited such uses from obsolete possessive suffixes (see earlier and further discussion in Sections 5.1.1 and 7.2, respectively). This connection can be tested in translation by substituting the demonstratives with possessive determiners, which indeed would still keep the same sense of bridging and situationally unique reference, at least with the animate antecedents, for instance, ‘his footprints’ in (180), ‘his corpse’ in (181) and ‘her body’ in (183).

An establishing referent usually concerns new information that is mentioned for the first time in the discourse but which points to a clarifying statement that is shortly after the mention revealed as a cataphoric relation, as in (186) and (187).

Northern Veps

- (186) ... *kaik mise tulde-iž măriče-m-ha hänel*
all so_that come-COND.3SG try_on-INF-ILL 3SG.ADES
nene predmeta-d ne: kolčšaš-t bašmako-d i šlāpā-d.
these object-PL DEM.PL ring-PL boot-PL and hat-PL
‘... so that everyone would come to try on the objects: rings, boots, and hats.’
(Onegina & Zajceva 1996: 34–51)

North Russian (Arkhangelsk)

- (187) *u nas doma darmoedov-to mnogo, sestra ostalas nezamužna,*
at 1PL home sponger.PL.GEN-DEM many sister stay.PST.FEM unmarried
brat es’ molože menja čerez brata, huden’koj,
brother be.3SG younger 1SG.GEN through brother.GEN slim
eščē dvoē. a ja, goorju, ja išo ostanus’ tret’ja.
still two but 1SG say.1SG 1SG PTCL become.1SG third
‘In our house, there are many of the spongers. My sister is still unmarried. My brother is younger and slimmer than me. There are still two. But I say: I will, you see, become the third [sponger].’
(RNC)

In both examples (188) and (189), the referents *nene predmetad ne* ‘the objects’ and *darmoedov-to* ‘the spongers’ are introduced to the discourse for the first time, before being elaborated on in the following phrase, to what the objects to be tried on are in (188), and as to who the spongers in the family are in (189). As Hawkins (1978: 150) and Himmelmann (1997: 93–101) consider this establishing use as decisive for identifying definite articles, this criterion, interestingly, is also met in Finnic varieties in the east and North Russian.

Table 34 summarises the occurrences of nominal phrases, in which postposed demonstratives are used in each variety. From Table 34, it is clear that the use of postposed demonstrative to with recognitional referents, absolutely unique referents, bridging referents, situationally unique referents, and establishing referents is observed mainly in the easternmost Finnic varieties, Lude and Veps, together with North Russian. This areal distribution, again, might be strongly related to the replacement of possessive suffixes by postposed demonstratives (see further discussion in Sections 7.1 and 7.2).

Language	Variety	Referential uses							Total
		Deictic	Recognitional	Absolutely Unique	Anaphoric	Bridging	Situationally unique	Establishing	
	Livonian	8	-	-	8	-	-	-	16
	South Estonian	7	-	-	5	-	-	-	12
	North Estonian	10	-	-	7	1	-	-	18
Votic	Western	-	-	-	13	-	-	-	13
	Eastern	-	-	-	4	-	-	-	4
Ingrian	Soikkola	1	-	-	4	-	-	-	5
	Heva	5	-	-	11	-	-	-	16
Karelian	Olonets	15	2	-	40	2	-	-	59
Lude	Northern	3	-	-	51	1	1	-	55
	Southern	30	5	3	65	3	2	1	109
Veps	Northern	16	1	-	88	-	9	1	115
	Central	41	6	2	213	5	29	-	296
	Southern	8	2	-	65	2	5	-	82
North Russian	Vologda	68	17	7	86	3	22	2	205
	Arkhangelsk	67	15	3	212	17	73	4	391

Table 34. Contexts of referential relation observed with postposed demonstratives

Considering the constituent order change assumed for Finnic varieties in the east, it is also interesting to compare the range of reference relations found with postposed vs. preposed demonstratives. In a similar manner as Section 6.3.3, referential uses in varieties which use postposed more than preposed demonstratives are compared in Table 35.

Language	Variety	Referential uses							Total
		Deictic	Recognitional	Absolutely Unique	Anaphoric	Bridging	Situationally unique	Establishing	
Lude	Southern	22	2	1	55	1	9	-	90
Veps	Northern	6	-	-	9	-	-	1	16
	Central	12	-	-	26	-	2	1	41
	Southern	3	-	-	6	-	-	-	9
North Russian	Vologda	9	-	-	6	-	2	-	17
	Arkhangelsk	31	-	1	24	-	-	-	56

Table 35. Contexts of referential relation observed with preposed demonstratives

The proportion of deictic vs. other non-deictic references co-occurring with preposed and postposed demonstratives are shown in Chart 11.

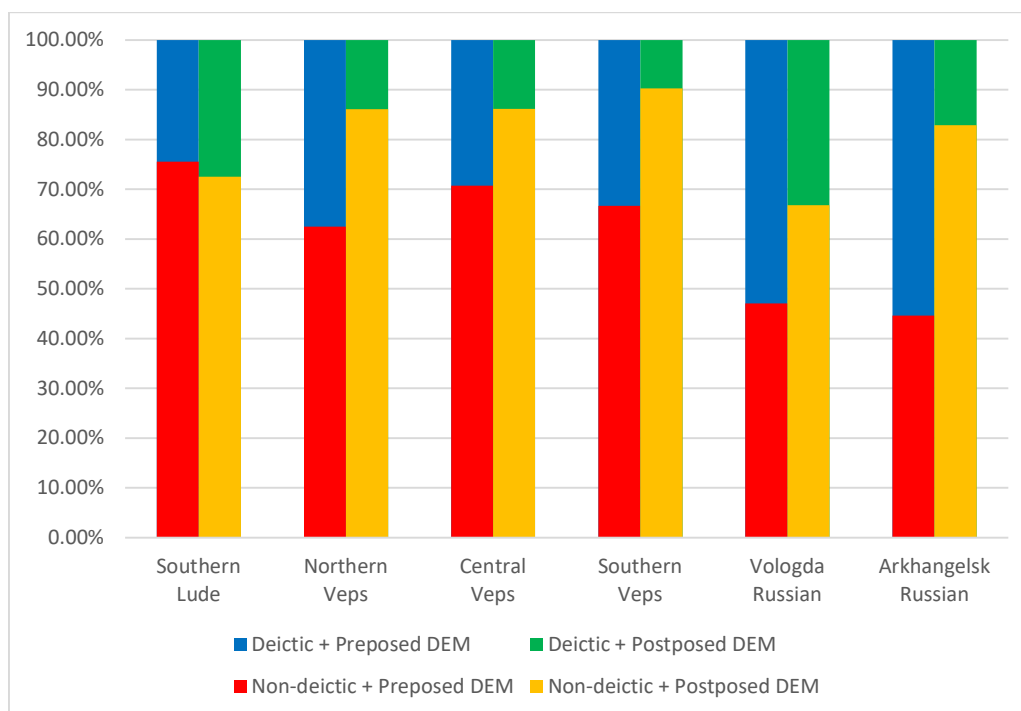


Chart 11 Proportions of deictic and non-deictic uses of preposed and postposed demonstratives

Fisher's exact test was conducted to see whether there is a significant correlation between deictic vs. referential uses and constituent orders, shown in Table 36.

Language	Variety	Preposed demonstratives		Postposed demonstratives		Statistical significance Fisher's exact 2-tail
		Deictic	Non-deictic	Deictic	Non-deictic	
Lude	Southern	22	68	30	79	$p = 0.75$ (not significant)
Veps	Northern	6	10	16	99	$p = 0.029$ (significant)
	Central	12	29	41	255	$p = 0.020$ (significant)
	Southern	3	6	8	74	$p = 0.074$ (not significant)
North Russian	Vologda	9	8	68	137	$p = 0.12$ (not significant)
	Arkhangelsk	31	25	67	324	$p = 0.000000037$ (significant)

Table 36. Deictic vs. other non-deictic uses of preposed and postposed demonstratives

Table 36 shows that the correlations between constituent orders and deictic vs. other non-deictic uses are significant ($p < .05$) in Northern and Southern Veps as well as North Russian, despite some deviation in Southern Veps and Vologda Russian due to insufficient data for preposed demonstratives. Nevertheless, the Southern Lude data is sufficient but still gives a negative result for this correlation, which might imply a

lower degree of pragmaticisation of postposed demonstratives. In any case, we can say that the shift from preposed to postposed demonstratives has decreased the functional load of spatial deixis of demonstratives, which indeed corresponds to the criteria for identifying grammaticalisation from a demonstrative to a definite article (given in C. Lyons 1977: 653–654; Hawkins 1978: 149–157; Anderson & Keenan 1985: 280; Diessel 1999a: 129; Dixon 2003, discussed in Sections 4.2 and 4.3).

This section has shown that based on referential uses observed in the data, the use of postposed demonstratives in Finnic varieties of the western and central areas do not fulfil the typological criteria of definite article in the classification by Becker (2018) and König (2018), while eastern Finnic and North Russian postposed demonstratives can appear in most reference relations characteristic of definite articles. More crucially, postposed demonstratives have shifted from the deictic domain, which remains the prototypical use for preposed demonstratives, to other non-spatial references. In any case, several contexts of use for postposed demonstratives discussed in this section can also be explained by non-referential uses, which will be discussed in the following section.

6.4.2 INFORMATION-STRUCTURAL USES

As shown in Section 6.1.6, postposed demonstratives in all Finnic and North Russian varieties under investigation are found in clause-second position, in which they mark a clause topic, resembling left dislocation constructions. As discussed in Section 4.4.2, various types of topic and topicalisation strategies are observed across the world's languages. This section focuses on the information flow, on whether the constituent co-occurring with a postposed demonstrative continues a referent currently active in the discourse, switches to a new referent, switches to a referent previously mentioned in the discourse, or starts a new narrative. This parameter has been previously applied to spoken Russian by Grenoble (1998: Ch. 5), and to Komi and North Russian by Leinonen (1998, 2006), among others. As the focus of this section is on topic marking, the first part of this section will only discuss demonstratives in clause-second position, generally considered the default place for topic markers in subject-prominent like those found in the like Finnic and Slavic languages, as discussed in Section 4.4.1.

A continuing topic continues the referential sphere by tracking the referent that remains active in the discourse, and the speaker thus assumes it to be expected by the addressee (discussed in Section 4.4.1). Accordingly, the distance between the antecedent and the referent is usually no longer than an interval of two clauses, usually less than 50 words, such as the close anaphoric reference seen in (175) to (178).

A switch-topic causes an interruption in the referential sphere in the ongoing discourse, and draws the addressee's attention to a referent that is not at the highest level of identifiability (discussed in Section 4.4.1). The referent can be old information that the speaker reintroduces to the discourse as an anaphoric referent. Often, a postposed demonstrative is used to help the addressee keep track of multiple protagonists in close context, which might become ambiguous over the course of a narrative or reported dialogue, as in (188).

Northern Veps

- (188) A bride has arrived to the house of the groom's godmother previously.

d'o nece ženih i tul-i. hän nühai-b i sanu-b:
 already this groom also come-PST.3SG 3SG sniff-3SG and say-3SG
"božuško, reska-n hengen duhh om." a adiv se
 godmother fresh-GEN breathe-GEN smell be.3SG but **girl** DEM
škap-ha peit-nu-ze. prihä se tul-i ka i sanub:
 closet-ILL hide-PST.PTCP-3SG **boy** DEM come-PST.3SG so and say-3SG
"reska-n henge-n duhh om silei." iga-ks žal'
 fresh-GEN breathe-GEN smell be.3SG 2SG.ALL all-TRANSL shame
liino-b, adiv se hivä ol-i. a božž se
 be.FUT-3SG **girl** DEM good be-PST.3SG but **godmother** DEM
i sanub: "ka tul-nu om."
 also say-3SG so come-PTCP.PST be.3SG

'This groom has also arrived. He sniffs and says: Godmother, there is a fresh smell. But the girl [bride] goes hiding in the closet. The boy [groom] came and said [to his godmother]: You really have a fresh smell. It would be a shame as the girl was a good person. But the godmother also says: She [the bride] has already arrived.' (Onegina & Zajceva 1996: 74–77)

This strategy observed in Finnic varieties in the east and North Russian is common in narratives where the narrator switches the from one protagonist to another. A similar use is also reported in from Komi (Leinonen 1998) and Forest Enets (Siegl 2013: 371–373), in which protagonists are marked with possessive suffixes.

The length of absence from the discourse (see Givón 1983: 11) does not seem to play a role. In the data, an antecedent can be found much further prior to the referent, with a distance over a hundred words in between the referent and its antecedent, as in (189) and (190).

Olonets Karelian

- (189) A daughter of the Tsar has been mentioned in the beginning of the story. The story runs for over intervening 100 words before the Tsar's daughter enters the scene again saying:

"vuata, sano-n jäll'es," tütär se sano-w,
 wait.IMP say-1SG after **daughter** DEM say-3SG
"vuata, sano-n jäll'es."
 wait.IMP say-1SG after

'Wait, I will tell you after – says **the daughter** – I will tell you later.'
 (Makarov & Rjagojev 1969: 45)

North Russian (Arkhangelsk)

- (190) A lady is telling about how they celebrated weddings in the village. The word ‘marriage’ was mentioned early in the discourse, but was idle for over 100 words before being mentioned again.

dak vot f svad’bu-tu bol’no krepko guljali,
so well in **marriage.ACC-DEM** painful hard walk.PST.PL
poplacēt, poplacēt,
cry.3SG cry.3SG
‘So, in the wedding, it was painfully tight for them to walk, she [the bride] cries and cries,
...’ (RNC)

In (189) and (190), ‘the daughter’ and ‘the wedding’ are reactivated to the discourse again, after having been idle. Nevertheless, the target of switching can also be new information that the speaker introduces to the discourse for the first time, such as the contexts with deictic, recognitional, and absolutely unique referents in (166) to (174).

Among the cases of switch-topic, a context of “highlighting parallels” in Krifka’s (2008) term (or “contrast marking” in Leinonen’s term 1998), in which several entities are set in contrast, is also observed in Heva Ingrian, eastern Finnic, and North Russian (see the similar uses of possessive suffixes in Section 5.3.2). This use resembles switch-topics, but there are always two or more referents in the immediate sequence as in (191) and (192).

Heva Ingrian

- (191) *metsä karhu-d ne on suure-d.*
forest.GEN bear-PL DEM be.3SG[/PL] big-PL
a kagra karhu-d ne on pikkarais-t. just nigu koir.
but oat.GEN bear-PL DEM be.3SG[/PL] small-3PL just like dog
‘Forest bears are big. But bears in the oat field are small. Just like a dog.’ (Laanest: 1966: 49)

North Russian (Vologda)

- (192) *tam dak, okopnját opjat’ na zăwstre-tu prigotovjat,*
there so dig_up.3PL again on **morning.LOC-DEM** cook.3PL
a vecerom-tu žè idut, my spat’ legëm.
but evening.INSTR-DEM PTCL go.3PL 1PL sleep.INF lay_down.1PL
‘There, they start digging [hay] again. In the morning, they make food, but in the evening,
they go [haymaking, while] we lay down to sleep.’ (Kasatkina 1991: 96)

In both (191) and (192), a contrast is created between two types of bears and two different times of day, respectively. This use of highlighting parallels is observed in several varieties, particularly in the east, but ultimately it may also be possible in the entire Finnic dialect continuum, given that it is used in the clause-second position.

A *beginning topic* is typically observed in the first clause of a narrative, such as in (193) and (194). For a North Russian example, see ‘the Victory Day’ in (173), which is the first word of the discourse.

Livonian

(193)	<i>brūni</i>	<i>se</i>	<i>um</i>	<i>seļļi</i>	<i>farb,</i>	<i>brūni</i>	<i>verm.</i>	<i>niemō-d</i>	<i>āt</i>
	brown	DEM	be.3SG	such	colour	brown	colour	cow-PL	be.3PL
	<i>brūnis</i>	<i>kāra-ks,</i>	<i>langō</i>	<i>painu-b</i>	<i>brūnizō-ks,</i>	<i>brūni</i>	<i>verm.</i>		
	brown	hair-TRANSL	thread	paint-3SG	brown- TRANSL	brown	colour		
	<i>brūni</i>	<i>kāra</i>	<i>ibīzō-n</i>	<i>pā-lō,</i>	<i>se</i>	<i>jumāl</i>	<i>um</i>	<i>tāmmō-n</i>	
	brown	hair	horse-DAT	up(on)-ADES	DEM	god	be.3SG	this-DAT	
	<i>andō-n</i>	<i>brūnis-t</i>	<i>arnō-d</i>	<i>ibīzō-n.</i>					
	give-PTCP.PST	brown-PL	cloth-PL	horse-DAT					
	‘Brown is such a colour. Cows have brown hair. Thread is painted in brown. Horses have brown hair. God has given brown clothes to it, to the horse.’ (Mägiste 2006: 126)								

South Estonian

(194)	<i>külä</i>	<i>sii</i>	<i>olʹ</i>	<i>yks</i>	<i>rikaśś</i>	<i>külä</i>			
	village	DEM	be.PST.3SG	one	rich	village			
	<i>joba</i>	<i>siin</i>	<i>olʹi-va,</i>	<i>joq,</i>	<i>pomba-kaavo-q,</i>	<i>jahh.</i>			
	already	here	be-EVT	well	pump_well-PL	yes			
	‘This village was a rich village where there were already pump wells.’ (EDC)								

The referent is new information being mentioned in the discourse for the first time and, therefore, has no alternative reading, in contrast to a switch-topic with a new referent, as seen in (166) to (174).

Table 37 summarises the distribution of topical functions in different varieties, as described with the examples above. Based on the results in Table 37, Chart 12 illustrates the proportions of topic continuing and switching functions across different varieties. Chart 12 shows that the division of labour does not differ across varieties in any significant way. However, Estonian and eastern Finnic varieties (Southern Lude and Veps), as well as North Russian do appear to use postposed demonstratives to mark topic switching more frequently. This implies that towards the east, in particular, postposed demonstratives have, in addition to continuing the flow of information, also become a device for gaining the addressee’s attention. The opposite, meanwhile, applies to central Finnic languages (Votic and Ingrian), in which postposed demonstratives generally serve as a device for carrying on the flow of discourse.

Topical functions		Continuing	Switching to old referent	Switching to new referent	Beginning	Total
Livonian		6	2	4	4	16
South Estonian		3	-	6	1	10
North Estonian		4	2	11	-	17
Votic	Western	9	4	-	-	13
	Eastern	1	1	-	-	2
Ingrian	Soikkola	3	1	1	-	5
	Heva	8	3	4	-	15
Karelian	Olonets	23	8	15	-	46
Lude	Northern	19	13	2	-	34
	Southern	22	14	12	-	48
Veps	Northern	26	32	14	-	72
	Central	63	70	53	1	187
	Southern	15	27	13	1	56
North Russian	Vologda	34	24	63	1	122
	Arkhangelsk	54	51	107	1	213

Table 37. *Topical functions across Finnic and North Russian varieties*

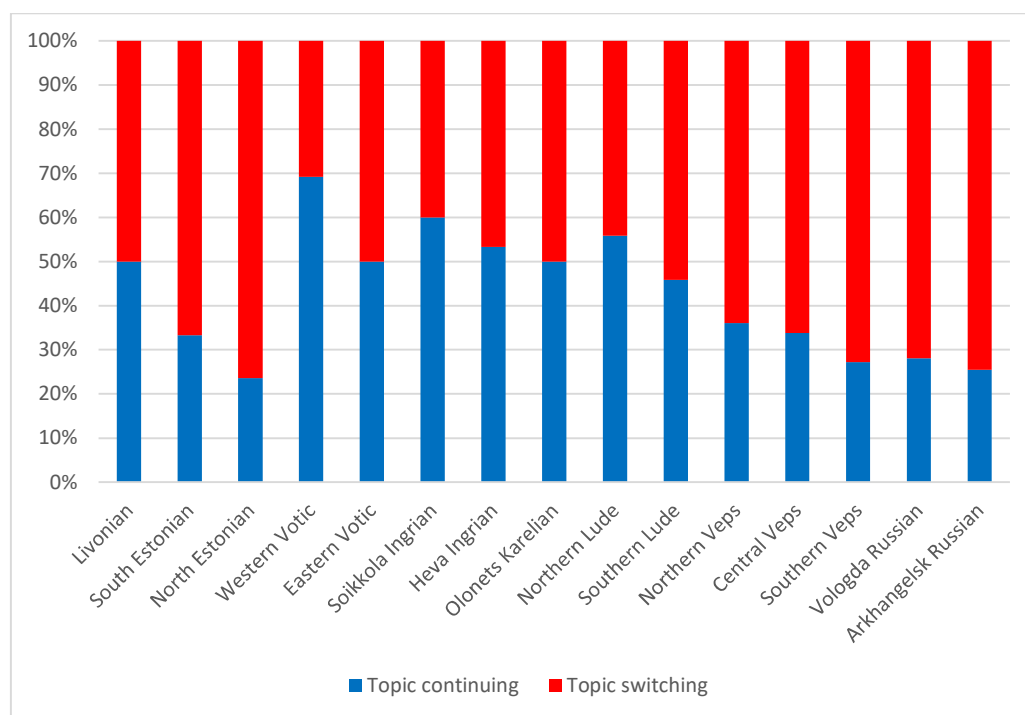


Chart 12 Proportions of topic continuing vs. topic switching

In addition to topic continuity, it is worth investigating whether given or discourse-old referents (discussed in Section 4.4.1) tend to co-occur with demonstratives more than demonstrative with new referents. Chart 13 illustrates the proportion of old and new referents in the topic position.

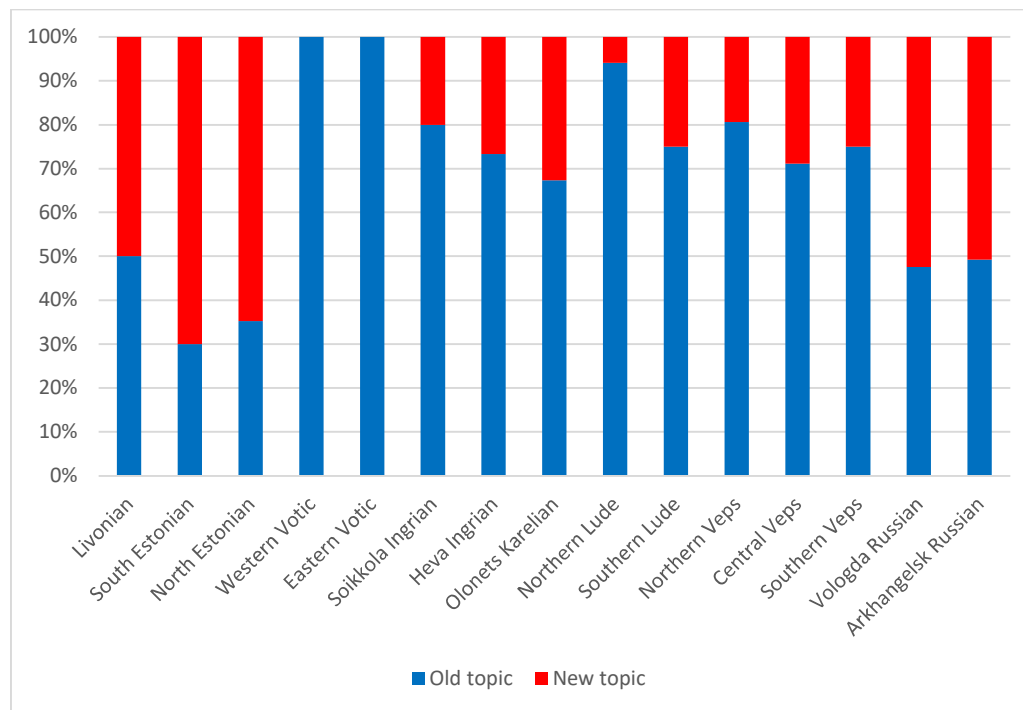


Chart 13 Proportions of old vs. new topics

Chart 13 shows that South Estonian and North Estonian more often use postposed demonstratives as a device for gaining the addressee's attention, when introducing a new topic. This is possibly due to the fact that the postposed demonstratives in South Estonian and North Estonian are very rare compared to other Finnic varieties (see Chart 4) and their use in such a marked construction of left dislocation is able to capture the addressee's attention more strongly (see also the discussion in Amon 2015: 126). At the same time, some other Finnic varieties predominantly use postposed demonstrative with old referents, regardless of whether this involves topic switching. Meanwhile, an almost equal proportion in North Russian implies that there is no predominant tendency of postposed demonstratives to more often mark new or old referents.

Based on the proportions in Chart 12 and Chart 13, we can claim that the primary information-structural functions of postposed demonstratives in Finnic varieties in the west (South Estonian and North Estonian) is for gaining the addressee's attention when switching or introducing a new topic, whereas they serve as a device for keeping the information flow in Finnic varieties of the central area (Votic and Ingrian), as well as Olonets Karelian and Northern Lude. Meanwhile, postposed demonstratives in eastern Finnic varieties (Southern Lude and Veps) along with North Russian are used

to gain the addressee's attention when switching topics (see a similar interpretation for standard Russian in Bolden 2008), most commonly between old referents in Finnic. There is no strong tendency of any particular function in Livonian, as the proportions of continuing and switch-topics, as well as new and old referents, are equal.

Thus far, the functions described above mainly concern demonstratives in clause-second position, which can also be alternatively viewed as a construction of left dislocation. Meanwhile, in clause-medial position, the speaker can also use postposed demonstratives when "listing" more than two things, as in (195) and (196).

Southern Lude

(195) An old man is telling that in Paloniemi, they always took water from the lake for various purposes.

... *lautuu-lpiä veži ot-tihe d'uo-da i, šüö-da i,*
 wharf-ABL water take-PASS.PST drink-INF and eat-INF and
st'itaa-da se, pes-ta se, kus-ta se, i kaik.
 wash(cloth)-INF DEM wash(body)-INF DEM pee-INF DEM and all
 '... from the wharf, we took water to drink, eat, wash clothes, wash up, to pee, and everything.
 (Pahomov 2011: 154)

North Russian (Vologda)

(196) *a kogdy malo wetb è bylo*
 and when little well this be.PST.NEU
tak mašyn-to tožo švejnyh-to, fsě rukami
 so machine.PL.GEN-DEM also sewing.PL.GEN-DEM all hand.PL.INSTR
šyli-to, fsě ètb-ot polotenca-tb, fsě rukami-te
 go.PST.PL-DEM all this-DEM towel.GEN-DEM all hand.PL.INSTR-DEM
 'And when there were only a few sewing machines, we did everything by hand, all these
 towels were made by hand.' (Kasatkina 1991: 97)

In (195) and (196), the speakers realise that they are telling too many things in series and do not want the addressee to miss any of the them, so they emphasise each lexical unit with a postposed demonstrative. Even though this use is only observed in Southern Lude, Central Veps, and Vologda Russian, it seems to be possible in all Veps and North Russian varieties.

Another information-structural strategy observed in the data is the "afterthought", or "right dislocation", that follows or falls in between a main clause and an additional phrase. This use concerns mainly demonstratives in non-second positions, i.e. clause-medial and clause-final. It can simply be "additional information" that the speaker wants to give as a clarification to the immediately preceding statement, as in (197) and (198). The preceding contexts in (197) and (198) are given that the speakers kept switching forth and back between two referents. At some point, they realise that the addressee might no longer be able to keep track of who the speakers are referring to. Therefore, the speakers clarify by inserting the subjects in the middle and at the end of the clauses, respectively.

Northern Lude

- (197) Once upon the time lived two brothers: one was rich and one was poor. The rich one had a lot of everything: good houses, good horses, a lot of cows, and the poor one only had a wife and children.
- | | | | | | | |
|--------------|---------------|-------------|-------------|------------------|--------------------|--------------|
| <i>mužik</i> | <i>elä-u,</i> | <i>köuh</i> | <i>mieš</i> | <i>še,</i> | <i>ei</i> | <i>tieda</i> |
| man | live-3SG | poor | man | DEM | NEG.3SG | know. CNG |
| <i>kus</i> | <i>suo-da</i> | <i>i</i> | <i>kui</i> | <i>lapš-ii-d</i> | <i>süö-ttä-dä.</i> | |
| where.INES | get-INF | and | how | child-PL-PTV | eat-CAUS-INF | |
- ‘A man, the poor man, does not know where to get [houses, good horses and cows] and how to feed his children.’ (Ojansuu et al. 1934: 149)

North Russian (Arkhangelsk)

- (198) A boy is going home, but his father has already gone to work in the office.
- | | | | | |
|----------|--------------------|--------------|----------------|-------------------|
| <i>i</i> | <i>vot</i> | <i>idët,</i> | <i>got,</i> | <i>igraët-to</i> |
| and | well | go.3SG | say | play.3SG-DEM |
| <i>v</i> | <i>garmošku-to</i> | <i>i</i> | <i>igraët,</i> | <i>paren'-ot.</i> |
| in | accordion.ACC-DEM | and | play.3SG | boy-DEM |
- ‘And well he goes, say, he keeps playing the accordion, that boy.’ (RNC)

This afterthought use more often reintroduces or elaborates on an old referent (anaphoric), or introduces a new referent related to the previously mentioned entity before the afterthought phrase (bridging and situationally unique referent).

In the afterthought use, the speakers can also use postposed demonstratives to mark a phrase on the left, which they then “repeat” as confirmation after recalling that the utterance is true, as in (199) and (200).

North Estonian

- (199) *ja noo maade-tiadus-t* *või me sellal üi-ti*
and well land-PL.GEN-science-PTV or 1PL then call-PASS.PST
keugrahvija ... keugrahvija sie ol'i siis kaa.
geography geography DEM be.PST.3SG thus so
‘And well, earth science or we called it then geography ... geography it was.’ (EDC)

Olonets Karelian

- (200) *jo nenga sova-t ol-lah: kowfta-t, fowftat net jongoi, ...*
already such cloth-PL be-3PL **blouse-PL blouse-PL DEM-PL** already
‘[At home,] there are already such clothes: blouses, the blouses, ...’
(Makarov & Rjagoev 1969: 208)

The repetition can also be a strategy to conclude the previous statement in order to carry on the story and, at the same time, may imply and entail a switch of discourse topic, such as switching from ‘the son going home’ to ‘the father’ in (201).

North Russian (Arkhangelsk)

- (201) ... *garmon'* *vzjal* *i* *pošël* *domoj.* *idët* *domoj-ta,*
 accordion take.PST.MASC and go.PST.MASC home go.3SG home-DEM
a *v* *kontore,* *otec-ot* *v* *kontore* *už* *robotuet,* ...
 but in office.LOC father-DEM in office.LOC already work.3SG
 ‘... he [the son] took the accordion and went home. He goes home, the father has already gone to work in the office, ...’ (RNC)

Postposed demonstratives in (199) to (201), from the referential aspect, can be simultaneously regarded as co-occurring with anaphoric referents.

In addition to the contexts above, the repetitive use of postposed demonstratives also appears in a pattern of double prepositional phrase observed in North Russian. In this construction, members of the same constituent are separated by an extra preposition, as in (202) and (203).

North Russian (Vologda)

- (202) *vyššli* *vot* *sb* *ëtovo* *z* *bolvta-ta* *tudy,* *i* *načevali.*
 exit.PST.PL well from this.GEN from swamp.GEN-DEM there and camp.PST.PL
 ‘We, well, came out from this swamp to there, and did camping.’ (Kasatkina 1991: 93)

North Russian (Arkhangelsk)

- (203) *vrode* *govorili,* *čto* *narod* *plohoj* *byl* *zdes'*
 sort_of say.PST.PL that people bad be.PST.MASC here
v *derevnjah-to* *v* *ètih.*
 in village.PL.LOC-DEM in this.PL.LOC
 ‘They sort of say that people were really bad in these villages.’ (RNC)

In (202), the second component of the noun phrase ‘swamp’ is followed by a postposed demonstrative, and *vice versa* in (203). On the one hand, these patterns can be regarded as marking of a repetition of the same noun phrase but, on the other hand, also as “repairing” expressions. The repairing use is observed when a speaker of Finnic first pronounces a word in Russian, but then corrects it to their native language, marking the repaired part with a postposed demonstrative, as in (204).

Central Veps

- (204) *pajoiž-idenke kävel-tas uliča-dme, kujo-dme se, kaike-n gulei-ba.*
 song-PL.COM walk-3PL street-PROL **street-PROL** DEM all-GEN walk-3PL
 ‘While singing, they walk along the street [Russian], the street [Veps], walk all the way.’
 (Gerd et al. 2002: 70–72)

Otherwise, the repairing expression can be a correction of something that the speaker has previously said incorrectly but realises afterwards, e.g., the correction of preposition choice from *ot* ‘from’ to *na* ‘on’ in (205).

North Russian (Arkhangelsk)

- (205) *iš’ vot eto žicja, žic’ka, èto ot motuški,*
 PTCL well this life life this **from** **thread.GEN**
žicja-ta na motuški da.
life-DEM on thread.LOC yes
 ‘You see, well, this is life, a life, it is from the thread, [I meant] the life on the thread, right.’ (RNC)

In both (204) and (205), the postposed demonstratives seem to highlight that the information in the afterthought phrase is not part of the main statement.

In terms of geographical distribution, the afterthought use is observed primarily in the eastern end of the Finnic dialect continuum, Lude and Veps, as well as in North Russian. It also sporadically appears in other Finnic varieties in the west, more likely to appear in truncated speech such, as in the following North Estonian example (206).

North Estonian

- (206) An old lady is telling about her life. After explaining many things, she concludes the story with a *c’est la vie* type of ending phrase.
 ... *sis nii kã-i-b se elamine se ...*
 so like_this go-PST-3SG DEM living **DEM**
 ‘... well, that was my story that was ...’ (EDC)

In (206), many events had been mentioned previously and, in the end, ‘the living’ summarises the story. However, the phrase ends with another demonstrative that marks the interruption of discourse, and that the speaker is running out of things to say and willing to hand the speech turn to the addressee.

For certain cases, however, explanations based on referentiality and information structure might not be satisfactory, so a missing link could be obtained from a discussion of evaluative uses of postposed demonstratives, discussed next.

6.4.3 EVALUATIVE USES

Beyond referentiality and information structure, postposed demonstratives can also be used to convey intersubjectivity and interaction between the interlocutors, such as by expressing subjective attitude of the speaker towards a mentioned state-of-affairs, or the speaker's intersubjective evaluation of the addressee's state of knowledge, in a similar mechanism to the "stance triangle" in Du Bois (2007), discussed in Section 4.5.1. As these evaluative uses greatly depend on the interpretation and context, and the prosodic dimension of the data also need to be explored. This preliminary investigation is restricted to a qualitative level, and explores possible contexts for different evaluative uses, leaving the possibility to perform a quantitative analysis for future studies. In this section, my interpretation of the implicature of utterances is quoted as bold text in parentheses, which accompanies the translation of language examples.

Concretely, a speaker can use postposed demonstratives to express their stance towards a state-of-affairs, which can be based on affection, judgement, or appreciation (White 2015). From the speaker's perspective, they can express their own stance, which can emerge from the speaker's subjective evaluation based on "affection" towards a state-of-affairs, as in (207) to (209).

Olonets Karelian

- (207) An old man is telling how hard working in the forest was in the old days. He concludes with the following statement.

ol'i jugei ruado se.
be.PST.3SG hard work DEM

'The work was hard (**I was really suffering**)!' (Makarov & Rjagoev 1969: 174)

Southern Veps

- (208) *el'-ii-maa ühtes pe'rti-š toš-tmu velje-mu i toraž-ii-maa*
live-PST-1PL together house-INES other-COM brother-COM and quarrel-PST-1PL
go'lu. mužik se mi-n sa-i, sen i jo-i.
much man DEM what-ACC get-PST.3SG it-ACC also drink-PST.3SG

'We were living together with my other brother, and we fought a lot. That jerk (**as far as I saw**), what he got, he drank it all.' (Kettunen & Siro 1935: 5)

North Russian (Vologda)

- (209) *a u nas otec-to vo vremja vojny ... ved' umer-to.*
but at 1PL father-DEM in time war.GEN indeed die.PST.MASC-DEM
'But our father during the war time ... indeed he died (**that made us very sad**).' (RNC)

In (207), the speaker expresses his frustration and suffering from a hard work in the forest. In (208), the speaker is annoyed by his brother, and even emphasises his irritation with a postposed demonstrative meanwhile addressing his brother as 'jerk.'

In (209), the speaker emphasises her sorrow by using a postposed demonstrative to mark the phrase where she is mourning the death of her father during the war.

Moreover, affection can also emerge from the speaker's "appreciation", typically through exclamation, such as the admiration of Saint Petersburg in (210), or through nostalgia in (211).

Central Veps

- (210) *oi om Piteri-š se ka čoma.*
 oh be.3SG Saint_Petersburg-INES DEM so beautiful
 'Oh, how beautiful Saint Petersburg is.' (Zajceva & Mullen 1969: 25–27)

North Russian (Arkhangelsk)

- (211) *kak horošo-to bylo ran'sè, ooj!*
 how good-DEM be. PST.NEU earlier oh
 'How good it was in the old days, oh (**I miss those days**)!' (RNC)

In (210), the speaker emphasises his exclamation about the beauty of Saint Petersburg he has witnessed himself. In (211), the speaker expresses a nostalgic feeling towards life in the old days by emphasising 'good' with a postposed demonstrative. In addition to affection, a speaker can also express their "surprise" towards a state-of-affairs as in (212) to (214).

Olonets Karelian

- (212) A story about three brothers who have departed from family and live their own separate life. One of the three brothers, Iivan Tuhkimus, is the one with the most fortune. Comparing individual lives of the three brothers, the narrator summarises the story:
ga tänäpäi-gi el'ä-w: mučoi-n nai da las-tu sa-i,
 so today-also live-3SG wife-GEN marry.PST.3SG and child-PTV get-PST.3SG
da tuhku-iivan se on bestrašiimoi
 and Tuhkimus-Iivan DEM be.3SG without_effort
 'So even today, he is alive: he married a woman, got a child, and that Iivan Tuhkimus (**I really wonder how on earth**) without effort, has also become rich!' (Makarov & Rjagojev 1969: 177)

Northern Lude

- (213) A conversation about etiquettes in the party. Ladies and gentlemen normally dance for a while before starting to kiss. Two gentlemen are discussing how to kiss the ladies.
 A: *ka mihi že čumoit-ettau?*
 well where.ILL PTCL kiss-PASS
 B: *ka käde-i.*
 well hand-ILL
 'Well, where do people kiss [ladies]? – Well, on the hand.'

