

RODOLFO BASILE

Inventive-locational constructions in Finnish:
A mixed methods approach



UNIVERSITY OF TARTU

Press

1632

University of Tartu, Institute of Estonian and General Linguistics

Dissertation accepted for the defence of the degree of Doctor of Philosophy on 12 March 2024 by the Committee of the Institute of Estonian and General Linguistics, Faculty of Arts and Humanities, University of Tartu.

Supervisors: Professor Liina Lindström (University of Tartu)
Professor Tuomas Huumo (University of Turku/ University of Tartu)
Associate Professor Ilmari Ivaska (University of Turku)

Opponent: PhD Seppo Kittilä (University of Helsinki)

Commencement: May 24, 2024 at 14.15 at the Senate's Hall (Ülikooli 18-204, Tartu)

This study has been supported by the Graduate School of Linguistics, Philosophy and Semiotics; funded by the European Regional Development Fund (University of Tartu ASTRA Project PER ASPERA).



European Union
European Regional
Development Fund



Investing
in your future

ISSN 1406-5657 (print)
ISBN 978-9916-27-504-7 (print)
ISSN 2806-237X (pdf)
ISBN 978-9916-27-505-4 (pdf)

Copyright: Rodolfo Basile, 2024

University of Tartu
www.tyk.ee

ACKNOWLEDGMENTS

My first thoughts go to my high school English teacher, Claudia Tabiano. Without her support and encouragement to pursue a degree in languages, I would have probably ended up being an unhappy engineer. When I started my BA at the University of Naples “L’Orientale” in 2011, I had chosen to study English and Mandarin before realizing there was a Finnish language course. I thought it would be a fun idea to take it as an elective course, since I had already been learning a few Finnish words and expressions. I certainly hadn’t imagined that this language would become a big part of my career and life. It was just perfect to study Finnish in Naples, where our teachers, Pirjo Nummenaho and Riikka Lindholm, created the friendliest and most comfortable environment our small class (only six people on the busy days!) could ever hope for. As a result, Finnish grammar was the only thing I diligently studied every day. I owe many thanks to both: to Pirjo for guiding me as my supervisor, and to Riikka not only for motivating me to speak Finnish and sending me to Finland to attend summer courses, but also for being a great friend. Even if Riikka is no longer with us, her memory will live on in her students’ hearts. Among these students, I would like to thank Marco Staiano, Marco Liazza, and Ennio Sapatello for making every moment precious during my MA studies.

In 2017 I got a scholarship that allowed me to go to Helsinki for a semester. From there, I was accepted to the student conference IFUSCO in Warsaw. There I met Merit Niinemägi and Anna Kuznetsova, who deserve a big thanks for telling me about the possibility of applying for a PhD in Tartu. They also mentioned they’d organize the same conference in Tartu in the spring of 2018. Fast forward, I went back to Naples, where I listened to a guest lecture by a Finnish scholar, Ilmari Ivaska. I told him about my plan to go to Tartu for the conference and maybe look for a supervisor there, but I wasn’t very hopeful. He suggested that I contact Liina Lindström, so I did. Liina was very supportive and agreed to become my supervisor together with Tuomas Huumo and Ilmari himself. When I got the grant and started my PhD, I couldn’t believe it. Something that started off as a big *maybe* was becoming real, as everything seemed to happen at just the right time. I owe my supervisors a big thanks for making this possible. Their expertise and professionalism have been a source of inspiration for all these years.

In Tartu I found a warm working environment, despite the harsh winters. All my colleagues from the department deserve my thanks for their interest in linguistics and for always providing constructive feedback. A special thanks goes to my dear friend Denys Teptiuk for the long talks about linguistic typology and the languages of the world, and to Gerson Klumpp for his class “Comparative Grammar of Finno-Ugric Languages”. The topics discussed in class with Gerson and Mariann Bernhardt helped me formulate the main theoretical concept behind my work which, without these inputs, probably wouldn’t have taken the comparative turn it eventually took. Gerson was the first person to encourage me to pursue my topic. I also owe my special thanks to Martin Haspelmath, who read several drafts

of my paper and suggested the term *inventive* (in lieu of my original *situative*) to name the constructions I study.

One of my favorite things about being enrolled in a PhD program is going to conferences and research facilities abroad to network with other scholars and see the world. I thank Marko Pantermöller and colleagues for hosting me at the University of Greifswald in November 2022. I got a lot of work done. I also thank John Beavers for a research visit to the University of Texas at Austin in December 2022, and all his colleagues for the warm welcome to their department. John gave me valuable feedback about my work. During my conference trips to the US, I also got to finally meet my cousin Mike Basile and his wife Patty, who I thank for welcoming me to the American side of the family that I thought I would never get acquainted with. I will also always treasure the week spent at the University of Hermosillo in November 2023 for a seminar on nonverbal predications and related constructions organized by Zarina Estrada Fernández. The hospitality of Albert Álvarez González and his students made me feel welcome both in the hot Sonoran Desert and on the beautiful coasts of the Seri people. Furthermore, many thanks to Chris Lasse Däbritz and Josefina Budzisch for including me in the organization of the workshop “Locative and existential predication – Core and periphery”, held in Athens at SLE2023. It was a very rewarding experience. A sincere thank you goes to the people from the “Communicating Science” course (Katarina Damčević, Roger Yallop, Djuddah Leijen) for raising awareness of academic writing, letting me be part of their project, and organizing wonderful writing retreats. My appreciation also goes to the people from the Livonian Institute in Riga, especially Gunta Kļava and Valts Ernštreits, for letting me drive a small bus to carry scholars around the Livonian coast while learning about Livonian language and culture in 2021. It was the highlight of the year.

A lot of people deserve to be thanked for commenting on my work and providing feedback over the years, as well as for showing interest in my topic. I’m probably forgetting someone, but here’s a non-exhaustive list: Anindita Sahoo, Anna Kampanarou, Ann Veismann, Anne Tamm, Antoine Guillaume, Bodo Winter, Bogáta Timár, Brian Joseph, Caterina Cacioli, Chingduang Yurayong, Daiki Horiguchi, Giacomo Bucci, Guglielmo Inglese, Ilja A. Seržant, Irantzu Epelde, Jeremy Bradley, Johannes Hirvonen, Joshua Wilbur, Karl Pajusalu, Kata Kubinyi, Kirsty Azzopardi, Livio Gaeta, Luca Brigada Villa, Marc Daniel Skipstað Volhardt, Maarja-Liisa Pilvik, Mari Aigro, Mariann Proos, Marili Tomingas, Marja-Liisa Helasvuo, Matti Miestamo, Michela Cennamo, Monica Kahumburu, Natalia Levshina, Petar Kehayov, Petra Sleeman, Pier Marco Bertinetto, Ray Fabri, Rita Eloranta, Rogier Blokland, Seppo Kittilä, Silvia Luraghi, Spike Gildea, Urtzi Etxeberria, Virve Vihman, and Zygmunt Frajzyngier, thank you all!

Thanks to Miina Norvik, Yrjö Lauranto, and Unni Leino for reviewing this thesis and helping me improve its quality, and to Uldis Balodis for proofreading it. Thanks to Kulla Mellov and Liina Lindström for proofreading the Estonian summary, and to Ilmari Ivaska for proofreading the Finnish one. Thanks to Tiia Margus and Meri Heinonen for all their help.

During these years I have made a lot of friends. In fact, too many to mention them all. Thanks to the Lai crew (Anna Branets, Merit Niinemägi, Allan Niinemägi, Līga Britāla, Andris Maskovs, Patrick O'Rourke) for having been good neighbors. Thanks to the band Kännu Peal Kābi (Britt-Kathleen Mere, Anna Venchakova, Dima Denisov, Merit Niinemägi, Heikki Mändmets, Anti Lillak, Bogáta Timár) for letting me be a rockstar. Thanks to Anni Polding, Deniss Savchenko, Eda-Riin Tuuling, Imar Koutchoukali, and Micke Mayer for the good times. Thanks to Michele Tita for giving me the chance to speak my home dialect even in Estonia, to Alessandra Dezi for the gossip sessions and to all the other Italian friends I met in Tartu (Emanuele, Jacopo, Marzia, Amedeo, Francesca, Alessandra, Mattia, Alessia). Thanks to Jesse Holmes for the inspiring topics we discuss, to James Baxenfield for all the jokes, and to Jay Zameska for letting me crash on his couch multiple times, for writing stories together and for reading every chunk of text I have ever sent him. Thanks to Tommy Stara and Milo Orlich for supporting me, especially before starting this journey, and for continuing to do so. Many thanks to Nicola Albarella for the great times, the inside jokes, the long walks during the pandemic, and for ranting together about our very different PhD projects. Thanks to my friends Marco, Elio, Silvia, Enrico, and Salvatore for always being there unconditionally, some of them for as long as I can remember.

A special thanks goes to Joanna Mae Cabuyadao, because everything is just so much easier when she is around. She is also a great Tagalog informant.

I was lucky to be born into a loving family. I thank all my uncles and cousins for being simply amazing, especially zio Peppe, who has always believed in me. I think this has played a fundamental role in my success since I was a kid. I thank my grandmom Emilia for being so understanding and youthful, and I dearly remember my grandpa Luigi, who passed away while this work was in the making. He has always been a source of wisdom and inspiration for the whole family. Thanks to my mom Annalisa and my dad Gerardo for being the best parents I could ever hope for. It's also thanks to their moral support and their positive attitude that I made it to this point. Thanks to our dog Chico for being himself. Finally, I want to thank my dear brother Riccardo who, despite being younger, always looks after me and knows how to talk sense into me during my darkest times. This thesis is dedicated to him.