A: *no ka miŋk kădei neiču-t čumoite-ttau?*
 but well why hand-ILL lady-PTV kiss-PASS
v'ed ei kăde-i se pidä-u huule-i čumoita-da se,
 well NEG.3SG hand-ILL DEM have_to-3SG lip-ILL kiss-INF DEM
mi'e om vailiše miärä-ttū se čumoitus se.
 how be.3SG only order-PTCP.PASS DEM kissing DEM
 'But, well, why do people kiss ladies on the hand? It should not be the hand, but rather
 kissing on the lips. How can kissing be defined only like that (**I am surprise to hear**)?'
 (Ojansuu et al. 1934: 294)

North Russian (Arkhangelsk) [identical to (2) and (164)]

(214) A villager is telling about her life, living with her husband, father-in-law, and two cows.
 The mother-in-law had passed away long ago. As the father-in-law is already old, they
 were not expecting him to remarry, until he actually did.

a eščě svėkor-ot-to eščě ženilsja na drugoj.
 but still father_in_law-DEM-DEM still marry.PST.MASC on other.LOC
 'But still, the father-in-law (**against all our expectations**) got married again.' (RNC)

In the three examples above, the speakers express their surprise towards unexpected information they have heard or experienced: the effortless good life of Iivan Tuhkimus in (212), a humble way people kiss in a party in (213), and a father-in-law's remarriage despite his old age in (214). The use of postposed demonstratives in these examples can be regarded as marking "inconsistency" between what the speaker's presupposition and what they actually experienced, in König's term (1991), or "mirativity" in DeLancey's term (1997, 2001). As the three contexts discussed above relate to the speaker's stance, the speaker is the one who takes the epistemic authority, subjectively without consideration of the addressee's stance.

Apart from the speaker's stance, a speaker can use postposed demonstratives in the intersubjective evaluation, considering also the addressee's attention and state of knowledge, such as highlighting "shared knowledge" between the speaker and the addressee. This use can be observed in a context where a speaker expresses their confidence in (215), concern in (216), or hedging in (217) in a situation, the outcome of which is also evident to the addressee.

In (215), the narrator took the role of the old man who emphasises his boast, while the rich man in (216) expresses his concern and tries to help the poor man to overcome his poverty problem. Meanwhile, the speaker in (217) does not want to humiliate the interviewer about being unfamiliar with the church practice, but formulates the information as a thing that the interviewer should nonetheless know. These cases may also be regarded as a speaker's face-saving strategy (see Fraser 1990, discussed in Section 4.5.1), corresponding to the evaluation type, in which the speaker maintains interactional relations with the addressee by establishing affiliation and inclusiveness of stancetaking (Hunston & Thompson 2000, discussed in Section 4.5.1).

Livonian

- (215) A story about a boy who left his house and went searching for his fortune: On his journey, he met an old man who promised to help the boy only if he would bring the old man to a sauna. After going to the sauna, the old man disappeared. After a long adventure, the boy found his fortune near the lake, a daughter of the king with who he later married. He was living happily in the palace doing no work, simply walking in the garden with the king's daughter all day long. One day he visited home, and everyone was very delighted about his fortune. The boy told his father: 'If you have a hard time living with these two brothers, go traveling with me.' His father replied: 'Thank you, son, I will just stay in this village, go with your fortune and live a happy life.' This was all thanks to the old man (who possibly had magical power). As an ending phrase to the story, the narrator quotes the speech of the old man who was there behind the scene:

sin vōŋ se lāb īŋō
 2SG.POSS fortune DEM go.3SG along

'Your fortune (**as I have told you earlier**) will go with you.' (Mägiste 2006: 191)

Olonets Karelian

- (216) A rich man is telling a poor man what he should do to overcome his poverty.

omu-a se pidä-š ruadu-a.
 own-PTV DEM must-COND work-PTV

'Your own work, you have to do it (**you surely know that**)!' (Makarov & Rjagoev 1969: 119)

North Russian (Arkhangelsk)

- (217) A villager is interviewed about how they perform baptism in their village's church.

A: *kak krestjat?*
 how baptise.3PL

'How do they baptise?'

B: *nu vot tak kak krestik-ot.*
 but well so how cross-DEM

'Well, with the cross (**you must have seen it before**).' (RNC)

When the information does not directly come from the speaker's private knowledge, nor from the addressee's, a speaker can create joint attention anew with the help of postposed demonstratives. In this context, the speaker intersubjectively evaluates whether the information to be uttered exceeds or conflicts with the addressee's expectation, i.e. "counter-expectation", as in (218) to (220).

Northern Veps

- (218) A baby was growing up till the age that she may start talking at any moment. That day has arrived, but the parents were not home and went out to the forest. The baby woke up and started crying so loud that even the goat nearby could hear her.

a kozeine se i kuul-išt'.

and goat DEM also hear-PST.3SG

‘And the goat even heard it (**she cried louder than you would think**).’

(Onegina & Zajceva 1996: 27–31)

Central Veps

- (219) *elo-tihe endo čelj kanz: tat da mam, oli*
live-PST.3PL earlier whole family father and mother be-PST.3SG

‘Once upon the time, there was a family: father and mother,’

hijou kuume poiga-d. no tat da mam se kou-dihe,
3PL.ADES three boy-PTV but father and mother DEM die-PST.3PL

‘They had three sons. But the father and mother passed away (**you might not expect that as the story has just begun**).’

a gä-i hiile kuume härgä-d.

and remain-PST.3SG 3PL.ALL three ox-PTV

‘And they still had three oxen.’ (Onegina & Zajceva 1996: 178–182)

North Russian (Vologda)

- (220) A lady was asked why she often comes to Karačunovo. The person asking may think she likes this place, but the truth is that she just goes there for grains.

da sjuda idti, v Karačunovo, замуž? da i dorogi-to net!

and here go.INF to Karačunovo married and also way.GEN-DEM no

my pryezžali sjuda za zernom byvalo.

1PL come.PST.PL here for grain.INSTR happen.PST.NEU

‘To come here to Karačunovo to get married? No way (**it is not what you think**)!

We used to come here just for grains.’ (RNC)

In (218) to (220), the proposed stance is promoted from the speaker’s evaluation, according to which the addressee would not expect such things to happen, so the speaker emphasises such pieces of information with postposed demonstratives.

As for joint attention, a stance can emerge from the speaker’s “hesitation” over a situation in the discourse, for which they seek confirmation from the addressee, as in (221) and (222). In (221) and (222), the speakers express their uncertainty, and indirectly seek confirmation, in case the addressee knows the answer to their doubt. The two examples can be considered as corresponding to “the interjective hesitation use” described by Hayashi and Yoon (2006: 507–513, presented in Section 4.5.2).

Southern Lude

- (221) An old gentleman was asked how many families were living in Joensuu earlier. He cannot remember exactly but tries to think about it.

ka eñ-t'ää mii ol'ii, ka ol'ii d'o.
 so NEG.1SG-know.CNG how_many be.PST.3SG so be.PST.3SG already
koumeküment se ol'ii naerno kod'i-d.
 thirty DEM be.PST.3SG probably house-PTV
 'So, I cannot remember, but there were about thirty households probably.'
 (Pahomov 2011: 200)

North Russian (Vologda)

- (222) A lady is telling about the process of weaving a rope, but she uses an old Turkish measuring unit, *arşın*, about which she is not sure what the correspondence in metric system is.

nè znaju mētrom-to uš on velik li dolok li budit.
 NEG know.3SG metre.INSTR-DEM really 3SG high Q long Q be.FUT.3SG
 'I do not know whether it [one arşın] will really be a metre high or long.' (Kasatkina 1991: 89)

In a similar sense, the speaker may create joint attention by presenting a “warning” in (223) and (224), or a “reminder” in (225) to the addressee about the situation, for which the speaker seeks confirmation to arrive at a shared stance.

Central Veps

- (223) *ala g'o neci-da vet. huba vezi se.*
 NEG.IMP drink-CNG this-PTV water.PTV bad water DEM
 'Do not drink this water. It is bad water.' (Zajceva & Mullonen 1969: 280–284)

Southern Veps

- (224) *Kost'a, nāhtaše-k, kondi se jo taganaa!*
 Kost'a look-IMP.2SG bear DEM already behind
 'Kost'a, look, the bear is already behind (you)!' (Zajceva & Mullonen 1969: 222–223)

North Russian (Arkhangelsk)

- (225) A villager is complaining that younger people nowadays just drink too much alcohol and go clubbing. It was better in the old days when there was still no club.

a rane-to it', klubu-tu i ne bylo, vsë na ulici.
 and earlier-DEM well club.GEN-DEM also NEG be.PST.NEU all on street.LOC
 'And in the old days, well, there was no club either, so everything [the entertaining activity] was happening on the street.' (RNC)

In (223), the speaker knows that ‘water’ is bad and, therefore, warns the addressee not to drink it, while in (224), the speaker urgently needs to establish joint attention, so

the addressee will not get attacked by a bear. In (225), the speaker reminds the addressee about the old lifestyle without alcohol and clubs (see also the encoding of warning and reminder with clause-second demonstratives in Finnish in Yurayong & Kittilä, forthcoming). In case of reminders, this evaluative use can also be compared to the recognitional use from the referential domain (an idea proposed in Hayashi & Yoon 2006), as the information does not come from the discourse but rather from the addressee's personal memory. In intersubjective contexts with the addressee's stance and joint stance, the speaker hands epistemic authority to the addressee to react and confirm the uttered statement.

Based on the observation from the data, Table 38 illustrates a qualitative survey on what kind of evaluative uses of postposed demonstratives are possible in each Finnic and North Russian variety. Given that the evaluative uses are crucially dependent on the text genres, and highly unequally distributed among the Finnic varieties, the distinction is only made on a binary scale: – (not observed) vs. + (observed). This means that the productivity with a more accurate rate of occurrence, is not considered, but rather only a single occurrence is regarded as + (observed).

Language	Variety	Subjective evaluation		Intersubjective evaluation			
		Speaker-oriented functions		Addressee-oriented functions	Creating joint attention		
		Affection	Surprise/ Admiration		Shared knowledge	Counter-expectation	Hesitation
Livonian		–	–	+	–	–	–
South Estonian		–	–	–	–	+	–
North Estonian		–	–	–	–	+	–
Votic	Western	–	–	–	–	–	–
	Eastern	–	–	–	–	–	–
Ingrian	Soikkola	–	–	–	–	–	–
	Heva	–	–	–	–	–	–
Karelian	Olonets	+	+	+	–	–	–
Lude	Northern	–	+	–	–	–	–
	Southern	+	+	+	–	+	+
Veps	Northern	–	+	+	+	+	+
	Central	+	+	+	+	+	+
	Southern	+	+	–	+	+	+
North Russian	Vologda	+	+	+	+	+	+
	Arkhangelsk	+	+	+	+	+	+

Table 38. *Evaluative uses of demonstratives observed in the Finnic and North Russian data*

As is clear from Table 38, evaluative uses of postposed demonstratives are very rare in Finnic varieties of the western and central areas. They are, meanwhile, more visible

in eastern Finnic and North Russian, in which the use of postposed demonstratives is morphosyntactically and pragmatically more flexible in terms of referential and information-structural uses, as has been discussed throughout this chapter so far.

Nevertheless, there is an issue of genre, because such extensive pragmatic functions are expected from dialogues, available in the current study mostly for Southern Lude (see Table 1 in Section 1.2). From the perspective of engagement, (Landaburu 2007; Evans et al. 2018a, 2018b, discussed in Section 4.5.2), it is unclear whether the postposed demonstratives, especially in those varieties in the east, primarily encode speaker-asymmetry (speaker's attention) or speaker-symmetry (addressee's attention), as contexts of both types occur in the data. This suggests that the interactional aspect of language use deserves more attention from the fields of Finnic and North Russian dialectology, and a systematic investigation missing from the majority of previous studies (discussed in Chapter 5) is definitely needed.

6.5 OVERVIEW ON THE OLD NOVGOROD DATA

Old Russian demonstratives *sb* (proximal) and *tb* (distal) occur attributively in 43 birch bark documents, with a total of 49 occurrences. Concerning constituent order, the demonstratives can be both prenominal as in (226) and (228), and postnominal as in (227) and (229).

Old Novgorod

- (226) *a pro sei čeloveko my jeho ne znajemo.*
 but about DEM.MASC man 1PL 3SG.ACC NEG know.1PL
 'But about this man, we do not know anything.' (Pskov 6 1260–1280, Zaliznjak 2004: 515)

- (227) *i krstb sb postavhъ.*
 and cross DEM.MASC put.AOR.1SG
 'And this cross, I erected it' (Sterž Cross 1133, Zaliznjak 2004: 344)

- (228) *u togo žb Žeběę Nosa prijevavšę ...*
 at DEM.MASC.GEN PTCL Žab'.GEN Nos.GEN come.IMP.3SG
 'From the very same Žab' Nos came ...' (document № 249 1380–1400, Zaliznjak 2004: 623)

- (229) *ceto esi prisale dova celoveka te pobegli*
 REL be.2SG send.PST.MASC two man.DU DEM.MASC.DU run_away.PST.PL
a kone ne vedaju gdje poimavoši.
 but horse NEG know.1SG where take.IMPF
 'Those two men whom you sent have run away; as for the horse, I do not know where they have taken to.' (document № 582 1280–1300, Zaliznjak 2004: 514)

In the majority of the cases, the demonstratives modify nouns, as in (226), (227) and (229), while cases of attributive use with a name are also observed as in (228). The use of demonstratives with non-nominal parts of speech, as mentioned previously for modern North Russian dialects, is completely absent from the Old Novgorod data. As for their morphosyntactic behaviour, the demonstratives always agree with their head nouns in case, number, and gender.

Table 39 summarises the occurrences of demonstratives *sb* and *tb* in the Novgorod birch bark documents according to their constituent orders.

Period (years)	Preposed		Postposed		Total	Number of documents with the occurrence of attributive demonstrative	Number of birch bark documents found in the period
	<i>sb</i>	<i>tb</i>	<i>sb</i>	<i>tb</i>			
1025–1050	1	-	-	1	2	1	4
1050–1075	-	-	-	-	-	-	12
1075–1100	1	-	-	-	1	1	15
1100–1120	-	3	-	2	5	3	29
1120–1140	-	-	-	1	1	1	43
1140–1160	-	-	-	-	-	-	74
1160–1180	-	3	-	1	4	4	105
1180–1200	-	-	-	1	1	1	117
1200–1220	-	4	-	-	4	3	38
1220–1240	-	3	-	-	3	2	25
1240–1260	-	1	-	-	1	1	49
1260–1280	-	-	1	-	1	1	22
1280–1300	-	-	-	1	1	1	47
1300–1320	-	-	-	-	-	-	32
1320–1340	-	-	-	-	-	-	38
1340–1360	1	2	1	2	6	5	34
1360–1380	-	4	-	1	5	5	96
1380–1400	-	5	1	-	6	6	60
1400–1410	1	1	-	1	3	3	28
1410–1420	-	1	-	1	2	2	25
1420–1430	-	2	-	-	2	2	19
1430–1450	-	1	-	-	1	1	9
1450–1500	-	-	-	-	-	-	1
Total	4	30	3	12	49	43	922

Table 39. Occurrences of adnominal demonstrative in Novgorod Birch Bark Documents

Based on the results in Table 39, Chart 14 shows the frequency of both preposed and postposed demonstratives *sb* and *tb* over different periods of time. Chart 14 shows no significant diachronic change in frequency. One noteworthy remark here is that postposed demonstratives are, in general, less frequent than preposed demonstratives, and they show no sign of becoming more frequent during these five centuries.

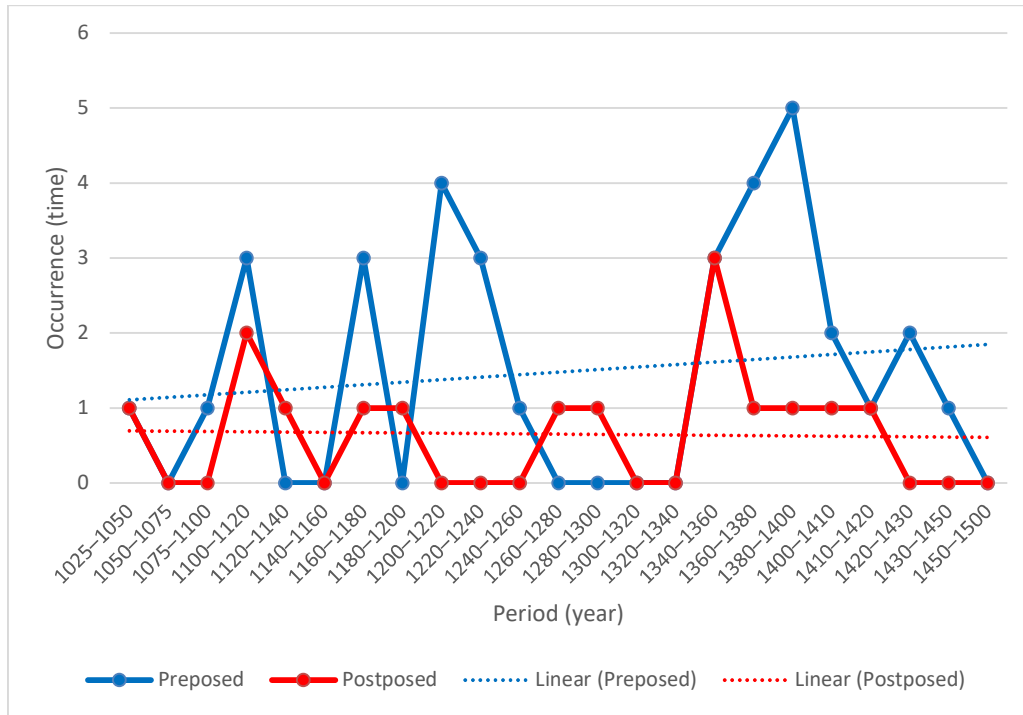


Chart 14 Occurrences of adnominal demonstratives in Novgorod birch bark documents

This is strong evidence supporting the view that the predominant use of demonstratives in a head-following position, as observed in modern North Russian dialects (shown in Section 6.1.5), has no link to the Old Novgorod period.

Out of 922 birch bark documents that belong to the data of the present study, we find the use of the clause particle *ti* in 74 documents (8.06%), the origin of which may be connected to the Proto-Slavic medial demonstrative root **t-* (Zaliznjak 2008: 32). In many documents, *ti* occurs multiple times in a clause, although its regular position is clause-second, i.e. Wackernagel's clitic. Unlike the Old Church Slavonic equivalent particle *ti* 'and' (Leskien [1871]1955: 115), this Old Novgorod particle is usually translated as 'then, in that case' (see Dekker 2018: 54), as its primary function is to topicalise a preceding lexical unit, which can a variety of different be parts of speech, including pronoun in (230) and verb in (231). Ultimately, Zaliznjak (2004: 196) claims that *ti* can follow any type of preceding word except imperative verb. Interestingly, this resembles the function of topic marking in modern North Russian as well as Finnic languages in general (discussed in Chapter 5 and earlier in this chapter).

Old Novgorod

(230) 'Gather your money and food.'

jazъ ti pridu po velikę dъni.
 1SG TI come.1SG after great.LOC day.LOC

'I will come after Easter.' (document № 380, 1140–1160, Zaliznjak 2004: 301)

(231) ‘Greetings from Danil to brother Ignat.’

bratb popečelišę o mne.

brother.VOC take_care.IMP about 1SG.LOC

hožju ti nago ni mętle ni inogo cego.

go.1SG TI naked neither cloak.GEN nor other.GEN what.GEN

‘Brother, take care of me. I have been going around as if I was naked, being without any cloak or anything else to put on. Send me a brown-red cloak and I will give you money.’

(document № 765, 1240–1260, Zaliznjak 2004: 480)

Table 40 shows the frequency of use for each period of time, as well as with the part of speech of the host words. The last column marked with “?” stands for such cases in which the interpretation is not possible, due to truncation or unreadability of the text on the birch bark.

Period (years)	Preceding word											Total	Documents with the occurrence of <i>ti</i> / Total documents found in the period	Chance of occurrence
	Noun	Pronoun	Adjective	Name	Verb	Adverb	Interrogative	Q.PTCL	RELPRON	CONJ	?			
1025–1050	-	-	-	-	1	-	-	-	-	-	-	1	1 / 4	25.00%
1050–1075	-	-	-	-	-	-	-	-	-	2	-	2	2 / 12	16.67%
1075–1100	-	-	-	-	-	-	-	-	-	-	-	-	- / 15	0.00%
1100–1120	1	3	1	-	1	-	1	-	-	3	-	10	5 / 29	17.24%
1120–1140	-	-	-	-	-	-	1	-	-	-	-	1	1 / 43	2.33%
1140–1160	1	1	-	-	2	-	1	-	-	4	-	9	9 / 74	12.16%
1160–1180	5	2	1	1	6	1	1	1	2	12	1	33	20 / 105	19.05%
1180–1200	2	-	3	1	4	1	9	1	-	5	-	26	17 / 117	14.53%
1200–1220	-	2	-	-	1	-	3	-	-	4	1	11	7 / 38	18.42%
1220–1240	-	-	-	-	-	1	-	-	-	-	-	1	1 / 25	4.00%
1240–1260	1	-	-	-	1	-	-	-	-	3	-	5	4 / 49	8.16%
1260–1280	1	-	-	-	-	-	-	-	-	1	1	3	2 / 22	9.09%
1280–1300	-	1	-	-	1	1	-	-	-	-	-	3	1 / 47	2.13%
1300–1320	-	-	-	-	-	-	-	-	-	-	-	-	- / 32	0.00%
1320–1340	-	-	-	-	-	-	-	-	-	-	-	-	- / 38	0.00%
1340–1360	-	-	-	-	-	-	-	-	-	1	-	1	1 / 34	2.94%
1360–1380	-	1	-	-	-	-	-	-	-	1	1	3	2 / 96	2.08%
1380–1400	-	-	-	-	-	-	-	-	-	1	-	1	1 / 60	1.67%
1400–1410	-	-	-	-	-	-	-	-	-	-	-	-	- / 28	0.00%
1410–1420	-	-	-	-	-	-	-	-	-	-	-	-	- / 25	0.00%
1420–1430	-	-	-	-	-	-	-	-	-	1	-	1	1 / 19	5.26%
1430–1450	-	-	-	-	-	-	-	-	-	-	-	-	- / 9	0.00%
1450–1500	-	-	-	-	-	-	-	-	-	-	-	-	- / 1	0.00%
Total	11	10	5	2	17	4	16	2	2	38	4	111	75 / 922	8.13%

Table 40. Occurrences of the clause particle *ti* in Novgorod birch bark documents

The results from Table 40 are converted to a timeline seen in Chart 15.

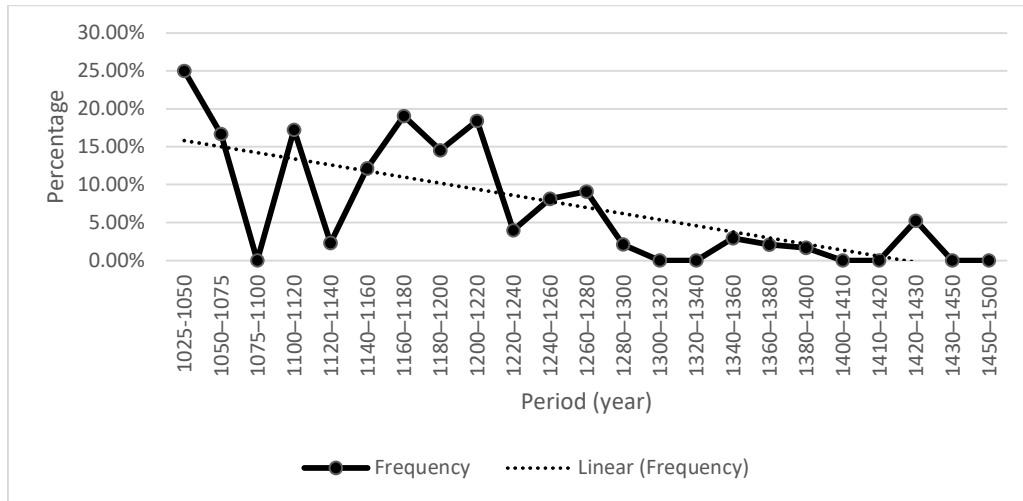


Chart 15 Occurrences of the clause particle *ti* in Novgorod birch bark documents

Chart 15 shows the striking fact that the use of clause particle *ti* was relatively frequent between the 11th and 13th centuries, after which the frequency radically drops, disappearing in the 15th century (as also pointed out earlier by Zaliznjak 2004: 197). For further studies, it would be interesting to investigate whether the decreasing frequency of the particle *ti* was due to any of the following factors. First, it could be related to replacement of *ti* by other clause-second particles, such as *že*. Second, it could be due to a change in syntax which led to the obsolescence of clause-second clitics, i.e. Wackernagel's clitics, in Middle Russian. Ultimately, it could be influenced by the development of literary standard in which discourse particles become less frequent.

To summarise the contribution of the Old Novgorod dialect, it provides historical evidence in favour and against several phenomena which occur in modern North Russian dialects. First, it shows that there was no default constituent order for demonstratives, as both preposed and postposed demonstratives are used, even though this could ultimately be controlled by information structure (cf. the Old Church Slavonic examples in Kurz 1939–1946; Večerka 1993, discussed in Section 5.2.1). Second, it was not possible to use a demonstrative with non-nominal parts of speech, unless the particle *ti* would ultimately be an element etymologically derived from a demonstrative. Third, the tendency to use a clause-second marker, in Wackernagel's sense (1892), was already present in Russian long ago, considering that Wackernagel reconstructs such syntax even to the Proto-Indo-European stage (see the discussion of demonstratives and Wackernagel's law in Slavic languages in Section 5.2.2). Nevertheless, the clause-second use of the particle *ti* might also have a literal function, facilitating reading by marking the beginning of a clause in the same manner as the particle *že* in Old Church Slavonic. This is true particularly given that written sources of these Ancient Slavic languages do not always put or mark interval between words, as illustrated in the Old Novgorod example with *ти* (*ti*) in Figure 24, and the Old Church Slavonic example with *ѣѣ* (*že*) in Figure 25.



Figure 24 Novgorod birch bark document №731 (ca 1160–1180)
(Russian Academy of Sciences:
gramoty.ru/birchbark/document/show/novgorod/731/)
Red = clause boundary, Blue = clause-second particle *ti*

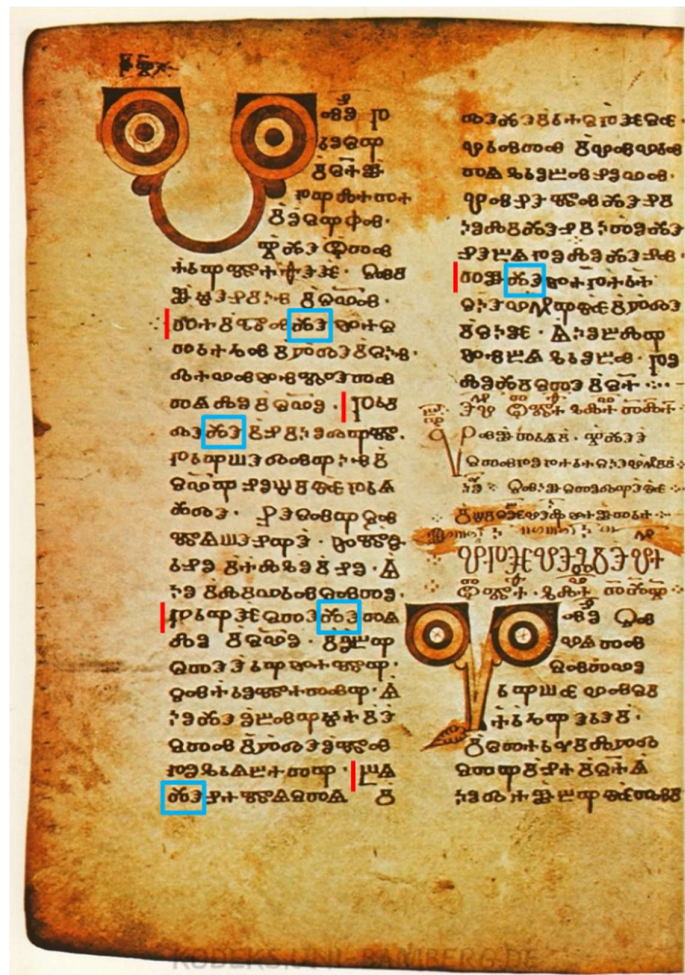


Figure 25 Codex Assemanianus (the early-11th century)
(kodeks.uni-bamberg.de/AKSL/media/Assemanianus/Assem07.jpg)
Red = clause boundary, Blue = clause-second particle *že*

In any case, the use of *ti* is generally not used in religious texts of the Novgorod birch bark document corpus, which suggests that this type of Wackernagel's particle was more common in the spoken language register, considering the colloquial nature of the Old Novgorod data (discussed in Section 2.1.3), unlike the literary standard, which does not favour the use of such clause particles.

In the end, if any association is to be made between Old Novgorod and modern North Russian dialects, it is the topical use of the particle *ti*, rather than the adnominal demonstratives *sb* and *tb*.

6.6 CONCLUSIONS OF THE DATA ANALYSIS

This chapter has shown many correlations across different categories. Such correlations can be cross-categorical interaction within the grammar, such as the effect of subjecthood in topicalisation, which equally applies across the dialect continuum. At the same time, variation among varieties is related to the geographical location of speaking areas, which further determines the usage of postposed demonstratives in each variety in an unequal but areally consistent degree, such as tendencies of head-following constituent order, non-second clausal positions, and the frequency of non-nominal hosts, which increases as one moves towards the east.

In all varieties under investigation, postposed demonstratives are primarily used for topic marking, but they also have other additional referential, focal, and evaluative functions, seen in Section 6.4. A number of factors that favour the use of postposed over preposed demonstratives have been identified for those varieties in which the use of postposed demonstratives is more frequent, as given below.

1. The use of preposed demonstratives is restricted to nominal parts of speech (noun, pronoun, name, adjective), whereas postposed demonstratives also co-occur with non-nominal parts of speech (verb, adverb, adposition).
2. Preposed demonstratives can only modify their following constituent noun phrase, whereas postposed demonstratives can mark an entire verb phrase.
3. Preposed demonstratives more often attach to noun phrases which mark adverbial phrases such as time and location, whereas postposed demonstratives co-occur with noun phrases which more often mark nuclear syntactic functions (subject, direct object, indirect object).
4. Deictic use is more frequently found in the use of preposed demonstratives, whereas anaphoric and other referential uses are more common for postposed demonstratives, suggesting a decreasing functional load of deixis for demonstratives that follow a head noun.
5. Topic-switching tends to entail the use of postposed demonstrative in clause-second position.

6. Postposed demonstratives in Finnic varieties in the east (Olonets Karelian, Lude, and Veps) along with North Russian are robustly employed with evaluative use, which is only sporadically observed in Finnic varieties in the west (Livonian, South Estonian, and North Estonian), and totally obsolete in Finnic varieties of the central area (Votic, Ingrian).

After comparing the Finnic and North Russian postposed demonstratives from different perspectives, only a few common features apply to the entire area under investigation. However, several distinct micro-areas can be identified. The only unified feature across the entire contact zone under discussion is that postposed demonstratives function as topicalisers in clause-second position (see also a confirmation from Finnish by Vilkuna 1989: 143–147). Nevertheless, unified patterns across categories can be found among (i) Finnic varieties around the Gulf of Riga (Livonian, South Estonian, and North Estonian), (ii) Finnic varieties in the central area (Votic and Ingrian), (iii) Finnic varieties of the Karelian type (Olonets Karelian and Northern Lude), and (iv) Finnic varieties in the eastern end (Southern Lude and Veps) along with North Russian dialects. This areal division will also participate in discussion of whether the demonstrative system can contribute to the classification of Finnic languages (see Section 7.2).

From a diachronic point of view, the Novgorod birch bark documents do not show a strong sign of continuation to modern North Russian dialects, as there are not many morphosyntactic and functional properties shared between Old Novgorod and modern North Russian dialects. This speaks in favour of the modern North Russian dialects being a descendant of Central Middle Russian (originating from the ancient Rostov-Suzdal' dialect), rather than originating from Old Novgorod (a hypothesis discussed in Section 2.2.2).

At the methodological level, the quantitative approach applied in this section appears fruitful, helping in maximising and solving data impairment issues. It makes possible comparison of languages with different sizes of corpora, and enables discovery of both language-internal as well as language-external factors influencing the use of postposed demonstratives. Without a quantitative approach, many important details would have been overlooked in the current study.

7 DISCUSSION

This chapter brings the results from the data analysis in Chapter 6 into discussion, addressing the main research questions introduced in Section 1.1:

1. How do postposed demonstratives function as grammatical markers?
2. What does the development of demonstratives tell us about the history of Finnic and Slavic languages?
3. Do postposed demonstratives result from a Finnic substratum in North Russian dialects?

The chapter is divided into three sections. Section 7.1 clarifies the status of Finnic and North Russian postposed demonstratives as grammatical markers by discussing three perspectives and a wider areal typological comparison with languages in Eurasia as tools. Section 7.2 applies the method of undoing grammaticalisation, to understand how postposed demonstratives could have evolved language-internally, and connects changes observed within demonstrative system with the taxonomical classification of Finnic. Finally, Section 7.3 attempts to provide various arguments for and against the idea of Finnic substratum in North Russian, which have been continuously discussed for many decades.

7.1 FUNCTIONAL CHARACTERISTICS OF POSTPOSED DEMONSTRATIVES

The data analysis in Chapter 6 showed that postposed demonstratives are grammatical markers which are functionally close not only to definite articles but also to topic, focus, and evaluative discourse markers. Depending on the perspective and variety, postposed demonstratives can be typologically classified in multiple ways. As the aim is to identify what the postposed demonstratives are in each variety, this section discusses the phenomena according to four theoretical perspectives: grammatical profiles (Section 7.1.1), information structure (7.1.2), stancetaking and evaluation (7.1.3), and areal typology (7.1.4).

7.1.1 PARALLELS BETWEEN DEMONSTRATIVES AND DEFINITE ARTICLES

This section focuses on grammatical changes in the use of postposed demonstratives, and whether they illustrate expected patterns of grammaticalisation from demonstratives to definite articles (discussed in Sections 4.2 and 4.3). Based on Diessel's criteria (1999a: 118), Table 41 compares the visibility of grammatical changes in each Finnic and North Russian variety under discussion.

Change in the use of demonstratives	Livonian	South Estonian	North Estonian	Votic		Ingrian		Karelian	Lude		Veps			North Russian	
				Western	Eastern	Soikkola	Heva		Northern	Southern	Northern	Central	Southern	Vologda	Arkhangelsk
1. They may have undergone a process of phonological reduction.	–	–	–	–	–	–	–	–	–	–	–	–	–	+	+
2. They may have coalesced with other free forms.	–	–	–	–	–	–	–	–	–	+	+	+	+	+	+
3. They are usually restricted to the distal	–	–	–	+	+	–	–	–	–	–	+	+	+	+	+
4. They may have lost their ability to inflect.	–	–	–	–	–	–	–	–	–	–	+	+	+	+	+
5. Their occurrence is often restricted to a particular syntactic context.	–	–	–	–	–	–	–	–	–	–	+	+	+	+	+
6. They are often obligatory to form a certain grammatical construction.	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–

Table 41. *Morphophonological and morphosyntactic changes observed in each variety (Diessel 1999a: 118)*

At first glance, the varieties of the east show a stronger sign of demonstratives becoming definite articles, in terms of grammatical changes. However, it is also worth discussing each point addressed in Table 41 at a deeper level.

In terms of phonology, there is no strong evidence of phonological reduction, as the Proto-Finnic demonstratives **se/ne(t)* remain in their full forms all of Finnic. Ultimately, only further investigation with suprasegmental details could reveal whether stress has been lost from demonstratives. In any case, postposed demonstratives in North Russian can be regarded as phonologically reduced, such as the masculine singular form *-ot* from *tot*. Moreover, they never receive stress in comparison to the stressed full form of the demonstrative *tot*. This also leads to the second point, that postposed demonstratives in Finnic varieties in the east along with North Russian have coalesced with other free forms, leading the majority of previous studies to attach demonstratives to their head with a hyphen (mentioned in Section 1.1).

Based on the description in Section 2.4.1, the Proto-Finnic demonstratives **se/ne(t)* can be synchronically regarded as proximal in South Estonian, neutral in Livonian and North Estonian, and as distal in Votic and Veps. Meanwhile, **se/ne(t)* in the rest of Finnic varieties, which retain a tripartite demonstrative system, is neither proximal nor distal. At the same time, modern North Russian *tot* clearly belongs to the distal category, both synchronically and diachronically in Old Russian. As for inflection, we can say that the Veps and North Russian postposed demonstratives have partially lost their ability to inflect, considering that the Veps postposed demonstratives can no longer agree with a head noun in case, but only in number, *se* for singular and *ne* for plural. North Russian postposed demonstratives no longer inflect in polysyllabic forms, such as the masculine/neuter genitive *togo* or dative *tomu*. Moreover, the vowel harmonisation with the head word in North Russian speaks in favour of a development towards suffixation. At the same time, postposed demonstratives in Votic, Ingrian, Olonets Karelian, and Lude can inflect in number and case, and there is no occurrence of postnominal demonstratives in other cases than nominative in Livonian, South Estonian, and North Estonian. Most importantly, in Southern Lude, Veps, and North Russian, the host attachment of postposed demonstratives has extended to non-nominal parts of speech (verb, adverb, and adposition), but the use of preposed demonstratives remains restricted to nominal classes, as in other Finnic varieties to the west (discussed in Section 6.2).

At the syntactic level, demonstratives in most Finnic varieties still occur more frequently before head nouns, although there is a clear tendency that demonstratives *se/ne* in Southern Lude and Veps predominantly occur after a head noun in a similar manner to North Russian (see Section 6.1.4). Moreover, postposed demonstratives in Southern Lude, Veps, and North Russian co-occur more frequently with nuclear syntactic functions (subject, direct object, indirect object), while preposed demonstratives are more often found in adverbial phrases (discussed in Section 6.3.3). As the majority of adverbial phrases observed in the data often involves deictic uses, i.e. gesturing of time and location, this division of labour also implies that the functional load of deixis in preposed demonstratives is higher than in postposed demonstratives, which are more commonly found in non-deictic uses, co-occurring in noun phrases representing nuclear functions. In any case, there is no evidence of obligatory use of postposed demonstratives to form a certain grammatical construction, one of the most decisive criteria for grammaticalisation in Greenberg's definition (1978: 61).