Tartu, 10.04.2024
Rodolfo Basile

CONTENTS

LIST OF PUBLICATIONS.....	9
LIST OF GLOSSING ABBREVIATIONS.....	10
INTRODUCTION.....	11
1 BACKGROUND	14
1.1 Linguistic typology and functionalist approaches to grammar	14
1.2 Inventive-locational constructions as identifiable construction- strategies.....	16
1.3 Earlier research on nonverbal clause constructions, copulas, and locational constructions.....	17
1.4 Subjecthood and partitives in Finnish Existential constructions.....	20
2 DATA AND METHODS	23
2.1 Data	23
2.2 A mixed methods approach.....	25
2.3 Quantitative methods in detail.....	26
2.3.1 Mixed-effect logistic regression models	26
2.3.2 Collostructional analysis	27
2.3.3 Expected Sample size.....	27
3 RESULTS	30
3.1 Finnish inventive-locational constructions.....	30
3.2 Inventive-locational constructions (AREA)	31
3.3 Nominative-partitive alternation in Finnish inventives (QUANTI)	35
3.4 Marginal FIND-based strategies in Finnish: Internet contexts (WEB)..	40
3.5 Finnish Existential Partitive constructions and Expected Sample sizes (PARTEX)	42
3.6 Discussion	43
4 CONCLUSION.....	47
REFERENCES.....	49
EESTIKEELNE KOKKUVÕTE.....	53
SUOMENKIELINEN TIIVISTELMÄ.....	59
PUBLICATIONS.....	65
CURRICULUM VITAE	174
ELULOOKIRJELDUS.....	176

LIST OF PUBLICATIONS

This doctoral dissertation consists of the following original research articles, here referred to by their abbreviations:

AREA (Article I): Basile, Rodolfo (Accepted for publication). Invenitive-locational constructions in the languages of Europe.

QUANTI (Article II): Basile, Rodolfo, and Ilmari Ivaska. 2021. Löytyä-verbin konstruktioiden yhteydessä esiintyvä subjektin sijanvaihtelu. *Eesti ja soome-ugri keeleteaduse ajakiri. Journal of Estonian and Finno-Ugric Linguistics*, 12(1), 11–39.

WEB (Article III): Basile, Rodolfo (Accepted for publication). ‘I am also found on Facebook’: Locuphoric ‘find’-based strategies in Finnish internet corpora. In Glaser, Sleeman, Strobel & Tamm (eds.). Partitive constructions and partitive elements within and across language borders in Europe Vol. – num. 5 (prov.) *Linguaggio e Variazione. Variation in Language*. Edizioni Ca’Foscari, Venice University Press.

PARTEX (Article IV): Basile, Rodolfo (Accepted for publication). The Finnish Existential Partitive Construction Analyzed: Comparing Two Applications of Collostructional Analysis. In Leuschner, Barðdal, Delaby & Vajnovszki (eds.). *How to Do Things with Corpora*. Berlin/Heidelberg: J.B. Metzler. (Linguistik in Empirie und Theorie/Empirical and Theoretical Linguistics).

Author’s contribution to the publications

In QUANTI, Basile manually coded the data, which was then quantitatively modeled by Ivaska. The qualitative analysis of the data and the main insights of the article are by Basile.

As for the rest of the articles, Basile is the sole author.

NB: The article AREA might undergo changes before the final version appears.

LIST OF GLOSSING ABBREVIATIONS

1	first person	GROUP	group
2	second person	ILL	illative
3	third person	INE	inessive
ABL	ablative	INF	infinitive
ADE	adessive	LOC	locative
ALL	allative	MM	middle marker
CONJ	conjunction	NOM	nominative
COP	copula	PART	partitive
DEF	definite	PL	plural
DET	determiner	PRS	present
ELA	elative	PRT	preteritum
EMPH	emphatic	PX	personal suffix
ENCL	enclitic	RELP	relative pronoun
EXV	existive	SG	singular
GEN	genitive	USER	username

INTRODUCTION

In natural languages, there are several ways to indicate the location of something or someone in space. In English, saying *the band is in town* is different than saying *there is a band in town*. In the first case, the band in question is a band we already know something about, one that is presumably familiar to us speaking as well as to those who are listening. This familiarity of the band is expressed using the definite article *the* and the sentence-initial position of the word *band*, which tells us this piece of information is salient, and that this word is the subject of the sentence. In the second case, the band does not represent a familiar piece of information: It is introduced by the indefinite article *a* and does not qualify as a subject in the same way. What we have here is a functional distinction: The function of the first sentence is that of telling us something about a certain band, while the second sentence has instead the function of telling us something about a certain town by placing a band in it. There are, however, several formal ways of expressing these functional differences. One could for example say that *the coffee has been sitting on the stove the whole morning*, and in this way say something about the coffee, which now *sits* instead of just *is* on the stove. The function of this sentence remains that of locating the coffee in a certain space, but this is conveyed in a formally distinct way, using a verb whose meaning is a little more complex than the verb ‘to be’. Yet another way of placing things in space is by using the verb ‘to find’, as in, *life jackets are found under your seats*. What this last sentence tells us is that our life jackets are located at all times in a certain place, and the fact that we can find them there in case our plane lands on the water is only a secondary piece of information. The present work deals precisely with these kinds of grammatical constructions where the existence or location of something or someone is expressed by *finding* it. This thesis provides both a theoretical framework and an analysis within the context of European languages, giving a central role to Finnish. As these locational constructions are based on the verb ‘to find’, I call them *inventive-locational constructions*, from the Latin verb *invenire* ‘to find’.

This doctoral dissertation is a contribution to several subfields of linguistic research, including language typology, corpus linguistics, and Fennistics, which here is intended as the study and analysis of Finnish (and, in a broader sense, of other Finnic languages, which this dissertation does not, however, take into account). This work is based on four publications brought together by their common purpose to understand, using mixed methods, the locational constructions I call *inventive-locational* constructions. Such constructions are common in European language, and in Finnish they are represented by the verb *löytyä* ‘to be found’. It will be shown that not only the use of this Finnish verb is widespread in different kinds of usage contexts, but that it also behaves differently compared to its counterparts in other European languages. One of the main differences with other languages consists of its use in existential constructions: The verb *löytyä* ‘to be found’ is, as the data show, the most common existential verb after *olla* ‘to

be'. The four publications introduce inventive-locational constructions from a cross-linguistic point of view, describing their use in European languages (AREA), and analyze inventive-locational constructions more specifically in Finnish (QUANTI; WEB), but also situate inventive-locational constructions within the framework of the various verbs used with the Finnish Existential Partitive construction, offering a language-specific outlook (PARTEX). This dissertation aims at answering the following research questions:

1. What are inventive-locational constructions and how do they differ from other locational constructions? (Q1)
2. Why are inventive-locational constructions so common in Finnish and what are their main features? (Q2)
3. What are the main features of the marginal locuphoric¹ inventive strategies used in Finnish? (Q3)
4. What are the most typically used existential verbs in Finnish? (Q4)

Q1. Inventive-locational constructions are presented and discussed in AREA. This article defines inventive-locational constructions as construction-strategies, which are specific formal configurations that functionally express a locational meaning. The article also provides an areal outlook on inventive-locational constructions in several European languages. Inventive-locational constructions are presented as recognizable constructions that can be used as a comparative concept for cross-linguistic surveys, that is, to investigate languages outside the scope of AREA.

Q2. There may be diachronic reasons why inventive constructions have become so common in Finnish, but the answer to Q2 is found in the synchronic perspective adopted in QUANTI, WEB, and PARTEX. The valuable insight offered by these three papers lies in the quantitative methods that, applied to linguistic corpus data, help us to understand the prominence and diffusion of inventive-locational constructions in Finnish. The articles also contribute to the literature about existential constructions in Finnish, which have historically been a complex topic, still an object of discussion for many linguists.

Q3. In QUANTI, we found that inventive constructions are largely used in Finnish. The verb *löytyä*, when indexed for third persons (allophoric forms), is used as a locational copula. Unlike in other languages of Europe, however, it is not productive when indexed for first and second persons (locuphoric forms). In fact, this strategy is rather marginal, as shown by QUANTI. In WEB, I look at this marginal locuphoric strategy, and describe what appears to be specialized use in internet contexts.

¹ The term locuphoric refers to first and second person indexes, while allophoric refers to third person indexes (Haspelmath 2013).

Q4. To understand why inventive constructions are so common in Finnish, we should compare their use with other lexical verbs used in Finnish Locational constructions, especially Existential constructions (also called *presentative* or *presentational*). The answer to Q4 is achieved again through quantitative methods in PARTEX, which proposes a novel way of sampling linguistic data, which are then processed using the method known as *collostructional analysis*.

This dissertation consists of the present introductory chapter and the four original publications described above. The introductory chapter is further divided into the sections described below.

Section 1 introduces the theoretical framework used. The present dissertation is typologically oriented and defines a novel category of locational construction-strategies. For this reason, the main trends in the typological literature are laid out, and concepts like *locational construction*, *locative construction*, and *existential construction* will be crucial for understanding the contribution of this work. Section 1 also relies on the centrality of *comparative concepts*, useful tools that explain why it is relevant to talk about linguistic constructions at all. Comparative concepts are then contrasted with language-specific categories, like the Finnish Existential construction, which will also be introduced as part of the theoretical background.

Section 2 presents the data and methods used in this dissertation. The data consist of either language material elicited from native speakers and other experts, or corpus-based language material. The latter consists of instances of the construction in question extracted from a variety of corpora such as the Corpora of Uralic Volga-Kama Languages or the Suomi24 corpora for Finnish. The methods used are described in the second part of the section and consist of mixed methods, i.e., both qualitative and quantitative methods. The qualitative methods are commonly associated with typological research: eliciting language material from native speakers, glossing it and drawing crosslinguistic conclusions. The quantitative methods include statistical analyses like logistic regression, the measure of dispersion of language phenomena in corpora, and collostructional analysis. Particular attention will be given to the novel proportional method proposed for sampling language material from corpora, a method based on what I call Expected Sample size (ES).

The aim of Section 3 is twofold. First, it sums up the results reported in the articles, presenting them separately. Second, it discusses the contact points between the articles, and answers the four research questions proposed above.

The concluding Section 4 sketches out solutions to provide future research with tools to investigate inventive-locational constructions and the like. Such constructions are presented as a comparative concept from a cross-linguistic point of view. This section also generalizes the results obtained and the importance they offer for the study of grammar in general and of Finnish in particular, giving special attention to the methodological innovations employed within this work.

1 BACKGROUND

This section provides a theoretical background for the articles presented below. Section 1.1 introduces linguistic typology and a general functionalist approach, also drawing from theoretical frameworks such as Cognitive Grammar and Construction Grammar, both functionally oriented. Section 1.2 presents inventive-locational constructions as identifiable construction-strategies. Section 1.3 reviews earlier research on nonverbal predication and copula clauses from a functional point of view. It also goes into more detail by presenting the main functional differences between predlocative² constructions and existential constructions, both contained under the more generic type of constructions called locational constructions. Section 1.4 relates predlocative and existential constructions to subjecthood and partitivity in Finnish.

1.1 Linguistic typology and functionalist approaches to grammar

Early cross-linguistic comparisons date back to ancient times, but it is only during the past century that the study and classification of natural languages has been made systematic. As reported by Song (2010), after the introduction of the term *Typologie* by Georg von der Gabelentz in the 19th century, the field was revitalized by Greenberg (1963). Since then, linguistic typology has been growing in popularity and now has an established tradition that unites scholars from all over the world (Croft 2003; Dryer 1992). Linguistic typology has the main aim of comparing the languages of the world in order to draw cross-linguistic conclusions and to classify languages (Comrie 1988). These comparisons and classifications can be made from the point of view of morphosyntax, phonology, semantics, or, for example, word formation, but recently the scope of cross-linguistic inquiry has broadened so that it now also includes, e.g., pragmatic typology, meaning the typology of languages in use (see De Vries 2008; Rossi 2020). It can be argued that linguistic typology wants to look at the differences between languages, rather than their similarities, hence it is the study of cross-linguistic variation (Daniel 2010). Typological methodologies can also be incorporated into two subfields that instead have at their core the study of similarities between languages: contact linguistics and historical linguistics. These two disciplines aim at finding a satisfactory explanation as to why certain phenomena have developed in the same way within a certain geographical area or throughout history and can adopt both synchronic and diachronic approaches.

When talking about the geographical distribution of a linguistic phenomenon limited to a specific area, we can also talk about areal typology (as done by, e.g.,

² Here, Haspelmath's (2022) term is used. A more common variant of this same concept is *locative construction*.

Aikhenvald & Dixon 1998; Arkadiev 2014; Dahl & Koptjevskaja-Tamm 2001; Idiátov 2018). In areal-typological studies, the aim is to investigate a circumscribed geographical area that may contain genetically unrelated languages, to see whether these languages share one or more grammatical, prosodic, or pragmatic features, usually through synchronic comparisons. The generalizations over those features can usually be explained through language contact hypotheses and open a door to further diachronic investigations. Areal-typological investigations often also result in finding significant differences between languages spoken in a particular area, even between languages that are genetically related. The study at hand draws from areal-typological methods to study a specific construction within the languages of Europe and formulate a new comparative concept (*invenitive locational construction*, AREA), that can be used in further cross-linguistic, typologically-oriented investigations of larger samples of languages.

A comparative concept is a concept used for cross-linguistic comparison, created *ad hoc* by linguists, and fundamentally different from language-specific descriptive categories. A comparative concept does not have to be exclusively semantic, but it usually contains a semantic component (Haspelmath 2010). Because of the centrality of semantics, and hence of meaning, working with comparative concepts is part of the functionalist approach, according to which language is seen from the point of view of the communicative and meaning-oriented function it conveys, instead of under a merely formal or compositional light. Examples of functionally defined comparative concepts are ‘partitive’ or ‘imperfect tense’, which do not consider the whole range of properties that language-specific Partitive³ cases or Imperfect tenses may have. For example, the Finnish Partitive case not only has the function of indicating the proportional relation of a subset to a superset (Seržant 2021), which can be considered the main function of the comparative concept named ‘partitive’, but it can also indicate aspect, a temporal function, a partial object, or a discourse-new syntactic argument (QUANTI). Similarly, the Finnish Imperfect tense not only expresses an aspectually imperfective function (like the Italian Imperfect tense does), but also an aspectually perfective function (like the English Simple Past). The functionalist approach is distinct from, e.g., the approach adopted by generative linguists, who often assume the universality of what are language-specific categories, adopting innatist views that include the formulation of a universal grammar possessed by every speaker (Chomsky 1965; Dąbrowska 2015).

Despite the centrality of a new comparative concept, language-specific terms are not completely irrelevant to the dissertation at hand, since QUANTI and PARTEX investigate the Finnish Existential construction based on the Fennistic tradition (Hakanen 1972; VISK; Miestamo 2005; Larjavaara 2019). For this reason, this work also draws from functionally oriented frameworks that complement its general typological aims. One of these approaches is Cognitive Grammar, which assumes, in contrast with generative theory, that “language is neither self-contained

³ Following Haspelmath (2010), among others, capitalized initials are used for language-specific categories, while non-capitalized initials for comparative concepts.

nor describable without essential reference to cognitive processing” (Langacker 2008: 29). In cognitive approaches, meaning is central and directly connected to the human sensorial and spatial experience of the world (Langacker 1987; Talmy 2000). Cognitive approaches intertwine with the performative nature of the speech act as a communicative tool, which argues in favor of a usage-based language theory (Diessel 2017). According to usage-based linguistics, the cognitive and interactive dimensions both participate in the production of language, going against the structuralist and generativist views that consider the study of a language system and the study of language use to be two separate endeavors. In general, this dissertation adheres to the view that meaning should be studied from a holistic perspective, a view that is also in line with another usage-based, functionalist approach to language, Construction Grammar (Goldberg 1995; 2006). This approach combines pragmatics and semantics and justifies linguistic structures, also called constructions, as having a meaning that is not the mere sum of its components. In other words, all the phonological, morphological, syntactic, lexical, and pragmatic parameters of a construction come equally into play in defining it, and changing one of these parameters can potentially lead to a completely different interpretation of the construction. Within this framework, constructions can be not only larger linguistic structures (also called *clause constructions*, see, e.g., Haspelmath 2022), but also smaller linguistic units like phrases, morphemes, or phonemes. Construction Grammar is used as part of the theoretical framework of QUANTI and PARTEX.

1.2 Invenitive-locational constructions as identifiable construction-strategies

The main goal of this dissertation is to define a class of linguistic clause constructions called *invenitive-locational constructions* (AREA). These structures are formally defined but fall within the functionally defined comparative concept of locational constructions. We have to keep in mind the difference between the two comparative concepts of construction-functions and construction-strategies (Croft 2022; Haspelmath 2021). Construction-functions are defined by their functions, construction-strategies by their formal properties. Construction-functions may be expressed by several strategies, e.g., a partitive construction can be expressed by a genitive strategy (Italian *delle ragazze* [GEN.DEF girl.PL] ‘[some] girls’), while in construction-strategies additional meanings and coexpression patterns come into play. Invenitive-locational constructions are indeed construction-strategies rather than construction-functions because they can be investigated from the point of view of the coexpression patterns that pertain to verbs with the meaning FIND. Moreover, there are cases in which a verb with the meaning FIND is the only strategy used in existential constructions (1).

- (1) Pite Saami (Uralic; Wilbur 2014: 234)
váre-n *gávdnu* *aj* *juomo*
 mountain-INE.SG exist.3PL.PRS also sorrel.NOM.PL
 ‘There is sorrel in the mountains, too.’

Pite Saami, alongside the copula *árrot*⁴ ‘to be’, also has an existential verb which derives from the verb *gávdnat* ‘to find’, much like its Swedish counterpart, where the verb *finna* ‘to find’ derives the existential *det finns* ‘there is’ (Joshua Wilbur, p.c.). Such cases of middle-marked FIND-verbs are also mentioned by Gaeta (2023: 113). Gaeta justifies such dynamic verbs with the meaning FIND, which, by inference, acquire a deictic value, as being used in existential constructions through “the conventionalization of a middle-passive or of a reflexive-anticausative construction containing change-of-place verbs (especially achievements)”.

Gaeta does not single out FIND-based constructions, but his differentiation of a semasiological versus an onomasiological approach for constructions in general (cf. Geeraerts 2010) is also valuable to understand why inventive-locational constructions can be identified as a comparative concept. In a nutshell, if we ask “Which morphosyntactic strategies express a locational function?” we can answer, e.g., “BE-copulas, posture verbs, and inventive verbs as well as other kinds of copulas and the so-called zero-copulas” – we now have a clearer idea of what a locational construction-function is. If, on the other hand, we ask “Which additional meanings do verbs with the meaning FIND have?” we can answer “Telic meaning, mirative meaning, and locational meaning when such verbs are morphologically or analytically marked, and their primary meaning is bleached” – this last point provides a clearer idea of what an inventive-locational construction-strategy is. In turn, the inventive construction is itself a comparative concept because it has a functional component and can be applied to cross-linguistic investigations. To understand why the inventive construction can function as a comparative concept, we first need to understand what kind of clause constructions are meant, starting with the broad group of constructions that fall under nonverbal predication.