Apart from the morphophonological and morphosyntactic aspects, referentiality is also a decisive domain of use for identifying the development of definite articles (discussed in Section 4.3.3). Applying the criteria in Becker (2018) and König (2018), Table 42 compares the contexts in which postposed demonstratives co-occur with different referential types, based on the data analysis in Section 6.4.1. As a definite article usually encodes anaphoric, bridging, situationally unique, and establishing referent functions (Becker 2018: 86), the data shows that postposed demonstratives behave more like definite articles as one moves towards the east. According to this parameter, Southern Lude and North Russian can use postposed demonstratives for all types of referential use mentioned in Table 42.

Referential use	Livonian	South Estonian	North Estonian	Votic		Ingrian		Karelian	Lude		Veps			North Russian	
				Western	Eastern	Soikkola	Heva	Olonets	Northern	Southern	Northern	Central	Southern	Vologda	Arkhangelsk
1. Deictic	+	+	+	–	–	+	+	+	+	+	+	+	+	+	+
2. Recognitional	–	–	–	–	–	–	–	+	–	+	+	+	+	+	+
3. Absolutely unique	–	–	–	–	–	–	–	–	–	+	–	+	–	+	+
4. Anaphoric	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
5. Bridging	–	–	+	–	–	–	–	+	+	+	–	+	+	+	+
6. Situationally unique	–	–	–	–	–	–	–	–	+	+	+	+	+	+	+
7. Establishing	–	–	–	–	–	–	–	–	–	+	+	–	–	+	+

Table 42. Contexts of referential uses observed with the use of postposed demonstratives

By excluding the establishing referential function, which is not necessarily decisive in Becker's classification, postposed demonstratives can be treated as functional equivalents to definite articles in the entire Lude and Veps dialect continuum. Postposed demonstratives in other Finnic varieties to the west, meanwhile, are used in a more restricted set of referential use. The reason for this might be due to the fact that in these varieties, the suspected functional extension of demonstrative to definite article only concerns preposed demonstratives, as observed in Estonian (Pajusalu 1999: 64), Votic (Agranat 2015), as well as in other West and South Slavic languages discussed in Section 5.2.1.

Based on the discussion of grammatical changes in this section, postposed demonstratives exhibit the characteristics of definite articles more strongly in Lude, Veps, and North Russian, where the grammatical changes and referential uses characteristic of definite articles (given in Diessel 1999a; Becker 2018; König 2018) are observed. Nevertheless, the use of postposed demonstratives in these varieties is optional, as omitting them does not make a clause ungrammatical (similarly to the case of Finnish prenominal demonstratives in Larjavaara 1990: 146; Chesterman 1991: 102–103; Juvonen 2000: 193, discussed in Section 6.4.1). Considering the criterion of the obligatory use in specific constructions (Greenberg 1978: 61; Diessel 1999a: 118), postposed demonstratives in Lude, Veps, and North Russian cannot be regarded as definite articles proper. For instance, the co-occurrences of demonstratives and absolutely unique referents as in (171) to (174) (shown in Section 6.4.1), which would obligatorily require markers of definiteness, are motivated by topicality rather than the definiteness of referents.

This observation implies that the definiteness effect in postposed demonstratives in eastern Finnic and North Russian has not conventionally emerged from the determiner use, such as in the case of other western Finnic, Mordvinic, and Slavic

languages (discussed in Sections 5.1 and 5.2). Instead, it has more likely emerged from information-structural uses as topic markers, which have later extended to contexts of use involving definite referents (see the relation between topicality and definiteness in Givón 2001: 474, and a similar idea for North Russian in Gvozdanović 2019: 126). To a certain degree, this also resembles Kasatkina's (2007, 2008) idea of "pseudoarticles" in North Russian dialects (see also the idea of "articloid" in the 17th century Russian texts in Mendoza 2011). This follows because the grammaticalisation of definite articles has not taken place completely, but what is observed instead is only a functional extension of demonstrative towards the marking of definiteness, without the bleaching of the erstwhile information-structural uses (see the definition of grammaticalisation in Section 3.1.2).

Unlike the proposed idea of preposed demonstratives as definite articles *in situ nascendi* (in Voge's term 1958) in a number of Finnic and Slavic languages (discussed in Chapter 5), I do not expect postposed demonstratives in Northwest Russia to develop further, into an obligatory definite article. Apart from the predominant information-structural uses that have not been weakening, sociolinguistic factors such as vitality as well as influence from perstrate literary languages do not favour, or may even halt the grammaticalisation process. First, the vitality of Lude and Veps as language communities is currently very much of concern (see Lallukka 1990, 2001; Puura et al. 2013, discussed in Section 2.3.2 and 3.2.1), and evidence suggest a rather pessimistic outcome, namely that language shift to Russian would happen before any further significant development in the language structure would be reached. Second, the emergence of literary languages for Finnic minorities in Russia as a perstratum-like scenario (discussed in Section 3.2.1) is also removing this language feature, as it is often negatively perceived as Russian interference. Instead, the model for developing standard language is more preferably taken from cognate state languages such as Estonian and Finnish, which do not use such multifunctional postposed demonstratives. Something similar can also be said for North Russian speakers, whose deviating language use becomes over time more and more levelled by the Central Russian standard (see, e.g., the levelling of the North Russian phonological system in a younger generation of speakers in Vaahtera 2009).

7.1.2 DEMONSTRATIVES AS TOPIC AND FOCUS MARKERS

As discussed in Section 6.3, postposed demonstratives can occur in clause-second position throughout the entire Finnic and North Russian dialect continuum. Their use resembles boundary markers previously reported from other languages, especially those in Southeast Asia (discussed in Section 4.4.4). As boundary markers functionally separate topics from comments, it is reasonable to claim that a general function of postposed demonstratives in all varieties under discussion here is the marking of topicality, rather than the marking of definiteness. This follows because definite referents do not obligatorily require the presence of postposed demonstratives in a phrase (discussed in Sections 6.4.1 and 7.1.1). Supporting this classification is the observation of functional dimensions that are typical for topic markers, discussed below.

As the data analysis in Section 6.4.2 has shown, postposed demonstratives can be used to mark both continuation and switch of topic in all varieties. Given these contexts, the use of postposed demonstratives as topic markers also fits into the simple model in Neeleman et al. (2009) as “aboutness topics” and “contrastive topics” (see Table 8). However, frequencies vary according to the areal distribution. Postposed demonstratives tend to mark switch of topic more often in South Estonian, North Estonian, Southern Lude, Veps, and North Russian but continuation of topic is preferred in other Finnic varieties (see Chart 12). Meanwhile, new referents are often marked as topics in South Estonian, North Estonian and North Russian, while old referents are more often observed before postposed demonstratives in other Finnic varieties (see Chart 13). This observation shows that clause-second demonstratives are more often used as devices to gain the addressee’s attention when switching topics in Southern Lude, Veps, and North Russian, as well as in South Estonian and North Estonian. At the same time, they function as devices to keep track of active topics in the discourse in other Finnic varieties. This functional characteristic could also explain more developed focal uses of postposed demonstratives of eastern varieties.

In terms of focus marking, “information focus” is not usually marked in Finnic and North Russian, likely due to newness, indefiniteness, and unidentifiability. In case focused referents are marked, they are frequently identifiable referents and, thus, are often regarded as a case of referential use of postposed demonstratives, rather than focus marked. Meanwhile, other focal uses of postposed demonstratives, adapted from the description in Krifka (2008) as examined in Section 6.4.2, are observed in the data as shown in Table 43.

Focal use	Livonian	South Estonian	North Estonian	Votic		Ingrian		Karelian	Lude		Veps			North Russian	
				Western	Eastern	Soikkola	Heva	Olonets	Northern	Southern	Northern	Central	Southern	Vologda	Arkhangelsk
1. Correction/Repair	–	–	–	–	–	–	–	+	–	+	+	+	+	+	+
2. Repetition	–	–	+	–	–	–	–	+	–	+	+	+	–	+	+
2. Highlighting of parallels	–	–	–	–	–	–	+	–	+	–	+	+	+	+	+
3. Afterthought (additional information/Delimitation of constituent)	–	+	–	–	–	–	–	–	+	+	+	+	+	+	+
4. Closed(/Opened) focus	–	–	–	–	–	–	–	–	+	+	+	+	+	–	+
5. Exhaustive focus	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
6. Scalar focus	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
7. Listing (multiple focus)	–	–	–	–	–	–	–	–	+	+	+	+	–	+	+

Table 43. *Focal uses of postposed demonstratives observed in each variety*

Considering the first four functions as characteristic of focus markers in Krifka's classification (discussed in Section 4.4.3), postposed demonstratives in Southern Lude, Veps, and North Russian function very similarly to proper focus markers (described in Section 6.4.2). Namely, they can be used to correct and confirm an utterance, to highlight statements in parallel, and to delimit the addressee's focus to a specific constituent. At the same time, focal usage of postposed demonstratives in Olonets Karelian and Northern Lude only occasionally occurs, and could be still regarded as underdeveloped. Focal use of postposed demonstratives is very rare in general in other Finnic varieties to the west.

The discussion in this section suggests that the use of postposed demonstratives in Finnic and North Russian can also be viewed as marking of topicality, which I regard as an erstwhile context that further motivates acquisition of the definite article-like uses (see the discussion in Sections 7.1.1 and 7.2.1). Furthermore, Southern Lude, Veps, and North Russian also frequently use postposed demonstratives for focus marking, given that their syntactic contexts cover not only the clause-second position, but also other clausal positions, a feature rare in other Finnic varieties to the west of Olonets Karelian (shown in Chart 6 and Chart 7 in Section 6.1.6).

7.1.3 DISCOURSE-PRAGMATIC USES OF DEMONSTRATIVES

A preliminary survey in Section 6.4.3 discusses the possibility of postposed demonstratives used as markers of stance in discourse. This type of function is closely related to the speaker's stancetaking and evaluation, which relate to the interlocutors' state of knowledge, attention, and intersubjectivity. Three types of interlocutor-related orientation can be expressed with postposed demonstratives, as shown in Table 44.

Evaluative use	Livonian	South Estonian	North Estonian	Votic		Ingrian		Karelian	Lude		Veps			North Russian	
				Western	Eastern	Soikkola	Heva		Northern	Southern	Northern	Central	Southern	Vologda	Arkhangelsk
1. Speaker-oriented	–	–	–	–	–	–	–	+	+	+	+	+	+	+	+
2. Addressee-oriented	+	–	–	–	–	–	–	+	–	+	+	+	+	+	+
3. Joint attention	–	+	+	–	–	–	–	–	–	+	+	+	+	+	+

Table 44. *Evaluative uses of postposed demonstratives observed in each variety*

In the evaluative domain, the speaker-oriented function, which includes the speaker's subjective evaluations such as affection and surprise, can be expressed by postposed demonstratives only in the varieties from Olonets Karelian eastwards. The same areal distribution applies to the addressee-oriented function, which concerns the

interlocutors' shared knowledge. At the same time, joint attention as the speaker's intersubjective evaluation can be created by postposed demonstratives in South Estonian and North Estonian, as well as Finnic varieties from Southern Lude eastwards. The North Russian postposed demonstratives can mark all of these evaluative uses.

The functional distribution in Table 44 may partially have its root in demonstrative systems of individual languages. For instance, the versatile evaluative uses of postposed demonstratives *se/ne* in Veps could be related to the loss of a proximity contrast (discussed in Section 2.4.1), which would no longer distinguish spheres of discourse, i.e. the speaker's sphere vs. the addressee's sphere, still being present in most other Finnic varieties. In South Estonian and North Estonian, however, the use of postposed demonstratives to establish joint attention is likely to be due to left dislocation, which pragmatically evokes a function similar to interjective hesitators (see further discussion in Section 7.2.1).

Based on the description above, postposed demonstratives in Finnic varieties in the east along with North Russian clearly show a tendency to function as discourse markers beyond the referential and information-structural domains. This versatile pragmatic use of postposed demonstratives in eastern varieties appears to correlate with their high degree of syntactic mobility, which is remarkably more restricted in other Finnic varieties in the west.

The multiple ways of interpretation from different perspectives, as discussed thus far, also appear to relate to the optionality of use, meaning that postposed demonstratives have not become fully accommodated and, are therefore not obligatory in any specific grammatical construction. In any case, it is interesting to note that the degree of development gradually decreases towards the west, confirming that the geographical proximity between Finnic and North Russian varieties, together with areal-particular contact situations, is one of the most significant factors responsible for the pragmaticisation of demonstratives in Finnic.

7.1.4 AREAL-TYPOLOGICAL CONSIDERATION ON A GRAMMATICAL CATEGORY OF POSTPOSED DEMONSTRATIVES

Even though functions and properties of postposed demonstratives in Finnic and North Russian can be defined (previously discussed in Sections 7.1.1, 7.1.2, and 7.1.3), it still remains unclear what kind of grammatical category postposed demonstratives typologically represent (see the issue of cross-linguistic comparability discussed in Haspelmath's approach 2010). For instance, the use of postposed demonstratives in Lude, Veps, and North Russian falls between borderlines, as they are capable of serving not only as markers of definiteness and topicality, but also as focus and discourse markers, the interpretation of which depends on both the contexts as well as the perspective chosen to analyse the data. For this reason, it is also worth comparing postposed demonstratives with grammatical elements in other languages across Eurasia, which reportedly behave in a similar way, and serve similar functions (see Chapters 4 and 5). This approach may potentially give a clearer picture of how to

define a grammatical category of postposed demonstratives in the languages of Northwest Russia.

For the sake of macro-areal comparison, this section will compare the profile of postposed demonstratives in Finnic and North Russian to the following grammatical markers.

Definite articles

- Scandinavian (Dahl & Kortmann 2004)
- Balkan: Bulgarian (Mladenova 2007; Lindstedt 2014), Macedonian (Koneski 1967), Romanian (Dobrovie-Sorin 2013)
- Mordvinic: Erzya (Cygankin et al. 2000), Moksha (Aljamkin et al. 2000)

Postposed demonstratives

- Western Finnic: Livonian, South Estonian, and North Estonian
- Central Finnic: Votic and Ingrian
- Karelian: Olonets Karelian and Northern Lude
- Eastern Finnic: Southern Lude and Veps
- North Russian

Possessive suffixes

- Komi (Leinonen 1998, 2006; Gerson Klumpp, p.c.)
- Ural-Altaic: Eastern Uralic (Nikolaeva 2004), Turkic (Stachowski 2010), Mongolic (Brosig et al. 2018)

Topic markers

- Japanese-Korean (Vermeulen 2009): Japanese (Martin 2004), Korean (Sohn 1999)

Demonstratives

- Mainland Southeast Asian: Burmese (Simpson 2008), Eastern Cham (Brunelle & Hăn 2019), Green Hmong (Kunyot 1984), Hakha Chin (Barnes 1998), Mon (Jenny 2014), Thai (Iwasaki & Ingkaphirom 2005), Vietnamese (Lê 2002; Adachi 2016, forthcoming), and Karenic languages (Kato 2003; Shee 2008)

Through a combination of consultation with language descriptions and personal knowledge of individual languages, selected features are compared cross-linguistically. Unless otherwise indicated, the source of information is based on the primary data of the present study. The division of four Finnic groups (western, central, Karelian, and eastern Finnic) is based on the conclusions of the data analysis in Section 6.6.

To facilitate data illustration, a computational-aided phylogenetic tool is employed here. Putting aside the matter of obligatory use in Greenberg (1978), the following 24 properties in Table 45 are investigated. This set of properties is converted to a NEXUS format with binary data: 0 vs. 1 (Maddison et al. 1997), and fed to SplitsTree4, version 4.16.1, built 19 May 2020 (Huson & Bryant 2006).

Features	Grammatical domains
1. Preposed bound position	Constituent order
2. Postposed bound position	
3. Free morpheme	
4. Clause-second position	Constituent position
5. Clause-middle position	
6. Clause-final position	
7. Phonological reduction and coalescence with head word	Form
8. Agreement with head word	
9. Full inflectional paradigm	
10. Hosted by verb	Host attachment
11. Hosted by adverb	
12. Phrase marking	
13. Marking new/unidentifiable referents	Referentiality
14. Marking bridging and/or situationally unique reference	
15. Marking establishing reference	
16. Continuing topic	Topicality
17. Switch-topic	
18. Marking left dislocation	
19. Correction and confirmation	Focality
20. Highlighting parallels	
21. Marking afterthought/right dislocation	
22. Marking the speaker's stance	Discourse
23. Maintaining the interlocutors' interactional relation	
24. Creating the interlocutors' joint attention	

Table 45. *Typological features used in the cross-linguistic comparison*

Each property is binarily evaluated: 0 = absent, 1 = present. The distribution of properties under discussion is shown in Table 46.

Language type	Label	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
Scandinavian	Scandinavian	0	1	0	1	1	1	1	1	0	0	0	0	0	1	1	1	1	0	0	1	0	0	0	0
Balkan	Balkan	0	1	0	1	1	1	1	1	0	0	0	0	0	1	1	1	1	0	0	1	0	0	0	0
Western Finnic	WFinnic	1	1	1	1	1	1	0	1	1	0	0	0	0	0	0	1	1	1	1	0	1	0	1	1
Central Finnic	CFinnic	1	1	1	1	0	1	0	1	1	0	0	0	1	0	0	1	1	1	0	1	0	0	0	0
Karelian	Karelian	1	1	1	1	1	1	0	1	1	0	0	0	1	1	0	1	1	1	1	1	1	1	0	0
Eastern Finnic	EFinnic	1	1	1	1	1	1	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
North Russian	NRussian	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Mordvinic	Mordvinic	0	1	0	1	1	1	1	1	0	0	0	0	0	1	1	1	0	0	0	1	0	0	0	0
Komi	Komi	0	1	0	1	1	1	1	0	0	1	1	1	1	1	0	1	1	1	0	1	1	1	1	1
Ural-Altaic	UralAltaic	0	1	0	1	1	1	1	1	0	1	1	1	1	1	0	1	1	1	0	1	1	1	1	1
Japanese-Korean	JapKor	0	1	0	1	0	0	1	0	0	1	1	1	1	1	1	1	1	1	0	1	1	0	0	0
Mainland Southeast Asian	MSEA	0	1	1	1	1	1	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

Table 46. *Typological profiles of grammatical markers in the selected groups of languages*

Operating on the NeighborNet algorithm (Bryant & Moulton 2004), the SplitsTree software visualises the distance between profiles of the grammatical markers in comparison as a network diagram in Figure 26.

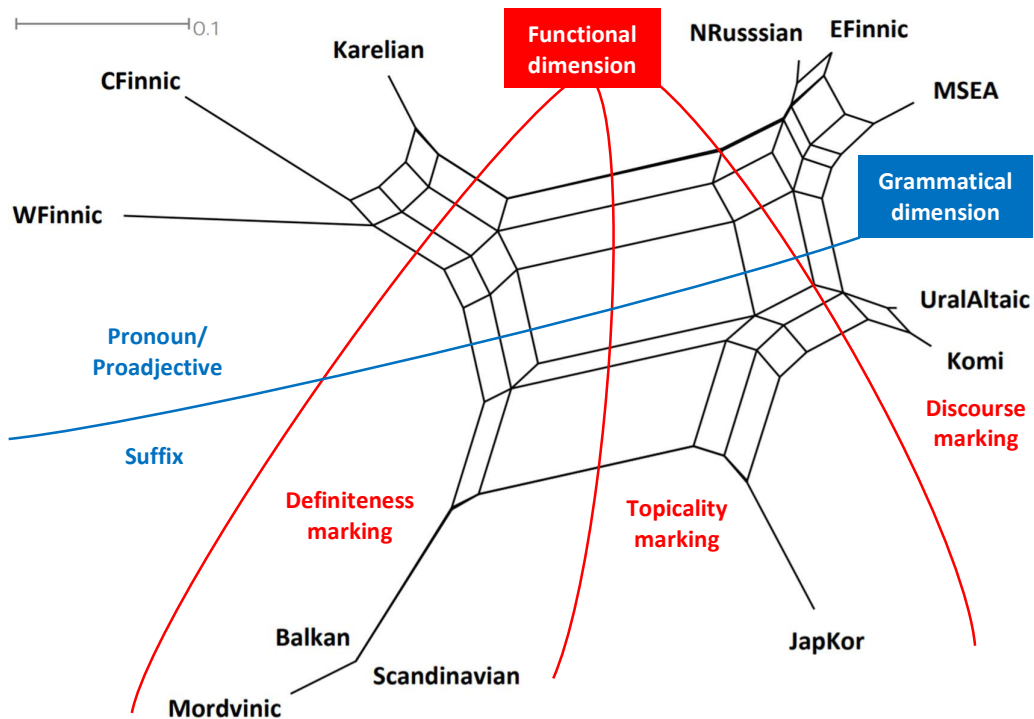


Figure 26 Functional distance between grammatical markers in different language groups

This network diagram shows that there are significant gaps lying on two dimensions, grammatical (blue) and functional (red) properties which determine the clustering of language types. The difference in grammatical properties 1–12 (vertical dimension) is due to the morphophonological and morphosyntactic behaviour of the markers, as the clades on top represent demonstratives which can be used as “pronouns” or “proadjectives”, whereas the clades on the bottom primarily include “suffixes”. In terms of functional properties 13–24 (horizontal dimension), the difference is determined by the functional range and domain in which each grammatical marker is predominantly used.

Based on this quantitative approach, several interpretations can be made about postposed demonstratives in Finnic and North Russian.

1. Postposed demonstratives in western and central Finnic, as well as Karelian Finnic, have not extended their functionality towards the domains of definiteness or discourse marking. Accordingly, the use of postposed demonstratives can be regarded as “dislocation markers”, which only mark the boundary between information chunks in the discourse, and are often employed in clause-second position to mark topics in left dislocation constructions.

2. The use of postposed demonstratives in eastern Finnic and North Russian is very close to that of the demonstratives in Mainland Southeast Asian languages, as they have developed versatile uses in the discourse beyond definiteness and topicality marking.
3. The use of postposed demonstratives in eastern Finnic and North Russian also resembles that of possessive suffixes in Komi and other Ural-Altaic languages, mainly in the areas of information-structural and discourse functions. The main difference between the postposed demonstratives and the possessive suffixes is due to grammatical properties, such as the ability to stand independently as free morphemes, phonological reduction, and inflection.
4. The use of “discourse markers” differs from “definiteness markers” in Scandinavian, Balkan, and Mordvinic types, and from “topicality markers” in the Japanese-Korean type. They differ from the Japanese-Korean type in terms of grammatical and discourse functions, and from markers of definiteness of the Scandinavian, Balkan, and Mordvinic type in that definite articles do not play a significant role in stancetaking. However, topic marking uses observed in “discourse markers” make them closer to “topicality markers” of the Japanese-Korean type.

From the typological point of view, this comparison has shown that postposed demonstratives in eastern Finnic and North Russian do not exclusively correspond to either 3rd person pronouns, definite articles or noun class markers, focus markers or boundary markers, all of which are regarded as potential typical targets of grammaticalised demonstratives (see Diessel 1999a: Ch. 6; Heine & Kuteva 2002: 106–116; Jenny 2009, discussed in Section 4.2). Instead, it seems that properties of these expected targets are combined onto the use of postposed demonstratives, which also resembles the phenomenon of “polygrammaticalisation” (Craig 1991), and partially the idea in computational pragmatics that ‘a single functional segment may express more than one dialogue act’ (Bunt 2017: 336). From the grammatical point of view, they behave as pronouns or proadjectives used to mark discourse-pragmatic functions related to stancetaking. Given that their functions are closely related to demonstratives in Mainland Southeast Asian languages and possessive suffixes in Ural-Altaic, which some previous studies (Leinonen 1998; Nikolaeva 2004; Adachi, forthcoming; Iwasaki & Dechapraturumwan, forthcoming) also label as “discourse markers”, it is likewise also reasonable to treat postposed demonstratives in eastern Finnic and North Russian as devices for discourse functions, particularly with information-structural and evaluative uses, in an areal-typological context.

The discussion in this section confirms the idea presented in Section 5.4 that the investigation of Finnic and North Russian postposed demonstratives should not be restricted to identifying them as definiteness markers (as had been previously done until the mid-20th century), but that their uses as topicality and discourse markers must also be taken into account. Accordingly, a comparison of postposed demonstratives

with grammatical elements which are functionally rather than formally similar could yield more fruitful results. Next, the discussion continues from synchrony to diachrony, concentrating on the question of how postposed demonstratives emerged and evolved.

7.2 HISTORY OF FINNIC DEMONSTRATIVES IN CONTACT WITH SLAVIC

This section is divided into two parts. The first part (Section 7.2.1) uses the method of morphosyntactic reconstruction, i.e. undoing grammaticalisation (DeLancey 1994; Harrison 2003; Heine 2003, discussed in Section 3.1.2), to postulate pathways of how postposed demonstratives emerged in different Finnic varieties, also with reference to parallel development in North Russian dialects. The second part (Section 7.2.2) uses an areal-linguistic method to capture areal diffusion (discussed in Section 3.3.2) among Finnic varieties by bringing together various features and changes in modern Finnic demonstrative systems. These shared features and changes which have been observed in the current study could shed light on the diversification and language contact which took place later across the branches. As result, the discussion given in Chapter 6 along with these subsections also propose an areal subgrouping of Finnic varieties into four areal genetic units: (i) western Finnic (Livonian, South Estonian, and North Estonian), (ii) central Finnic (Votic and Ingrian), (iii) Karelian Finnic (Olonets Karelian and Northern Lude), and (iv) eastern Finnic (Southern Lude and Veps).

7.2.1 GRAMMATICAL CHANGES IN THE FINNIC DEMONSTRATIVES IN COMPARISON TO NORTH RUSSIAN

Given that all Finnic varieties possess this feature, I propose that the first functional extension of the postposed Finnic demonstratives **se/ne(t)* started from a “dislocation” construction (more often left dislocation), which places demonstrative pronouns after a head word as a boundary-marking element. As a device for the syntactic dislocation is closely related to such uses like avoidance and hesitation (discussed in Section 4.5.2), this property appears to develop demonstratives further, to functions involving stancetaking and evaluation. Meanwhile, typical uses of non-proximal demonstratives in this domain are maintenance of interactional relations between the interlocutors, and the organisation the discourse (shown in Section 4.5.2). Consequently, this characteristic would seem to give rise to a boundary-marking function between topic and comment, stabilising the use of postposed demonstratives in clause-second position. Until this stage, the use of postposed demonstratives as “topic markers” is able to be found across the entire Finnic dialect continuum. Beyond this stage, however, differences between Finnic varieties start to emerge.

The most remarkable change in the following stage is that Karelian Finnic and eastern Finnic start using multiple postposed demonstratives in a clause to divide constituents in focus. This extends the syntactic range to clause-medial and clause-

final, indicating the ability of demonstratives to move to a particular position according to the speaker's information-structural strategy. This movability is a syntactic characteristic of "focus", in contrast to topic, which is usually tied to a specific position (Kiss 1995: 6), such as clause-initial position in Uralic and Slavic languages more generally (see, e.g., Vilkuna 1998). The extension from demonstrative to focus marking is expectable in a number of languages (Heine & Kuteva 2002: 111–112, discussed in Section 4.2), and postposed demonstratives have acquired under this extension additional functions characteristic of focus, such as delimiting a scope and establishing contrast in a set of alternatives (see Section 4.4.1). Simultaneously, the development of focal uses seems to also reinforce the usage domain of "discourse marking", as we observe more versatile evaluative uses in Karelian Finnic (Olonets Karelian and Northern Lude) and eastern Finnic (Southern Lude and Veps), compared to western Finnic (Livonian, South Estonian, and North Estonian) and central Finnic (Votic and Ingrian), where the evaluative uses of demonstratives are remarkably less visible (discussed in Section 7.1.3).

Given that topics are generally definite (Li & Thompson 1976: 461, discussed in Section 4.4.1), the combination of definiteness with focality-oriented functions delimiting a scope and establishing contrast in a set of relevant alternatives seems to motivate the furthest stage, in which postposed demonstratives arrive from the topicality domain to "definiteness marking". In a relative chronology, this development would have taken place during the weakening and obsolescence of the use of possessive suffixes in eastern Finnic varieties. In any case, the functional extension towards definiteness marker is less advanced in Karelian Finnic, while remaining unfinished in eastern Finnic, as it has not reached a point in which the use of definiteness marker would be obligatory in any specific grammatical construction. The grammaticalisation of obligatory definiteness markers is likely halted by the other non-referential (information-structural and evaluative) uses, which maintain optionality in a number of contexts of use (discussed in Section 7.1.1). Given this language-structural obstacle, I suggest that unless these varieties start to employ other grammatical elements for non-referential uses and reserve postposed demonstratives only for referential uses involving definiteness marking, a complete semantic extension to definiteness markers is unlikely.

Based on the description above, I propose that the functional extension from demonstrative pronouns towards definiteness markers in Finnic varieties has advanced according to the following scheme in Figure 27. As a remark for Figure 27, I adopt the idea of pragmaticisation by Norde (2009: 23) that 'movement towards discourse is genuinely different from movement towards grammar, and the two are therefore best kept separate,' and in this vein treat the functional extension of demonstratives towards a marker of definiteness (grammar) and towards a marker of discourse as two parallel development paths. Furthermore, the progress in each Finnic group and North Russian is also indicated under the scale, which can be seen as a formation of different concentric circles in Güldemann's framework (2008, discussed in Section 3.3.1), including an "areal hotbed" (EFin, NRus), a "core circle" (Kar), and the "periphery" (CFin, WFin) (see discussion of individual isoglosses in Section 7.2.2).

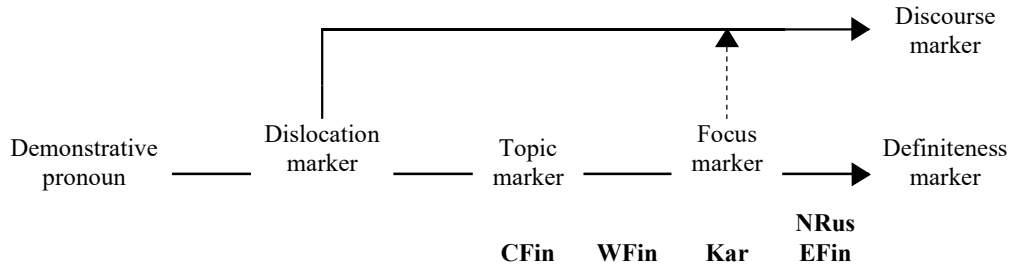


Figure 27 Functional extension of postposed demonstratives towards markers of definiteness

This grammaticalisation path is by no means intended to be a description of a universal tendency, but rather a language-specific tendency in the context of Finnic and North Russian postposed demonstratives. For instance, the function of definiteness marking in Finnic preposed demonstratives has not emerged through this mechanism, but rather from the adnominal use as a determiner (cf. Laury 1996, 1997; Pajusalu 1997a; Juvonen 2000; Agranat 2015, discussed in Section 7.1.1). Instead, this scenario of postposed demonstratives in eastern Finnic varieties is an empirical example of the mechanism on how markers of definiteness can develop from the non-referential domain of topic and focus marking (see a similar idea in Givón 2001: 474, versus the opposite development direction from referential to non-referential uses of definite articles in the Philippine type of Austronesian languages in Nagaya 2011). Hereby, I also propose this scheme as a model of polysemy copying shared between eastern Finnic and North Russian, in which the Finnic-Slavic contact has been the most intense and continued the longest (as described in Section 2.3.2).

Ultimately, the topicality effect in the Russian *-to* could have gone back to the connective use in the construction *X, to Y* ‘(if/when) *X*, then *Y*’ (see Gvozdanović 2019: 126, presented in Section 5.2.2), which I regard as one of several functions in the stage of dislocation use. This development from a device for dislocation in conditional or temporal constructions to a topic marker corresponds to Haiman’s idea (1987, discussed in Section 4.4.1) that conditionals are semantically comparable to topics. In any case, this connective use of demonstratives is not observed in Finnic, not even in the easternmost varieties, Lude and Veps, which may speak in favour of the proposal above, that the pronominal use as a clause connector is only one of several pragmatic functions of a dislocation marker, prior to the rise of a topic marker.

Beyond the scheme in Figure 27, it is also worth discussing changes and behaviours of postposed demonstratives in each Finnic variety, which appear to be connected with various changes that have taken place in demonstrative systems from the Proto-Finnic stage.

In western Finnic (Livonian, South Estonian, and North Estonian), the Proto-Finnic proximal series **tämä/nämä(t)* shifted to 3rd person pronouns, and pulled the Proto-Finnic medial series **se/ne(t)* to fill the proximal domain. The new proximal series was used alongside the distal **too/noo(t)*, before the new proximal started replacing the distal series and eventually taking over all deictic spheres (Pajusalu

1995, 1998, 2006). However, the new proximal series *se/ne*, when used in postnominal position, does not show referential uses typical for definite articles, such as bridging and establishing referential relations. Instead, they only function as dislocation markers when occurring after a head noun (see Amon 2015), compared to the article-like uses of preposed demonstratives *see* in Livonian and North Estonian, and *tuu* in South Estonian (see Pajusalu 1997a, 1998). In terms of evaluative uses, it seems that the erstwhile semantics of proximal deixis restricts the scope to the speaker's sphere, and thereby gives epistemic authority to the speaker. In this vein, the earlier proximal postposed demonstratives in western Finnic varieties are primarily used to establish joint attention with the addressee by offering a stance from the speaker's sphere (shown in Section 6.4.3). Ultimately, the additional pragmaticisation of postposed demonstratives in western Finnic may also be disfavoured by the coexistence of other discourse markers frequent in spoken language. Such discourse markers in competition are, for example, Finnic manner demonstratives *siis* 'so, thus', *nii(n)* 'in this way, like this', or a speaker-symmetry particle *ju* 'of course', which is likely a Germanic borrowing, either previously from Gothic *ju*, or later from Swedish *ju*, consider a particle *jo* in Votic, Finnish, Ingrian, Karelian, and Veps (Metsmägi et al. 2012).

The encoding of joint attention observed in western Finnic is absent from Votic, despite the fact that a similar change from proximal demonstratives to 3rd person pronouns has also occurred in Votic. This is possibly due to the renewal of the demonstrative system according to the Russian bipartite model which has taken place in Votic and Veps (see Table 3). In this renewed system, newly emerged compound demonstratives *kase/kane* in Votic have filled the empty proximal slot, while keeping the Proto-Finnic medial series **se/ne(t)* away from becoming proximal in the model of western Finnic. Therefore, the use of Votic *se/ne* rather belongs to the central Finnic type, in which postposed demonstratives are only involved in clause-second use as a marker of left dislocation, marking the continuation of topic. Given that the demonstratives *se/ne(t)* are non-proximal both in Votic and Ingrian as well as in Karelian and Lude, the focus of attention falls on the addressee rather than the speaker. This addressee-oriented property seems to favour the use of maintaining interactional relations between the interlocutors, by highlighting the information in the addressee's sphere. Therefore, continuing topics, which are generally more identifiable and active in the addressee's attention, can often co-occur with postposed demonstratives in central Finnic and also Karelian Finnic, as well as in Finnish (see Etelämäki 2009).

Unlike less advanced pragmatic uses of postposed demonstratives in western and central Finnic, Karelian Finnic and eastern Finnic have reached the stage of definiteness marking. The main difference that splits these two groups is due to deixis, which plays a crucial role in the evaluative uses. As Proto-Finnic demonstratives **se/ne(t)* remain as medial with a contrast with distal **too/noo(t)* in both Karelian and Lude, their evaluative uses are more restricted than in Veps. Section 6.1.4 showed that postposed demonstratives *se/ne* in eastern Finnic primarily occur after the head word, which may have been affected by a more relaxed constituent order in noun phrases. Eastern Finnic and North Russian, in particular, use postnominal determiners considerably more often than other cognate languages (see Ojanen 1985: 230–235,

and also a similar case of constituent order shift in eastern Saami discussed in Section 5.1.2). This structural renewal suggests that the Veps demonstrative system is gradually becoming unipartite, with a single series of proper demonstratives *ñece/ñene*, which will eventually turn postposed demonstratives *se/ne* into deictically neutral demonstratives (as estimated by Grünthal 2015a: 277, discussed in Section 2.4.1). Ultimately, I consider the obsolescence of distal *se/ne* and neutralisation of the deictic contrast as a consequence of the weakening of deictic uses in *se/ne*, due to the shift towards more pragmatic uses. This reorganisation of the demonstrative paradigm could explain a wider range of evaluative uses in eastern Finnic, compared to the Karelian Finnic (observed in Section 6.4.3).

The description above is mainly drawn on Finnic-internal evidence, but influence from North Russian dialects as an external factor of language change is discussed further in the following subsections.

7.2.2 CLASSIFICATION AND DIVERSIFICATION OF FINNIC LANGUAGES IN THE LIGHT OF DEMONSTRATIVES

The diversification of Finnic languages is a result in distribution of innovations in an earlier Finnic dialect continuum. As development in the demonstrative systems has also yielded a number of innovations in particular groups of Finnic, it can contribute to the improvement of the Finnic family tree. A number of dialectal variations discussed in Chapter 6 can be included to several stages of branching off. The following discussion concerns nine isoglosses, of which Isoglosses 1–5 relate to the organisation of demonstrative systems, while the use of postposed demonstratives is discussed in connection to Isoglosses 6–9.

Isogloss 1: The development of the demonstratives **tämä/nämä(t)* to 3rd person pronouns

This isogloss is an areal feature shared between the western Finnic languages (Livonian, South Estonian, North Estonian, Votic), as shown in Table 3 (presented in Section 2.4.1) and in (134) (presented in Section 6.1.1). It could be dated to a stage after the split of Central Finnic from other North Finnic languages, in which the semantic shift from proximal to personal pronoun is not observed widely, with the exception of Finnish, which has developed the 3rd person deixis involving all demonstrative series *tämä*, *se* and *tuo* (see, e.g., Seppänen 1998; Varteva 1998; A. Hakulinen et al. 2004: §1425; Priiki 2017). The areal status of this semantic shift can also be supported by typological uncommonness (see a parameter for contact-induced change by Saarikivi 2000, in Table 6), as the source of 3rd person pronouns is typologically more often a distal demonstrative (Greenberg 1978: 61; Givón 2001: 226; Diessel 1999a: 161; Bhat 2013), such as in the majority of Slavic languages (discussed in Section 2.4.2).