1.3 Earlier research on nonverbal clause constructions, copulas, and locational constructions

Nonverbal clause constructions are often also called nonverbal predication or copula clauses (Declerck 1988; Hengeveld 1992; Mikkelsen 2011; Roy 2013) and are constructions that lack a typical verb (Haspelmath 2022). A typical verb can be defined as a semantically full verb that describes a happening or an action. Conversely, nonverbal predication usually features a copula – an atypical verb that only indicates a stative link between two arguments. However, as Haspelmath (2022) discusses, nonverbal clause constructions can also lack an overt

⁴ When inflected, this verb has a suppletive *l-* stem (Wilbur 2014).

copula, and the class of ‘atypical verbs’ is not necessarily limited to copulas. In my opinion, this class can be extended also to include certain verbs that have undergone semantic bleaching and do not express their full lexical meaning. These verbs include what I call *inventive verbs* (AREA). Haspelmath (2022) divides nonverbal clause constructions into seven types, of which only two, predlocative and existential constructions, are relevant to the thesis at hand. Paraphrasing Haspelmath, predlocative constructions are defined as clause constructions that have a subject, also called the *locatum*, which is said to be located in a place expressed by a locative phrase. On the other hand, in existential constructions the locatum is not a subject, it is discourse-new and is also called the *existent*. Existential constructions also have the function of locating this argument in a place expressed by a locative phrase. These two types of clause constructions can be considered part of a larger group of constructions called locational constructions. Haspelmath (2022) also singles out another type of construction called a *hyparctic construction*, which expresses “pure” existence, without expressing a locational function. This can therefore be considered a special case of what other authors call *existential constructions*.

An example of a predlocative construction (2) and existential construction (3) follow. Some existential constructions can also feature a specific kind of form called an *existive* (Haspelmath 2022), which does not appear in other forms of locational predication (4).

- (2) Estonian (personal knowledge)

Karu on metsas
 bear be.3SG forest.INE
 ‘The bear is in the forest.’

- (3) Estonian

Metsas on karusid
 forest.INE be.3SG bear.PART.PL
 ‘There are bears in the forest.’

- (4) Tagalog (Austronesian; consultation with a native speaker)

May kare-kare sa mesa
 EXV⁵ kare kare⁶ on table
 ‘There is kare kare on the table.’

Locative and existential constructions have been theorized by several authors (e.g., McNally 2016; Creissels 2014), and sometimes the thin line between comparative concepts and the formal configurations of these constructions has been left blurry. One example of this is provided by Creissels (2019), who refers to the first type as *plain-locational predication (PLP)* and the second type as *inverse-*

⁵ *Existive* (Haspelmath 2022) – this term is not yet established within the linguistic community.

⁶ A traditional Filipino stew.

locational predication (ILP). These two terms are proposed as comparative concepts, but they are fundamentally different from the terms *predlocative* and *existential clause construction* used by Haspelmath (2022). In fact, Creissels' terms suggest a formal distinction in the syntax of the two types of constructions, in that the difference in word order alone seems to speak in favor of a semantic difference between the two constructions. Creissels' terms hence refer to very transparent syntactic configurations, but do not fully correspond to Haspelmath's (2022), who argues that the main functional difference lies in whether the locatum is definite or not. Creissels' (2019) term *inverse-locational predication* instead allows for constructions that do not exhibit the so-called *definiteness effect* (Abbott 1992; 1993). Similarly to Koch (2012), who speaks of the two types of constructions from an information structural point of view, referring to PLP as *thematic location* and to ILP as *rhematic location*, Creissels (2019) adopts the semantic concept of *perspectivization*, which relates to *figure-ground relationships* (Talmy 1983; 2000: 311). The ground is an entity that is usually fixed in space, while the figure is more mobile. Some authors call the figure a *pivot* (Milsark 1977; Bentley et al. 2013) and the ground a *coda* (Bentley et al. 2013). Because of the figure's mobility, Creissels (2019) assumes that it has a more salient nature than the ground, making the 'figure>ground' perspectivization (PLP) the prototypical, unmarked one. This perspectivization is contrasted with the marked 'ground>figure' perspectivization of inverse-locational predication. Assuming that one of the constructions is prototypical and the other not, however, is problematic for one main reason. This assumption is in fact correlated with the assumption that, cross-linguistically, predlocative constructions are more common than existential constructions, that languages always prefer the first type of construction, and that the second type of construction is always marked. However, Haspelmath (2022) shows that in languages like Wambaya, spoken in Australia, there is no difference in marking between predlocative and existential constructions. In (5), both locative phrases are clause-final, and there is no morphological marking on the locati.

(5) Wambaya (Mirndi; Haspelmath 2022; Nordlinger 1998: 177)

- a. *Janji inyaga jalyu-ni!*
 dog.NOM that.NOM bed-LOC
 'The dog is on the bed!'
- b. *Garnguji julaji-rdarra gayangga darranggu-ni.*
 many.NOM bird-GROUP.NOM high tree-LOC
 'There are lots of birds up in the trees.'

Language-specificity comes into play in other ways, too. Some languages tend not to vary their word order, regardless of the definiteness of the locatum. This means that such languages may not allow ILP at all simply because of grammaticality constraints on how the information structural patterns are conveyed, but it does not mean that they do not allow existential constructions, as both Haspelmath and Creissels show for Mandinka (6).

- (6) Mandinka (Mande; Haspelmath 2022; Creissels 2019: 51)
Wúlôo bé yíròo kótò
 dog.DET LOC.COP tree.DET under
 ‘The dog is under the tree. / There is a dog under the tree.’

What Haspelmath (2022) considers existential constructions are therefore also constructions that do not have a marked ILP-word-order such as (7a), which has an indefinite locatum. Consequently, despite having a marked ILP-word-order, (7b) can be considered a predlocative construction because its locatum is definite.

- (7) Italian (personal knowledge)
 a. *Uccelli volano nel cielo*
 bird.PL fly.3PL in.DEF sky
 ‘Birds fly in the sky. / There are birds flying in the sky.’
 b. *Nel cielo volano gli uccelli*
 in.DEF sky fly.3PL DEF.PL bird.PL
 ‘In the sky the birds are flying. / The birds are flying in the sky.’

Definiteness will have a central role in the next Section 1.4, in which Finnish locational constructions are introduced. In Finnish, indefinite locati are in fact often marked for partitive, and this type of existential NP can also be in clause-initial position.

1.4 Subjecthood and partitives in Finnish Existential constructions

Just like the other Finnic languages (Estonian, Ingrian, Karelian, Livonian, Ludic, South Estonian, Veps, Votic), the Saami languages, and Hungarian, Finnish belongs to the Finnic genus of the Uralic language family, spoken mainly in Russia. Its geographical position allowed for an ongoing language contact situation with its European neighbors including Swedish, Russian, German, and the Baltic languages, which meant plenty of innovations for this Uralic language. However, because of its fundamentally different morphosyntactic typology, Finnish also presents some features that are typical of Finnic languages. One of these features is a grammaticalized Partitive case, which has several functions. It can in fact be used not only to indicate part-whole relations (Seržant 2021), but also expresses, e.g., temporal adverbials, a partial object vs. a total object indicated by an accusative or genitive-marked argument, and a discourse-new referent in existential and presentational clauses (Gast & Haas 2011; cf. Gaeta 2023). Presentational clauses are a formally defined clause type that has the function of introducing a new referent; however, in the Finnish tradition these clauses usually fall under the category of the Finnish Existential construction. An example of the Finnish Existential construction is given in (8a) and is contrasted with both its predlocative counterpart (8b) and a Finnish Existential construction (8c), which features a verb different from *olla* ‘to be’.

- (8) Finnish (personal knowledge)
- a. *Järvessä on kaloja*
 lake.INE be.3SG fish.PL.PART
 ‘There are fish in the lake.’
- b. *Kalat ovat järvessä*
 fish.PL be.3PL lake.INE
 ‘The fish are in the lake.’
- c. *Järvessä ui kaloja*
 lake.INE swim.3SG fish.PL.PART
 ‘There are fish swimming in the lake.’

The Finnish Existential construction has often been defined from a formal rather than a functional point of view, although what is common across these definitions is that it has the function of introducing a new referent (Ikola 1954; Hakanen 1972; VISK § 893). The formal criteria to which the Finnish Existential construction adheres constitute the following prototype (translated from VISK § 893):

- the verb ‘to be’;
- a locative phrase in thematic position;
- a subject in the partitive;
- a partitive-marked subject argument in case of negative polarity;
- absence of verb agreement;
- a discourse-new subject.

Existential constructions in Finnish do not always fully adhere to this prototype and can present one or more variations of the six criteria above. This becomes a problem when talking about the Finnish Existential sentence from a functional point of view. By changing the first criterion, the Finnish Existential sentence can have any other intransitive verb. Because the list of intransitives that can feature in this construction is very long (see Larjavaara 2019), it becomes a definitional matter whether Finnish Existential constructions are indeed always Existential constructions or something else, like presentationals. Even hands-on applied research like Ivaska (2011: 81) has shown that, in contexts such as that of second-language acquisition, the difference between Existential constructions and other clause types can be quite unclear. In those cases, criteria like subjecthood, word order, case marking, or verb agreement are not solid enough and do not suffice to effectively classify some borderline instances of Existential constructions. Tweaking one or the other criterion can however lead to more or less existential readings (Ivaska 2010; 2011).

Another problem is the very definition of subject given in the list of criteria established for the prototype. Can a “subject in partitive” be called a subject? Or should it be called something else? The discourse-new referents introduced by partitive-marked arguments have often simplistically been called subjects in the literature, even though they do not exactly behave like proper subjects in a syntactic sense. This is shown for example by Huumo (2003), who argues that in Existential constructions “[t]he participation of individual entities in the activity

is backgrounded”, and that these constructions can be interpreted from a holistic point of view. From this, it follows that there is no proper subject in existential predication, but rather an *e-theme* or *existential theme* (Huumo 2003: 462). In recent accounts, Huumo (2023) has adopted the term *existential S argument*, getting rid of the notion of subjecthood altogether. According to Huumo and Helasvuo (2015: 37), the Finnish subject usually has an unmarked, nominative morphology, triggers agreement in the predicate, and occupies a preverbal, thematic position in the sentence. This is why they suggest a new language-specific term for designating the kind of syntactic unit Existential constructions feature, and they call it an *e-NP* or *existential Noun Phrase* (Huumo & Helasvuo 2015; Helasvuo 1996). However, the term *partitive subject* is still sometimes used as a convenience term (Huumo & Lindström 2014). What all these studies agree upon is that these syntactic units have the function of introducing new indefinite referents in the discourse, and hence cannot qualify as prototypical subjects. In line with this feature, the inventive verb *löytyä* ‘to be found’ can also be used with non-prototypical subjects, i.e., Partitive-marked e-NPs.

2 DATA AND METHODS

This section describes the data and methods used throughout this dissertation. It is further divided into three parts: data, qualitative methods, and quantitative methods. First, I will introduce the data used in the articles. These data consist of language samples for the research conducted in AREA and of Finnish corpus data for the studies conducted in QUANTI, WEB, and PARTEX. Second, I will describe the purely qualitative methods I used. These are typically used in typological research and applied to areal typological studies (AREA). Third, I will introduce the quantitative methods used in QUANTI and PARTEX. These methods are well established in corpus linguistics and use statistics to gain insight into how languages work.

2.1 Data

AREA provides a sketch of the morphosyntactic and semantic characteristics of inventive-locational constructions in the small sample of 13 European languages shown in Table 1. The languages have been chosen based on the availability of informants. Utterances were collected through elicitation from native speakers, and in some cases collected from the internet and subsequently confirmed as valid by native speakers or language experts.

Table 1 – Language sample used in Article I

Language	Language family, genus
Albanian	Indo-European, isolate
Basque	Isolate
English	Indo-European, Germanic
Estonian	Uralic, Finnic
Finnish	Uralic, Finnic
German	Indo-European, Germanic
Greek	Indo-European, Hellenic
Hungarian	Uralic, Ugric
Italian	Indo-European, Romance
Latvian	Indo-European, Baltic
Maltese	Afroasiatic, Semitic
Russian	Indo-European, Slavic
Sardinian	Indo-European, Romance

AREA shows that inventive-locational constructions are worth considering but does not make crosslinguistic claims about such constructions. That is, even though it looks at inventive-locational constructions from a typological perspective, it is not a typological study per se, since it does not employ the classical sampling criteria used in crosslinguistic surveys (Miestamo et al. 2016). The sample includes at least one representative language from each major Indo-European language group of Europe. It also has three Uralic languages (Estonian, Finnish, and Hungarian), one Semitic language (Maltese), and the language isolate Basque. The sample has also been reduced in size, because inventive-locational constructions in closely related languages (e.g., Polish and Russian) turned out to behave in similar ways, and expanding the sample size would not have added much novel information to the analysis. The examples contained in AREA have been collected from native-speaking informants or from the web. Whenever examples were collected online, their acceptability and grammaticality was checked by subsequently asking at least one language expert. Many examples are also based on my own native intuition in Italian, which takes up a large part of the analysis. All the examples used in the analysis are available in the text of AREA, as well as in an appendix at the end of the paper.

As for QUANTI, WEB, and PARTEX, I have used data from Finnish corpora. QUANTI focuses on the verb *löytyä* ‘to be found’ in Finnish journalistic texts contained in the corpus *Kansalliskirjaston lehtikokoelman (KLK) suomenkieliset lehdet*. This corpus comprises newspapers and magazines published in Finnish and collected by the National Library of Finland. The corpus search was made using the infrastructure Korp (Borin et al. 2012), and it was limited to texts from the beginning of 1990 to the end of 2000. This subcorpus, totaling around 149 million words, was chosen because the nature of this research is fundamentally synchronic. After collecting 500 random affirmative sentences with an SV word order and 500 with a VS word order, the sample was polished. This was a necessary step because not all the sentences were relevant. For example, since QUANTI investigates the nominative-partitive alternation in e-NPs contained in Finnish Locational constructions, all the negative occurrences were eliminated. Negative polarity, in fact, almost always automatically triggers the appearance of the partitive case, and this represents a bias towards how that partitive marking is analyzed. The final corpus of sentences considered for analysis consists of 779 sentences, of which 387 have SV word order and 392 VS word order.

For WEB, I used data from the corpus Finnish Web 2014 (fiTenTen2014), which I collected on the platform Sketch Engine (Kilgarriff et al. 2014). The corpus is quite large as it contains ca. 1.7 billion tokens coming from various internet text types. WEB focuses on one marginal strategy encountered in QUANTI, and hence can be considered its continuation. The sample consists of a total of 449 sentences containing the verb *löytyä* ‘to be found’ indexed for locuphoric forms, i.e., first and second persons.

In PARTEX, I collected material from Suomi24:2017, an automatically annotated corpus of around 168 million words, through the Korp infrastructure. This corpus contains messages published in 2017 in the online discussion forum

Suomi24. At the time of conducting this research, 2017 was the most recent year for which the corpus was collected, and I chose it because it represented the closest version of web-forum Finnish available. This is motivated by the synchronic nature of the article. Web-forum language is also different from journalistic texts: Because it is produced by users and is unedited, we can expect it to present a lot more variation, and to provide interesting insights into the Finnish Existential construction. For this study, I chose 10 intransitive Finnish verbs that can appear in Existential constructions, and I downloaded 10 random samples of 1000 sentences each, totaling 10,000 sentences. I then applied two different sampling methods to this sample, one based on the real frequencies of Partitive Existential constructions, and the other based on Expected Sample sizes. I will discuss these methods in Section 2.3.

2.2 A mixed methods approach

This dissertation contributes to different fields of linguistic research, and hence adopts a mixed methods approach, i.e., an approach that considers multiple perspectives for the same research questions (Johnson et al. 2007: 13; Ivaska 2020: 17). In the specific case of this dissertation, I have combined qualitative and quantitative methods. These two main types of methods complement each other: Quantitative methods are often used to provide insights on a qualitative level, too. The mixed nature of the methods employed in this dissertation does not only concern the differentiation between qualitative and quantitative methods, but also the variety of quantitative methods used. These are the following:

1. Descriptive statistics to emphasize the marginal nature of the constructions analyzed in WEB;
2. Mixed-effect logistic regression to investigate which grammatical features influence the nominative-partitive alternation in QUANTI;
3. Collostructional analysis to analyze the strength of interaction between verbs and constructions in PARTEX;
4. The novel sampling method referred to as Expected Sample size to improve the results obtained through collostructional analysis in PARTEX;
5. State-of-the-art data visualization in WEB and QUANTI.

Despite all these possible quantitative research directions, what I consider the main contribution of this dissertation has a comparative nature. AREA introduces a novel class of locational construction-strategies called inventive constructions, and employs the qualitative methods typically used in typological research. This means that a large part of this work is descriptive and based on qualitative observations from a small set of data, which I collected from informants and language experts. I then analyzed the examples collected by considering their morpho-syntactic characteristics, to sketch a typology of inventive constructions in the languages of my sample.

Similarly, the sample used in WEB is analyzed from a qualitative point of view. After carrying out the corpus search, I analyzed the sentences collected by looking at their pragmatic and morphosyntactic features. Since the sample was quite small in comparison to the large corpus from where it was taken, the quantitative part of this article consists of a descriptive statistical outlook of the occurrence of the several verb forms analyzed.

2.3 Quantitative methods in detail

In this section, I will show in more detail the quantitative methods used in this dissertation, apart from the descriptive statistics used in WEB. Quantitative methods were used in QUANTI and PARTEX, both relating to the Finnish language and therefore dealing with language-specific constructions. QUANTI employs quantitative statistical methods to infer qualitative analyses and follows an already established trend in linguistic research (Levshina 2015; Winter 2019; Stefanowitsch 2020; Gries 2021; Ivaska 2022). In particular, the method used in QUANTI is a mixed-effect logistic model (Gries 2015). PARTEX is almost purely methodological and provides a novel sampling method based on an Expected Sample size, which is then applied to *collostructional analysis* (Stefanowitsch & Gries 2003; Gries & Stefanowitsch 2004).

The methods mentioned above will be described in the following subsections. All the statistical analyses present in QUANTI and PARTEX, as well as the visualization of the descriptive statistics present in WEB, were carried out in the R environment (R Core Team 2018).

2.3.1 Mixed-effect logistic regression models

Statistical models, especially various regression models, can be applied to large corpora of language data to gain insights into how language works, changes, develops, or varies. This line of work is established in researching grammar, especially with regard to grammatical constructions and it is in constant expansion (see, e.g., Klavan et al. 2015; Varjo & Suomalainen 2018; Klavan 2020; Lindström et al. 2021; Pook 2021; Varjo 2022; Ivaska 2022).

An overview of simple logistic models in corpus linguistics is Speelman (2014). In regression analysis, a numeric response variable is predicted based on one or more predictor variables. In simple logistic regression analysis, the response variable can also be a categorical variable, which makes this method suitable for linguistic analysis. Multiple logistic regression models predict the response variable based on the simultaneous effect of two or more fixed-effect predictor variables. In this case, the predictors are often categorical. Finally, mixed-effect logistic regression models also introduce random effects in the picture, alongside fixed-effect predictors. In QUANTI, we used mixed-effect logistic regression to model the subject case alternation in Finnish sentences containing the verb *löytyä* ‘to be found’, as a function of both fixed-effect and random-effect predictors.