Isogloss 2: The neutralisation of deictic distinctions by the demonstratives **se/ne(t)*

This isogloss is a shared areal pattern in the Estonian area (see also Pajusalu 1996a; Tomingas 2018), as shown in Table 3. Hypothetically, the Proto-Finnic medial demonstratives **se/ne(t)* must have replaced the proximal series, which subsequently became 3rd person pronouns (Isogloss 1). While maintaining the proximal status, the Proto-Finnic distal series **too/noo(t)* started to disappear from the language, and has eventually freed the entire deictic sphere for **se/ne(t)* to take over (see also Pajusalu 2006).

From an areal perspective, this isogloss seemingly spread from North Estonian (possibly the Insular dialect) to Livonian, at the latest, towards the end of 19th century through intense contact between fishermen across Gulf of Riga (see Kettunen 1938; Ariste 1981: 78; Grünthal 2015b: 127–137). This separate contact scenario that later took place between Livonian and North Estonian is also supported by studies on shared lexicon (see Koponen 1990, and comments in Viitso 1990b) and a lesser degree of contact between Courland Livonian and South Estonian (see Ariste 1954: 260).

The neutralisation of the demonstrative paradigm in Livonian and North Estonian resembles that of Slovincian as an intermediate stage, before the application of deictic intensifiers in most other West Slavic languages, as well as Belarusian and Russian, in which only the Proto-Slavic medial series **tb* is used throughout the paradigm (discussed in Section 2.4.2). Given these parallels, the Livonian and North Estonian systems might have been influenced or motivated by the model in earlier West Slavic, or western East Slavic languages spoken along the southern coast of Baltic Sea. This potential areal influence is also supported by the fact that this isogloss is absent from South Estonian spoken inland further from the maritime trade routes, as well as from East Slavic languages in the south (Rusyn and Ukrainian), which maintain other Proto-Slavic demonstrative series.

Moreover, contact with Swedish, which also uses deictic intensifiers in the neutralised demonstrative paradigm *det här* ‘this’ vs. *det där* ‘that’, a phenomenon which is not observed in Danish or Norwegian, should also be taken into account. In regard to contact history, Scandinavian-speaking populations, initially Viking traders later peasants and fishermen, have been present in the islands of the Gulf of Riga and the northwestern coast of Estonia since the second half of the 1st millennium until the 20th century, which resulted in contact with North Estonian and Livonian neighbours (see Grünthal 2015b: 118–126, 139).

Ultimately, this areal feature could have extended even as far as to Votic and Veps. In these languages, namely, the intermediate stage before arriving to the bipartite system (Isogloss 4) with compound demonstratives (Isogloss 3) could have resembled a unipartite system in Livonian and North Estonian. This change could happen in a very similar pattern to the Slavic parallel, in which only the Proto-Slavic demonstrative stem **tb* is retained, and serves as the base for new compound demonstratives. Nevertheless, this contact hypothesis requires further empirical investigation.

Isogloss 3: The emergence of compound demonstratives

This development is only observed in Votic and Veps, likely a consequence of the renewal of the demonstrative paradigm into a bipartite system (Isogloss 4) according to the Russian model (discussed in Section 2.4.1). In any case, the Votic and Veps cases are not mutually related, as deictic intensifiers derived from different lexical sources, Votic *ka* ‘well’ vs. Veps *näged* ‘you see’ (see Kettunen 1943: 403). In parallel to this, the modern Russian proximal *ëtot* emerged in the language only in the 16th–17th century, suggesting that the emergence of the Votic and Veps compound demonstratives as replications of the Russian model should not have taken place before the 16th century, or possibly even later, after *ëtot* was accepted to the literary norm in the 18th century (see Vlasto 1986: 129, discussed in Section 2.4.2).

In a similar scenario to Isogloss 2, the emergence of compound demonstratives may also have involved Swedish contact. At the same time, the occurrence of *ëtot* in Russian could have spread from Belarusian (or even further from a West Slavic language), given that the Standard (Moscow) Russian, with an admixture of the western (Smolensk) dialect evolved in the 16th century (discussed in Section 2.2.2), matching the earliest attestation of *ëtot* in Russian.

Meanwhile, the occurrence of *neče/nenne* in Olonets Karelian and *ñeče/ñeñe* in Lude can be seen as a spread from Veps. In any case, the dating is still uncertain, but could be thought to have taken place after the split of Ladogan Finnic (cf. postposed demonstratives as focus and definiteness markers in Isogloss 7).

Isogloss 4: The renewal of deictic distinction of demonstratives to a bipartite system

As discussed in Section 2.4.1, the renewal towards a bipartite demonstrative system in Votic and Veps was potentially motivated by the Russian model, which only uses the Proto-Slavic medial series **tъ* with a modification by the intensifier *è-*. Consequently, the newly emerged compound demonstratives in Votic and Veps replace the obsolete Proto-Finnic proximal series **tämä/nämä(t)*, while the Proto-Finnic medial series **se/ne(t)* remains as non-proximal, so they have become distal in the newly rearranged paradigm. In any case, the distribution of constituent orders (see Table 19 and Chart 4 in Section 6.1.4) shows the tendency for the Veps distal demonstratives *se/ne* to be predominantly postposed, and more often used for other extended pragmatic uses (Isogloss 7) than as determiners. Consequently, the use of compound demonstratives *ñece/ñene* is gradually taking over the prenominal position, which may eventually lead to the neutralisation of deictic distinctions, towards a unipartite system (as stated in Grünthal 2015a: 277, discussed in Section 2.4.1), parallel to Livonian and North Estonian (Isogloss 2). Elsewhere, tripartite systems are retained, while Livonian and North Estonian have neutralised the Finnic series **se/ne(t)* to initially cover proximal, and later all deictic spheres. This can also be regarded as an areal feature, as discussed in Isogloss 2.

Worth mentioning is also the South Estonian system, in which the demonstratives *taa/naaq* are currently reported to be disappearing from the spoken language, reducing

the paradigm to a bipartite system, with proximal *see* and distal *too* based on the Standard Estonian model (Pajusalu 2015; Tammekänd 2015). Meanwhile, the issue of South Estonian proximal *sjoo/njooq* and distal *tuu/nuuq*, resembling Old East Slavic (Old Novgorod and Old Rus') demonstrative stems *sb* and *tb* in forms and meanings, is also interesting, as pointed out by Pajusalu (1996a: 150–151). Despite the fact that their etymology is problematic, the series *taa/naaq* has been proposed as a borrowing from Baltic to Late-Proto-Finnic (Larjavaara 1986: 73–75). In such a scenario, the series *taa/naaq* can be considered as a later addition to the Inland Finnic system, in which the Proto-Finnic proximal and distal series had been reorganised into a bipartite system according to a similar model observed in Old Novgorod and Old Rus'.

These parallels of bipartite demonstrative systems might imply that Slavic contact with South Estonian took place earlier than contact with Veps and Votic, and the dating could even trace back to the first wave of migration of Slavs in the 5th–8th century (discussed in Section 2.3.1). For instance, Ariste (1981: 85–86) suggests that Livonian and South Estonian were the first Finnic groups to receive Slavic loanwords, before spreading them to North Estonian and elsewhere. This also relates to evidence that old Slavic loanwords are found in Estonian more than in other Finnic languages (see also Must 2000). Whether this parallel to Old East Slavic languages is due to a very old contact-induced change remains for future studies to investigate.

Isogloss 5: Distribution of frequency of use among different demonstrative series

Based on the information in Chart 1 (presented in Section 6.1), Livonian shares the similar distribution with North Estonian but not with South Estonian (Isogloss 5.1), which might also strengthen the claim about their intense contact as in Isogloss 2. Meanwhile, Lude is clearly divided into two groups: the northern variety forms a parallel with Olonets Karelian (Isogloss 5.2), whereas the southern variety with Veps (Isogloss 5.3). This also supports the idea concerning the formation of Lude discussed earlier by Pahomov (2011: 10–12) that Northern Lude emerged from Karelian while Southern Lude has a Veps base (discussed in Section 2.2.1).

Isogloss 6: The clause-second use of the demonstratives **se/ne(t)* after topics

This feature is observed in all Finnic varieties under investigation, seen in (188), (189), (191), (193), and (194) presented in Section 6.4.2. A previous investigation by Vilkuna (1989: 145–147) also confirms this tendency in Finnish. As the development of demonstratives into dislocation markers, eventually topicality markers, is also typologically expectable (Podlesskaya 2010), it is possible that this feature was already present as early as in Proto-Finnic. Given that this tendency of Wackernagel's clause-second clitic is widely present in Indo-European languages since ancient times, including Old Church Slavonic and Old Novgorod (illustrated in Section 6.5), I propose that an early date can also apply to the Finnic case (see also a discussion on Wackernagel's law in Uralic languages in Nevis 1988: 104–108).

Isogloss 7: The use of the demonstratives **se/ne(t)* after focused and definite referents

This feature is present only in Finnic languages which allow the use of postposed demonstratives also in clause-medial and clause-final positions (shown in Sections 6.4.1 and 6.4.2). This includes Olonets Karelian, Lude, and Veps, all of which derive from Proto-Ladogan-Finnic and have undergone intense contact with North Russian dialects. The extension of functional range, as discussed in Section 7.2.1, seemingly correlates with the degree of Russian influence and adjacency to the nexus of the Finnic-Slavic contact in Vologda: Karelian < Lude < Veps (see T. Itkonen 1971; V. Koivisto 1990: 20, discussed in Section 2.3.2). Chronologically, the development of these pragmatic uses should have taken place after the branching of Proto-Ladogan-Finnic, but no later than the late 19th century, when this was first observed by field linguists (e.g., Kettunen 1943: 399–403), approximately between the 16th and 18th centuries (see further discussion in Section 7.3.2).

Isogloss 8: The co-occurrence of postposed demonstratives with non-nominal parts of speech

The results in Table 26 and Table 27 (presented in Section 6.2) show that most Finnic varieties to the west of Lude only use postposed demonstratives with nominal parts of speech (nouns, pronouns, names, and adjectives). Meanwhile, eastern Finnic (Lude and Veps), together with North Russian, also allow postposed demonstratives to occur after non-nominal parts of speech (verbs, adverbs, and adpositions). This points to a later innovation, a shared areal feature between eastern Finnic and North Russian.

Isogloss 9: The co-occurrence of postposed demonstratives with adverbials, determiners and predicatives

The results in Table 28 and Table 29 (presented in Section 6.3.1) show that postposed demonstratives in western and central Finnic varieties (Livonian, Estonian, Votic, and Ingrian) can mainly co-occur with subjects and direct objects. Meanwhile, Karelian and eastern Finnic varieties (Olonets Karelian, Lude, and Veps) together with North Russian, also allow postposed demonstratives to co-occur with non-nuclear syntactic functions (adverbial, determiner, and predicative). This feature seemingly points to the branching of Proto-North-Finnic. However, Finnish does not allow this possibility either, so the feature must be a later innovation, a shared areal feature among eastern Finnic varieties spoken in Russia along with North Russian dialects.

Based on the discussion above, each isogloss can be placed onto the Finnic family tree (Figure 28) according to the following distribution.

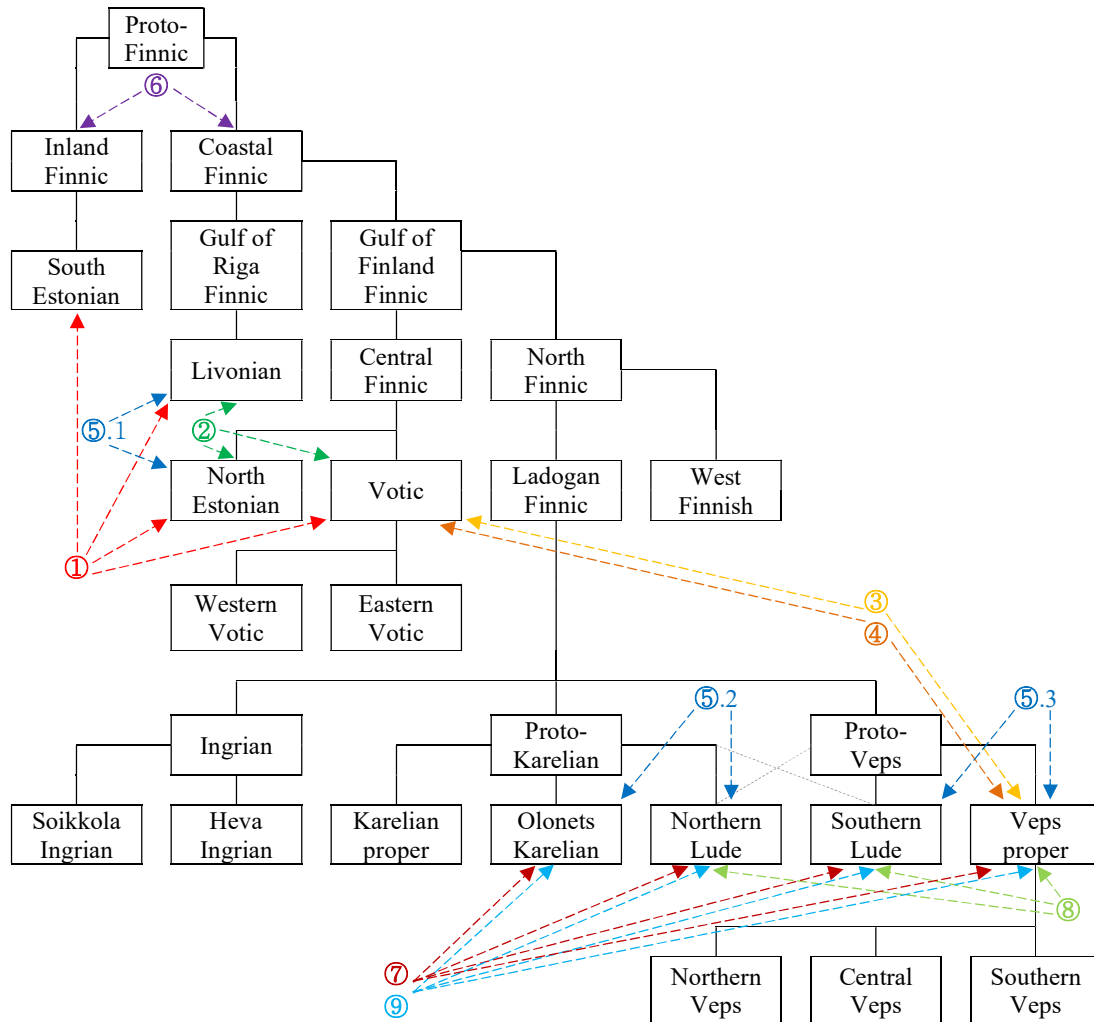


Figure 28 Diversification of Finnic languages and innovations in demonstrative systems (a tree modified from the basis of Kallio 2014: 163)

Figure 28 shows that most isoglosses concern Finnic sub-branches, suggesting that changes in the Finnic demonstrative systems must have taken place more recently as areal features and, thus, do not affect structure of the proposed Finnic tree. This speaks in favour of the idea that demonstratives are an unstable area of grammar genealogically (see a similar tendency of divergence among Saami languages in Ylikoski 2020), prone to the environment of speech communities and contact-induced changes (see also Sidnell & Enfield 2017). Some of the isoglosses discussed above are clearly due to Russian influence, which provides a model for the development in demonstrative systems. Given that Isoglosses 3 and 4 only concern Votic and Veps, which supposedly came into contact with Slavs earlier than Karelian, these cases likely predate Isoglosses 7, 8, and 9 which only concern Olonets Karelian, Lude, and Veps.

In terms of areal distribution, these isoglosses can be place to a cartographic illustration, seen in Figure 29.

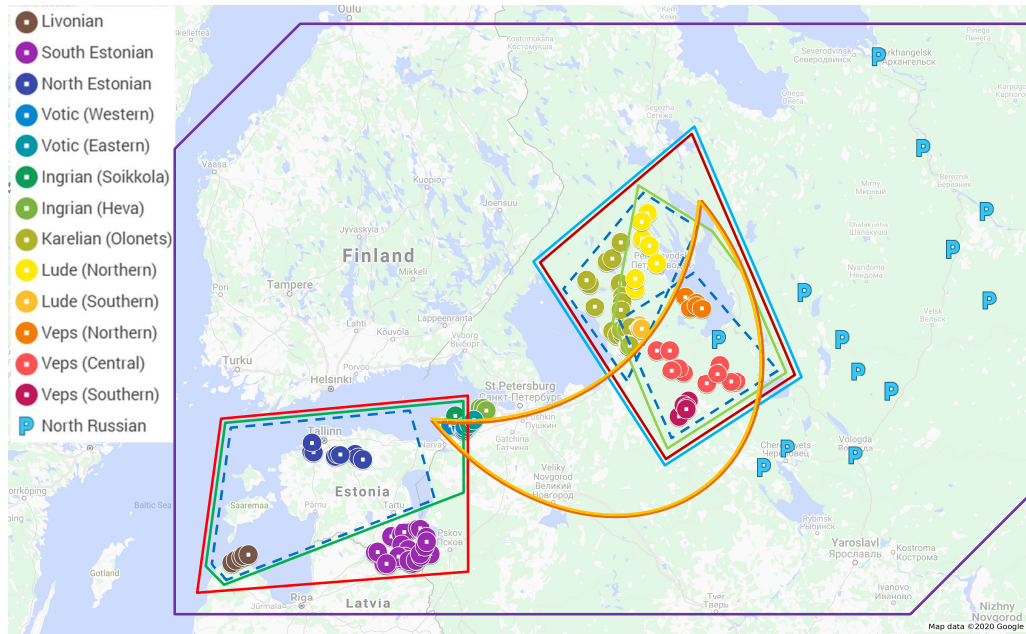


Figure 29 Geographical distribution of isoglosses concerning development of the Finnic demonstrative systems

The map shows that the development of demonstrative systems is clearly divided into two main micro-areas, west vs. east, the boundary which could go back as far as the Ancient Novgorod period (see Figure 4), correlating with the division of the western and eastern Church (discussed in Section 2.3.1). As for the isoglosses which concern the use of postposed demonstratives (6, 7, 8, and 9), they clearly concentrate in the eastern zone, forming a “core circle” which includes Olonets Karelian, Lude, and Veps. The rest of Finnic varieties in the west (Livonian, South Estonian, North Estonian, Votic, and Ingrian), meanwhile, remain in the “periphery” (in Güldemann’s term 2008). Another notable remark is that Ingrian can be considered relatively conservative, as it shares but one isogloss (6) with other Finnic languages, even with Votic, despite their geographical adjacency. This is likely due to a later migration of Ingrian speakers from the northeast, after the diversification of Proto-Ladogan-Finnic.

Geographically, several isoglosses cross through the isthmus between Lake Onega and Lake Ladoga, corresponding to an areal division based on toponymic evidence, which suggests a later migration of Karelians from the west towards the Onega isthmus during the 14th century (discussed in Pimenov 1965: 184–185; Mullonen 1989, 1990). Moreover, onomastic evidence from census books of Novgorod during the 15th–16th centuries discussed by Raunamaa and Kanner (forthcoming) also suggests a similar geographical division along the border between historical Vodskaja and Obonežskaja Pjatinas (Russian *pjatina* ‘fifth’, referring to the five administrative divisions in the Ancient Novgorod principality).

The relation between demonstrative systems and the Finnic taxonomy discussed above shows that this study not only provides relevance for descriptive and areal-typological linguistics, but also historical-comparative studies concerning later contact

between the Finnic sub-branches. In the end, most isoglosses do not look old enough that they would go back to intermediate proto-languages, but rather in the majority of cases points to areal features.

7.3 DO POSTPOSED DEMONSTRATIVES RESULT FROM A FINNIC SUBSTRATUM IN NORTH RUSSIAN DIALECTS?

One of the main goals of the current study is to discuss whether postposed demonstratives are a Finnic substratum in North Russian dialects, as claimed in previous studies (Veenker 1967; Kiparsky 1969). The following subsections discuss two approaches for identifying whether postposed demonstratives should be regarded as a substrate feature (Section 7.3.1), or as an areal feature (Section 7.3.2).

7.3.1 EVALUATION OF THE FINNIC SUBSTRATE HYPOTHESIS

This section will apply four criteria to identify substrate features in Saarikivi (2000, discussed in Section 3.2.3), which examine whether a feature under investigation is (i) characteristic or marked construction in a shift-target language, (ii) common among cognate languages of the shift-target language, (iii) expected language-internal development, and (iv) typologically unique (shown in Table 47 towards the end of this section).

First, it is crucial to identify the characteristics of postposed demonstratives present in North Russian but not in Central (Moscow) Russian. These are cases in which influence from Finnic can be considered. Comparing to the indeclinable *-to* in Central Russian, North Russian postposed demonstratives express a higher degree of coalescence with their head word through vowel harmonisation (see Kuz'mina 1993: 184–187; Mendoza 2011: 249, and the discussion in Section 6.2). Otherwise, postposed demonstratives in both varieties behave in the same way morphosyntactically, particularly with a wide range of host attachment and clausal positions (see Sections 6.1.6, 6.2 and 6.3). The main difference, however, is found in their pragmatic functions. Central Russian primarily uses *-to* as a device to convey thematization or topicality, contrast, and turn-taking (Rathmayr 1985; Bonnot 1986, 1987, 1990, 1991; Zybatow 1990; Grenoble 1998: Ch. 5; McCoy 2001, 2003; Bolden 2008, discussed in Section 5.2.2), equivalent to topic and focus markers in many other languages (illustrated in Section 7.1.4). This use is also observed in North Russian (see Section 6.4.2). Evaluative uses still need to be investigated more in Central Russian and other varieties, however, in order to make a holistic comparison. In any case, there has not been any report that postposed demonstratives in Central Russian would cover all the referential uses characteristic of definite articles, which North Russian postposed demonstratives are able to do (see Section 6.4.3). A usage-based approach in Talmy's terms (1985: 62, discussed in Section 3.1.2) suggests that the use of postposed demonstratives in North Russian can be considered, to a certain degree, marked construction. It deviates from the context of Russian in general, given that

postposed demonstratives are (i) colloquial in style rather than literary, (ii) frequent in speech rather than occasional, and (iii) pervasive with a wide range of semantic notions expressed in this type. However, the fact that postposed demonstratives are optional indicates that they have not been fully grammaticalised into part of a specific construction. As the criteria discussed above do not strongly speak in favour of postposed demonstratives as a characteristic of Russian, I regard the degree of markedness as moderate.

Second, there are a number of properties in postposed demonstratives which distinguish North Russian from other Slavic languages. However, the same properties also distinguish eastern Finnic from other Finnic varieties to the west. For instance, the head-initial constituent order of demonstrative violates the canonical head-final syntax of Finnic and Slavic languages (see Dryer 2013a, discussed in Section 4.1). However, postnominal demonstratives are not completely unknown to Slavic languages, which also tolerate interchangeable constituent orders in the noun phrase (see Kurz 1939–1946; Večerka 1993: 82–85, discussed in Section 5.2.1). In any case, the flexibility of demonstratives to co-occur with non-nominal hosts (adverbs, verb, and adpositions), breaks and extends the domain of use from determiner to topicality and discourse marking, a phenomenon which is not observed in other Finnic and Slavic languages. In any case, this functional extension of postposed demonstratives is also observed in Central Russian, and it clearly distinguishes Russian from other Slavic cognate languages, including the closely related Belarusian and Ukrainian. At the same time, a similar development of postposed demonstratives into markers of definiteness has its parallel in enclitic definite articles of the Balkan Slavic languages, so it is not totally unique in the Slavic context (see Section 5.2.1). Meanwhile, the development from prenominal demonstratives to markers of definiteness is also reported in Finnic varieties in the west (see Section 5.1.1). According to the discussion above, particular properties of postposed demonstratives are not always common in Slavic and Finnic cognate languages. Based on this observation, I consider postposed demonstratives moderately common among cognate languages of both substrate (Finnic) and shift-target languages (Slavic).

Third, the development from demonstrative into a marker of definiteness and topicality is not rare among the world's languages (see Heine and Kuteva 2002: 109–112, discussed in Section 4.2). In some areas, this development may be related to language contact, such as the cases of the Balkan *Sprachbund*, or Western Continental Europe more generally. In other cases, however, the functional extension of demonstratives to topicality and discourse marking may not necessarily result from language contact (cf. Mainland Southeast Asia, Papua New Guinea, and Finnic varieties in the west, see Sections 4.4, 4.5 and 5.1.1). Given the parallels in other areas, I regard the possibility of postposed demonstratives being a language-internal development as relatively high.

Fourth, there are many languages in non-adjacent or remote areas to Northwest Russia, in which postposed demonstratives are morphosyntactically and pragmatically employed in a similar manner to eastern Finnic and North Russian, such as in Mainland Southeast Asia (see Sections 4.4 and 4.5). Similar pragmatic functions can also be coded with other grammatical elements such as possessive suffixes, widely

observed in Northern Eurasia (see Section 4.5 and 5.3). This shared property among grammatical elements indicates that the phenomenon under discussion is not typologically rare among the world's languages.

Based on the discussion above, the four criteria are evaluated as given in Table 47.

More likely to be substrate		Moderate	Less likely to be substrate	
Marked feature in the shift-target language (Russian)		+		Characteristic feature in the shift-target language (Russian)
Common among the substrate language and its cognate languages (Finnic)		+		Common among the shift-target language and its cognate languages (Slavic)
Rare language-internal development			+	Expected language-internal development
Typologically rare among the world's languages			+	Typologically common among the world's languages

Table 47. *Likelihood of the Finnic substrate influence on the North Russian postposed demonstratives (applied to the parameter by Saarikivi 2000: 398–399)*

Based on the criteria introduced in Table 47, it can be argued that the use of postposed demonstratives in North Russian displays more characteristics of an adstrate feature shared with Finnic. This suggests that the North Russian postposed demonstratives have not purely developed from the Central Russian *-to*, through language shift by the erstwhile eastern Finnic-speaking population to Russian in a vertically-oriented substrate scenario. More plausible is that the development has taken place in a horizontally-oriented adstrate scenario, in which Finnic-Russian bilingualism was stable, pointing to areal diffusion being a more important factor.

7.3.2 UNDERSTANDING THE EMERGENCE OF POSTPOSED DEMONSTRATIVES IN AN AREAL CONTEXT

Given the arguments presented in Section 7.3.1, it is not likely that postposed demonstratives can be simply claimed as a Finnic substrate feature in North Russian dialects. Instead, the story definitely involves a larger areal tendency in a particular contact zone of Northwest Russia. Recalling each factor that may give rise to similarities between languages (given by Aikhenvald and Dixon 2001, presented in Section 3.3.2), several points are worth discussing on the basis of the data used in the present study, particularly the issues of universality, borrowability, and convergence.

First, it is reasonable to begin the discussion by eliminating cases of similarities by chance or genealogical retention. For the former issue, the quantitative analysis in Chapter 6, and the decoding of the functional extension discussed in Section 7.2.1, have shown that the further east a Finnic variety is spoken, the more its demonstrative system resembles that of North Russian. This geographical correlation suggests that eastern Finnic and North Russian must have definitely undergone mutual reinforcement under long-standing intense language contact (discussed in Section

2.3), which results in postposed demonstratives only resembling each other in function but not in form, i.e. pattern borrowing (in Sakel's term 2007, discussed in Section 3.3.2). In other words, these languages in contact have copied polysemy from a shared structural model (Weinreich [1953]1974; Tkačenko 1989; Heine & Kuteva 2005; Wiemer & Wälchli 2012, discussed in Section 3.2.2). As for the latter, the discussion in Section 7.2 illustrated that the use of postposed demonstrative in modern Finnic varieties and North Russian could not go back much further than several centuries ago, so genealogical retention is unlikely. This leaves three factors for further consideration: (i) universality, (ii) borrowability, and (iii) convergence.

The earlier discussion in Section 7.1.4 has shown that Eurasia manifests a tendency of marking definiteness, topicality, and discourse functions with several specific grammatical elements, particularly demonstratives and possessive suffixes. Even though this tendency might not be ultimately common worldwide, it is present in this particular macro-area of the world. The realisation of this areal tendency is also divided into micro-areas, such as postposed demonstratives in Northwest Russia, possessive suffixes in Central Russia, and demonstratives in Mainland Southeast Asia (discussed in Chapters 4 and 5).

As for identifying the source of borrowing, the situation becomes challenging, particularly when the feature under investigation is not common among cognate languages, such as the case of multifunctional postposed demonstratives (previously discussed in Section 7.3.1). Regarding the direction of borrowing, the data presented in Chapter 6 does not imply a strong unilateral influence from Finnic to Russian, proposed in some previous studies (cf. Kiparsky 1967; Veenker 1969), nor does it imply the opposite direction (cf. Vahros 1951; Larjavaara 1986). It points to a more complex story, in which postposed demonstratives in Northwest Russia can be regarded a result of areal diffusion.

The interpretation of areal diffusion suggested in the current study is motivated by the observation in Chapter 6 and discussion in Chapter 7, that such inflectable and pragmatically robust postposed demonstratives are not productively observed in other western and central Finnic, nor in Russian varieties of other dialectal zones. Despite the fact that this contact-induced feature should have emerged in the adstrate environment through balanced and stable bilingualism between Finnic and Russian prior to the Russian Empire and Soviet era (discussed in Section 2.3.2), it is not impossible to propose Russian as the source language, given that the connective use in conditional clauses as a potential ultimate source construction was already attested as early as the 11th century's Old Rus' (Vlasto 1986: 237). As there is a general tendency for at least the neuter form of the postposed demonstrative *-to* to be currently used across the entire Russian dialect continuum (*tb* in southern and Siberian dialects), this Russian conceptual template should have allowed additional pragmaticisation more easily than the Finnic system, in which the use of demonstrative after a head noun, other than for clause-second topic marking, was relatively rare prior to Russian contact.

Such a claim would not have to entail that the development of postposed demonstratives completely originated and was totally transmitted from Russian. North Russian postposed demonstratives are pragmatically relatively different from those of

Central Russian (discussed in Section 7.3.1). There are several properties that North Russian postposed demonstratives resemble parallels in eastern Finnic subsequently evolved without Central Russian interference, such as the ability to inflect, or vowel harmonisation and the functional extension to contexts in which they co-occur with definite referents (discussed in Chapter 6). The deviation from the Central Russian model, and convergence with eastern Finnic, speaks in favour of a development through mutual reinforcement, favoured by the long-lasting advanced multilingual practice of inhabitants in the area of contact (discussed in Sections 2.1 and 2.3).

As an answer to the key research question of the current study, the data of the current study does not support the idea of postposed demonstratives being a Finnic substrate feature in North Russian, but rather speaks in favour of a shared adstrate feature between eastern Finnic and North Russian, illustrated in Figure 30.

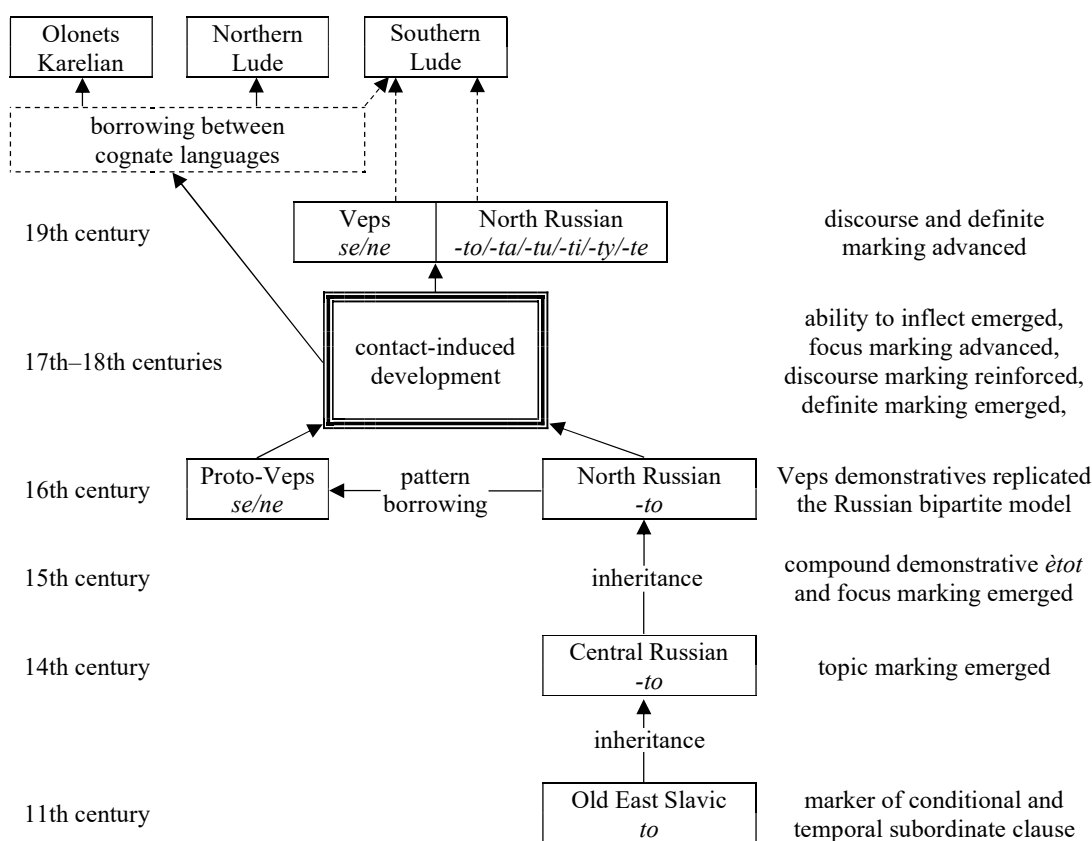


Figure 30 Borrowing and the development of postposed demonstratives in the Northwest Russian context

To elaborate on Figure 30, I regard this whole phenomenon as having its root at least in the indeclinable Russian postposed demonstrative *-to*, inherited in both form and function from Central (Middle) Russian to North Russian. Nevertheless, the earliest attestation of topical use of *-to* in Central Russian, and the ultimate origin of the Old

East Slavic connective demonstrative *to* in conditional constructions, remain open for further investigation.

Regarding the stability of bilingual practices which were potentially also responsible for the renewal of a demonstrative system in Veps according to the Russian model (discussed in Section 2.4.1), I consider the use of postposed demonstratives as a borrowing of the Russian conceptual template with the distal demonstrative *tot*, substituted by the Veps distal demonstratives *se/ne*. This conceptual template and lexical substitution may also relate to the tendency of constituent order shifts from prenominal to postnominal determiners, which favours the use of postposed over preposed demonstratives, observed in the Southern Lude, Veps, and North Russian data illustrated in Section 6.1.4 (see also Ojanen 1985: 230–235, discussed in Section 7.2.1). Chronologically, this scenario of pattern borrowing from North Russian to (Proto-)Veps should have taken place no earlier than the 15th–16th centuries for two reasons. First, Central Russian speakers were imported to the erstwhile Old Novgorod-speaking areas primarily after the fall of Novgorod State in 1447, and the invasion of Ivan IV of the Grand Duchy of Moscow in 1570 (discussed in Section 2.2.2). Second, provided that the emergence of postposed demonstratives in Veps is closely related to the renewal of the demonstrative system according to the Russian bipartite model, with the application of intensifiers applied to the Proto-Slavic medial series **tb(tb)* (discussed in Section 7.2.2), the emergence of Russian *ètot* itself did not take place until the 16th century, so it is unlikely that pattern borrowing into (Proto-)Veps would predate this change in Russian (see Vlasto 1986: 129, discussed in Section 2.4.2).

In this particular language ecology, Veps and North Russian had started to develop further referential uses as a “shared drift” (in Janhunen’s term 2014b) that were not present in Central Russian. In Güldemann’s classification (2008), the particular contact between Veps and North Russian can be regarded as an “areal hotbed” of innovation concerning the versatile use of postposed demonstratives. Given that this language phenomenon was already observed in Lude and Veps in the late 19th and early 20th centuries by a number of linguists (e.g., Lauri Kettunen and Heikki Ojansuu) and it was likely emerging in the Finnic-Russian adstrate scenario (as discussed thus far), the evolution of multifunctional postposed demonstratives in eastern Finnic should have taken off in the 17th century at the earliest. This dating is chronologically situated in between the fall of Novgorod State, and the rise of the Russification policy of the Russian empire in the 18th century, which turned the Russian adstratum into a superstratum of Finnic.

North Russian becoming an established regional variety of Russian, Veps, and potentially also North Russian introduced this language feature to adjacent Finnic varieties, Olonets Karelian and Lude. In Güldemann’s parameter (2008), Karelian and Lude would not belong to the areal hotbed but rather the “core circle”, given that the functional extension of multifunctional postposed demonstratives is less advanced in Olonets Karelian and Northern Lude (shown in Section 7.2.1). The borrowing from Veps, thus, must have taken place at the earliest not long after the 17th century, in which advanced referential uses were still in the process of development. The topicality marking was inherited from Central Russian, and could ultimately also be a

pre-existing tendency of Wackernagel's clitics in Finnic (see Nevis 1988: 104–108, discussed in Section 7.2.2). Meanwhile, postposed demonstratives in Southern Lude have acquired as many extended functions as in Veps and North Russian. On the one hand, this suggests that Southern Lude has continued more intense contact with Veps and North Russian, due to the geographical adjacency of their speaking areas. On the other hand, this could also be used as an argument for the idea that Southern Lude was a Proto-Veps variety which later became Karelianised (Pahomov 2011). As it currently stands, the former case of longer, intense contact between Southern Lude, Veps, and North Russian seems more plausible. Ultimately, further studies on dating the diversification of Olonets Karelian and Lude varieties from Proto-Karelian and Proto-Veps could also bring us closer to the absolute chronology of postposed demonstratives.

Concerning the potential influence of Avvakum's Russian vernacular in the 17th century (Kiparsky 1969: 25–26, discussed in Section 5.2.2), I do not see a direct connection to modern North Russian dialects. In terms of functions, postposed demonstratives of Avvakum primarily involve the domain of referentiality, and information-structural and evaluative uses, such as those observed in modern North Russian, have not been reported (cf. Voge 1958, Mendoza 2011). In case there would have been a connection with Avvakum's vernacular, the chronology of Avvakum's texts would change nothing, merely supporting my dating of mutual reinforcement in Veps and North Russian to the 17th–18th centuries, as proposed above.