2.3.2 Collostructional analysis

Collostructional analysis is a quantitative method introduced twenty years ago and revised a couple of times during the last two decades (Stefanowitsch & Gries 2003; see also Gries & Stefanowitsch 2004; Gries 2019). This method is based on the core ideas of Construction Grammar (e.g., Lakoff 1987; Goldberg 1995), according to which a linguistic construction is “a pairing of form with meaning/use such that some aspect of the form or some aspect of the meaning/use is not strictly predictable from the component parts or from other constructions already established to exist in the language” (Goldberg 1996: 68). While most authors usually consider constructions to be only clause constructions in a strict sense (i.e., larger syntactic units), Construction Grammar postulates that any linguistic pattern can be a construction, even words or morphemes, but also that constructions are unique in that their meaning cannot be predicted based on the smaller linguistic units that form them.

In corpus linguistics, collocational analysis is the investigation of either a certain node word in its context (e.g., Oh 2000) or of the frequent words that can appear in the vicinity of a node word, i.e., its collocates (e.g., Kennedy 1991). Collostructional analysis is a type of collocational analysis that measures the degree to which different lexemes (i.e., words) are attracted to or repelled by a certain construction. Its strong point lays in the capacity to distinguish construction-specific typicality from construction genericity by relating the number of co-occurrences of the word and the construction to the general number of occurrences of both of them. Collostructional analysis considers specific constructions, differentiating between the several possible functional configurations that can be assigned to a certain formal configuration. More concretely, collostructional analysis can distinguish between, e.g., the English past participle construction and the past-tense construction, both represented by the formal configuration [V-*ed*], while collocational analysis cannot (Stefanowitsch & Gries 2003: 214). This is precisely the reason why I chose collostructional analysis as a method of investigation for PARTEX, which focuses on the Finnish Existential Partitive construction. In Finnish, the Partitive case is polyfunctional and the fact that a partitive-marked NP can appear as a collocation of a certain existential verb does not guarantee that said NP is the e-NP (Huumo & Helasvuo 2015) of the construction studied. As my study showed, many of the NPs resulting from the corpus search in PARTEX had indeed other functions, such as temporal functions, and required a significant amount of manual polishing of the corpus data.

2.3.3 Expected Sample size

The ideal sample of observations extracted from a corpus and analyzed by means of collostructional analysis is a sample of observations that all represent a certain construction. In the case of PARTEX, the studied construction is the Finnish Existential Partitive construction. Ideally, if we could instruct the corpus to search only for examples in which there is a partitive-marked NP that functions as an

e-NP, i.e., the subject-like element of the Finnish Existential construction, we could have a clear picture of how many of these sentences there are on a corpus level, and we could see which verbs (i.e., *collexemes*) are more strongly associated with the construction at hand. This is, however, impossible to achieve, since the Finnish partitive case can express a variety of functions that do not necessarily appear in the Finnish Existential construction. Because of this problem, the collected samples must be small enough to be manually polished, to rule out all the example sentences that do not represent the construction at hand.

In PARTEX, I chose 10 Finnish verb that can appear in the Finnish Existential Partitive construction. For each verb, I have extracted from the corpus a random sample of 1000 sentences that satisfied the typical criterion of the Finnish Existential Partitive construction: absence of verb agreement (verb in 3SG) and affirmative polarity. This search yielded both sentences with nominative-marked NPs – as either subjects of the verb or e-NPs – and sentences with partitive-marked e-NPs. After isolating the sentences that contained partitive-marked e-NPs and eliminating the ones where partitive-marked NPs belonged to different constructions, I have applied collostructional analysis to this reduced sample size. However, on a whole-corpus level, the frequencies of the 10 verbs differed greatly from one another, and it is self-evident that the reduced sample provided a biased view of the Finnish Existential Partitive construction, especially since the aim of the research at hand is to determine which existential verbs are more strongly associated with said construction. The corpus-level information about the occurrence frequencies of the single verbs used in the studied construction inevitably slipped away. To overcome this obstacle, I have devised a sampling method based on simulated occurrence frequencies of each verb, considering their corpus-level presence. I based this sampling method on real observed frequencies of the verbs, taken from the reduced sample (1000 sentences). I called the resulting number the verb's Occurrence Frequency (OF). Each verb also has an Absolute Frequency (AF), which is the total number of occurrences of the verb in the whole corpus, and a Relational Frequency (RF), which is the number of occurrences of the verb within a specific corpus search, which has been tailored for yielding occurrences of the verb in a specific construction (the Finnish Existential Partitive construction). For example, the verb *löytyä* 'to be found', has an Absolute Frequency of 116,068 observations, a Relational Frequency of 72,276 observations (out of 116,068, where the formal search criteria of the construction at hand were fulfilled), and an Occurrence Frequency of 490 observations (out of the sample of 1000 observations, which all fulfilled the formal search criteria of the construction). From these frequencies, we can calculate the Expected Sample size (ES), which results from the following proportion:

$$AF : RF = ES : OF$$

And hence:

$$ES = AF \times OF / RF$$

In the case of *löytyä*, $ES = 116,068 \times 490 / 72,276 = 787$ (rounded). This resulting number tells us that 787 is the ideal sample size of sentences we should analyze to get the Occurrence Frequency observed for *löytyä*. This sample size is entirely hypothetical, but its strong point is that it considers the corpus-level occurrence of the verb analyzed, eliminating at least some of the bias provided by the first, reduced sampling. After running these numbers for all the verbs of the sample, I applied collocation analysis again by including their ES instead of their biased sample size.

3 RESULTS

In this section, I will introduce the features of the Finnish inventive-locational constructions, as well as the main results of the four articles that make up this study. This section ends with a discussion, which brings together the articles and highlights their interconnections.

AREA introduces a novel class of locational construction-strategies that I call inventive-locational constructions. The main goal of the study is to survey several languages spoken in Europe and prove that inventive-locational constructions are worth considering as a comparative concept through the description of the two main types of these constructions, thus, responding to Q1.

QUANTI was published before AREA was even beginning to take shape, so the terminology contained in the latter is not used. This second paper analyzes FIND-based intransitive strategies in Finnish using quantitative methods. It is the first paper to propose that these strategies be treated as copulas and argues in favor of the interchangeability of usage of the Finnish verb *löytyä* ‘to be found’ and the locational copula *olla* ‘to be’ in certain contexts. One of the main criteria adopted in the definition of inventive-locational constructions (AREA) is semantic bleaching. QUANTI dives deep into Finnish, aiming at responding to Q2.

WEB expands over one of the findings of QUANTI and responds to Q3. This study confirms that locuphoric occurrences of *löytyä* ‘to be found’ are indeed marginal, as claimed in QUANTI, and that they are mostly confined to the internet environment.

PARTEX is a corpus study that investigates the Finnish Existential Partitive construction from a quantitative point of view (Q4). The main goal is to compare the usage of the established copula-like verb *löytyä* ‘to be found’ with other traditionally considered existential verbs in Finnish. Because of problems arising from the application of collocation analysis to restricted corpus samples, it proposes an alternative sampling method. It confirms what is claimed in QUANTI, in that the verb *löytyä* ‘to be found’ is the verb most strongly associated with the Finnish Existential construction.

3.1 Finnish inventive-locational constructions

In Finnish, the verb *löytyä* ‘to be found’ is commonly used as a locational copula, hence as an inventive-locational construction. It is an intransitive form obtained from the verb *löytää* ‘to find’ through the Uralic middle-marking morpheme *-U*⁷, a productive morpheme that can have different meanings depending on the verb. The inventive verb *löytyä*, similarly to the copula *olla* ‘to be’ and to other verbs traditionally considered existential verbs (such as *istua* ‘to sit’, *juosta* ‘to run’), can trigger nominative-partitive alternation in an argument that functions as a subject or an e-NP. It is mostly indexed for allophoric (third person) forms and marginally for locuphoric (first and second person, i.e., speech-act participant) forms.

⁷ The morpheme is capitalized because of vowel harmony.

3.2 Invenitive-locational constructions (AREA)

In the typological tradition, much attention has been brought to the distinction between language-specific categories and comparative concepts, especially through the work of linguists like Martin Haspelmath (2010; 2011; 2016; 2018), who strongly argues for the necessity of said distinction (criticized by, e.g., Lander and Arkadiev 2016). Invenitive-locational constructions are defined through a set of features that are applicable to any linguistic sample and aimed at verifying that invenitive-locational constructions exist in the languages within said sample. This makes invenitive-locational constructions a comparative concept: They have a functional component, but on a fundamental level they are formally defined.

Defining the *invenitive* verb characterizing invenitive-locational constructions is the first step towards recognizing them. This verb satisfies two criteria (borrowed from AREA):

- I. It has a root with the meaning FIND which undergoes semantic bleaching;
- II. It features a morphological or periphrastic valency- or voice-changing marker.

Additionally, when an invenitive verb is found in a locational construction, it satisfies a third criterion (also borrowed from AREA):

- III. It expresses a locational function without semantically marking a specific posture.

An example of an invenitive-locational construction is given below (9).

- (9) Italian (AREA)
- | | | | | | |
|-----------|--------------|-----------|--------------|--------------|---------------|
| <i>Il</i> | <i>gatto</i> | <i>si</i> | <i>trova</i> | <i>sull'</i> | <i>albero</i> |
| DEF | cat | MM.3SG | find.3SG | on.DEF | tree |
- ‘The cat is in the tree.’