On the theoretical level, this contact scenario has shown that the language sociology and ecology of this micro-area, in particular stable bilingualism, was a crucial sociolinguistic factor, playing an important role in the development of such a marked language feature more generally.

7.4 SUMMARY OF THE DISCUSSION

This chapter has presented a number of observations, arguments and conclusions, regarding the use and history of postposed demonstratives in Finnic and North Russian. The discussion from various perspectives has formulated answers to the three main research questions of the present study, as given below.

1. How do postposed demonstratives function as grammatical markers?

Throughout all Finnic and North Russian varieties, postposed demonstratives are used in clause-second position to mark topics. However, in eastern Finnic and North Russian, they have also extended the host they attach to non-nominal parts of speech, and extended their use to contexts involving definite referents and evaluation, resembling markers of definite and evaluative discourse markers, respectively. Even though postposed demonstratives in these eastern varieties are found in all referential relations characteristic of definite articles, the functional extension towards definite articles stays as a (permanently) unfinished process for a few reasons. First, the use of postposed demonstratives remains optional, and has not been narrowed down

exclusively to the marking of definiteness. Postposed demonstratives still retain robust information-structural and evaluative uses. From an areal-typological perspective, eastern Finnic and North Russian postposed demonstratives show many similar properties with the use of demonstratives in Mainland Southeast Asian languages, as well as with the possessive suffixes in languages of the Ural-Altaic type, and can functionally be regarded as a type of discourse marking.

2. What does the development of demonstratives tell us about the history of Finnic and Slavic languages?

The development of multifunctional postposed demonstratives is primarily concentrated on the eastern part of the Finnic dialect continuum (Veps), whereas the functional extension of prenominal demonstrative determiners to definite articles takes place in the west (South Estonian, North Estonian, Votic, and Finnish). A number of isoglosses observed in modern Finnic demonstrative systems also point to four areal genetic units: (i) western Finnic (Livonian, South Estonian, and North Estonian), (ii) central Finnic (Votic and Ingrian), (iii) Karelian Finnic (Olonets Karelian and Northern Lude), and (iv) eastern Finnic (Southern Lude and Veps). Regarding the use of postposed demonstratives as an innovation, eastern Finnic and North Russian form an areal hotbed in the core circle which also extends to Karelian Finnic, while western and central Finnic remain unaffected in the periphery. One of the most important factors that influence the development of demonstrative systems in Finnic languages is language contact with Russian and the degree of Russification, which is higher towards the east, and correlates with geographical proximity to the Russian-speaking area. This areal grouping also corresponds to results from studies on onomastics and toponyms of Northwest Russia.

3. Do postposed demonstratives result from a Finnic substratum in North Russian dialects?

Postposed demonstratives, as multifunctional markers of definiteness, information structure, and evaluation, are not a Finnic substrate influence in North Russian dialects, but rather a shared adstrate feature, absent from western and central Finnic, and less developed in Central and Southern Russian dialects. This phenomenon has its root in the indeclinable postposed demonstrative *-to* as an information-structuring marker, inherited from Central (Middle) Russian to North Russian dialects, which provided a model to (Proto-)Veps. Together with Veps, North Russian dialects have developed two properties not observed in the equivalent form in Central or South Russian dialects: (i) the ability to inflect, and (ii) co-occurrence with definite referents. During this process of mutual reinforcement, Veps as well as potentially North Russian have introduced the extended use of postposed demonstratives to neighbouring Finnic languages, Lude and Karelian.

These ideas are the main findings of the present study, which will hopefully provide a good ground for future studies to continue the discussion.

8 CONCLUSIONS

In the present study, I have analysed the use of postposed demonstratives in a dozen Finnic and North Russian varieties spoken in Northwest Russia. By applying both a qualitative as well as quantitative approach to the language data, I have described the typological profiles of the postposed demonstratives in each variety from different grammatical perspectives, which serves as the basis for reconstructing morphosyntactic changes and semantic extensions (in the similar spirit to DeLancey 1994; Harrison 2003; and Heine 2003). Through qualitative comparison and quantification of their typological profiles, I identified an areal nexus in Vologda, the direction from which the linguistic influence has arrived from North Russian to the Finnic speech communities. The functional profiles of postposed demonstratives in eastern Finnic resemble North Russian the most, and the degree of similarities gradually decreases as one moves towards the west, confirming the idea of different levels of Russian influence across the Finnic dialect continuum (as suggested in T. Itkonen 1971, and V. Koivisto 1990: 20). Based on this finding, I have shown that geographical proximity is the single most overriding factor which explains the degree of language change and development in each variety under investigation. Further studies might also test whether other measures, such as travel time (as proposed by Gooskens 2005), would be significant for the context of Northwest Russia.

Based on the comparison of typological profiles of postposed demonstratives in Finnic, I divide the Finnic dialect continuum into four areal units: (i) western Finnic (Livonian, South Estonian, and North Estonian), (ii) central Finnic (Votic and Ingrian), (iii) Karelian Finnic (Olonets Karelian and Northern Lude), and (iv) eastern Finnic (Southern Lude and Veps). Unlike the classification that applies more conventional comparative methods to historical phonology and morphology (as in Sammallahti 1977; Viitso 1985, 2000; Koponen 1991; and Kallio 2014), this classification does not indicate the branching of Proto-Finnic in a genealogical sense, but instead focuses on areal diffusion, i.e. areal genetic units (in Helimski's term 2003) of different Finnic sub-branches from the perspective of demonstratives.

Diachronically, I have also shown that there is no direct continuation from Old Novgorod to modern North Russian in terms of the demonstrative system. Most importantly, the Novgorod birch bark data do not have more postposed than preposed demonstratives, and do not show contexts of use in which the demonstratives co-occur with non-nominal parts of speech (verb, adverb, and adposition), all of which are possible in modern North Russian dialects. This ultimately supports the idea of Old Novgorod being an extinct language which left traces as a substratum in modern North Russian dialects, which rather have their root in the Central (Moscow) dialect of Middle Russian (cf. the description by Gorškova 1972: 153–154, and Vlasto 1986: 363).

As a contribution to the ongoing discussion of the Finnic substratum in North Russian dialects, the analysis shows how postposed demonstratives have emerged and later acquired various syntactic and pragmatic functions. Through the methodological

triangulation that includes historical, contact, and areal linguistics, I have arrived to the conclusions about three matters: (i) the grammatical category of postposed demonstratives, (ii) the functional extension of the postposed demonstratives, and (iii) contact-induced change through mutual reinforcement between Finnic and North Russian dialects.

From a typological perspective, I identify the Finnic and North Russian postposed demonstratives as grammatical markers which can express various functions beyond prototypical deictic uses of demonstratives. Across the Finnic varieties, demonstratives occur in clause-second position serving as markers of topic, developing from a left dislocation construction. In western Finnic, Karelian Finnic, and eastern Finnic, the topical use more often marks a switch of topic, while central Finnic demonstratives primarily mark a continuation of topic. This topical use is also observed in North Russian and Central Russian, speaking in favour of Wackernagel's law being active in both Finnic and Russian (supporting the claim by Nevis 1988: 104–108).

This functional extension involved a morphosyntactic change, when postposed demonstratives started to move around the clause in Karelian Finnic and eastern Finnic, eventually becoming clitics which can follow nearly any part of speech in eastern Finnic (as seen in Table 25 presented in Section 6.2). This syntactic mobility and free host attachment extend the contexts of use from clause-initial topic to clause-medial and clause-final focused referents, emerging from right dislocation constructions. The topic is usually definite (as stated by Li & Thompson 1976: 461), while focus typically delimits a scope and establishes a contrast in a set of relevant alternatives (defined by Krifka 2008: 247) as well as intensifies a referent (as shown in König 1991: 181). This functional combination that postposed demonstratives inherited from earlier contexts of use have extended from information-structural uses to contexts where they co-occur with definite referents (similarly to the emergence of definiteness markers from topic constructions, as given in Givón 2001: 474). This development has allowed postposed demonstratives to appear with various types of identifiability, most notably anaphoric, bridging, situationally unique, and establishing referents, which are characteristic of definite articles (as given in Becker 2018, and König 2018).

The functional extension from non-referential uses to definiteness marking has been preliminarily speculated for North Russian (as discussed in Gvozdanović 2019: 126). In other case studies, referential uses have extended to non-referential uses, such as in the Philippine type of Austronesian languages (shown by Nagaya 2011). At the same time, preposed demonstratives in other western and central Finnic languages have acquired definiteness marking from adnominal use of determiners (as shown in Laury 1996, 1997; Pajusalu 1997a; Juvonen 2000; and Agranat 2015). Therefore, the present study leaves the question for further typological studies on whether a reverse functional extension from non-referential (topic and focus marking) to referential uses is also found in connection to other grammatical elements, and in other areal contexts.

In any case, the functional extension towards marking definiteness has not been completed in Karelian Finnic, eastern Finnic, or North Russian, as the most important property that the postposed demonstratives lack is that they must be obligatory in

certain constructions (cf. a criterion of Greenberg 1978: 61, and Diessel 1999a: 118). This is a familiar case previously also observed in prenominal demonstratives of Estonian, Finnish, and Votic, in which the demonstratives cannot be regarded as obligatory grammaticalised definiteness markers (as argued in Pajusalu 1997a: 173; Juvonen 2000: 196; and Agranat 2015: 51). In any case, I do not consider the postposed demonstratives to be able to develop much further, given both the structural context and the sociolinguistic situation of eastern Finnic and North Russian. In terms of language structure, the use of postposed demonstratives with information-structural functions remains considerably more productive than the co-occurrence with definite referents. Furthermore, absolutely unique referents as a characteristic context of definite articles do not evoke obligatory marking in these varieties.

From a sociolinguistic perspective, language vitality and language planning of eastern Finnic are not going into a direction that would favour the complete grammaticalisation. Considering the language sociology of Finnic minorities in Russia (as described in Lallukka 1990, 2001, and Puura et al. 2013), it would appear more likely that language shift to Russian would happen first, with the language authority meanwhile constantly trying to remove Russian-like structures from the written language of Finnic minorities. At the same time, North Russian dialects as a non-standard variety continually become assimilated to the Central Standard Russian norm (such as the case study of phonology examined by Vaahtera 2009). Given the circumstances, the current study shows that not only are languages endangered, but particular language features can also be in danger of vanishing or levelled by the norm of literary language as a perstratum.

Considering the use of postposed demonstratives as an innovation, the areal-genetic classification of Finnic can be placed in three concentric circles (in Güldemann's framework 2008), in which eastern Finnic and North Russian form an areal hotbed in a core circle which also includes Karelian Finnic, but this innovation does not reach western and central Finnic which lie more in the periphery (confirming the observation by Larjavaara 1986: 307–310, and Yurayong 2020). Further studies on other grammatical elements would likewise provide more evidence for tracing contact between Finnic sub-branches.

Regarding the role of language contact in the development of postposed demonstratives (or perhaps the demonstrative systems in general), the Russian language both in the present as well as in the past has been responsible for a number of changes in Finnic languages. More recently, the bipartite demonstrative paradigm with compound demonstratives in Russian has provided a model for Votic and Veps. Further back in time, the Old East Slavic (Old Novgorod and Old Rus') model with the proximal *sb* and distal *tb* could have also influenced the South Estonian system, which similarly employs *sjo* as proximal and *tuu* as distal (as speculated by Pajusalu 1996a: 150–151). This awaits further investigation, however, on tracing their ultimate etymological connection. In contrast, the functional extension of preposed demonstratives towards definiteness marking in Finnic varieties in the west does not necessarily involve language contact in a similar fashion as in the situation in the east, but the change could have well emerged from a language-internal grammaticalisation

path which subsequently became reinforced under contact with Germanic (as proposed in Metslang 2011, and Nordlund et al. 2013).

In Veps, the Russian influence continued to the area of postposed demonstrative, a grammatical element which North Russian has inherited from the Central (Moscow) dialect of Middle Russian, before introducing it to (Proto-)Veps. In intense contact with stable multilingualism, North Russian *-to* and Veps *se/ne* started to converge through polysemy copying (in the same model described by Weinreich [1953]1974, among others) and developed additional properties which significantly deviate from the Central Russian *-to*, such as the ability to inflect or co-occurrence with definite referents in an article-like manner. Given that such multifunctional postposed demonstratives are not common in other Finnic languages, I have ruled out the possibility of this phenomenon emerging from Finnic, or to be precise, from the Veps substratum in North Russian (cf. the substratum explanation in Veenker 1967: 88–90, and Kiparsky 1969: 25–26). Instead, this contact scenario rather points to a contact-induced change by mutual reinforcement (in the same vein as Leinonen 1998; Stadnik-Holzer 2006; and Kasatkina 2007, 2008), a similar contact phenomenon to what has been previously identified in the linguistic areas of Amdo-Qinghai (Janhunen 2007), the Balkans (Kopitar 1829 and Sandfeld 1926), India (Masica 1976), Japan-Korea (Yurayong & Szeto 2020), Mainland Southeast Asia (Enfield 2005), and Meso-America (Campbell et al. 1986), among others.

In the abovementioned mechanism of convergence, Veps and North Russian have been approaching a common structural model (as described in Weinreich [1953]1974) or conceptual templates in their grammar (as defined in Heine & Kuteva 2001: 410). Most notably, the constituent order shift from prenominal to postnominal determiners has been active over a relatively long period of time during the contact between eastern Finnic and North Russian (as reported by Ojanen 1985: 230–235). In regard to this matter, it would also be interesting to conduct further quantitative investigation on how extensively this syntactic shift has affected the adjectival system, and other different semantic categories of adjectives in eastern Finnic and North Russian, not to mention Saami languages spoken in Russia. This structural convergence may eventually reach the stage of intertranslatability (in Romaine's term 1988: 80–81), unless complete structural convergence would simultaneously mean language death through language shift (such as in the case study of mood and modality by Kehayov 2017).

As an epilogue to close the current study, I hope that the treatment of a language-specific and area-specific phenomenon in Northwest Russia can provide empirical evidence for relevant topics in the theoretical framework of historical, contact, and areal linguistics. Equally important, I anticipate that the results in the current study could also motivate new research questions in Uralic and Slavic linguistics, which would challenge and refine our understanding of language history and contact in the Uralic and Slavic linguistic spheres.

REFERENCES

- Abbott, Barbara. (2004). Definiteness and Indefiniteness. In Laurence R. Horn & Gregory L. Ward (Eds.), *The Handbook of Pragmatics*, 122–150. Oxford: Blackwell Publishing.
- Abondolo, Daniel. (1998). Introduction. In Daniel Abondolo (Ed.), *The Uralic Languages*, 1–42. London: Routledge.
- Adachi, Mayumi. (2016). *Betonamugo no shijishi to shijishi ni yuraisuru bunmatsushi, kandoushi* [Demonstratives, demonstrative-derived sentence-final particles, and interjections in Vietnamese] (Doctoral dissertation). Tokyo: The University of Tokyo.
- Adachi, Mayumi. (Forthcoming). Demonstratives as sentence-final particles in Vietnamese conversation. In Foong Ha Yap (Ed.), *Demonstrative in interaction* (Studies in Language).
- Agranat, Tatiana. (2015). The definite article in Votic: the process of grammaticalisation. *Journal of Estonian and Finno-Ugric Linguistics*, 6(1), 41–53.
- Ahlborn, Svetlana. (2018). The post-positive particle *-to* in the tales from Novgorod region (Paper presented at a conference Grammar of non-standard varieties in the East of the Circum-Baltic area, Tartu, Estonia, 1–3 February 2018).
- Ahlqvist, Arja. (1998). Merjalaiset - suurten järvien kansaa [The Merya - the Great Lakes people]. *Virittäjä*, 102(1), 24–55.
- Ahlqvist, Arja. (2006). Ancient Lakes in the Former Finno-Ugrian Territories of Central Russia: An Experimental Onomastic Palaeo-geographical Study. In Juhani Nuorluoto (Ed.), *The Slavicization of the Russian North: Mechanisms and Chronology* (Slavica Helsingiensia 27), 11–49. Helsinki: University of Helsinki.
- Aikhenvald, Alexandra Y. (2004). *Evidentiality*. Oxford: Oxford University Press.
- Aikhenvald, Alexandra Y. & Dixon, Robert M. W. (2001). Introduction. In Alexandra Aikhenvald & Robert M. W. Dixon (Eds.), *Areal Diffusion and Genetic Inheritance: Problems in Comparative Linguistics*, 1–26. Oxford: Oxford University Press.
- Aikio, Ante. (2004). An essay on substrate studies and the origin of Saami. In Irma Hyvärinen, Petri Kallio & Jarmo Korhonen (Eds.), *Etymologie, Entlehnungen und Entwicklungen: Festschrift für Jorma Koivulehto zum 70. Geburtstag* (Mémoires de la Société Néophilologique de Helsinki 63), 5–34. Helsinki: Société Néophilologique.
- Aissen, Judith. (2003). Differential object marking: Iconicity vs. economy. *Natural Language & Linguistic Theory*, 21(3), 435–483.
- Alatyrev, V. (Ed.). (1970). *Grammatika sovremennogo udmurtskogo jazyka: Sintaksis prostogo predloženiya* [A grammar of modern Udmurt: Syntax of the simple clause]. Iževsk: Udmurtija.
- Alhoniemi, Alho. (1988). Das Tscheremissische. In Denis Sinor (Ed.), *The Uralic Languages: Description, History and Foreign Influences*, 84–95. Leiden: Brill.
- Aljamkin, Nikolaj S. et al. (2000). *Mokšen' kjal': Morfologija* [The Moksha language: Morphology]. Saransk: Krasnyj Oktjabr'.
- Amon, Marri. (2015). *Initial and final detachments in spoken Estonian: A study in the framework of Information Structuring* (Doctoral dissertation). Tartu: Tartu Ülikooli kirjastus.

- Anderson, Stephen R. & Keenan, Edward L. (1985). Deixis. In Timothy Shopen (Ed.), *Language Typology and Syntactic Description, Vol. 3: Grammatical Categories and the Lexicon*, 259–308. Cambridge: Cambridge University Press.
- Andrews, Avery D. (2007). The major functions of the noun phrase. In Timothy Shopen (Ed.), *Language Typology and Syntactic Description, Vol. 1: Clause Structure*, 132–223. Cambridge: Cambridge University Press.
- Angere, Johannes. (1956). *Die uralo-jukagirische Frage: ein Beitrag zum Problem der sprachlichen Urverwandtschaft*. Stockholm: Almqvist & Wiksell.
- The Archives of Estonian Dialects and Kindred Languages* (Eesti murrete ja sugulaskeeelte digitaalne arhiiv). Tartu University. <http://www.murre.ut.ee/arhiiv/>.
- Arcihovskij, A. V. & Borkovskij, V. I. (1963). *Novgorodskie gramoty na bereste (iz raskopok 1956–1957. gg.)* [Novgorod birch bark documents (from excavations between 1956–1957)]. Moskva: Nauka.
- Ariel, Mira. (1988). Referring and accessibility. *Journal of Linguistics*, 24, 65–87.
- Ariel, Mira. (1990). *Accessing Noun Phrase Antecedents*. London: Routledge.
- Ariste, Paul. (1954). K voprosu o razvitii livskogo jazyka [On the question of development of the Livonian language]. *Trudy Instituta jazykoznanija AN SSSR*, 4, 254–307.
- Ariste, Paul. (1958). Vene laensõnadest vanemas eesti kirjakeeles [On Russian loanwords in Old Literary Estonian]. *Keel ja kirjandus*, 1, 25–33.
- Ariste, Paul. (1981). *Keelekontaktid: eesti keele kontakte teiste keeltega* [Language contacts: contacts of Estonian with other languages]. Tallinn: Valgus.
- Ascoli, Graziadio Isaia. (1870). *Lezioni di fonologia comparata del sanscrito, del greco e del latino*. Torino: Ermanno Loescher.
- Avanesov, R. I. (1958). Voprosy istorii russkogo jazyka v èpohu formirovanija i dal'nejšego razvitija russkoj (velikoruskoj) narodnosti [Questions of the history of the Russian language in the stage of formation and further development of the Russian nation]. In *Voprosy formirovanija russkoj narodnosti i natsii. Sbornik statej*, 155–191. Moskva: Izdatel'stvo AN SSSR.
- Avanesov, R. I., & Orlova, V. G. (Eds.). (1965). *Russkaja dialektologija* [Russian dialectology]. Moskva: Nauka.
- Babič, Vanda & Šivic-Dular, Alenka & Zor, Janez. (2003). *Učbenik stare cerkvene slovanščine* [A textbook of Old Church Slavonic]. Ljubljana: Filozofska fakulteta Univerze v Ljubljani, Oddelek za slavistiko.
- Bacz, Barbara. (1991). On some article-like uses of the demonstrative *ten* [this] in Polish. Could *ten* become an article?. *Langues et Linguistique*, 17, 1–16.
- Bakker, Peter & Matras, Yaron. (2013). Introduction. In Peter Bakker & Yaron Matras (Eds.), *Contact Languages: A Comprehensive Guide*, 1–14. Göttingen: Mouton de Gruyter.
- Barford, Paul M. (2001). *The early Slavs: Culture and society in early medieval Eastern Europe*. Ithaca, NY: Cornell University Press.
- Barnes, Jonathan. (1998). Tsuu khaa tii hla?: deixis, demonstratives and discourse particles in Lai Chin. *Linguistics of the Tibeto-Burman Area*, 21(1), 53–86.
- Bartens, Angela. (2013). Creole Languages. In Peter Bakker & Yaron Matras (Eds.), *Contact Languages: A Comprehensive Guide*, 65–158. Göttingen: Mouton de Gruyter.
- Bartens, Raija. (1996). Über die Deklinationen im Mordwinischen. *Finnisch-ugrische Forschungen*, 53(1–3), 1–113.
- Bartens, Raija. (1999). *Mordvalaiskielten rakenne ja kehitys* [Structure and development of the Mordvinic languages]. Helsinki: Société Finno-Ougrienne.

- Bartens, Raija. (2000). *Permiläisten kielten rakenne ja kehitys* [Structure and development of the Permic languages]. Helsinki: Société Finno-Ougrienne.
- Bartnik, Artur. (2015). Demonstrative or article? The case of *ten* in Polish. *Roczniki humanistyczne*, 63(11), 7–20.
- Barwise, Jon & Perry, John. (1983). *Situations and Attitudes*. Cambridge, MA: MIT Press.
- Batalova, Raisa M. (1975). *Komi-permjackaja dialektologija* [Komi-Permyak dialectology]. Moskva: Nauka.
- Becker, Laura. (2018). *Articles in the world's languages* (Doctoral dissertation). Leipzig: Universität Leipzig.
- Bereczki, Gábor. (1990). *Chrestomathia ceremissica*. Budapest: Tankönyvkiadó.
- Bereczki, Gábor. (2007). Der Sprachbund des Wolga-Kama-Gebietes. *Incontri Linguistici*, 30, 1000–1018.
- van den Berg, René. (1997). Spatial deixis in Muna (Sulawesi). In Senft, Gunter (Ed.), *Referring to space: studies in Austronesian and Papuan languages*, 197–220. Oxford: Clarendon Press.
- Berger, Tilman. (1999). Die Gebrauchsbedingungen des bestimmten Artikels im älteren Obersorbischen. *Lětopis*, 46, 7–23.
- Bergsland, Knut. (1950). Remarques sur les pronoms démonstratifs lapons. In *Commentationes Fenno-Ugricae in honorem Y. H. Toivonen* (Mémoires de la Société Finno-Ougrienne 98), 27–35. Helsinki: Société Finno-Ougrienne.
- Bhat, D. N. S. (2013). Third Person Pronouns and Demonstratives. In Matthew S. Dryer & Martin Haspelmath (Eds.), *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. Retrieved 02-02-2020, from <http://wals.info/chapter/43>.
- Biber, Douglas & Finegan, Edward. (1989). Styles of stance in English: Lexical and grammatical marking of evidentiality and affect. *Text-interdisciplinary journal for the study of discourse*, 9(1), 93–124.
- Bickel, Balthasar & Nichols, Johanna. (2006). Oceania, the pacific rim, and the theory of linguistic areas. In *Proceedings of the annual meeting of the Berkeley linguistics society*, 32, 3–15.
- Birnbaum, Henrik. (1979). *Common Slavic: progress and problems in its reconstruction*. Bloomington: Slavica Publishers.
- Birner, Betty J. (2006). Inferential relations and noncanonical word order. In Betty J. Birner & Gregory Ward (Eds.), *Drawing the Boundaries of Meaning: Neo-Gricean Studies in Pragmatics and Semantics in Honor of Laurence R. Horn*, 31–51. Amsterdam: John Benjamins Publishing Company.
- Birner, Betty J. & Ward, Gregory. (1998). *Information Status and Noncanonical Word Order in English*. Amsterdam: John Benjamins Publishing Company.
- Bjørnflaten, Jan Ivar. (2006). Chronologies of the Slavicization of Northern Russia Mirrored by Slavic Loanwords in Finnic and Baltic. In Juhani Nuorluoto (Ed.), *The Slavicization of the Russian North: Mechanisms and Chronology* (Slavica Helsingiensia 27), 50–77. Helsinki: University of Helsinki.
- Blokland, Rogier. (2005). *The Russian loanwords in literary Estonian*. Wiesbaden: Harrassowitz Verlag.
- Blokland, Rogier. (2012). Borrowability of pronouns: Evidence from Uralic. *Finnisch-Ugrische Mitteilungen*, 35, 1–34.
- Blommaert, Jan. (2010). *The sociolinguistics of globalization*. Cambridge: Cambridge University Press.

- Bolden, Galina. B. (2008). Reopening Russian conversations: The discourse particle *-to* and the negotiation of interpersonal accountability in closings. *Human Communication Research*, 34 (1), 99–136.
- Bondarenko, O. E. & Horunžaja, T. M. & Cypanov, E. A. (2000). *Komi nem* [Komi language]. Syktyvkar: Komi nebög ledžanin.
- Bonnot, Christine. (1986). Emplois de la particule *-to*. In *Les Particules Enonciatives en Russe Contemporain, Vol. 1*, 21–30. Paris: Institut d'études slaves.
- Bonnot, Christine. (1987). *-to* particule de rappel et de thématization. In *Les particules énonciatives en russe contemporain, Vol. 2*, 113–171. Paris: Institut d'études slaves.
- Bonnot, Christine. (1990). La particule *-to* et la polémique cachée en russe moderne: à propos du statut énonciatif du thème. *Revue des études slaves*, 62(1/2), 67–75.
- Bonnot, Christine. (1991). La particule de thématization *-to* en russe moderne. *Revue des études slaves*, 63(4), 853–861.
- Bopp, Franz. (1827). Über J. Grimms Deutsche Grammatik II. In *Berliner Jahrbuch für wissenschaftliche Kritik February and May*, 251–303, 725–759. Göttingen: Dieterichsche Buchhandlung.
- Breu, Walter. (2002). Der definite Artikel in der obersorbischen Umgangssprache. In Marion Krause & Christian Sappock (Eds.), *Slavistische Linguistik 2002*, 9–57. München: Verlag Otto Sagner.
- Brosig, Benjamin & Gegentana & Yap, Foong Ha. (2018). Evaluative uses of postnominal possessives in Central Mongolian. *Journal of Pragmatics*, 135, 71–86.
- Brugmann, Karl. (1904). *Demonstrativpronomina der indogermanischen Sprachen*. Leipzig: Teubner.
- Brunelle, Marc & Hân, Phú Vãn. (2019). Colloquial Eastern Cham. In Alice Vittrant & Justin Watkins (Eds.), *The Mainland Southeast Asia Linguistic Area*, 523–558. Berlin: Walter de Gruyter.
- Bryant, David & Vincent Moulton. (2004). Neighbor-Net: An agglomerative method for the construction of phylogenetic networks. *Molecular Biology and Evolution*, 21(2), 255–265.
- Buchler, Justus. (Ed.). (1940). *Philosophical Writings of Peirce*. London: Routledge & Kegan Paul.
- Bühler, Karl. (1934). *Sprachtheorie: Die Darstellungsfunktion der Sprache*. Jena: Fischer.
- Bunt, Harry. (2017). Computational Pragmatics. In Yan Huang (Ed.), *The Oxford Handbook of Pragmatics*, 326–345. Oxford: Oxford University Press.
- Bybee, Joan L. & Perkins, Revere Dale & Pagliuca, William. (1994). *The evolution of grammar: Tense, aspect, and modality in the languages of the world*. Chicago: University of Chicago Press.
- Campbell, Lyle. (1998). *Historical Linguistics. An Introduction*. Edinburgh: Edinburgh University Press.
- Campbell, Lyle. (2000). What's wrong with grammaticalization?. *Language sciences*, 23(2–3), 113–161.
- Campbell, Lyle & Kaufman, Terrence & Smith-Stark, Thomas C. (1986). *Meso-America as a linguistic area*. *Language*, 62(3), 530–570.
- Carlson, Gregory. (2004). Reference. In Laurence R. Horn & Gregory L. Ward (Eds.), *The Handbook of Pragmatics*, 74–96. Oxford: Blackwell Publishing.
- Castrén, Matthias Alexander. (1839). *Dissertatio academica de affinitate declinationum in lingua fennica, esthonica et lapponica*. Helsinki: Frenckell.

- Castrén, Matthias Alexander. (1850). *De affixis personalibus linguarum altaicarum*. Helsinki: Frenckell.
- Castrén, Matthias Alexander. (1854). *Nordische Reisen und Forschungen: Grammatik der samojedischen Sprachen*. Saint Petersburg: Buchdruckerei der Kaiserlichen Akademie der Wissenschaften.
- Cegłowski, Piotr & Tajsner, Przemysław. (2006). Topicalisation and object fronting in Polish: A view from a minimalist perspective. In Katarzyna Dziubalska-Kołaczyk (Ed.), *IFAtuation: A Life in IFA: A Festschrift for Professor Jacek Fisiak on the Occasion of His 70th Birthday*, 99–129. Poznań: Adam Mickiewicz University.
- Chafe, Wallace. (1976). Givenness, contrastiveness, definiteness, subjects, topics, and point of view. In Charles N. Li (Ed.), *Subject and topic*, 25–55. New York, NY: Academic Press.
- Chesterman, Andrew. (1991). *On definiteness: A study with special reference to English and Finnish*. Cambridge: Cambridge University Press.
- Chomsky, Noam. (1965). *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.
- Christoffersen, Paul. (1939). *The Articles: A Study of Their Theory and Use in English*. Copenhagen: Munksgaard.
- Clark, Herbert H. (1975). Bridging. In Roger C. Schank & Bonnie L. Nash-Webber (Eds.), *Theoretical issues in natural language processing*, 169–174. New York, NY: Association for Computing Machinery.
- Clark, Herbert H. & Marshall, Catherine R. (1981). Definite reference and mutual knowledge. In Aravind K. Joshi, Bonnie L. Webber & Ivan A. Sag (Eds.), *Elements of discourse understanding*, 10–63. Cambridge: Cambridge University Press.
- Cleary-Kemp, Jessica. (2007). Universal uses of demonstratives: Evidence from four Malayo-Polynesian languages. *Oceanic Linguistics*, 46(2), 325–347.
- Collinder, Björn. (1957). *Survey of the Uralic languages*. Stockholm: Almqvist & Wiksell.
- Collinder, Björn. (1960). *Comparative grammar of the uralic languages*. Stockholm: Almqvist & Wiksell.
- Comrie, Bernhard. (1978). Definite direct objects and referent identification. *Pragmatics Microfiche*, 3.1, 1–19.
- Comrie, Bernard & Corbett, Greville. (1993). *The Slavonic Languages*. London: Routledge.
- Craig, Colette G. (1991). Ways to Go in Rama: A case study in polygrammaticalization. In Elizabeth C. Traugott & Bernd Heine (Eds.), *Approaches to Grammaticalization, Vol. 2: Focus on types of grammatical markers*, 455–492. Amsterdam: John Benjamins Publishing Company.
- Csúcs, Sándor. (1988). Die wotjakische Sprache. In Denis Sinor (Ed.), *The Uralic Languages: Description, History and Foreign Influences*, 131–146. Leiden: Brill.
- Csúcs, Sándor. (1998). Udmurt. In Daniel Abondolo (Ed.), *The Uralic languages*, 276–304. Routledge: London.
- Cummins III, George M. (1998). Definiteness in Czech. *Studies in Language*, 22(3), 567–596.
- Curl, Traci S. (1999). The Lakhota definite articles and topic marking. *Kansas Working Papers in Linguistics, Studies in Native American Linguistics*, 24(2), 1–16.
- Curnow, Timothy Jowan. (2001). What Language Features Can Be ‘Borrowed’?. In Alexandra Aikhenvald & Robert M. W. Dixon (Eds.), *Areal Diffusion and Genetic*

- Inheritance: Problems in Comparative Linguistics*, 412–436. Oxford: Oxford University Press.
- Cygankin, D. V. et al. (2000). *Èrzjan' kel': Morfemika, valon' teevema dy morfologija* [Erzya language: morphemic, word formation and morphology]. Saransk: Krasnyj Oktjabr'.
- Dahl, Östen & Kortmann, Bernd. (2004). Definite articles in Scandinavian: Competing grammaticalization processes in standard and non-standard varieties. In Bernd Kortmann (Ed.), *Dialectology meets Typology: Dialect Grammar from a Cross-Linguistic Perspective*, 147–180. Berlin: Mouton de Gruyter.
- Dalrymple, Mary & Nikolaeva, Irina. (2011). *Objects and information structure*. Cambridge: Cambridge University Press.
- Darwin, Charles. (1859). *On the origin of species*. London: John Murray.
- Darwin, Charles. (1871). *The Descent of Man*. London: John Murray.
- Dasinger, Lisa Kay. (1995). *The Development of Discourse Competence in Finnish Children: The expression of definiteness* (Doctoral dissertation). Berkeley, CA: University of California, Berkeley.
- Dekker, Simeon. (2018). *Old Russian birchbark letters: a pragmatic approach*. Leiden: Brill.
- DeLancey, Scott. (1994). Re: 5.1393 Using the comparative method in syntax. Posting 5.1433, message 1, to LINGUIST (7 December 1994).
- DeLancey, Scott. (1997). Mirativity: The grammatical marking of unexpected information. *Linguistic Typology*, 1, 33–52.
- DeLancey, Scott. (2001). The mirative and evidentiality. *Journal of pragmatics*, 33(3), 369–382.
- Derksen, Rick. (2008). *Etymological dictionary of the Slavic inherited lexicon*. Leiden: Brill.
- Dik, Simon C. (1989). *The theory of functional grammar. Part I: The structure of the clause*. Dordrecht: Foris Publications.
- Diessel, Holger. (1999a). *Demonstratives: Form, Function and Grammaticalization*. Amsterdam: John Benjamins Publishing Company.
- Diessel, Holger. (1999b). The morphosyntax of demonstratives in synchrony and diachrony. *Linguistic Typology*, 3, 1–49.
- Diessel, Holger. (2013a). Distance Contrasts in Demonstratives. In Matthew S. Dryer & Martin Haspelmath (Eds.), *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. Retrieved 02-02-2020, from <http://wals.info/chapter/41>.
- Diessel, Holger. (2013b). Where does language come from? Some reflections on the role of deictic gesture and demonstratives in the evolution of language. *Language and Cognition*, 5(2–3), 239–249.
- Dixon, Robert M. W. (1997). *The rise and fall of languages*. Cambridge: Cambridge University Press.
- Dixon, Robert M. W. (2003). Demonstratives: A cross-linguistic typology. *Studies in Language*, 27(1), 61–112.
- Dobrovie-Sorin, Carmen. (2013). *A Reference Grammar of Romanian, Vol. 1, the noun phrase*. Amsterdam: John Benjamins Publishing Company.
- Dobrovský, Josef. (1792). *Geschichte der böhmischen Sprache und Litteratur*. Prag: Bey Johann Gottfried Calve.
- Dobrovský, Josef. (1818). *Geschichte der böhmischen Sprache und Litteratur* (2nd ed.). Prag: Gottlieb Haase.
- Dobrovský, Josef. (1822). *Institutiones linguae slavicae dialecti veteris*. Vindobonae: sumptibus et typis Antonii Schmid.

- Dolgova, Irina & Maksimova, Elena. (1996). Deixis in Time and Space, The Fate of the Russian Demonstrative *sej*. In Edna Andrews & Yishai Tobin (Eds.), *Toward a Calculus of Meaning: Studies in Markedness, Distinctive Features and Deixis*, 217–233. Amsterdam: John Benjamin Publishing Company.
- Dryer, Matthew S. (2013a). Order of Demonstrative and Noun. In Matthew S. Dryer & Martin Haspelmath (Eds.), *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. Retrieved 02-02-2020, from <http://wals.info/chapter/88>.
- Dryer, Matthew S. (2013b). Definite Articles. In Matthew S. Dryer & Martin Haspelmath (Eds.), *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. Retrieved 02-02-2020, from <http://wals.info/chapter/37>.
- Du Bois, John W. (1980). Beyond definiteness: The trace of identity in discourse. In Wallace L. Chafe (Ed.), *The pear stories: Cognitive, cultural, and linguistic aspects of narrative production* (Advances in discourse processes 3), 203–274. Norwood, NJ: Ablex.
- Du Bois, John W. (2007). The stance triangle. In Robert Englebretson (Ed.), *Stancetaking in discourse: Subjectivity, evaluation, interaction*, 139–182. Amsterdam: John Benjamins Publishing Company.
- Duvallon, Outi. (2014). Le marqueur discursif finnois *-hAn*: double repérage énonciatif. *Faits de langues*, 43(1), 87–114.
- Duvallon, Outi & Peltola, Rea. (2012). Prises de distance en finnois: le mode jussif et le marqueur discursif *-hAn*. *Linx. Revue des linguistes de l'université Paris X Nanterre*, 66–67, 101–116.
- EDC = *Estonian Dialect Corpus* (Eesti murrete korpus). University of Tartu. <http://www.murre.ut.ee/estonian-dialect-corpus/>.
- Egedi, Barbara. (2014). The DP-cycle in Hungarian and the functional extension of the noun phrase. In Katalin É. Kiss (Ed.), *The Evolution of Functional Left Peripheries in Hungarian Syntax*, 56–82. Oxford: Oxford University Press.
- Enfield, Nicholas J. (2003). The definition of what-d'you-call-it: Semantics and pragmatics of recognitional deixis. *Journal of Pragmatics*, 35(1), 101–117.
- Enfield, Nicholas. J. (2005). Areal linguistics and mainland Southeast Asia. *Annual Review of Anthropology*, 34, 181–206.
- Englebretson, Robert. (2003). *Searching for structure: The problem of complementation in colloquial Indonesian conversation*. Amsterdam: John Benjamins Publishing Company.
- Englebretson, Robert. (2007). Grammatical resources for social purposes: Some aspects of stancetaking in colloquial Indonesian conversation. In Robert Englebretson (Ed.), *Stancetaking in discourse: Subjectivity, evaluation, interaction*, 69–110. Amsterdam: John Benjamins Publishing Company.
- Epstein, Richard. (1997). Viewpoint and the definite article. In Adele E. Goldberg (Ed.), *Conceptual structure, discourse and language*, 99–112. Stanford: CSLI.
- Epstein, Richard. (2001). The definite article, accessibility, and the construction of discourse referents. *Cognitive linguistics*, 12(4), 333–378.
- Eriksen, Love. (2011). Nature and culture in prehistoric Amazonia using G.I.S. to reconstruct ancient ethnogenetic processes from archaeology, linguistics, geography, and ethnohistory (Doctoral dissertation). Lund: Lunds universitet.
- Ernanda & Yap, Foong Ha. (Forthcoming). A phrasal alternation test for disambiguating referential and non-referential uses of versatile demonstratives: The case of (*i*)*neh* in Kerinci Malay. In Foong Ha Yap (Ed.), *Demonstrative in interaction* (Studies in Language).

- Erteschik-Shir, Nomi. (1986). Wh-questions and focus. *Linguistics and Philosophy*, 9(2), 117–149.
- Erteschik-Shir, Nomi. (1997). *The Dynamics of Focus Structure*. Cambridge: Cambridge University Press.
- Erteschik-Shir, Nomi. (2007). *Information structure: The syntax-discourse interface*. Oxford: Oxford University Press.
- Etelämäki, Marja. (2006). *Toiminta ja tarkoite: Tutkimus suomen pronominista tämä* [Action and referent: A study of Finnish pronoun *tämä*] (Doctoral dissertation). Helsinki: University of Helsinki.
- Etelämäki, Marja. (2009). The Finnish demonstrative pronouns in light of interaction. *Journal of Pragmatics*, 41(1), 25–46.
- Evans, Nicholas & Bergqvist, Henrik & San Roque, Lila. (2018a). The grammar of engagement I: Framework and initial exemplification. *Language and Cognition*, 10(1), 110–140.
- Evans, Nicholas & Bergqvist, Henrik & San Roque, Lila. (2018b). The grammar of engagement II: Typology and diachrony. *Language and Cognition*, 10(1), 141–170.
- Evans, Nicholas & Levinson, Stephen C. (2009). The myth of language universals: Language diversity and its importance for cognitive science. *Behavioral and brain sciences*, 32(5), 429–448.
- Feist, Timothy R. (2015). A grammar of Skolt Saami (Mémoires de la Société Finno-Ougrienne 273). Helsinki: Société Finno-Ougrienne.
- Feoktistov, Alexandr P. (1963). *Kategorija pritjažatel'nosti v mordovskix jazykax* [The category of possessivity in Mordvinic languages]. Saransk: Mordovskoe knižnoe izdatel'stvo.
- Fernandez-Vest, M. M. Jocelyne (2009). Typological evolution of Northern Sami: spatial cognition and Information Structuring. In Jussi Ylikoski (Ed.), *The Quasiquicentennial of the Finno-Ugrian Society* (Mémoires de la Société Finno-Ougrienne 258), 33–55. Helsinki: Société Finno-Ougrienne.
- Filkova, Penka D. (1973). *Istorija russkogo literaturnogo jazyka XI – XVIII vv.* [History of Literary Russian between the 11th–18th centuries]. Sofia: Nauka i iskusstvo.
- Fischer, Andreas. (2003). Lexical Borrowing and the History of English: A Typology of Typologies. In Dieter Kastovsky & Arthur Mettinger (Eds.), *Language Contact in the History of English*, 97–115. Frankfurt am Main: Peter Lang.
- Fisher, Ronald A. (1922). On the interpretation of χ^2 from contingency tables, and the calculation of p . *Journal of the Royal Statistical Society*, 85(1), 87–94.
- Fox, Anthony. (1995). *Linguistic Reconstruction: An Introduction to Theory and Method*. Oxford: Oxford University Press.
- Fraser, Bruce. (1990). Perspectives on politeness. *Journal of pragmatics*, 14(2), 219–236.
- Fraurud, Kari. (2001). Possessives with extensive use. In Irène Baron, Michael Herslund & Finn Sørensen (Eds.), *Dimensions of possession*, 243–268. Amsterdam: John Benjamins Publishing Company.
- Frey, Werner. (2001). About the whereabouts of indefinites. *Theoretical Linguistics*, 27, 137–161.
- Frog & Saarikivi, Janne. (2015). De situ linguarum fennicarum aetatis ferreae, Pars I. *The Retrospective Methods Network Newsletter*, 9, 64–115.
- Fritts, Harold C. (1976). *Tree rings and climate*. London: Academic Press.

- Gebaura, J. & Gebauer, J. (1874). O příčinách proměn jazykových, zvláště slovanských [On the causes of language changes, especially in Slavic]. *Listy filologické a paedagogické*, 1(1), 43–56.
- Gerd, A. & Savijärvi, M. & de Graaff, T. (2002). *Jazyk i narod: teksty i komentarii razgovornoj reči pribaltijsko-finskijh jazykov i dialektov ruskogo jazyka na Severo-zapade Rossii* [Language and nation: texts and commentaries on Finnic and Russian dialects in Northwest Russia]. Saint Petersburg.
- Gerland, Doris. (2014). Definitely not possessed? Possessive suffixes with definiteness marking function. In Thomas Gamerschlag, Doris Gerland, Rainer Osswald & Wiebke Petersen (Eds.), *Frames and concept types*, 269–292. Cham: Springer.
- Gernsbacher, Morton Ann & Robertson, Rachel R. W. (2002). The definite article *the* as a cue to map thematic information. In Max Louwerse & Willie van Peer (Eds.), *Thematics: Interdisciplinary studies*, 119–136. Amsterdam: John Benjamins Publishing Company.
- Gibbs, Raymond W. (2017). Experimental pragmatics. In Yan Huang (Ed.), *The Oxford Handbook of Pragmatics*, 310–325. Oxford: Oxford University Press.
- van Gijn, Rik. (2014). The Andean foothills and adjacent Amazonian fringe. In *The native languages of South America. origins, development, typology*, 102–125. Cambridge: Cambridge University Press.
- van Gijn, Rik & Muysken, Pieter. (In press). Highland-lowland language interactions. In *Rethinking the andes-amazonia 'divide': a cross-disciplinary exploration*. London: UCL Press.
- van Gijn, Rik & Wahlström, Max. (Forthcoming). Linguistic areas. In Anja Hasse, Hanna Ruch, Rik van Gijn & Max Wahlström (Eds.), *Language contact: Bridging the gap between individual interactions and areal patterns*.
- Gippius, A. A. (1997). К истории сложения текста Новгородской первой летописи [Towards the history of writing texts of the First Novgorod Chronicle]. *Novgorodskij istoričeskij sbornik*, 6(16), 3–72.
- Givón, Talmy. (1983). An Introduction. In Talmy Givón (Ed.), *Topic Continuity in Discourse: A Quantitative Cross-language study*, 1–42. Amsterdam: John Benjamins Publishing Company.
- Givón, Talmy. (2001). *Syntax, Vol. I*. Amsterdam: John Benjamins Publishing Company.
- Givón, Talmy, & Li, Charles N. (1976). *Topic, Pronoun and Grammatical Agreement*. In Charles N. Li (Ed.), *Subject and Topic*, 149–188. New York, NY: Academic Press.
- Gooskens, Charlotte. (2005). Traveling time as a predictor of linguistic distance. *Dialectologia et Geolinguistica*, 13, 38–62.
- Gorškova, K. V. (1972). *Istoričeskaja dialektologija ruskogo jazyka* [Historical dialectology of the Russian language]. Moskva: Prosveščenie.
- Greenberg, Joseph H. (Ed.). (1963). *Universals of language*. Cambridge, MA: MIT Press.
- Greenberg, Joseph H. (1978). How does a language acquire gender markers?. In Joseph H. Greenberg, Charles A. Ferguson & Edith A. Moravcsik (Eds.), *Universals of human language 3. Word structure*, 47–82. Stanford: Stanford University Press.
- Grenoble, Lenore. A. (1998). *Deixis and information packaging in Russian discourse*. Amsterdam: John Benjamins Publishing.
- Grønbech, Kaare. ([1936/1979]1997). *The Structure of the Turkic Languages* (Uralic and Altaic Series 136). Richmond: Curzon Press Ltd.

- Grković-Major, Jasmina. (2011). The development of predicative possession in Slavic languages. In Motoki Nomachi (Ed.), *The Grammar of Possessivity in South Slavic Languages: Synchronic and Diachronic Perspectives* (Slavic Eurasian Studies 24), 35–54. Sapporo: Slavic Research Center, Hokkaido University.
- Grünthal, Riho. (1997). *Livistä liiviin: Itämerensuomalaiset etnonymit* [From Livvi to Livonian: Finnic ethnonyms] (Castrenianumin toimitteita 51). Helsinki: Société Finno-Ougrienne.
- Grünthal, Riho. (2015a). *Vepsän kielioppi* [A grammar of Veps]. Helsinki: Société Finno-Ougrienne.
- Grünthal, Riho. (2015b). Livonian at the crossroads of language contacts. In Santeri Junttila (Ed.), *Contacts between the Baltic and Finnic languages* (Uralica Helsingiensia 7), 97–150. Sastamala: Vammalan Kirjapaino Oy.
- Güldemann, Tom. (2008). The Macro-Sudan Belt: Towards Identifying a Linguistic Area in Northern Sub-Saharan Africa. In Bernd Heine & Derek Nurse (Eds.), *A Linguistic Geography of Africa*, 151–185. Cambridge: Cambridge University Press.
- Gundel, Jeanette K. (1978). Stress, pronominalization and the given-new distinction. *University of Hawai'i Working Papers in Linguistics*, 10, 1–13.
- Gundel, Jeanette K. (1985). “Shared knowledge” and topicality. *Journal of Pragmatics*, 9, 83–107.
- Gundel, Jeanette K. (1988). Universals of topic–comment structure. In Michael Hammond, Edith Moravcsik & Jessica R. Wirth (Eds.), *Studies in Syntactic Typology*, 109–239. Amsterdam: John Benjamins Publishing Company.
- Gundel, Jeanette K. (1999). On different kinds of focus. In Bosch and van der Sandt (Eds.), *Focus: Linguistic, cognitive, and computational perspectives*, 293–305. Cambridge: Cambridge University Press.
- Gundel, Jeanette K. & Fretheim, Thorstein. (2004). Topic and focus. In Laurence R. Horn & Gregory L. Ward (Eds.), *The Handbook of Pragmatics*, 175–196. Oxford: Blackwell Publishing.
- Gvozdanović, Jadranka. (2019). Standard Average European Revisited in the Light of Slavic Evidence. In Andrii Danylenko & Motoki Nomachi (Eds.), *Slavic on the Language Map of Europe*, 113–144. Berlin: Mouton de Gruyter.
- Haarmann, Harald. (1977). Arealtypologie der finnisch-ugrischen Sprachen. *Finnisch-ugrische Mitteilungen*, 1, 3–14.
- Haiman, John. (1987). Conditionals are topics. *Language*, 54(3), 564–589.
- Hajdú, Péter. (1982). *Chrestomathia Samoiedica*. Budapest: Tankönyvkiadó.
- Häkkinen, Jaakko. (2012). After the protolanguage: Invisible convergence, false divergence and boundary shift. *Finnisch-Ugrische Forschungen*, 61(1–3), 7–28.
- Häkkinen, Kaisa. (1984). Wäre es schon an der Zeit, den Stammbaum zu fällen? Theorien über die gegenseitigen Verwandtschaftsbeziehungen der finnisch-ugrischen Sprachen. *Ural-Altaische Jahrbücher*, 32, 1–4.
- Hakulinen, Auli. (1985). On cohesive devices in Finnish. In Emel Sözer (Ed.), *Text Connexity, Text Coherence: Aspects, Methods, Results*, 337–362. Hamburg: Helmut Buske Verlag.
- Hakulinen, Auli & Korhonen, Riitta & Vilks, Maria & Koivisto, Vesa. (2004). *Iso suomen kielioppi* [The large grammar of Finnish]. Helsinki: Finnish Literature Society.
- Hakulinen, Lauri. (1968). *Suomen kielen rakenne ja kehitys* [Structure and development of the Finnish Language] (3rd ed.). Helsinki: Otava.
- Hakulinen, Lauri. (1999). *Luennot suomen kielen partikkeleista* [Lectures on Finnish particles]. Yrjö Lauranto & Tapani Lehtinen (Eds.). Helsinki: University of Helsinki.

- Halliday, Michael A. K. (1966). Intonation system in English. In Angus McIntosh & Michael A. K. Halliday (Eds.), *Patterns of Language: Papers in General, Descriptive and Applied Linguistics*, 111–133. London: Longman.
- Halliday, Michael A. K. (1994). *An Introduction to Functional Grammar* (2nd ed.). London: Edward Arnold.
- Halm, Tamás. (2018). From possessive suffix to affective demonstrative suffix in Hungarian: a grammaticalization analysis. *Morphology*, 28(4), 359–396.
- Harris, Alice C. & Campbell, Lyle. (1995). *Historical Syntax in Cross-Linguistic Perspective*. Cambridge: Cambridge University Press.
- Harrison, Sheldon P. (1976). *Mokilese Reference Grammar*. Honolulu: University of Hawai'i Press.
- Harrison, Sheldon. P. (2003). On the limits of the comparative method. In Brian D. Joseph & Richard D. Janda (Eds.), *The handbook of historical linguistics*, 213–243. Oxford: Blackwell Publishing.
- Haspelmath, Martin. (1998). How young is Standard Average European? *Language sciences*, 20(3), 271–287.
- Haspelmath, Martin. (2001). The European linguistic area: Standard Average European. In Martin Haspelmath (Ed.), *Language typology and language universals* (Handbücher zur Sprach- und Kommunikationswissenschaft), 1492–1510. Berlin: Mouton de Gruyter.
- Haspelmath, Martin. (2010). Comparative concepts and descriptive categories in crosslinguistic studies. *Language*, 86(3), 663–687.
- Haspelmath, Martin. (2019). Differential place marking and differential object marking. *Sprachtypologie und Universalienforschung*, 72(3), 313–334.
- Hausenberg, Anu-Reet. (1998). Komi. In Daniel Abondolo (Ed.), *The Uralic Languages*, 305–326. London: Routledge.
- Hayashi, Makoto & Yoon, Kyung-eun. (2006). A cross-linguistic exploration of demonstratives in interaction. With particular reference to the context of word-formulation trouble. *Studies in Language*, 30(3), 485–540.
- Hawkins, John A. (1978). *Definiteness and Indefiniteness. A Study in Reference and Grammaticality Prediction*. London: Croom Helm.
- Heine, Bernd. (2003). Grammaticalization. In Brian D. Joseph & Richard D. Janda (Eds.), *The handbook of historical linguistics*, 575–601. Oxford: Blackwell Publishing.
- Heine, Bernd & Kuteva, Tania. (2001). Convergence and divergence in the development of African languages. In Alexandra Aikhenvald & Robert M. W. Dixon (Eds.), *Areal Diffusion and Genetic Inheritance: Problems in Comparative Linguistics*, 393–411. Oxford: Oxford University Press.
- Heine, Bernd & Kuteva, Tania. (2002). *World lexicon of grammaticalization*. Cambridge: Cambridge University Press.
- Heine, Bernd & Kuteva, Tania. (2005). *Language Contact and Grammatical Change*. Cambridge: Cambridge University Press.
- Helmski, Eugene. (1986). O pribaltijsko-finskom jazykovim materiale v novgorodskih beresjanyh gramotah. In Valentin L. Janin & Andrej A. Zaliznjak (Eds.), *Novgorodskie gramoty na bereste iz raskopok 1977–1983 godov. Kommentarij i slovoukazatel' k berestjanyh gramotam (iz raskopok 1951–1983 gg.)*, 252–259. Moskva.
- Helmski, Eugene. (1998). Nganasan. In Daniel Abondolo (Ed.), *The Uralic Languages*, 480–515. London: Routledge.
- Helmski, Eugene. (2003). Areal groupings (Sprachbünde) within and across the borders of the Uralic language family: A survey. *Nyelvtudományi közlemények*, 100, 156–167.

- Helinski, Eugene. (2006). The “Northwestern” Group of Finno-Ugric Languages and its Heritage in the Place Names and Substratum Vocabulary of the Russian North. In Juhani Nuorluoto (Ed.), *The Slavization of the Russian North: Mechanisms and Chronology* (Slavica Helsingiensia 27), 109–127. Helsinki: University of Helsinki.
- Hesselbäck, André. (2001). *Tatariska och tjuvasjiska kodkopior i mariskan* [Tatar and Chuvash code-copies in Mari] (Doctoral dissertation). Uppsala: Uppsala universitet.
- Himmelmann, Nikolaus. P. (1996). Demonstratives in narrative discourse. In Barbara A. Fox (Ed.), *Studies in anaphora*, 201–254. Amsterdam: John Benjamins Publishing Company.
- Himmelmann, Nikolaus. P. (1997). *Deiktikon, Artikel, Nominalphrase: Zur Emergenz syntaktischer Struktur*. Tübingen: Max Niemeyer Verlag.
- Hint, Helen & Nahkola, Tiina & Pajusalu, Renate. (2017). With or without articles? A comparison of article-like determiners in Estonian and Finnish. *Lähivõrdlusi*, 27, 65–106.
- Honti, László. (1988). Die wogulische Sprache & Die ostjakische Sprache. In Denis Sinor (Ed.), *The Uralic Languages: Description, History and Foreign Influences*, 147–196. Leiden: Brill.
- Hopper, Paul. J. & Traugott, Elizabeth C. (1993). *Grammaticalization*. Cambridge: Cambridge University Press.
- Hornborg, Alf. (2005). Ethnogenesis, regional integration, and ecology in prehistoric amazonia: toward a system perspective. *Current Anthropology*, 46(4), 589–620.
- Huang, Yan. (2000). *Anaphora: A cross-linguistic approach*. Oxford: Oxford University Press.
- Hunston, Susan & Thompson, Geoffrey. (2000). Evaluation: An introduction. In Susan Hunston & Geoffrey Thompson (Eds.), *Evaluation in text: Authorial stance and the construction of discourse: Authorial stance and the construction of discourse*, 1–27. Oxford: Oxford University Press.
- Huson, Daniel H. & David Bryant. (2006). Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution*, 23(2), 254–267.
- Huszcza, Ryszard. (2000). Nie ma żeby nie było – o segmentalnych wykładnikach tematycznorematicznej struktury zdania w polszczyźnie [On segmental exponents of theme-rheme structure in Polish]. *Poradnik językowy*, 8, 1–9.
- Huomo, Tuomas & Sivonen, Jari. (2020). Suurtyö kielen kohdennusjärjestelmässä [A great work on the targeting system of language]. *Virittäjä*, 124(1).
- Ihsane, Tabea & Puskás, Genoveva. (2001). Specific is not definite. *GG@ G*, 2, 39–54.
- Itkonen, Erkki. (1966). *Kieli ja sen tutkimus* [Language and its research]. Porvoo: WSOY.
- Itkonen, Terho. (1958). *Koltan- ja kuolanlapin sanakirja/Wörterbuch des kolta- und kotalappischen, Vol. 1* (Lexica Societatis Fenno-Ugricae 15). Helsinki: Société Finno-Ougrienne.
- Itkonen, Terho. (1971). Aunuksen äänneopin erikoispiirteet ja aunukselaismurteiden synty [Some characteristics of Olonetsian phonology and the origin of Olonetsian dialects]. *Virittäjä*, 75(2), 153–185.
- Itkonen, Terho. (1998). Reflections on pre-Uralic and the “Saami-Finnic protolanguage”. *Finnisch-Ugrische Forschungen*, 54(3), 229–266.
- Ivanov, Valerij V. (1990). Genealogičeskaja klassifikacija jazykov [Genealogical classification of languages]. In Viktorija N. Jarceva (Ed.), *Lingvističeskij ènciklopedičeskij slovar’*. Moskva: Sovetskaja ènciklopedija.

- Iwasaki, Shoichi & Ingkaphirom, Preeya. (2005). *A reference grammar of Thai*. Cambridge: Cambridge University Press.
- Iwasaki, Shoichi & Dechapratumwan, Parada. (Forthcoming). Versatility of Thai demonstratives. In Foong Ha Yap (Ed.), *Demonstrative in interaction* (Studies in Language).
- Jackendoff, Ray S. (1972). *Semantic interpretation in generative grammar*. Cambridge, MA: MIT Press.
- Jacobs, Joachim. (1983). *Fokus und Skalen. Zur Syntax und Semantik der Gradpartikeln im Deutschen*. Tübingen: Niemeyer.
- Jacobs, Joachim. (1988). Fokus-Hintergrund-Gliederung und Grammatik. In Hans Altmann (Ed.), *Intonationsforschungen*, 89–134. Tübingen: Niemeyer.
- Jakobson, Roman. (1931). Über die phonologischen Sprachbünde. *Travaux du cercle linguistique de Prague*, 4, 234–240.
- Janda, Gwen E. (2015). Northern Mansi possessive suffixes in non-possessive function. *Journal of Estonian and Finno-Ugric Linguistics*, 6(2), 243–258.
- Janda, Richard D. (1995). From agreement affix to subject “clitic” and bound root: *-mos* > *-nos* vs. *(-)nos(-)* and *nos-otros* in New Mexican and other regional Spanish dialects. *Papers from the Parasession on Clitics*, 118–139. Chicago: Chicago Linguistic Society.
- Janda, Richard D. (2000). Beyond “pathways” and “unidirectionality”: On the discontinuity of language transmission and the counterability of grammaticalization. *Language sciences*, 23(2–3), 265–340.
- Janhunen, Juha. (1977). Samoyed-Altaic Contacts. Present state of research. In *Altaica: Proceedings of the 19th annual meeting of the Permanent International Altaistic Conference, held in Helsinki, 7–11 June 1976* (Mémoires de la Société Finno-Ougrienne 158), 123–129. Helsinki: Société Finno-Ougrienne.
- Janhunen, Juha. (1982). On the structure of Proto-Uralic. *Finnisch-Ugrische Forschungen*, 44, 23–42.
- Janhunen, Juha. (2001). Indo-Uralic and Ural-Altaic: on the diachronic implications of areal typology. In Christian Carpelan, Asko Parpola & Petteri Kallioikoski (Eds.), *Early contacts between Uralic and Indo-European: linguistic and archaeological considerations* (Mémoires de la Société Finno-Ougrienne 242), 207–220. Helsinki: Société Finno-Ougrienne.
- Janhunen, Juha. (2007). Typological interaction in the Qinghai linguistic complex. *Studia Orientalia*, 101, 85–102.
- Janhunen, Juha. (2009). Proto-Uralic—what, where, and when. In Jussi Ylikoski (Ed.), *The Quasiquicentennial of the Finno-Ugrian Society* (Mémoires de la Société Finno-Ougrienne 258), 57–78. Helsinki: Société Finno-Ougrienne.
- Janhunen, Juha. (2014a). *A legkeletibb uráliak* [The easternmost Uralians]. Budapest: Magyar Tudományos Akadémia.
- Janhunen, Juha. (2014b). Ural-Altaic. In Martine Robbeets & Walter Bisang (Eds.), *Paradigm Change: In the Transeurasian languages and beyond*, 311–336. Amsterdam: John Benjamins Publishing Company.
- Jantunen, Santra. (2019). Syntactic and aspectual functions of Latvian verbal prefixes in Livonian. In Sofia Björklöf & Santra Jantunen (Eds.), *Multilingual Finnic. Language contact and change* (Uralica Helsingiensia 14), 15–53. Helsinki: Société Finno-Ougrienne.
- Jantunen, Santra. (Forthcoming). Causative strategies in Livonian: augmentation, lexical inheritance, and Latvian influence.
- Jelisejev, J. S. (1961). Vanhin itämerensuomalainen kielenmuistomerkki [The oldest Finnic language monument]. *Virittäjä*, 65(1), 134.

- Jenny, Mathias. (2009). Deixis and information structure in Mon: the Multifunctional particle *kəh*. *Journal of Southeast Asian Linguistics Society*, 2, 53–72.
- Jenny, Mathias. (2014). Mon. In Paul Sidwell & Mathias Jenny (Eds.), *The handbook of Austroasiatic languages*, 554–600. Leiden: Brill.
- Johanson, Lars. (1998). The structure of Turkic. In Lars Johanson & Eva Csató (Eds.), *The Turkic Languages*, 31–66. London: Routledge.
- Joseph, Brian D. (2000). Is there such a thing as “grammaticalization?”. *Language sciences*, 23(2–3), 163–186.
- Jun, Youngchul. (2015). Focus, Topic, and Contrast. In Lucien Brown & Jaehoon Yeon (Eds.), *The handbook of Korean linguistics*, 179–195. Oxford: Blackwell Publishing.
- Junttila, Santeri. (2016). *Tiedon kumuloituminen ja trendit lainasanatutkimuksessa. Kantasuomen balttilaislainojen tutkimushistoria* [The accumulation of information and tendencies in the study of loanwords. The research history of Baltic loans in Proto-Finnic] (Doctoral dissertation). Helsinki: University of Helsinki.
- Juvonen, Päivi. (2000). *Grammaticalizing the Definite Article. A study of definite adnominal determiners in a genre of spoken Finnish* (Doctoral dissertation). Stockholm: Stockholms universitet.
- Kaiser, Elsi & Hiietam, Katrin. (2003). A comparison of the referential properties of the third person pronouns in Finnish and Estonian. *Nordlyd*, 31(4), 654–667.
- Kalima, Jalo. (1952). *Slaavilaisperäinen sanastomme. Tutkimus itämerensuomen vanhimmista slaavilaisperäisistä lainasanoista* [Our Slavic vocabularies. A study of the oldest Slavic loanwords in Finnic] (Suomalaisen Kirjallisuuden Seuran Toimituksia 243). Helsinki: Finnish Literature Society.
- Kalima, Jalo. (1956). *Die Slavischen Lehnwörter im Ostseefinnischen*. Veröffentlichungen der Abteilung für Slavische Sprachen und Literaturen des Osteuropa-Instituts (Slavisches Seminar) an der Freien Universität Berlin. Berlin: In Kommission bei Otto Harrassowitz (Wiesbaden).
- Kallio, Petri. (2006). On the earliest Slavic loanwords in Finnic. In Juhani Nuorluoto (Ed.), *The Slavization of the Russian North: Mechanisms and Chronology* (Slavica Helsingensia 27), 154–166. Helsinki: University of Helsinki.
- Kallio, Petri. (2007). Kantasuomen konsonanttihistoriaa [History of Proto-Finnic consonants]. In Jussi Ylikoski & Ante Aikio (Eds.), *Sámit, sánit, sátnehámit. Riepmočála Pekka Sammallahitii miessemánu 21. beaivve 2007* (Mémoires de la Société Finno-Ougrienne 253), 229–250. Helsinki: Société Finno-Ougrienne.
- Kallio, Petri. (2014). The Diversification of Proto-Finnic. In Joonas Ahola & Frog (Eds.), *Fibula, Fabula, Fact: The Viking Age in Finland* (Studia Fennica Historica 18), 155–168. Helsinki: Finnish Literature Society.
- Kallio, Petri. (2020). Se. In Sampsa Holopainen, Juha Kuokkala, Janne Saarikivi & Susanna Virtanen (Eds.), *Ёмас смылу нэкве вёртур ётност самын патум. Scripta miscellanea in honorem Ulla-Maija Forsberg* (Mémoires de la Société Finno-Ougrienne 275), 171–183. Helsinki: Société Finno-Ougrienne.
- Kallio, Petri & Laakso, Johanna. (2020). Uralic and Slavic. In Marc L. Greenberg, Lenore A. Grenoble, Stephen M. Dickey, Masako Ueda Fidler, René Genis, Marek Łaziński, Anita Peti-Stantić, Björn Wiemer & Nadežda V. Zorixina-Nilsson, (Eds.), *Encyclopedia of Slavic Languages and Linguistics Online*. Retrieved 20-10-2020, from http://dx.doi.org/10.1163/2589-6229_ESLO_COM_032004.
- Kangasmaa-Minn, Eeva. (1998). Mari. In Daniel Abondolo (Ed.), *The Uralic languages*, 219–248. London: Routledge.
- Kappeler, Andreas. (2001). *The Russian Empire: A Multiethnic History*. Harlow: Pearson Education.

- Karjalainen, Heini. (2019). Borrowing morphology: The influence of Russian on the Veps system of indefinite pronouns. In Sofia Björklöf & Santra Jantunen (Eds.), *Multilingual Finnic. Language contact and change* (Uralica Helsingiensia 14), 55–87. Helsinki: Société Finno-Ougrienne.
- Karlsson, Fred. (1975). Suomen kielen tulevaisuus [The future of the Finnish language]. *Sananjalka*, 17, 51–66.
- Karttunen, Lauri. (1969). *Problems of Reference in Syntax* (Doctoral dissertation). Bloomington, IN: Indiana University.
- Karttunen, Lauri. (1976). Discourse referents. In J. D. McCawley (Ed.), *Syntax and Semantics*, Vol. 7, 1–38. New York, NY: Academic Press.
- Kasatkin, Leonid L. (1999). *Russkie dialekty. Karty* [Russian dialects. Maps] (Russkie Monografiya Instituta ètnologii i antropologii RAN). Moskva: Nauka.
- Kasatkina, Rozalija F. (1991). *Russkie narodnye govory, Zvuščaja hrestomatija, Čast' 1, Severnorusskie govory* [Russian dialects, Phonological chrestomathy, Vol. 1, North Russian dialects]. Moskva: Institut Russkogo Jazyka.
- Kasatkina, Rozalija F. (2007). Psevdoartikl' i predartikl' v severnorusskih govorah [Pseudo-article and pre-article in North Russian dialects]. In Juhani Nuorluoto (Ed.), *Topic on the Ethnic, Linguistic and Cultural Making of the Russian North* (Slavica Helsingiensia 32), 101–110. Helsinki: University of Helsinki.
- Kasatkina, Rozalija F. (2008). Artikli, predartikli i psevdartikli v slavjanskijh jazykah [Articles, pre-articles and pseudo-articles in Slavic languages]. In *Slavenskoe jazykoznanie, XIV Meždunarodnyj s'ezd slavistov (Ohrid, 10–16 sentyabrya 2008 g.)*, Doklady rossijskoj delegacii, 305–321. Moskva: Indrik.
- Kato, Atsuhiko. (2003). Pwo Karen. In Graham Thurgood & Randy J. LaPolla (Eds.), *The Sino-Tibetan Languages*, 632–648. London: Routledge.
- Keenan, Edward L. (1976). Towards a universal definition of “subject”. In Charles N. Li (ed.), *Subject and Topic*, 303–333. New York, NY: Academic Press.
- Keevallik, Leelo. (2010). The interactional profile of a placeholder: the Estonian demonstrative see. In Nino Amiridze, Boyd H. Davis & Margaret MacLagan (Eds.), *Fillers, Pauses, and Placeholders* (Typological Studies in Language 93), 139–172. Amsterdam: John Benjamins Publishing Company.
- Kehayov, Petar. (2017). *The fate of mood and modality in language death: Evidence from Minor Finnic*. Berlin: Mouton de Gruyter.
- Kel'makov, Valej. (1996). Formy sub'ektivnoj ocenki imeni suščestvitel'nogo v udmurtskom jazyke [Affectionate forms in Udmurt]. In Heikki Leskinen, Sándor Maticsák & Tõnu Seilenthal (Eds.), *Congressus Octavus Fenno-Ugristarum. 10–15.08.95. Pars III. Sessiones sectionum. Phonologia, Morphologia*, 131–134. Jyväskylä: Moderatones.
- Kel'makov, Valej. (2001). O kategorii opredelennosti/noepredelennosti v udmurtskom jazyke (K postanovke voprosa) [On the category of definiteness/indefiniteness in Udmurt]. In Bernard Comrie et al. (Eds.), *International symposium on deictic systems and quantification in languages spoken in Europe and North and Central Asia* (Abstracts), 179–181. Izhevsk: Udmurt State University.
- Keresztes, László. (1999). *Development of Mordvinic definite conjugation*. Helsinki: Société Finno-Ougrienne.
- Kert, G. M. (1971). *Saamskij jazyk (Kil'dinskij dialekt)* [Saami language (Kildin dialect)]. Leningrad: Nauka.
- Kettunen, Lauri. (1938). *Livisches Wörterbuch mit grammatischer Einleitung* (Lexica Societatis Fenno-Ugricae 5). Helsinki: Société Finno-Ougrienne.
- Kettunen, Lauri. (1943). *Vepsän murteiden lauseopillinen tutkimus* [A syntactic study of Veps dialects]. Helsinki: Société Finno-Ougrienne.

- Kettunen, Lauri. (1960). *Suomen lähisukukielten luonteenomaiset piirteet* [Characteristic features of closely related languages of Finnish] (Mémoires de la Société Finno-Ougrienne 119). Helsinki: Société Finno-Ougrienne.
- Kettunen, Lauri & Posti, Lauri. (1932). *Näytteitä vatjan kielestä* [Samples of Votic language] (Mémoires de la Société Finno-Ougrienne 63). Helsinki: Société Finno-Ougrienne.
- Kettunen, Lauri & Siro, Paavo. (1935). *Näytteitä vepsän murteista* [Samples of Veps dialects] (Mémoires de la Société Finno-Ougrienne 65). Helsinki: Société Finno-Ougrienne.
- Khaidakov, S. M. (1966). The dialect divisions of Lak. *Studia Caucasica*, 2, 9–18.
- Khrunin, Andrey V. et al. (2013). A genome-wide analysis of populations from European Russia reveals a new pole of genetic diversity in Northern Europe. *PLoS ONE*, 8(3). Retrieved 02-02-2020, from <http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0058552>.
- Kim, Alexandra. (1987). *Vyraženie kategorii pritjažatel'nosti v dialektax sel'kupskogo jazyka* [The category of possessivity in the Selkup dialects] (Doctoral dissertation). Tartu: University of Tartu.
- Kim, Ronald I. (2019). North Slavic -ě vs. South Slavic -ę: A Problem of Forward Reconstruction. *Journal of Slavic Linguistics*, 27(1), 1–26.
- Kiparsky, Valentin. (1967). *Russische historische Grammatik, Vol. 2: Die Entwicklung des Formensystems*. Heidelberg: Carl Winter Universitätsverlag.
- Kiparsky, Valentin (1969). *Gibt es ein finnougrieches Substrat im Slavischen?* (Suomalaisen tiedeakatemian toimituksia sarja B 153/4). Helsinki: Finnish Academy of Science and Letters.
- Kiss, Katalin É. (1995). *Discourse Configurational Languages*. Oxford: Oxford University Press.
- Kiss, Katalin. É. (1998). Identificational focus versus information focus. *Language*, 74(2), 245–273.
- Kiuru, Silja (1990). Mistä vanhaan kirjasuomeen tulivat jälkipronominit *se* ja *hän*? [Where did suffixal pronouns *se* and *hän* in Old Literary Finnish come from?]. *Virittäjä*, 94, 278–305.
- Klaas-Lang, Birute & Norvik, Miina. (2014). Balti areaali tüpoloogilisi sarnasusi morfosüntaksi valdkonnas [Typological similarities of the Circum-Baltic area in the field of morphosyntax]. *Keel ja Kirjandus*, 2014(8–9), 590–608.
- Klumpp, Gerson. (2009). Variation in Komi object marking. In Andreas Dufter, Jürg Fleischer & Guido Seiler (Eds.), *Describing and modeling variation in grammar*, 325–359. Berlin: Mouton de Gruyter.
- Klumpp, Gerson. (2012). Differential object marking and information structure: on the function of two different pronominal accusatives in Komi and Khanty dialects. *Journal of Estonian and Finno-Ugric Linguistics*, 3(1), 343–372.
- Koivisto, Andreas. (2006). Trade Routes and their Significance in the Christianization of Karelia. In Juhani Nuorluoto (Ed.), *The Slavization of the Russian North: Mechanisms and Chronology* (Slavica Helsingiensia 27), 167–178. Helsinki: University of Helsinki.
- Koivisto, Vesa. (1990). Venäjän vaikutuksesta itäisten itämerensuomalaisten kielten refleksiiviverbeihin [Russian influence in reflexive verbs of eastern Finnic languages]. In Sari Vaula (Ed.), *Itämerensuomalaiset kielikontaktit/ Läänemeresoome keelekontaktid*, 20–27. Helsinki: Institute for the Languages of Finland.
- Koivulehto, Jorma. (1983). Suomalaisten maahanmuutto indoeurooppalaisten lainasanojen valossa [The immigration of Finns in the light of Indo-European loanwords]. *Journal de la Société Finno-Ougrienne*, 78, 107–132.