In (9), the cat in question is a cat nobody is necessarily looking for. However, the strategy employed, very common for speakers of Italian, uses a middle-marked form of a verb with the meaning FIND. This is functional for expressing the location of the cat in the tree. It should be noted that no information is conveyed about the posture of the cat (unlike posture verbs such as ‘to sit’, ‘to lie’, and so on). The reasons why this specific strategy is used here are beyond the scope of AREA but shed light on the possibility of the beginning of a grammaticalization process. Desemanticization, also referred to as “bleaching” or semantic reduction, is in fact a key feature of grammaticalization, and is usually accompanied by extension, or use in new contexts (Heine 2017). Invenitive-locational constructions make use of a FIND-based strategy in the new context of its use as a copula. Moreover, the development of FIND-based strategies into invenitive-locational constructions adheres to the so-called overlap model (Heine 1993: 48– 53). This model describes grammaticalization as a chain-like three-stage process, in which an expression is recruited for grammaticalization and then it acquires a secondary use pattern, which then replaces the original use pattern. As specified by Heine (2017: 579), in some instances of grammaticalization this secondary use pattern

does not replace the original one, resulting in ambiguity. This is why it is sometimes hard to distinguish between instances of inventive-locational constructions and instances of FIND-based strategies in which the verb is used in its original meaning, especially, it seems, when the referent being located is a movable entity (10).

- (10) Finnish
Olut löytyy jääkaapista
 beer find.MM.3SG fridge.ELA
 ‘The beer is (found) in the fridge.’

(10) is interpretable as both an instance of locational predication and a sentence in which the verb retains its original meaning FIND. In the latter case, the pragmatic context plays a big role. The sentence above might just be the answer to the question *Where is the beer (to be found)?*, which in Finnish would employ the same middle-marked form of the verb ‘to find’. For the purpose of such a study on inventive-locational constructions, these thoughts about pragmatic contexts are important, but further corpus-driven research will be needed in the future to try and assess where to draw the line between (inventive-)locational constructions and other constructions. It is interesting to consider that locational constructions are pragmatically speaking marked, because predicating about the position of something (or someone) in space essentially has the function of communicating to an interlocutor where they can find that something (or that someone). This inverted point of view regarding locational constructions could be the key to understanding why inventive-locational constructions exist in the first place, but would require substantial further research, both from a synchronic and diachronic point of view.

It is clear from the criteria above that such constructions are defined in a way that distinguishes them from other instances of locational constructions, e.g., locational constructions that employ a BE-copula (10a), a posture verb such as Finnish *maata* ‘to lie’ (10b), a specialized existive (10c), or the absence of any linking element whatsoever (sometimes also called a *zero-copula* [10d]).

- (10) a. Finnish
Olut on jääkaapissa
 beer be.3SG fridge.INE
 ‘The beer is in the fridge.’
- b. Finnish
Käärme makaa auringossa
 snake lie.3SG sun.INE
 ‘The snake is lying in the sun.’
- c. Spanish (personal knowledge)
Hay gente en la calle
 EXV people in the street
 ‘There are people in the street.’

In AREA, I investigated a small sample of European languages with the aim of analyzing their inventive-locational constructions. Table 2 groups inventive-locational constructions into two types, while the results of the analysis are shown in Table 3.

Table 2 – Inventive-locational construction types (AREA)

	Type 1: animate locatum	Type 2: inanimate locatum
Criterion 1 (semantic): concreteness, locative phrase	concrete/abstract	concrete
Criterion 2 (morpho-syntactic): indexing, verb	both locuphoric and allophoric forms (language-specific differences)	allophoric forms
Language-specific subtypes: mirativity	mirative reading possible (construction specialization)	N/A

Table 3 – Features of inventive-locational constructions in the sample (AREA)

Language	Animate locatum ⁸	Inanimate locatum	Concrete location	Abstract location	Locuphoric forms	Allophoric forms	Dedicated mirative strategy
Albanian	(+)	(+)	(+)	(+)	(+)	(+)	(-)
Basque	(-) only mirative	(+)	(+)	(+)	(-) only mirative	(+)	(+)
English	(-) only mirative	(+)	(+)	(+)	(-) only mirative	(+)	(+)
Estonian	(+) except sg	(+)	(+)	(-)	(-)	(+)	(+)
Finnish	(+) except sg	(+)	(+)	(-)	(+) marginal	(+)	(+)
Greek	(+)	(+)	(+)	(+)	(+)	(+)	(-)
German	(+)	(+)	(+)	(+)	(+)	(+)	(-)
Hungarian	(-)	(+)	(+)	(-)	(-)	(+)	(-)
Italian	(+)	(+)	(+)	(+)	(+)	(+)	(-) ⁹
Latvian	(+)	(+)	(+)	(+)	(+)	(+)	(-)
Maltese	(+)	(+)	(+)	(-)	(+)	(+)	(+)
Russian	(+)	(+)	(+)	(+)	(+)	(+)	(-)
Sardinian	(+)	(+)	(+)	(+)	(+)	(+)	(-)

⁸ In this column, Basque and English are marked (-) because they allow for an animate locatum only in their dedicated mirative strategies (e.g., *I found myself at the hospital*, **I am found at the hospital*). Similarly, Estonian and Finnish only allow animate locati if these are marked for (partitive) plural.

⁹ Even though the verb *ritrovarsi* ‘to end up (in/doing sth)’ always has a mirative or non-volitional reading, and hence could be thought of as a dedicated strategy, Italian inventives do not in principle need the additional morphemic marking *ri-* to express a mirative or non-volitional meaning. For this reason, I do not consider Italian as having a(n obligatory) dedicated mirative strategy.

Table 2 presents the two types of inventive-locational constructions, differentiated by the animacy of the locatum, or the referent being located. Such binary classification showed the most variation in the data, with the two types of constructions arranging according to two main criteria. The first criterion establishes whether the locative phrase employed is concrete or abstract, while the second criterion concerns indexing, which refers to the grammatical person expressed on the verb. Languages can either present both types of constructions or only Type 2, which is the most common type. In this type, an inanimate locatum is located in a concrete location, and the verb will consequently always present third person (allophoric) indexing. Type 2 does not contemplate mirativity, as it is a category correlated with animate referents. Type 1 has animate referents, which can consequently be located in both concrete and abstract locations (Criterion 1), and can include pronominal subjects, which correlate with indexing for both locuphoric and allophoric forms (Criterion 2). There are some remarks to be made on Type 1 because it can present language-specific variation. Finnish inventive-locational constructions are represented by the verb *löytyä* ‘to be found’. Contrary to, e.g., Italian, in which the same strategy is used for both types of inventive-locational constructions, this Finnish verb does not always behave in the same way. Within Type 1, the Finnish inventive verb is mostly indexed for third (allophoric) persons, but rarely for first or second (locuphoric) persons. These findings are confirmed by QUANTI and WEB. Moreover, mirative marking is rendered by employing a different analytical strategy altogether, which employs the transitive verb *löytää* ‘to find’ and a reflexive pronoun (11). It should also be noted that, within Type 2, compared to many other European languages, Finnish does not allow for immovable referents, such as geographical referents (cf. 12).

(11) Finnish

Löysin itseni sairaalasta
 find.PRT.1SG self.1PX hospital.ELA
 ‘I found myself at the hospital.’

(12) a. Italian

La Finlandia si trova in Europa
 DEF Finland MM find.3SG in Europe
 ‘Finland is located in Europe.’

b. Finnish

* *Suomi löytyy Euroopasta*
 Finland find.MM.3SG Europe.ELA
 * ‘Finland is located in Europe.’

Example (12b) is pragmatically odd. The only context in which it would work would be if two people were looking at a map and one of them uttered (12b) after seeing that the other person is unable to find Finland on said map (Tuomas Huomo, p.c.).

Language-specific subtypes that correlate with the mirativity of the action are not a core defining criterion of inventive-locational constructions for two main

reasons. First, when mirative readings are possible, they may employ a completely different, specialized grammatical strategy compared to their non-mirative counterparts. This is evident in languages such as Maltese, where mirative readings arise from analytical reflexive constructions, whereas languages like Sardinian do not employ any specialized strategy to distinguish between non-mirative and mirative readings. The classification of mirative-marked strategies as inventive-locational constructions is debatable, since verbs with the meaning FIND already have some degree of mirativity in their semantics: the event of finding involves a low degree of control. Finding oneself could be interpreted in a more literal sense. Nonetheless, I included such mirative-marked examples because they can formally be identical to their non-mirative counterparts. Whether they can be considered purely locational or not is indeed problematic but can be left for further research to establish. The main goal of AREA is to show that such construction-strategies do exist and that they have, at least to some extent, a locational function.

The second reason why mirativity is not a core criterion in the definition of inventive-locational constructions is the following. It can be argued that mirative readings cannot occur in Type 2 because this type features inanimate referents. Mirative readings in inventive-locational constructions concern animate referents, who suddenly realize they are in a certain location, facing a certain problem, or undergoing a certain state of mind. For this reason, if a mirative reading were to occur with an inanimate referent, this would acquire semantic properties that are typical of humans and animals. This could be the case with anthropomorphized inanimate entities in fairytales and such. One of the shortcomings of AREA concerns the way animacy is treated: as a logistic (yes/no) property rather than a hierarchy with many possible degrees on a spectrum. The discussion concerning whether, e.g., fungi and plants are to be considered animates or not is left to future research (for a reference on animacy see, e.g., Kittilä et al. 2011). In AREA, the semantic category of animates pertains to entities that have cognitive abilities and are, essentially, movable, such as people and animals. This movability could also justify why mirative readings arise.

3.3 Nominative-partitive alternation in Finnish inventives (QUANTI)

In Finnish, the inventive verb *löytyä* ‘to be found’ is extremely common in locational predication. This study substantiates this by showing quantitative evidence, by analyzing the nominative-partitive marking alternation in the NP that represents the located element, here referred to as the subject. Partitive marking is thought to trigger an existential reading, while nominative NPs can appear in both locative and existential constructions. When they appear in existential constructions, they are usually postverbal, while partitive-marked NPs may also be preverbal and yet trigger an existential reading.

Our quantitative research has shown that different variables contribute to the case marking on the (e-)NP. It also, however, showed that some variables are

associated with each other, and that such interaction modeled the nominative-partitive alternation (the response variable) better than the same variables alone. Let us look at Figure 1.