- Koivulehto, Jorma. (2006). Wie alt sind die Kontakte zwischen Finnisch-Ugrisch und Balto-Slavisch?. In Juhani Nuorluoto (Ed.), *The Slavization of the Russian North: Mechanisms and Chronology* (Slavica Helsingiensia 27), 179–196. Helsinki: University of Helsinki.
- Kokla, Paul. (1963). *Pritjažatel'nye suffiksy v marijskom jazyke* [Possessive suffixes in Cheremis]. Tallinn: Akademija Nauk Estonskoj SSR.
- Koneski, Blaže. (1967). *Grammatika na makedonskiot literaturni jazik, Vol. 1 & Vol. 2* [A grammar of Literary Macedonian, Vol. 1 & Vol. 2]. Skopje: Kultura.
- König, Ekkehard. (1991). *The Meaning of Focus Particles: A Comparative Perspective*. London: Routledge.
- König, Ekkehard. (2012). Le rôle des déictiques de la manière dans le cadre d'une typologie de la deixis. *Bulletin de la Société de Linguistique de Paris*, 107, 11–42.
- König, Ekkehard. (2015). Manner deixis as source of grammatical markers in Indo-European languages. In Carlotta Viti (Ed.), *Perspectives on Historical Syntax*, 35–57. Amsterdam: John Benjamins Publishing Company.
- König, Ekkehard. (2017). The deictic identification of similarity. In Yvonne Treis & Martine Vanhove (Eds.), *Similitive and equative constructions: A cross-linguistic perspective*, 143–164. Amsterdam: John Benjamins Publishing Company.
- König, Ekkehard. (2018). Definite articles and their uses. Diversity and patterns of variation. In Daniël Olmen, Tanja Mortelmans & Frank Brisard (Eds.), *Aspects of Linguistic Variation*, 165–184. Berlin: Mouton de Gruyter.
- König, Ekkehard & Umbach, Carla. (2018). Demonstratives of manner, of quality and of degree. In Marco Coniglio, Andrew Murphy, Eva Schlachter & Tonjes Veenstra (Eds.), *Atypical demonstratives: Syntax, Semantics and Pragmatics*, 285–328. Berlin: Mouton de Gruyter.
- Konkka, Unelma S. (1963). *Karel'skie narodnye skazki* [Karelian folk tales]. Moskva: Izdatel'stvo Akademii nauk SSSR.
- Kopitar, Jernej Bartholomäus. (1808). *Grammatik der slavischen Sprache in Krain, Kärnten und Steyermark*. Leibach: Wilhelm Heinrich Korn.
- Kopitar, Jernej Bartholomäus. (1829). Albanische, walachische und bulgarische Sprache. *Jahrbücher der Literatur* (Wien), 46, 59–106.
- Kopitar, Jernej Bartholomäus. (1857). *Kleinere Schriften: Sprachwissenschaftlichen, geschichtlichen, ethnographischen u. rechtshistorischen Inhalts*. Franz Miklosich (Ed.). Wien: Friedrich Beck's Universitäts Buchhandlung.
- Koponen, Eino. (1990). Viron ja liivin yhteissanaston tarkasteluja [Revision of Estonian and Livonian common vocabularies]. In Sari Vaula (Ed.), *Itämerensuomalaiset kielikontaktit/Läänemeresoome keelekontaktid*, 35–39. Helsinki: Institute for the Languages of Finland.
- Koponen, Eino. (1991). Itämerensuomalaisen marjannimistön kehityksen päälinjoja ja kantasuomen historiallista dialektologiaa [The development of berry names in the Finnic languages and historical dialectology of Proto-Finnic]. *Journal de la Société Finno-Ougrienne*, 83, 123–161.
- Kopotev, Mikhail. (2006). Ancient Lakes in the Former Finno-Ugric Territories of Central Russia: An Experimental Onomastic Palaeo-geographical Study. In Juhani Nuorluoto (Ed.), *The Slavization of the Russian North: Mechanisms and Chronology* (Slavica Helsingiensia 27), 197–209. Helsinki: University of Helsinki.
- Koptjevskaja-Tamm, Maria & Wälchli, Bernhard. (2001). The Circum-Baltic languages: An Areal-typological Approach. In Östen Dahl & Maria Koptjevskaja-Tamm (Eds.), *The Circum-Baltic languages: typology and contact, Vol. 2: Grammar and typology*, 615–750. Amsterdam: John Benjamins Publishing Company.

- Korpus vepsskogo jazyka* [The corpus of Veps language]. Institut jazyka, literatury i istorii Karel'skogo naučnogo centra RAN. <http://vepsian.krc.karelia.ru/about/>.
- Kortlandt, Frederik. (1983). Demonstrative pronouns in Balto-Slavic, Armenian, and Tocharian. *Studies in Slavic and General Linguistics*, 3, 311–322.
- Kozlova, Elena. (2002). Vova da muköd [Vova and friends]. *Vojvyv kodzuv*, 2002(5), 3–23.
- Krámský, Jiří. (1972). *The Article and the Concept of Definiteness in Language*. The Hague: Mouton & Co. N. V.
- Kratochvíl, František. (2011). Discourse-structuring functions of Abui demonstratives. In Foong Ha Yap, Karen Grunow-Hårsta & Janick Wrona (Eds.), *Nominalization in Asian languages: Diachronic and typological perspectives* (Typological studies in language 96), 757–788. Amsterdam: John Benjamins Publishing Company.
- Krifka, Manfred. (2008). Basic notions of information structure. *Acta Linguistica Hungarica*, 55(3–4), 243–276.
- KTTGN = Komissija toponomiki i transkripcii geografičeskikh nazvanij. (1970). *Pinež'e: Očerki po morfologii pinežskogo govora* [Pinega: Studies on morphology of Pinega variety]. Leningrad: Geografičeskoe obščestvo SSSR.
- Kujola, Juho. (1910). Sananen lyydiläisistä ja heidän kielestään [Sanane Ludes and their language]. In Iivo Härkönen (Ed.), *Karjalan kirja. II osa*, 162–171. Porvoo: WSOY.
- Künnap, Ago. (2004). About the non-personal definite function of the Uralic 3rd person possessive suffix. *Linguistica Uralica*, 40(1), 1–4.
- Kunyt, Taweesak. (1984). *General characteristics of Hmong Njua grammar* (Master's thesis). Bangkok: Mahidol University.
- Kurz, Josef. (1937–1938). K otázce členu v jazycích slovanských, se zvláštním zřetelem k staroslověnině [On the question of articles in Slavic languages, with a special focus on Old Church Slavonic]. *Byzantinoslavica*, 7, 212–340.
- Kurz, Josef. (1939–1946). K otázce členu v jazycích slovanských, se zvláštním zřetelem k staroslověnině (Dokončení) [On the question of articles in Slavic languages, with a special focus on Old Church Slavonic (Completion)]. *Byzantinoslavica*, 8, 172–288.
- Kushniarevich, Alena et al. (2015). Genetic heritage of the Balto-Slavic speaking populations: A synthesis of autosomal, mitochondrial and Y-chromosomal data. *PLoS ONE*, 10(9). Retrieved 02-02-2020, from <http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0135820>.
- Kusmenko, Jurij. (2008). *Der samische Einfluss auf die skandinavischen Sprachen: Ein Beitrag zur skandinavischen Sprachgeschichte* (Berliner Beiträge zur Skandinavistik 10). Berlin: Nordeuropa-Institut der Humboldt-Universität.
- Kusters, Wouter. (2003). *Linguistic complexity: The influence of social change on verbal inflection*. Utrecht: Landelijke Onderzoekschool Taalwetenschap.
- Kuzmin, Denis. (2014). *Vienan Karjalan asutushistoria nimistön valossa* [Settlement history of White Sea Karelia in the light of onomastics] (Doctoral dissertation). Helsinki: University of Helsinki.
- Kuz'mina, Irina B. (1993). *Sintaksis russkikh govorov v lingvogeografičeskom aspekte* [Syntax of Russian dialects in the lingvo-geographic aspect]. Moskva: Nauka.
- Kuz'mina, Irina B. & Nemčenko, E. V. (1962). K voprosu o postpositivnyx časticax v russkix govorax [Towards the question of postpositive particles in Russian dialects]. In V. G. Orlova (Ed.), *Materialy i issledovanija po russkoj dialektologii. Novaja serija. Tom. 3*, 5–32. Moskva: Nauka.
- Kuznecova, Ariadna I. & Helimskij, Evgenij A. & Gruškina, Elena V. (1980). *Očerki po sel'kupskomu jazyku* [Studies on the Selkup language]. Moskva: MGU.

- Laakso, Johanna. (1999). Vielä kerran itämerensuomen vanhimmista muistomerkeistä [Once again about the oldest Finnic language monument]. *Virittäjä*, 103(4), 531–531.
- Laanest, Arvo. (1966). *Isuri murdetekste* [Ingrian dialect texts]. Tallinn: Eesti NSV Teaduste Akademia.
- Laanest, Arvo. (1982). *Einführung in die ostseefinnischen Sprachen*. Hamburg: Buske Verlag.
- Labov, William. (1972). *Language in the inner city*. Philadelphia: University of Pennsylvania Press.
- Labov, William. (1994). *Principles of linguistic change Vol. 1: Internal factors*. Oxford: Blackwell Publishing.
- Laitinen, Lea. (2002). From logophoric pronoun to discourse particle. In Ilse Wischer & Gabriele Diewald (Eds.), *New reflections on grammaticalization*, 327–344. Amsterdam: John Benjamins Publishing Company.
- Laitinen, Lea. (2005). *Hän*, the third speech act pronoun in Finnish. In Ritva Laury (Ed.), *Minimal reference. The use of pronouns in Finnish and Estonian discourse*, 75–106. Helsinki: Finnish Literature Society.
- Lallukka, Seppo. (1990). *The East Finnic minorities in the Soviet Union: An appraisal of the erosive trends*. Helsinki: Finnish Academy of Science and Letters.
- Lallukka, Seppo. (1995). *Komipermjakit-perämaan kansa: syrjäytyminen, sulautuminen ja postkommunistinen murros*. Helsinki: Venäjän ja Itä-Euroopan Instituuti.
- Lallukka, Seppo. (2001). Finno-Ugrians of Russia: vanishing cultural communities? *Nationalities Papers*, 29(1), 9–39.
- Lallukka, Seppo. (2003). *From Fugitive Peasants to Diaspora: The Eastern Mari in Tsarist and Federal Russia* (Humaniora 328). Helsinki: Annales Academiae Scientiarum Fennicae.
- Lambrecht, Knud. (1994). *Information Structure and Sentence Form: Topic, Focus and the Mental Representation of Discourse Referents*. Cambridge: Cambridge University Press.
- Landaburu, Jon. (2007). La modalisation du savoir en langue Andoke (Amazonie Colombienne). In Zlatka Guentchéva & Jon Landaburu (Eds.), *L'énonciation médiatisée II – Le traitement épistémologique de l'information: illustrations amérindiennes et caucasiennes*, 23–47. Paris: Édition Peeters.
- Lang, Valter. (2016). Early Finnic-Baltic contacts as evidenced by archaeological and linguistic data. *Journal of Estonian and Finno-Ugric Linguistics*, 7(1), 11–38.
- Larjavaara, Matti. (1986). *Itämerensuomen demonstratiivit, 1, Karjala, aunus, lyydi ja veps* [Finnic demonstratives, 1, Karelian, Olonets Karelian, Lude and Veps]. Helsinki: Finnish Literature Society.
- Larjavaara, Marri. (1990). *Suomen deiksis* [Deixis of Finnish]. Helsinki: Finnish Literature Society.
- Laryëna, A. A. (2016). The linguistic status of the category of definiteness/indefiniteness in the Belarusian language. *Vesci Nacyjanal'naj akadëmii navuk Belarusi. Seryja humanitarnyh navuk*, 4, 68–75.
- Laury, Ritva. (1991). On the development of the definite article *se* in Spoken Finnish. In Maria Vilkuna & Arto Anttila (Eds.), *Suomen kielitieteellisen yhdistyksen vuosikirja 1991*, 93–121. Helsinki: The Linguistic Association of Finland.
- Laury, Ritva. (1996). *Sen kategoriasta - onko suomessa jo artikkeli?* [The category of *se* - is there already an article in Finnish?]. *Virittäjä*, 100(2), 162–181.
- Laury, Ritva. (1997). *Demonstratives in interaction: The emergence of a definite article in Finnish*. John Benjamins Publishing Company.

- Lê, Thị Hoài Dương. (2002). *Tiểu từ tình thái cuối câu tiếng Việt và việc dạy tiểu từ tình thái cuối câu tiếng Việt cho người nước ngoài* [Sentence-final modal particles in Vietnamese and teaching sentence-final modal particles in Vietnamese to foreigners] (Master's thesis). Hanoi: Vietnam National University.
- Lehmann, Christian. (1982). *Thoughts on Grammaticalization: A Programmatic Sketch, Volume I* (Arbeiten des Kölner Universalien-Projekts 48). Köln: Universität Köln.
- Lehmann, Christian. (1985). Grammaticalization: Synchronic variation and diachronic change. *Lingua e Stile*, 20, 303–318.
- Leinonen, Marja. (1998). The postpositive particle *-to* of Northern Russian dialects, compared with Permic languages (Komi Zyryan). *Studia Slavica Finlandensia*, 15, 74–90.
- Leinonen, Marja. (2002). Influence of Russian on the syntax of Komi. *Finnisch-ugrische Forschungen*, 57(1–3), 195–358.
- Leinonen, Marja. (2006). Omistussuhteen ulokkeita: komin possessiivisuffiksin ei-possessiivisista funktioista. *Journal de la Société Finno-Ougrienne*, 91, 93–114.
- Leinonen, Marja. (2009). Russian influence on the Ižma Komi dialect. *International Journal of Bilingualism*, 13(3), 309–329.
- Leskien, August. ([1871]1955). *Handbuch der Altbulgarischen Sprache* (7th ed.). Heidelberg: Carl Winter Universitätsverlag.
- Leskien, August. (1879). Der bulgarische Artikel mit *v* oder *n*. *Archiv für slavische Philologie*, 3, 522–523.
- Levinson, Stephen C. (1983). *Pragmatics*. Cambridge: Cambridge University Press.
- Levinson, Stephen C. (2004). Deixis. In Laurence R. Horn & Gregory L. Ward (Eds.), *The Handbook of Pragmatics*, 97–121. Oxford: Blackwell Publishing.
- Li, Charles N. & Thompson, Sandra A. (1976). Subject and topic: A new typology of language. In Charles N. Li (Ed.), *Subject and topic*, 457–489. New York, NY: Academic Press.
- Liefländer-Koistinen, Luise. (1989). Zum deutschen *doch* und finnischen *-han*. Beobachtungen zur Übersetzbarkeit der deutschen Abtönungspartikel. In Harald Weydt (Ed.), *Sprechen mit Partikeln*, 185–195. Berlin: Mouton de Gruyter.
- Lind, John. (2006). Problems of ethnicity in the interpretation of written sources on Early Rus'. In Juhani Nuorluoto (Ed.), *The Slavization of the Russian North: Mechanisms and Chronology* (Slavic Helsingiensia 27), 246–258. Helsinki: University of Helsinki.
- Lindstedt, Jouko. (2000). Linguistic Balkanization: Contact-induced change by mutual reinforcement. *Studies in Slavic and General Linguistics*, 28, 231–246.
- Lindstedt, Jouko. (2009). Balkanin kieliliitto [Balkan Sprachbund]. In Jyrki Kalliokoski, Lari Kotilainen & Päivi Pahta (Eds.), *Kielet kohtaavat*, 75–95. Helsinki: Finnish Literature Society.
- Lindstedt, Jouko. (2014). Markeri na opredelenostta na imennata grupa v istorijata na bălgarskija ezik [Markers of definiteness in noun phrase in the history of the Bulgarian language]. In *Treti Meždunaroden kongres po bălgaristika*, 59–68. Sofia: Universitetsko izd-vo Sv. Kliment Okhridski.
- Lindstedt, Jouko. (2016). Multilingualism in the Central Balkans in late Ottoman times. In Max Wahlström & Maxim Makartsev (Eds.), *In search of the center and periphery: Linguistic attitudes, minorities, and landscapes in the Central Balkans* (Slavica Helsingiensia 49), 51–67. Helsinki: University of Helsinki.
- Lindstedt, Jouko & Salmela, Elina. (2020). Migrations and language shifts as components of the Slavic spread. In Tomáš Klír & Vít Boček (Eds.), *Language Contact and the Early Slavs*. Heidelberg: Carl Winter Universitätsverlag.

- Liu, Danqing. (1997). Sūzhōu fāngyán de dòngcí wèiyǔ jù [Verb predicate sentences in the Suzhou dialect]. In Rulong Li & Song Hing Chang (Eds.), *Dòngcí wèiyǔ jù*, 1–20. Guangzhou: Ji'nan University Press.
- Longacre, Robert E. (1996). *The grammar of discourse*. New York, NY: Springer Science & Business Media.
- Lorentz, Friedrich. (1903). *Slovinzische Grammatik*. Saint Petersburg: Kaiserliche Akademie der Wissenschaften.
- Lyons, John. (1977). *Semantics*. Cambridge: Cambridge University Press.
- Lyons, Christopher. (1999). *Definiteness*. Cambridge: Cambridge University Press.
- Maddison, David R. & David L. Swofford & Wayne P. Maddison. (1997). Nexus: An extensible file format for systematic information. *Systematic Biology*, 46(4), 590–621.
- Mägiste, Julius. (1962). *Äldre ryska lånord i estniskan, särskilt i det gamla estniska skriftspråket* (Lunds universitets årsskrift, Avd. 1, Band 55 (1)). Lund: Lunds universitet.
- Mägiste, Julius. (2006). Muistoja Liivinrannasta: Liivin kieltä Ruotsista [Memories from the Livonian coast: Livonian language from Sweden]. Translated into Finnish and published by Anneli Honko. (Mémoires de la Société Finno-Ougrienne 250). Helsinki: Société Finno-Ougrienne.
- Makarov, Grigorij N. & Rjagoev, Vladimir D. (1969). *Obrazcy karel'skoj reči* [Samples of Karelian dialects]. Leningrad: Nauka.
- Marti, Roland. 2012. (Un)articulated definiteness. Nominal determination in Sorbian: Moving from an articleless language to a language with articles?. *Sprachtypologie und Universalienforschung*, 65(3), 296–316.
- Martin, Samuel E. (2004). *A reference grammar of Japanese*. Honolulu: University of Hawai'i Press.
- Marty, Anton (1884). Über subjektlose Sätze und das Verhältnis der Grammatik zu Logik und Psychologie. *Vierteljahresschrift für wissenschaftliche Philosophie*, 8, 161–192, 292–340.
- Marušič, Franc & Žaucer, Rok. (2006). The 'definite article' *ta* in colloquial Slovenian. In *Formal Approaches to Slavic Linguistics*, 189–204. Ann Arbor, MI: Michigan Slavic Publications.
- Marušič, Franc & Žaucer, Rok. (2012). On slovenian demonstrative reinforcers and the internal structure of demonstratives. In S. Halupka-Rešetar, M. Marković, T. Milićev, and N. Milićević (Eds.), *Selected papers from SinFonJA 3*, 76–104. Cambridge: Cambridge Scholars Publishing.
- Masica, Colin P. (1976). *Defining a linguistic area: South Asia*. Chicago: Chicago University Press.
- Mathiassen, Terje. (1985a). Slavisk – Baltisk – Østersjøfinske syntaktiske isoglosser og spørsmål om et sprachbund i den østlige del av østersjøområdet [Slavic – Baltic – Finnic syntactic isoglosses and questions of *Sprachbund* in the eastern part of the Baltic area]. *Meddelanden från Åbo Akademi forskningsinstitut*, 102, 123–149.
- Mathiassen, Terje. (1985b). A discussion of the notion 'sprachbund' and its application in the case of the languages in the eastern Baltic area (Slavic, Baltic, and West Finnish). *International Journal of Slavic Philology*, 21/22, 273–281.
- Matić, Dejan. (2003). *Topics, presuppositions, and theticity: An empirical study of verb-subject clauses in Albanian, Greek, and Serbo-Croat* (Doctoral dissertation). Köln: Universität zu Köln.
- Matić, Dejan. (2015). Information structure in linguistics. *International Encyclopedia of the Social & Behavioral Sciences* (2nd ed., vol. 12), 95–99. Amsterdam: Elsevier Ltd.

- Matveev, Aleksandr K. (1996). Substratnaja toponimija Russkogo Severa i merjanskaja problema [Substrate toponyms in the Russian North and the Meryan problem]. *Voprosy jazykoznanija*, 1996/1, 3–23.
- Matveev, Aleksandr K. (1998). Merjanskaja toponimija na Russkom Severe – fantom ili fenomen? [Meryan toponyms in the Russian North – phantom or phenomenon?]. *Voprosy jazykoznanija*, 1998/5, 90–105.
- Matveev, Aleksandr K. (2001). Merjanskaja problema i lingvističeskoe kartografirovanie [The Meryan problems and linguistic cartography]. *Voprosy jazykoznanija*, 2001/5, 32–59.
- Mazzitelli, Lidia Federica. (2015). *The expression of predicative possession: A comparative study of Belarusian and Lithuanian*. Berlin: Mouton de Gruyter.
- McAnallen, Julia. (2011). *The History of Predicative Possession in Slavic: Internal Development vs. Language Contact* (Doctoral dissertation). Berkeley, CA: University of California, Berkeley.
- McCoy, Svetlana G. (2001). *Colloquial Russian particles -to, že, and ved' as set-generating ("kontrastive") markers: A unifying analysis* (Doctoral dissertation). Durham: University of New Hampshire.
- McCoy, Svetlana G. (2003). Connecting Information Structure and Discourse Structure through “Kontrast”: The Case of Colloquial Russian Particles -to, že, and ved'. *Journal of Logic, Language and Information*, 12(3), 319–335.
- McWhorter, John H. (1998). Identifying the creole prototype. Vindicating a typological class. *Language*, 74(4), 788–818.
- McWhorter, John H. (2001). The world's simplest grammars are creole grammars. *Linguistic Typology*, 5, 125–166.
- McWhorter, John H. (2005). *Defining Creole*. Oxford: Oxford University Press.
- Meillet, Antoine. (1912) L'évolution des formes grammaticales. *Scientia (Rivista di Scienza)*, 12, No. 26,6.
- Meillet, Antoine. (1921). *Linguistique historique et linguistique générale*. Paris: La Société de Linguistique de Paris.
- Meillet, Antoine. (1928). *La méthode comparative en linguistique historique*. Paris: Champion.
- Mendoza, Imke. (2005). Polish demonstrative pronouns as “markers of value”. *Glossos*, 6, 1–20.
- Mendoza, Imke. (2011). On demonstrative pronouns and the question of the article in 17th century Russian. *Russian linguistics*, 35(2), 245–265.
- Meščerskij, Nikita A. (1972). *Russkaja dialektologija* [Russian dialectology]. Moskva: Vysšaja Škola.
- Metslang, Helle. (2011). Some grammatical innovations in the development of Estonian and Finnish: Forced grammaticalization. *Linguistica Uralica*, 47(4), 241–256.
- Metsmägi, Iris & Sedrik, Meeli & Soosaar, Sven-Erik (Eds.). (2012). *Eesti etümoloogiasõnaraamat* [The etymological dictionary of Estonian]. Tallinn: Institute of the Estonian Language. <https://www.eki.ee/dict/ety/>.
- Miestamo, Matti & Sinnemäki, Kaius & Karlsson, Fred. (2008). Introduction: The problem of language complexity. In Matti Miestamo, Kaius Sinnemäki & Fred Karlsson (Eds.), *Language complexity: Typology, contact, change*, vii–xiv. Amsterdam: John Benjamins Publishing Company.
- Mikkola, Jooseppi Julius. (1894). *Berührungen zwischen den westfinnischen und slavischen Sprachen I. Slavische Lehnwörter in den westfinnischen Sprachen* (Mémoires de la Société Finno-Ougrienne 8). Helsinki: Société Finno-Ougrienne.

- Miklosich, Franz. (1883). *Vergleichende Syntax der slavischen Sprachen*. Wien: Wilhelm Braumüller.
- Miletič, Ljubomir. (1887). Členāta vā bālgarskija ezikā. *Periodičesko spisanie na Bālgarskoto spisanie na Bālgarskoto knižovno družestvo v Sredec*, 21–22, 305–331.
- Milroy, Lesley & Gordon, Matthew. (2003). *Sociolinguistics: Method and Interpretation*. Oxford: Blackwell Publishing.
- Mladenova, Olga. M. (2007). *Definiteness in Bulgarian: Modelling the processes of language change*. Berlin: Mouton de Gruyter.
- Molnár, Valeria. (2006). On different kinds of contrast. In Valeria Molnár & Susanne Winkler (Eds.), *The architecture of focus* (Studies in Generative Grammar 82), 197–233. Berlin: Mouton de Gruyter.
- Mullonen, Irma I. (1989). Vepsy Prionež’ja po dannym toponimii [Veps of Lake Onega in the present toponym]. *Problemy istorii i kultury vepsskoj narodnosti*, 84–90. Petrozavodsk: Karel’skij filial AN SSSR.
- Mullonen, Irma I. (1990). Mitä paikannimistö kertoo vepsäläisten asettumisesta Äänisen, Laatokan ja Valkeajärven väliselle alueelle? [What does toponym reveal about the settlement of Veps in the area between Onega, Ladoga and White Lake?]. In Sari Vaula (Ed.), *Itämerensuomalaiset kielikontaktit/Läänemeresoome keelekontaktid*, 84–90. Helsinki: Institute for the Languages of Finland.
- Mullonen, Irma I. (2002). *Toponimija Prisvir’ja: problemy ètnojazykovogo kontaktirovanija* [Toponyms of Prisvir’e: problems of ethnolinguistic contacts]. Petrozavodsk: Petrozavodskij gosudarstvennyj universitet.
- Mullonen, Irma I. (2007). Ancient Place Names of Obonež’e in the Context of Ethnic and Linguistic Contacts. In Juhani Nuorluoto (Ed.), *Topics on the Ethnic, Linguistic and Cultural Making of the Russian North* (Slavica Helsingiensia 32), 163–175. Helsinki: University of Helsinki.
- Muriel Norde. (2009). *Degrammaticalization*. Oxford: Oxford University Press.
- Must, Mari. (2000). *Vene laensõnad eesti murretes*. Tallinn: Eesti Keele Sihtasutus.
- Muysken, Pieter. (2008). Introduction: Conceptual and methodological in areal linguistics. In Pieter Muysken (Ed.), *From Linguistic Areas to Areal Linguistics* (Studies in Language Companion Series 90), 1–24. Amsterdam: John Benjamin Publishing Company.
- Myznikov, Sergej A. (2003). *Leksika finno-ugorskogo proishoždenja v russkikh govoraх severo-zapada (ètimologičeskij i lingvogeografičeskij analiz)* [Lexicon of Finno-Ugric origin in Northwest Russian dialects (etymological and lingvogeographic analysis] (Doctoral dissertation’s summary). Sankt-Peterburg: Rossijskaja akademija nauk, Institut lingvističeskikh issledovanij.
- Nagaya, Naonori. (2011). Rise and fall of referentiality: articles in Philippine languages. In Foong Ha Yap, Karen Grunow-Hårsta & Janick Wrona (Eds.), *Nominalization in Asian Languages: Diachronic and typological perspectives*, 589–626. Amsterdam: John Benjamins Publishing Company.
- Nau, Nicole. (1995). *Möglichkeiten und Mechanismen kontaktbewegten Sprachwandels: unter besonderer Berücksichtigung des Finnischen*. München: LINCOM Europa.
- Neeleman, Ad & Titov, Elena & van de Koot, Hans & Vermeulen, Reiko. (2009). A syntactic typology of topic, focus and contrast. In Jeroen van Craenenbroek (Ed.), *Alternatives to cartography*, 15–51. Berlin: Mouton de Gruyter.
- Neeleman, Ad, & Vermeulen, Reiko. (2012). The syntactic expression of information structure. In Ad Neeleman & Reiko Vermeulen (Eds.), *The syntax of topic, focus, and contrast: An interface-based approach*, 1–38. Berlin: Mouton de Gruyter.
- Nettle, Daniel. (1999). *Linguistic diversity*. Oxford: Oxford University Press.

- Nevis, Joel A. (1988). *Finnish particle clitics and general clitic theory*. New York, NY: Garland.
- Newmeyer, Frederick J. (2000). Deconstructing grammaticalization. *Language sciences*, 23(2-3), 187–229.
- Nichols, Johanna. (1992). *Linguistic Diversity in Space and Time*. Chicago: University of Chicago Press.
- Nichols, Johanna. (1997). Modeling ancient population structures and movement in linguistics. *Annual Review of Anthropology*, 26(1), 359–384.
- Nichols, Johanna. (1998). The Eurasian spread zone and the Indo-European dispersal. In Roger Blench & Matthew Spriggs (Eds.), *Archaeology and Language II: Correlating archaeological and linguistic hypotheses* (One World Archaeology 29), 220–266. London: Routledge.
- Nichols, Johanna. (2003). Diversity and stability in language. In Brian D. Joseph & Richard D. Janda (Eds.), *The handbook of historical linguistics*, 283–310. London: Blackwell Publishing.
- Nickel, Klaus Peter & Sammallahhti, Pekka. (2011). *Nordsamisk grammatikk* [A grammar of North Saami]. Karasjok: Davvi Girji.
- Nikolaeva, Irina. (2001). Secondary Topic as a Relation in Information Structure. *Linguistics*, 39, 1–49.
- Nikolaeva, Irina. (2003). Possessive affixes as markers of information structuring: Evidence from Uralic. In Pirkko Suihkonen & Bernard Comrie (Eds.), *International symposium on deictic systems and quantification in languages spoken in Europe and North and Central Asia. Collection of papers*, 130–145. Izhevsk & Leipzig: Udmurt State University & Max Planck Institute of Evolutionary Anthropology.
- Nikolaeva, Irina & Tolskaya, Maria. (2001). *A grammar of Udihe*. Berlin: Mouton de Gruyter.
- Nordlund, Taru, Laury, Ritva, Pajusalu, Renate & Habicht, Külli. (2013). Kielikontaktit, kieliopillistuminen ja kieltenvälinen vertailu [Language contacts, grammaticalisation and cross-language comparison]. In Leena Kolehmainen, Matti Miestamo & Taru Nordlund (Eds.), *Kielten vertailun metodiikka* [Methodology of language comparison], 251–290. Helsinki: Finnish Literature Society.
- Nosov, Eugenie & Ribina, Elena & Janin, Valentin L. (1993). Novgorod. In Michael Andersen & Frank Birkebæk (Eds.), *Vikingerne Rusland - Staraja Ladoga og Novgorod*, 27–75. Roskilde: Roskilde Museums Forlag.
- Nunberg, Geoffrey. (1993). Indexicality and deixis. *Linguistics and philosophy*, 16(1), 1–43.
- Nuorluoto, Juhani. (2006). Is there a Sound Change of “e > o” in Russian?. In Juhani Nuorluoto (Ed.), *The Slavicization of the Russian North: Mechanisms and Chronology* (Slavica Helsingiensia 27), 293–308. Helsinki: University of Helsinki.
- Nuorluoto, Juhani. (2007). The Interchangeability of the Graphemes <o> and <ъ> in Old Russian Birchbark Documents: A Graphical Effect or a Reflection of Sound Change in Progress?. In Juhani Nuorluoto (Ed.), *Topics on the Ethnic, Linguistic and Cultural Making of the Russian North* (Slavica Helsingiensia 32), 176–195. Helsinki: University of Helsinki.
- Ojanen, Muusa. (1985). *Adjektiivikategoria venäläis-lyydiläisissä kontakteissa: Lingvistinen interferenssitutkimus* [Adjective category in the Russian-Lude contacts: A study of linguistic interference] (Mémoires de la Société Finno-Ougrienne 188). Helsinki: Société Finno-Ougrienne.

- Ojanen, Muusa. (1988). Suomalais-karjalais-venäläisistä kontakteista. In Muusa Ojanen & Ilkka Savijärvi (Eds.), *Kielikontakteja I. Tutkimuksia venäjän ja itämerensuomalaisten kielten alalta*, 11–30. Joensuu: University of Joensuu.
- Ojansuu, Heikki. (1922). Itämerensuomalaisten kielten pronominioppia [A study of Finnic pronouns]. *Annales universitatis fennicae Aboensis Series, Sarja B, Osa I, N:o 3*.
- Ojansuu, Heikki & Kujola, Juho & Kalima, Jalo & Kettunen, Lauri. (1934). *Lyydiläisiä kielennäytteitä* [Language samples of Lude] (Mémoires de la Société Finno-Ougrienne 59). Helsinki: Société Finno-Ougrienne.
- Onegina, N. F. & Zajceva, M. I. (1996). *Vepsän rahvhan sarnad* (Вепсские народные сказки) [Veps folktales]. Petrozavodsk.
- Onishi, Masayuki (2001). Non-canonically marked subjects and objects: Parameters and properties. In Alexandra Y. Aikhenvald, Robert M. W. Dixon & Masayuki Onishi (Eds.), *Non-canonical marking of subjects and objects*, 1–51. Amsterdam: John Benjamins Publishing Company.
- Pahomov, Miikul. (2011). *Kuujärven lyydiläistekstejä* [Lude texts from Kuujärvi] (Mémoires de la Société Finno-Ougrienne 263). Helsinki: Société Finno-Ougrienne.
- Pahomov, Miikul. (2017). *Lyydiläiskysymys. Kansa vai heimo, kieli vai murre?* [The question of Lude. Nation or tribe, language or dialect?] (Doctoral dissertation). Helsinki: University of Helsinki.
- Pajusalu, Renate. (1995). Pronominit *see*, *tema* ja *ta* viron puhekielessä [Pronouns *see*, *tema* and *ta* in spoken Estonian]. *Sananjalka*, 37, 81–93.
- Pajusalu, Renate. (1996a). Pronoun systems of common Estonian and Estonian dialects in a contrastive perspective. *Estonian: Typological Studies*, 1, 145–164.
- Pajusalu, Renate. (1996b). *See tädi ja too onu*: eesti, vene ja soome demonstratiivpronomeenid adjektiivatribuudina [This aunt and that uncle: Estonian, Russian and Finnish demonstrative pronouns as adjective attributes]. In Raili Pool (Ed.), *Emakeel ja teised keeled II. 18.-20. oktoober 1996. Ettekanded*, 91–101. Tartu: University of Tartu.
- Pajusalu, Renate. (1997a). Is there an article in (spoken) Estonian?. *Estonian typological studies*, 2, 146–177.
- Pajusalu, Renate. (1997b). Eesti pronomeneid I. Ühiskeelee *see*, *too* ja *tema/ta* [Estonian pronouns I. Literary language *see*, *too* and *tema/ta*]. *Keel ja Kirjandus*, 1, 106–115.
- Pajusalu, Renate. (1998). Eesti pronomeneid II. Võro *sjoo*, *taa* ja *tuu* [Estonian pronouns II. Võro *sjoo*, *taa* and *tuu*]. *Keel ja Kirjandus*, 3, 159–172.
- Pajusalu, Renate. (1999). *Deiktikud eesti keeles* [Deictic words in the Estonian language] (Doctoral dissertation). Tartu: Tartu Ülikooli kirjastus.
- Pajusalu, Renate. (2001). Definite and indefinite determiners in Estonian. In Enikő Nemeth (Ed.), *Pragmatics in 2000, Selected Papers from the 7th International Pragmatics Conference*, 458–469. Antwerp: International Pragmatics Association.
- Pajusalu, Renate. (2005). Anaphoric pronouns in spoken Estonian. In Ritva Laury (Ed.), *Minimal reference. The use of pronouns in Finnish and Estonian discourse*, 107–134. Helsinki: Finnish Literature Society.
- Pajusalu, Renate. (2006). Death of a demonstrative: person and time. The case of Estonian *too*. *Linguistica Uralica*, 42(4), 241–253.
- Pajusalu, Renate. (2015). Võro demonstratives: changing or disappearing?. *Journal of Estonian and Finno-Ugric Linguistics*, 6(2), 167–190.
- Pakendorf, Brigitte. (2007a). The ‘non-possessive’ use of possessive suffixes in Sakha (Yakut). *Turkic Languages*, 11, 226–234.

- Pakendorf, Brigitte. (2007b). *Contact in the prehistory of the Sakha (Yakuts): Linguistic and genetic perspectives* (Doctoral dissertation). Leiden: Universiteit Leiden.
- ParaSol = *A Parallel Corpus of Slavic and other languages*. Universität Regensburg. <http://parasolcorpus.org/>.
- Paul, Hermann. (1880). *Prinzipien der Sprachgeschichte*. Halle: Max Niemeyer.
- Payne, Thomas Edward. (1997). *Describing morphosyntax: A guide for field linguists*. Cambridge: Cambridge University Press.
- Pilarský, Jiří. (2001). *Donausprachbund. Das arealistische Profil einer Sprachlandschaft* (Doctoral dissertation, Habilitationsschrift). Debrecen: University of Debrecen.
- Pimenov V. V. (1965). *Vepsy. Očerok ètničeskoj istorii i genezisa kul'tury* [The Veps. A study of ethnohistory and genesis of culture]. Moskva: Nauka.
- Pintzuk, Susan. (2003). Variationist Approaches to Syntactic Change. In Brian D. Joseph & Richard D. Janda (Eds.), *The Handbook of Historical Linguistics*, 509–528. Oxford: Blackwell Publishing.
- Pixabaj, Can, & England, Nora. (2011). Nominal topic and focus in K'ichee'. In Rodrigo Guitérrez-Bravo, Line Mikkelsen, & Eric Potsdam (Eds.), *Representing Language: Essays in Honor of Judith Aissen*, 15–30. Santa Cruz, CA: Booksurge Publishing.
- Podlesskaya, Vera I. (2010). Parameters for typological variation of placeholders. In Nino Amiridze, Boyd H. Davis & Margaret MacLagan (Eds.), *Fillers, Pauses, and Placeholders* (Typological Studies in Language 93), 11–32. Amsterdam: John Benjamins Publishing Company.
- Polański, Kazimierz. (1993). Polabian. In Bernard Comrie & Greville Corbett (Eds.), *The Slavonic Languages*, 795–824. London: Routledge.
- Pool, Raili. (1999). About the use of different forms of the first and second person singular personal pronouns in Estonian cases. *Estonian: Typological Studies*, 3, 158–184.
- Popov, A. I. (1973). *Nazvanija narodov SSSR: Vvedenie v ètnonimiku* [Names of nations in USSR: Introduction to ethnonyms]. Leningrad: Nauka.
- Priiki, Katri. (2015). *Se oli iha hullu se hammaslääkäri*. Kaakkois-Satakunnan henkilöviitteiset *se, hän, tää ja toi* eteenpäin lohkeavan konstruktion osina ja ensimainintoina. *Puhe ja kieli*, 35(2), 47–71.
- Priiki, Katri. (2017). *Hän, se, tää vai toi?: Vuorovaikutussosiolingvistinen tutkimus henkilöviittauksista Kaakkois-Satakunnan nykypuhekielessä* [Hän, se, tää or toi? An interactional sociolinguistic study of third-person references to people in contemporary spoken Finnish in southwestern Satakunta] (Doctoral dissertation). Turku: University of Turku.
- Prince, Ellen F. (1992). The ZPG Letter: subjects, definiteness, and information-status. In William C. Mann & Sandra A. Thompson (Eds.), *Discourse Description: Diverse Analyses of a Fundraising Text*, 295–325. Amsterdam: John Benjamins Publishing Company.
- Pronk-Tiethoff, Saskia. (2013). *The Germanic loanwords in Proto-Slavic*. Amsterdam: Rodopi.
- Puura, Ulriikka et al. (2013). *The Veps Language in Russia: ELDIA Case Specific Report*. Retrieved 02-02-2020, from https://phaidra.univie.ac.at/detail_object/o:315545.
- Rahkonen, Pauli. (2013). *The South-Eastern Contact Area of Finnic Languages in the Light of Onomastics* (Doctoral dissertation). Jyväskylä: University of Helsinki.
- Rathmayr, Renate. (1985). *Die russischen Partikeln als Pragmalexeme*. München: Verlag Otto Sagner.

- Raun, Alo. (1988). The Mordvin Language. In Denis Sinor (Ed.), *The Uralic Languages: Description, History and Foreign Influences*, 96–110. Leiden: Brill.
- Raunamaa, Jaakko & Kanner Antti. (Forthcoming). Computer, languages, names and people in the Medieval Novgorod: Personal names in the census book of Vodskaja pjatina.
- Ravila, Paavo. (1966). *Johdatus kielihistoriaan* [Introduction to language history] (3rd ed.). Helsinki: Finnish Literature Society.
- Rédei, Károly. (1988). Die syrjänische Sprache. In Denis Sinor (Ed.), *The Uralic Languages: Description, History and Foreign Influences*, 111–130. Leiden: Brill.
- Reile, Maria, Taremaa, Piia, Nahkola, Tiin, Liisa Marjatta & Pajusalu, Renate. (2019). Reference in the borderline of space and discourse: a free production experiment in Estonian, Finnish and Russian. *Linguistica Uralica*, 60(3), 185–208.
- Reinhart, Tanya. (1982). Pragmatics and linguistics. An analysis of sentence topics. *Philosophica*, 27, 53–94.
- Rießler, Michael et al. (2005+). *Kola Saami Documentation Project*. The Language Archive (TLA). Nijmegen: Max Planck Institute for Psycholinguistics. Retrieved 02-02-2020 from <http://dobes.mpi.nl/projects/sami/>.
- Rießler, Michael. (In press). Kildin Saami. In Marianne Bakró-Nagy, Johanna Laakso & Elena Skribnik (Eds.), *The Oxford Guide to the Uralic Languages*, 00–00. Oxford: Oxford University Press.
- RNC = *Russian National Corpus* (Nacional’nyj korpus russkogo jazyka). Russian Academy of Sciences. <http://ruscorpora.ru/>.
- Rochemont, Michael Shaun. (1986). *Focus in Generative Grammar*. Amsterdam: John Benjamins Publishing Company.
- Romaine, Suzanne. (1988). *Pidgin and creole Languages*. London: Longman.
- Róna-Tas, András. (1988). Turkic Influence on the Uralic languages. In Denis Sinor (Ed.), *The Uralic Languages: Description, History and Foreign Influences*, 742–780. Leiden: Brill.
- Rooth, Mats. (1985). Association with focus (Doctoral dissertation). Amherst, MA: University of Massachusetts.
- 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: Regularity and irregularity in language change*, 180–217. Oxford: Oxford University Press.
- Ross, Malcolm D. (1998). Sequencing and dating linguistic events in Oceania: the linguistics/archaeology interface. In Roger Blench & Matthew Spriggs (Eds.), *Archaeology and Language II: Correlating archaeological and linguistic hypotheses* (One World Archaeology 29), 141–173. London: Routledge.
- Ruoppila, Veikko (1986). *Venäläisperäistä sanastoa Suomen murteissa* (Suomi 136). Helsinki: Finnish Literature Society.
- Rutkowski, Paweł. (2006). From demonstratives to copulas: A cross-linguistic perspective and the case of Polish. *Journal of Universal Language*, 7(2), 147–175.
- Saarikivi, Janne. (2000). Kontaktilähtöinen kielenmuutos, substraatti ja substraattinimistö [Contact-induced language change, substratum and substrate toponyms]. *Virittäjä*, 104(3), 393–415.
- Saarikivi, Janne. (2004). Über die saamischen Substratennamen des Nordrusslands und Finnlands. *Finnisch-Ugrische Forschungen*, 58, 162–234.
- Saarikivi, Janne. (2006). *Substrata Uralica: Studies on Finno-Ugrian substrate in northern Russian dialects* (Doctoral dissertation). Tartu: University of Helsinki.
- Saarikivi, Janne. (2007a). Finnic Personal Names on Novgorod Birch Bark Documents. In Juhani Nuorluoto (Ed.), *Topics on the Ethnic, Linguistic and*

- Cultural Making of the Russian North* (Slavica Helsingiensia 32), 196–246. Helsinki: University of Helsinki.
- Saarikivi, Janne. (2007b). On the Uralic substrate toponymy of Arkhangelsk region: problems of research methodology and ethnohistorical interpretation. In Janne Saarikivi & Ritva-Liisa Pitkänen (Eds.), *Borrowing of place names in Uralian languages* (Onomastica Uralica 4), 45–109. Debrecen: University of Debrecen.
- Saarikivi, Janne. (2009). Itämerensuomalais-slaavilaisten kontaktien tutkimuksen nykytilasta [On the current state of studies in Finnic-Slavic contacts]. In Jussi Ylikoski (Ed.), *The Quasiquicentennial of the Finno-Ugrian Society* (Mémoires de la Société Finno-Ougrienne 258), 109–160. Helsinki: Société Finno-Ougrienne.
- Saarikivi, Janne. (2017). The spread of Finnic pre-Christian anthroponyms in toponyms. In Terhi Ainiala & Janne Saarikivi (Eds.), *Personal name systems in Finnic and beyond* (Uralica Helsingiensia 12, pp 13–32). Helsinki: Société Finno-Ougrienne.
- Saarikivi, Janne. (2018). Finnic and other Western Uralic borrowings in Permian. In Sampsa Holopainen & Janne Saarikivi (Eds.), *Peri orthotētos etymon. Uusiutuva uralilainen etymologia* (Uralica Helsingiensia 11), 269–356. Tallinn: Printon.
- Saarikivi, Janne. (In press). Dispersal of Proto-Uralic. In Marianne Bakró-Nagy, Johanna Laakso & Elena Skribnik (Eds.), *The Oxford Guide to the Uralic Languages*, 00–00. Oxford: Oxford University Press.
- Saarikivi, Janne & Grünthal, Riho. (2005). Itämerensuomalaisten kielten uralilainen tausta [The Uralic background of Finnic languages]. In Johanna Vaattovaara, Toni Suutari, Hanna Lappalainen & Riho Grünthal (Eds.), *Muuttuva muoto : kirjoituksia Tapani Lehtisen 60-vuotispäivän kunniaksi*, 111–146. Helsinki: University of Helsinki.
- Saarikivi, Janne & Lavento, Mika. (2012). *Linguistics and Archaeology: A Critical View of an Interdisciplinary Approach with Reference to the Prehistory of Northern Scandinavia*. In Charlotte Damm & Janne Saarikivi (Eds.), *Networks, Interaction and Emerging Identities in Fennoscandia and Beyond*, 177–216. Helsinki: Société Finno-Ougrienne.
- Sahkai, Heete. (2016). Demonstratives and Information Structure in Spoken Estonian. In M. M. Jocelyne Fernandez-Vest & Robert D. Van Valin (Eds.), *Information Structuring of Spoken Language from a Cross-linguistic Perspective*, 211–235. Berlin: Mouton de Gruyter.
- Šahmatov, A. A. (1915). Očerok drevnejšego perioda istorii russkogo jazyka [A study of the earlier period of the Russian language history]. *Ėnciklopedija slavjanskoj filologii*, 11.
- Sakel, Jeanette. (2007). Types of loan: Matter and pattern. In Yaron Matras & Jeanette Sakel (Eds.), *Grammatical borrowing in cross-linguistic perspective*, 15–29. Berlin: Mouton de Gruyter.
- Salminen, Tapani (1998). Pohjoisten itämerensuomalaisten kielten luokittelun ongelmia [Problems in the classification of North Finnic languages]. In Riho Grünthal & Johanna Laakso (Eds.), *Oekeeta asijoo. Commentationes Fenno-Ugricae in honorem Seppo Suhonen sexagenarii 16.V.1998* (Mémoires de la Société Finno-Ougrienne 228), 390–406. Helsinki: Société Finno-Ougrienne.
- Salminen, Tapani. (1999). Euroopan kielet muinoin ja nykyisin [Languages of Europe in the past and today]. In Paul Fogelberg (Ed.), *Pohjan poluilla. Suomalaisten juuret nykytutkimuksen mukaan*, 13–26. Helsinki: Finnish Society of Sciences and Letters.
- Salminen, Tapani. (2002). Problems in the taxonomy of the Uralic languages in the light of modern comparative studies. *Lingvističeskij bespredel. Sbornik statej k 70-letiju A. I. Kuznecovoj*, 44–55. Moskva: Izdatel'stvo Moskovskogo universiteta.

- Sammallahti, Pekka. (1977). Suomalaisen esihistorian kysymyksiä [Questions of Finnish prehistory]. *Virittäjä*, 81, 119–136.
- Sammallahti, Pekka. (1999). Saamen kielen ja saamelaisten alkuperästä [On the origin of Saami language and people]. In Paul Fogelberg (Ed.), *Pohjan poluilla. Suomalaisen juuret nykytutkimuksen mukaan*, 70–90. Helsinki: Finnish Society of Sciences and Letters.
- Sandfeld, Kristian. (1926). *Balkanfilologien: En oversigt over dens resultater og problemer* [Balkan philology: An overview on its results and problems]. København: B. Lunos Bogtryk.
- Sapir, Edward. (1921). Language as a historical product: drift (Chapter 7). Chapter in Edward Sapir, *Language: An Introduction to the Study of Speech*, 147–170. New York, NY: Harcourt, Brace & World.
- Sarhimaa, Anneli. (1992). Karelian Sprachbund? Theoretical basis of the study of Russian/Baltic-Finnic contacts. *Finnisch-Ugrische Forschungen*, 50(3), 209–219.
- Sarhimaa, Anneli. (1995). Karjalan kansat ja kielet kontakteissa: asutushistoriallista taustaa ja lingvistisiä seurauksia. *Virittäjä*, 99, 191–223.
- Sarhimaa, Anneli. (1999). *Syntactic transfer, contact-induced change, and the evolution of bilingual mixed codes: Focus on Karelian-Russian language alternation*. Helsinki: Finnish Literature Society.
- Saukkonen, Pauli. (1967). Persoonapronominien *hän: se, he: ne* distinktiivinen oppositio. *Virittäjä*, 71(3), 286–292.
- Saurio, Päivi & Nuorluoto, Jussi. (2016). *Kortfattad rysk språkhistoria* [A concise Russian language history]. Uppsala: Uppsala Universitet.
- Schaeken, Jos. (2012). *Stemmen op berkenbast. Berichten uit middeleeuws Rusland: dagelijks leven en communicatie*. Leiden: Leiden University Press.
- Schapper, Antoinette & San Roque, Lila. (2011). Demonstratives and non-embedded nominalisations in three Papuan languages of the Timor-Alor-Pantar family. *Studies in Language*, 35(2), 380–408.
- Schenker, Alexander M. (1995). *The dawn of Slavic: An introduction to Slavic philology*. New Haven, CT: Yale University Press.
- Schlachter, Wolfgang. (1960). *Studien zum Possessivsuffix des Syrjänschen*. Berlin: Akademie-Verlag.
- Schmerling, Susan F. (1976). *Aspects of English Sentence Stress*. Austin, TX: University of Texas Press.
- Schroeder, Christoph. (2006). Articles and Article Systems in Some Areas of Europe. In Giuliano Bernini & Marcia L. Schwartz (Eds.), *Pragmatic Organization of Discourse in the Languages of Europe* (Empirical Approaches to Language Typology 8), 545–611. Berlin: Mouton de Gruyter.
- Schwarz, Florian. (2013). Two kinds of definites cross-linguistically. *Language and Linguistics Compass*, 7(10), 534–559.
- Sebestyén, Irene. (1976). Possessive Personalsuffixe SG., PL.1. und SG., PL.2. im juraksamojedischen. *Acta Linguistica Academiae Scientiarum Hungaricae*, 26(3–4), 307–356.
- Sedov, V. V. (1990). Pribaltijsko-finskie plemena nakanune obrazovaniya drevnerusskogo gosudarstva [Finnic tribes before the formation of Old Rus' state]. In *Finnish v Evropi VI–XV veka I*, 8–35. Москва.
- Selkirk, Elizabeth O. (1984). *Phonology and Syntax: The Relation between Sound and Structure*. Cambridge, MA: MIT Press.
- Seppänen, Eeva-Leena. (1998). *Läsnäolon pronominit: tämä, tuo, se ja hän viittaamassa keskustelun osallistujiaan* [Pronouns of presence: tämä, tuo, se and hän as pointing to participants of the conversation] (Doctoral dissertation). Helsinki: University of Helsinki.

- Serebrennikov, Boris A. (1956). Problema substrata [Problems of substratum]. *Institut jazyžnaniya. Doklady i soobščeniya*, 9, 33–56.
- Serebrennikov, Boris A. (1963). *Istoričeskaja morfologija permskih jazykov* [The historical morphology of Permic languages]. Moskva: Nauka.
- Seržant, Ilja A. (2012). The so-called possessive perfect in North Russian and the Circum-Baltic area: A diachronic and areal account. *Lingua*, 122(4), 356–385.
- Seržant, Ilja A. (2015). An approach to syntactic reconstruction. In Carlotta Viti (Ed.), *Perspectives on Historical Syntax*, 117–154. Amsterdam: John Benjamins Publishing Company.
- Setälä, Eemil Nestor (1922). Lyydiläismurteet [The Lude dialects]. In *Tietosanakirja. 11. osa. Täydennysosa*, 729–731. Helsinki: Tietosanakirja-Osakeyhtiö.
- Setälä, Eemil Nestor. (1929). Itämerensuomalaisten kielten kosketukset slaavilaisten kanssa [Contacts of Finnic languages with Slavs]. *Journal de la Société Finno-Ougrienne*, 43, 30–39.
- Shee, Naw Hsar. (2008). *A descriptive grammar of Geba Karen* (Master's thesis). Chiang Mai: Payap University.
- Sidnell, Jack & Enfield, Nicholas J. (2017). Deixis and the interactional foundations of reference. In Yan Huang (Ed.), *The Oxford Handbook of Pragmatics*, 217–239. Oxford: Oxford University Press.
- Siegl, Florian. (2013). *Materials on Forest Enets, an Indigenous Language of Northern Siberia* (Mémoires de la Société Finno-Ougrienne 267). Helsinki: Société Finno-Ougrienne.
- Silverstein, Michael. (1976). Shifters, linguistic categories, and cultural description. In Keith H. Basso & Henry A. Selby (Eds.), *Meaning in anthropology*, 11–55. Albuquerque: University of New Mexico Press.
- Šimík, Radek. (2009). The syntax, semantics, and pragmatics of the focus particle *to* in Czech. In Gerhild Zybatow, Denisa Lenertová, Uwe Junghanns & Petr Biskup (Eds.), *Studies in Formal Slavic Phonology, Morphology, Syntax, Semantics and Information Structure: Proceedings of FDSL 7, Leipzig 2007*, 327–340. Frankfurt am Main: Peter Lang.
- Šimík, Radek. (2016). On pragmatic demonstratives: The case of pragmatic discourse anaphora in Czech. In *Proceedings of Sinn und Bedeutung, Vol. 20*, 640–657.
- Simpson, Andrew. (2008). The grammaticalization of clausal nominalizers in Burmese. In María José López-Couso & Elena Seoane (Eds.), *Rethinking grammaticalization: New perspectives*, 265–288. Amsterdam: John Benjamins Publishing Company.
- Singnoi, Unchalee. (2004). Discourse Functions of Thai Demonstratives: a Case On Pragmatically Controlled Irregular functions. In Somsonge Burusphat (Ed.), *Papers from the Eleventh Annual Meeting of the Southeast Asian Linguistics Society*, 645–657. Tempe, AZ: Arizona State University.
- Skalička, Vladimír. (1933–1934). Zur Charakteristik des eurasischen Sprachbundes. *Archiv Orientální*, 6, 272–274.
- Skribnik, Elena. (2001). Pragmatic Structuring in Northern Mansi. In Tõnu Seilenthal (Ed.), *Congressus nonus internationalis fenno-ugristarum, 7.–13.8.2000 Tartu. Pars IV* (Dissertationes sectionum: Linguistica 3), 222–239. Tartu: University of Tartu.
- Sohn, Ho-Min. (1999). *The Korean language*. Cambridge: Cambridge University Press.
- Sosa, Sachiko. (2017). *Functions of morphosyntactic alternations, and information flow in Surgut Khanty Discourse* (Doctoral dissertation). Helsinki: University of Helsinki.

- Sperber, Dan & Wilson, Deirdre. (1986). *Relevance: Communication and cognition*. Oxford: Blackwell Publishing.
- Spriggs, Matthew. (1998). From Taiwan to the Tuamotus: absolute dating of Austronesian language spread and major sub-groups. In Roger Blench & Matthew Spriggs (Eds.), *Archaeology and Language II: Correlating archaeological and linguistic hypotheses* (One World Archaeology 29), 115–127. London: Routledge.
- Stachowski, Marek. (1998). An example of Nganasan-Dolgan linguistic contact. *Turkic Languages*, 2, 126–129.
- Stachowski, Marek. (2010). On the article-like use of the Px2Sg in Dolgan, Nganasan and some other languages in an areal Siberian context. *Finnisch-Ugrische Mitteilungen*, 32(33), 587–593.
- Stadnik-Holzer, Elena. (2004a). Areallinguistik und Rekonstruktion von Sprachgeschichte (Am Beispiel des Slawischen). *East Central Europe*, 31(1), 61–76.
- Stadnik-Holzer, Elena. (2004b). Grundsätzliches zur Erforschung des eurasischen Sprachbundes (Methodologisches, Theoretisches, Historisches). *Zeitschrift für slavische Philologie*, 63(1), 9–21.
- Stadnik-Holzer, Elena. (2006). Zur Frage nach der Herkunft der sog. postponierten Partikel in den nordrussischen Dialekten. In Juhani Nuorluoto (Ed.), *The Slavization of the Russian North: Mechanisms and Chronology* (Slavica Helsingiensia 27), 347–354. Helsinki: University of Helsinki.
- Stolz, Thomas. (1991). *Sprachbund in Baltikum? Estnisch und Lettisch im Zentrum einer sprachlichen Konvergenzlandschaft* (Bochum-Essener Beiträge zur Sprachwandelforschung 13). Bochum: Universitätsverlag Dr. N. Brockmeyer.
- Strawson, Peter F. (1964). Identifying reference and truth-values. *Theoria*, 30(2), 96–118.
- Stubbs, Michael. (1986). A Matter of Prolonged Field Work: Notes Towards a Modal Grammar of English. *Applied Linguistics*, 7(1), 1–25.
- Svane, Gunnar. 1961. O sintaksičeskom primenenii bolgarskogo člena v XIII veke. A. Čergedskie teksty [On a syntactic function of the Bulgarian article in 13th century. A. Čerged texts]. *Scando-Slavica*, 7, 233–251.
- Svane, Gunnar. 1962. O sintaksičeskom primenenii bolgarskogo člena v XIII veke. B. Dobrejšovo evangelie [On a syntactic function of the Bulgarian article in 13th century. B. Dobreyshovo gospel]. *Scando-Slavica*, 8, 224–238.
- Suhonen, Seppo. (1990). Über fremden Einfluss auf die Struktur der ostseefinnischen Sprachen. In Sari Vaula (Ed.), *Itämerensuomalaiset kielikontaktit/ Läänemeresoome keelekontaktid*, 94–104. Helsinki: Institute for the Languages of Finland.
- Sundbäck, Irene. (1995). Suomen kielessäkö artikkeli? [An article in the Finnish language?]. *Kielikello*, 2, 25–28.
- Tabakowska, Elżbieta. (1989). On pragmatic functions of the particle *to* in Polish. In Harald Weydt (Ed.), *Sprechen mit Partikeln*, 535–545. Berlin: Mouton de Gruyter.
- Tammekänd, Liina. (2015). Demonstratives in Võru and Estonian oral narratives. *Journal of Estonian and Finno-Ugric Linguistics*, 6(2), 191–216.
- Talmy, Leonard. (1985). Lexicalization patterns: semantic structure in lexical forms. In Timothy Shopen (Ed.), *Language Typology and Syntactic Description Vol. 3: Grammatical Categories and the Lexicon*, 57–149. Cambridge: Cambridge University Press.
- Tauli, Valter. (1966). *Structural tendencies in Uralic languages*. The Hague: Mouton de Gruyter.
- Tereščenko, Natalja M. (1979). *Nganasanskij jazyk* [The Nganasan language]. Leningrad: Nauka.

- Thomason, Sarah Grey. (2001). *Language Contact. An Introduction*. Edinburgh: Edinburgh University Press.
- Thomason, Sarah Grey. (2010). Contact Explanations in Linguistics. In Raymond Hickey (Ed.), *The Handbook of Language Contact*, 31–47. Oxford: Blackwell Publishing.
- Thomsen, Vilhelm. (1893). *Berøringer mellem de finske og de baltiske (litauisk-lettiske) sprog: en sproghistorisk undersøgelse* [Contacts between Finnic and Baltic (Lithuanian-Latvian) languages: a study of language history]. København: Lunos.
- Timberlake, Alan. (1974). *The Nominative Object in Slavic, Baltic, and West Finnic*. München: Verlag Otto Sagner.
- Timberlake, Alan. (2013). Culture and the spread of Slavic. In Balthasar Bickel, Lenore A. Grenoble, David A. Peterson & Alan Timberlake (Eds.), *Language Typology and Historical Contingency. In honor of Johanna Nichols*, 331–355. Amsterdam: John Benjamins Publishing Company.
- Tkačenko, Orest B. (1989). *Sopostavitel'no-istoričeskaja frazeologija slavjanskih i finno-ugorskih jazykov* [Comparative-historical phraseology of Slavic and Finno-Ugric languages]. Kiev: Naukova dumka.
- Toivonen, Yrjö H. ([1955]1987). *Suomen kielen etymologinen sanakirja, Vol. 1* [An etymological dictionary of the Finnish language] (5th ed.). Helsinki: Société Finno-Ougrienne.
- Tomingas, Marili. (2018). Demonstrative proadjectives in Livonian—morphosyntactic use and semantic functions. *Journal of Estonian and Finno-Ugric Linguistics*, 9(2), 241–255.
- Tommola, Hannu. (1998). Grammatičeskij stroj russkogo jazyka v tipologičeskom osveščanii [Grammatical system of the Russian language in the light of typology]. *Studia Slavica Finlandensia*, 15, 146–167.
- Tournadre, Nicolas, & LaPolla, Randy J. (2014). Towards a new approach to evidentiality: Issues and directions for research. *Linguistics of the Tibeto-Burman area*, 37(2), 240–263.
- Trenkić, Danijela. (2004). Definiteness in Serbian/Croatian/Bosnian and some implications for the general structure of the nominal phrase. *Lingua*, 114(11), 1401–1427.
- Trovesi, Andrea. (2004). *La genesi di articoli determinativi. Modalità di espressione della definitezza in ceco, serbo-lusaziano e sloveno*. Pavia: Franco Angeli.
- Trubetzkoy, Nikolai S. (1923). *Vavilonskaja bašnja i smešenie jazykov* [The tower of Babel and the confusion of languages]. *Evrasijskij vremennik*, 3, 107–124.
- Trubetzkoy, Nikolai S. (1928). Proposition 16. In *Actes de 1er Congrès international de linguistes*, 17–18. Leiden: A. W. Sijthoff's Uitgevermaatschappij.
- Trubinskij, V. I. (1970). Ob učastii postpozitivnoj časticy *to* v postroenii složnogo predloženiya (po materialam pinežskih govorov [On the role of postpositive particle *to* in the formation of complex clause construction (based on materials of Pinega varieties)]. *Severnorusskie govory*, 1, 48–66.
- Trudgill, Peter. (2011). *Sociolinguistic typology: Social determinants of linguistic complexity*. Oxford: Oxford University Press.
- Turunen, Aimo. (1946). *Lyydiläismurteiden äännehistoria I. Konsonantit* [Historical phonology of Lude dialects I. Consonants] (Mémoires de la Société Finno-Ougrienne 89). Helsinki: Société Finno-Ougrienne.
- Turunen, Aimo. (1947). Lyydin asema vepsän ja karjala-aunuksen välimurteistona. *Virittäjä*, 51, 1–105.

- Turunen, Aimo. (1950). *Lyydiläismurteiden äännehistoria II. Vokaalit* [Historical phonology of Lude dialects II. Vowels] (Mémoires de la Société Finno-Ougrienne 99). Helsinki: Société Finno-Ougrienne.
- Tužarov, G. M. (1987). *Grammatičeskie kategorii imeni suščestvitel'nogo v marijskom jazyke* [Grammatical categories of nouns in Mari]. Joškar-Ola: Marijskoe knižnoe izdatel'stvo.
- Tvauri, Andres. (2007). Migrants or natives? The research history of long barrows in Russia and Estonia in the 5th–10th centuries. In Juhani Nuorluoto (Ed.), *Topics on the Ethnic, Linguistic and Cultural Making of the Russian North* (Slavica Helsingiensia 32), 247–285. Helsinki: University of Helsinki.
- Umbach, Carla. (2004). On the notion of contrast in information structure and discourse structure. *Journal of Semantics*, 21(2), 155–175.
- Umbach, Carla, & Ebert, Cornelia. (2009). German demonstrative 'so' – intensifying and hedging effects. *Sprache und Datenverarbeitung* (International Journal for Language Data Processing), 2009/1-2, 153–168.
- Vaahtera, Jouni. (2009). *Ėvoljucija sistemy glasnyh fonem v nekotoryh russkih govoraх Vologodskoj oblasti* [Evolution of phonological system in several Russian dialects of Vologda Oblast] (Doctoral dissertation, Slavica Helsingiensia 37). Helsinki: University of Helsinki.
- Vahros, Igor. (1951). Der Artikel im Russischen. *Neuphilologische Mitteilungen*, 52(1), 30–44.
- Vaillant, André. (1958). *Grammaire comparée des langues slaves II: Morphologie*. Lyon: I.A.C.
- Vaillant, André. (1977). *Grammaire comparée des langues slaves V: La syntaxe*. Paris: Klincksieck.
- Valkhoff, Marius. (1932). *Latijn, Romaans, Roemeens. Openbare les gegeven bij de opening van Zijn Kolleges als privaats-docent in de roemeense taal en letterkunde ann de Universiteit van Amsterdam op 1 novembre 1932 door Amersfoort*. Amersfoort: Valjhoff et Co.
- Vallduví, Enric. (1992). *The Informational Component*. New York, NY: Garland.
- Vallduví, Enric & Engdahl, Elisabet. (1996). The Linguistic Realization of Information Packaging. *Linguistics*, 34, 459–519.
- Vallduví, Enric & Vilks, Maria. (1998). On rheme and kontrast. In Peter Culicover & Louis McNally (Eds.), *Syntax and semantics* (The Limits of Syntax 29), 79–108. New York, NY: Academic Press.
- Varteva, Annukka. (1998). Pronominit *hän* ja *tämä* tekstissä [Pronouns *hän* and *tämä* in the text]. *Virittäjä*, 102(2), 202–223.
- Vasmer, Max. ([1950–1958]1987). *Ėtimologičeskij slovar' russkogo jazyka, Tom IV* [An etymological dictionary of the Russian language, Vol. 4]. Moskva: Progres. (A Russian translation from the original book in German, *Russisches etymologisches Wörterbuch*. Heidelberg: Carl Winter Universitätsverlag).
- Večerka, Radoslav. (1993). *Altkirchenslavische (altbulgarische) Syntax II: Die innere Satzstruktur*. Freiburg im Breisgau: U. W. Weiher.
- Veenker, Wolfgang. (1967). *Die Frage des finnougriischen Subtrats in der russischen Sprache* (Uralic and Altaic Series 82). Bloomington, IN: Indiana University Press.
- Vermeer, Willem. (1986). The rise of the North Russian dialect of Common Slavic. *Studies in Slavic and General Linguistics*, 8, 503–515.
- Vermeer, Willem. (1991). Aspects of the oldest Finnic poem and some related texts (Novgorod birchbark documents 292, 56 and 403). *Studies in Slavic and General Linguistics*, 16, 315–369.
- Vermeulen, Reiko. (2009). On the syntactic typology of topic marking: A comparative study of Japanese and Korean. *UCL Working Papers in Linguistics*, 21, 335–363.

- Viitso, Tiit-Rein. (1985). Kriterien zur Klassifizierung der Dialekte der ostseefinnischen Sprachen. In Wolfgang Veenker (Ed.), *Dialectologia Uralica* (Veröffentlichungen der Societas Uralo-Altaica 20), 89–96. Wiesbaden: Societas Uralo-Altaica.
- Viitso, Tiit-Rein. (1990a). On the earliest Finnic and Balto-Slavic contacts. In Sari Vaula (Ed.), *Itämerensuomalaiset kielikontaktit/Läänemeresoome keelekontaktid*, 140–147. Helsinki: Institute for the Languages of Finland.
- Viitso, Tiit-Rein. (1990b). Kommenttipuheenvuoro Eino Koposen esitelmään ‘Viron ja liivin yhteissanaston tarkastelua’ [Comments to Eino Koponen’s presentation ‘Revision of Estonian and Livonian common vocabularies’]. In Sari Vaula (Ed.), *Itämerensuomalaiset kielikontaktit/Läänemeresoome keelekontaktid*, 40–43. Helsinki: Institute for the Languages of Finland.
- Viitso, Tiit-Rein. (2000). Finnic Affinity. In *Congressus Nonus Internationalis FennoUgristarum I: Orationes plenariae & Orationes publicae*, 153–178. Tartu: OÜ Paar.
- Viitso, Tiit-Rein. (2008). *Liivi keel ja läänemeresoome keelemaastikud* [The Livonian language and the Finnic linguistic landscapes]. Tartu: Eesti Keele Sihtasutus.
- Vilkuna, Maria. (1989). *Free word order in Finnish: Its syntax and discourse functions* (Doctoral dissertation). Helsinki: University of Helsinki.
- Vilkuna, Maria. (1992). *Referenssi ja määräisyys suomenkielisten tekstien tulkinnessa* [Reference and definiteness in the interpretation of Finnish texts]. Helsinki: Finnish Literature Society.
- Vilkuna, Maria. (1998). Word order in European Uralic. In Anna Siewierska (Ed.), *Constituent order in the languages of Europe*, 173–233. Berlin: Mouton de Gruyter.
- Virtanen, Susanna. (2013). *Transitivity in Eastern Mansi - An Information Structural Approach* (Doctoral dissertation). Helsinki: University of Helsinki.
- Virtaranta, Pertti. (1972). Die Dialekte des Karelischen. Gruppierung (mit Kurzcharakteristika) und einige Fragen des Ursprungs. *Sovetskoje finno-ugrovedenije*, 8, 7–27.
- Vlasto, Alexis Peter. (1986). *A linguistic history of Russia to the end of the eighteenth century*. Oxford: Clarendon Press.
- Voge, Noel. (1958). The post-positive article of Avvakum and the problem of norms in seventeenth-century Russian. *The Slavic and East European Journal*, 16, 115–119.
- Vostrikov, O. V. (1990). *Finno-ugorskij substrat v russkom jazyke* [Finno-Ugric substratum in the Russian language]. Sverdlovsk: Ural’skij gosudarstvennyj universitet.
- Wackernagel, Jakob. (1892). Über ein Gesetz der indogermanischen Wortstellung. *Indogermanischen Forschungen*, 1, 333–436.
- Wagner-Nagy, Beáta & Szeverényi, Sándor & Várnai, Zsuzsa. (2002). *Chrestomathia nganasanica* (SUA Supplementum10). Szeged-Budapest: SZTE Finnugor Tanszék, MTA Nyelvtudományi Intézet.
- Wälchli, Bernhard. (Forthcoming). Topics in subordinate clauses.
- Warotamasikkhadit, Udom. (1997). Fronting and backing topicalization in Thai. *Mon-Khmer Studies*, 27, 303–306.
- von Wartburg, Walther. (1951). *Die Ausgliederung der romanischen Sprachräume*. Bern: A. Francke Ag. Verlag.
- Weinreich, Uriel. ([1953]1974). *Language Contact: Findings and Problems* [8th printing]. The Hague: Mouton de Gruyter.

- Westerståhl, Dag. (1985). *Determiners and context sets*. In Alice ter Meulen & Johan van Benthem (Eds.), *Generalized quantifiers in natural language* (Groningen-Amsterdam Studies in Semantics 4), 45–72. Berlin: Mouton de Gruyter.
- Wexler, Paul. (1976). On the non-lexical expression of determinedness (with special reference to Russian and Finnish). *Studia Linguistica (Lund)*, 30(1), 34–67.
- White, Peter R. R. (2015). Appraisal theory. *The international encyclopedia of language and social interaction*, 1–7. Oxford: Wiley Blackwell.
- Whorf, Benjamin Lee. ([1939]1956). The Relation of Habitual Thought and Behavior to Language. In J. B. Carroll (Ed.), *Language, Thought & Reality. Selected Writings of Benjamin Lee Whorf*. Cambridge, MA: MIT Press.
- Wiemer, Björn & Wälchli, Bernhard. (2012). Contact-induced grammatical change: Diverse phenomena, diverse perspectives. In Björn Wiemer, Bernhard Wälchli & Björn Hansen (Eds.), *Grammatical Replication and Borrowability in Language Contact* (Trends in Linguistics 242), 3–64. Göttingen: Mouton de Gruyter.
- Wilbur, Joshua. (2014). *A Grammar of Pite Saami* (Studies in Diversity Linguistics 5). Berlin: Language Science Press.
- Wilson, Patricia R. (1980). *Ambulas Grammar*. Ukarumpa, Papua New Guinea: Summer Institute of Linguistics.
- Wilson, John & Henry, Alison. (1998). Parameter setting within a socially realistic linguistics. *Language in Society*, 27(1), 1–21.
- Winkler, Eberhard. (2002). Der livische Wortschatz russischer Provenienz. In Eugen Helinski & Anna Widmer (Eds.), *Wúsa wúsa – sei begrüßt! Beiträge zur Finnougristik zu Ehren von Gert Sauer dargebracht zu seinem siebzigsten Geburtstag*, 379–385. Wiesbaden: Harrassowitz Verlag.
- Zacharski, Ron. (1993). *A Discourse Pragmatics Model of English Accent* (Doctoral dissertation). Minneapolis, MN: University of Minnesota.
- Yap, Foong Ha. (2011). Referential and non-referential uses of nominalization constructions in Malay. In Foong Ha Yap, Karen Grunow-Hårsta & Janick Wrona (Eds.), *Nominalization in Asian Languages: Diachronic and typological perspectives*, 627–658. Amsterdam: John Benjamins Publishing Company.
- Ylikoski Jussi. (2016). Degrammaticalization in North Saami: Development of adpositions, adverbs and a free lexical noun from inflectional and derivational suffixes. *Finnisch-ugrische Mitteilungen*, 40, 113–173.
- Ylikoski, Jussi. (2020). Remarks on the Saami demonstrative pronouns. In Sampsa Holopainen, Juha Kuokkala, Janne Saarikivi & Susanna Virtanen (Eds.), *Ёмас самыу нэ́кве вёртур этност самын патум. Scripta miscellanea in honorem Ulla-Maija Forsberg* (Mémoires de la Société Finno-Ougrienne 275), 398–417. Helsinki: Société Finno-Ougrienne.
- Yurayong, Chingduang. (2016). Adnominal possession with kinship terms in Macedonian dialects around Lakes Ohrid and Prespa. In Max Wahlström & Maxim Makartsev (Eds.), *In search of the center and periphery: Linguistic attitudes, minorities, and landscapes in the Central Balkans* (Slavica Helsingiensia 49), 143–161. Helsinki: University of Helsinki.
- Yurayong, Chingduang. (2019). Predicative possession in Novgorod Birch Bark documents in the Ural-Altaic context. *Journal de la Société Finno-Ougrienne*, 97, 183–233.
- Yurayong, Chingduang. (2020). Postposed demonstrative: An innovation from contact between North Russian and Central Veps dialects. In Yoshiyuki Asahi (Ed.), *Proceedings of Methods XVI. Papers from the sixteenth international conference on Methods in Dialectology, 2017* (Bamberger Beiträge zur Englischen Sprachwissenschaft 59), 47–64. Bern: Peter Lang.

- Yurayong, Chingduang & Kittilä, Seppo. (Forthcoming). Demonstrative and pronoun as enclitic discourse markers in Finnish and other Finnic languages. In Foong Ha Yap (Ed.), *Demonstrative in interaction* (Studies in Language).
- Yurayong, Chingduang & Szeto, Pui Yiu. (2020). Altaicization and de-Altaicization of Japonic and Koreanic. *International Journal of Eurasian Linguistics*, 2(1), 108–148.
- Zajceva, Maria I. (1981). *Grammatika vepsskogo jazyka: fonetika i morfologija* [A grammar of the Veps language: phonetics and morphology]. Leningrad: Nauka.
- Zajceva, Maria I. & Mullonen, Maria I. (1969). *Obrazcy vepsskoj reči* [Samples of Veps dialects]. Leningrad: Nauka.
- Zajceva, Nina. (1995). *Vepsän kelen grammatik: I. Nimiden kändluz* [A grammar of the Veps language. I Inflection of nouns]. Petroskoi: Petroskoin valtionyliopiston kustantamo.
- Zaliznjak, Andrej A. (2004). *Drevnenovgorodskij dialekt* [The Old Novgorod dialect]. Moskva: Jazyki slavjanskoj kul'tury.
- Zaliznjak, Andrej A. (2008). *Drevnerusskie ènklitiki* [Old Russian enclitics]. Moskva: Jazyki slavjanskoj kul'tury.
- Zayzon, Réka. (2015). Observations on non-possessive usages of personal markers (possessive suffixes) in Nganasan. *Journal of Estonian and Finno-Ugric Linguistics*, 6(2), 259–278.
- Zeps, Valdis. (1962). *Latvian and Finnic Linguistic Convergence* (Uralic and Altaic Series 9). Bloomington, IN: Indiana University Press.
- Žuravlev, A. F. (1994). *Leksiko-statističeskoe modelirovanie systemy slavjanskogo jazykovogo rodstva* [Lexico-statistic modelling of the system of Slavic linguistic relationship]. Moskva: Indrik.
- Zybatow, Lew. (1990). *Was die Partikeln bedeuten. Eine kontrastive Analyse Russisch-Deutsch* (Slavistische Beiträge 254). München: Verlag Otto Sagner.