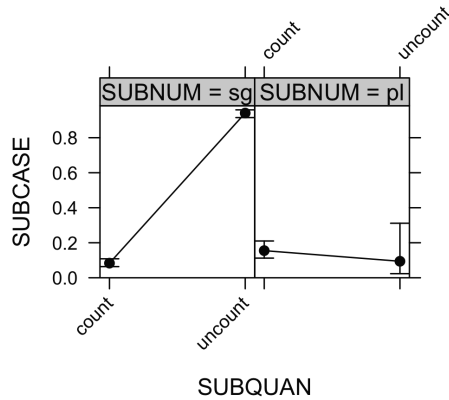


Figure 1. Interaction between number (SUBNUM) and countability (SUBQUAN)

Figure 1 shows that there is a significant interaction between the number and countability of the subject in predicting the nominative-partitive alternation (SUBCASE), coded on the y-axis. Here we can see a scale that ranges from 0 to 1, where 0 corresponds to the Nominative case and 1 to the Partitive case. This means that values at points that are closer to 0 predict for the tendency of that variable to trigger the appearance of the Nominative case, while points that are closer to 1 will predict for the Partitive case. Count nouns in the singular are more likely to trigger the appearance of the Nominative case, as they also do in the plural, while uncount nouns are more likely to do the same only in the plural, while in the singular they usually trigger the Partitive case. It should be mentioned that there were only 10 cases of uncount nouns in the plural, and that these usually indicate a definite quantity of a mass noun (e.g., *the old man's money* above).

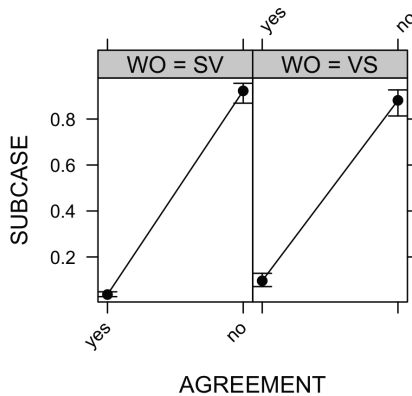


Figure 2. Interaction between agreement and word order (WO)

Figure 2 describes the interaction between agreement and word order (SV = Subject Verb; VS = Verb Subject). This interaction is not surprising, as the absence of agreement is one of the defining criteria of the Finnish Existential construction and is usually correlated with the appearance of the Partitive case. As expected, we can see that in VS sentences, the absence of agreement usually means the appearance of the Partitive case, but it also allows for more Nominative appearances compared to its SV counterpart. VS sentences with no agreement correspond to what is usually undoubtedly an Existential construction (in a language-specific context), which prototypically has a postverbal e-NP, be it flagged for the Nominative or Partitive. In (13), both a Nominative-flagged and Partitive-flagged e-NP make their appearance.

- (13) *Siitä löytyy myös urheiluseurojen ja kylätoimikuntien yhteistiedot sekä joitakin maalaiskunnan nähtävyyksiä*
 3SG.ELA find.MM.3SG also sport.club.PL.GEN and
 village.committee.PL.GEN contact.information.PL and
 some.PL.PART.ENCL rural.community attraction.PL.PART
 ‘There is also contact information for sports clubs and village committees as well as some attractions in the rural community.’ (Basile & Ivaska 2021: 28)

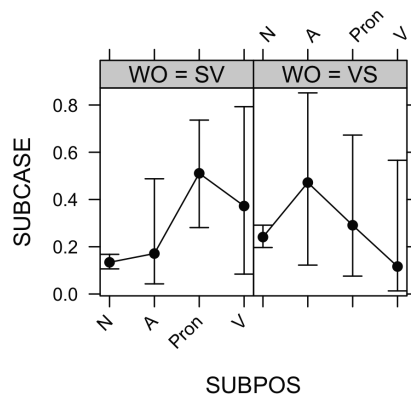


Figure 3. Interaction between part of speech (SUBPOS) and word order

Because most of the subjects (or e-NPs) of the sample are represented by nouns, part of speech and word order do not correlate in a statistically significant way. Despite this, it is interesting to notice that, for example, when the subject is a pronoun, it is more likely to appear flagged for the Partitive case in SV sentences compared to VS sentences. This could be due to the fact that pronouns represent something already mentioned before in discourse, and with a typical VS configuration in Existential constructions, which have the function of introducing a new referent into discourse, they would probably appear in the Nominative and be followed by a relative clause (14). On the other hand, a newly introduced pronominal referent in SV sentences appears flagged for the Partitive, as it could be

interpreted as referring to an open, indefinite subset of previously mentioned referents. This subset is perceived as new and triggers an existential reading (15).

- (14) *Varmasti löytyy joku, joka lähtee sitä viljelemään.*
surely find.MM.3SG someone RELP go.3SG 3SG.PART farm.INF.ILL
‘Surely there will be someone who starts farming it.’ (Basile & Ivaska 2021: 29)
- (15) *Niitä löytyy myös Sri Lankasta.*
3PL.PART find.MM.3SG also Sri Lanka.ELA
‘They are also found in Sri Lanka.’ (Basile & Ivaska 2021: 29)

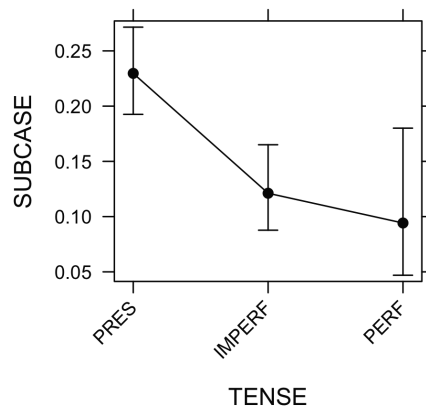


Figure 4. The role of tense in predicting case marking

Looking at Figure 4, we can draw some predictable conclusions. Tenses that refer to past events are usually used in narrative contexts, where the events depicted are finished, telic, and hence definite (VISK §1531; §1538). On the contrary, the Present tense tends to simply describe an ongoing, atelic action and, in the case of a verb like *löytyä*, which is similar to a grammaticalized copula, the presence of something or its location in space. This fact raises the likelihood of having an existential reading, which correlates with the likelihood of having a Partitive-flagged NP. This also means that, when used in the Imperfect or Perfect tense, the verb *löytyä* tends to retain its original meaning and express the result of an actual search (16).

- (16) *Etsinnöissä hänen pyöränsä löytyi läheisen koulun pihasta.*
search.PL.INE 3SG.GEN bicycle.3PX find.MM.PRT nearby.GE
school.GEN yard.ELA
‘During the search, her bicycle was found in the yard of the nearby school.’ (Basile & Ivaska 2021: 31)

We also found that the lemma of the subject or e-NP has a statistically significant role in predicting its Nominative-Partitive alternation. For example, two of the most frequent lemmas in the sample, *ruumis* ‘body, cadaver’ (8 times) and *mies* ‘man’ (6 times), appear in the Nominative. This is not surprising, since these two lemmas usually appear in contexts that describe a finite event (e.g., the finding of a dead body) rather than simply conveying a locational function (17). As for Partitive-marked NPs, two of the most common ones unsurprisingly refer to uncount nouns: *työ* ‘work’ (9 times) and *raha* ‘money’ (6 times). In such cases, it is semantically more likely for the verb *löytyä* to simply convey the presence or the location of certain referents (18).

- (17) a. *isän ruumis löytyi ratin takaa*
 father.GEN body find.MM.PRT.3SG steering.wheel.GEN behind.ELA
 ‘The body of the father was found behind the wheel.’ (Basile & Ivaska 2021: 33)
- b. *Kuollut mies löytyy puistonpenkiltä*
 dead man find.MM.3SG park.bench.ABL
 ‘The dead man is on the park bench.’ (Basile & Ivaska 2021: 33)
- (18) a. *Töitä löytyy keittiön puolelta.*
 work.PL.PART find.MM.3SG kitchen.GEN side.ABL
 ‘There is work in the kitchen sector.’ (Basile & Ivaska 2021: 33)
- b. *palloiluun kyllä löytyy vain lisää rahaa.*
 ball.game.ILL indeed find.MM.3SG only more money.PART
 ‘There is indeed only more money for ball games.’ (Basile & Ivaska 2021: 33)

In this paper, the verb *löytyä* is found often to work as a substitute for the copula *olla* ‘to be’. Since this work was published before I got the idea for AREA, where notions such as semantic bleaching and the copular function of FIND-based strategies are central, we do not refer to *löytyä* as an inventive verb. What follows is that not all the examples included in the paper can be considered inventive-locational constructions (19).

- (19) *Poliisi on vihdoin päässyt murtomiehen jäljille,*
 police be.3SG finally reach.PART burglar.GEN trace.PL.ALL
ja Ukon rahatkin löytyvät.
 and old.man.GEN money.PL.ENCL find.MM.3PL
 ‘The police have finally traced the burglar, and the old man’s money was found.’
 (Basile & Ivaska 2021: 24)

Example (19) is open to interpretation, as it could have a locational function. However, given the context, it is more likely that the verb here retains its original meaning. The police in the sentence are in fact actively searching for hints and end up finding the money. The verb at the end of the example represents a middle-marked impersonal construction with a telic reading rather than an inventive-locational construction.

3.4 Marginal FIND-based strategies in Finnish: Internet contexts (WEB)

This paper is tied to the previous one, in that its starting point is a marginal example of *löytyä*-based constructions. This example represents a hapax in the sample used in QUANTI (20).

- (20) *Mistä löyd-y-t prinsini, 44–50-v. fiksu, pitkäkö,*
 where.ELA find-MM-2SG prince.1PX 44–50-y.o. smart tallish
ulkonäkö ok, pilke silmäkulmassa, lenkkeilet ja
 appearance ok twinkle eye.corner.INE jog.2SG and
tanssit
 dance.2SG
 ‘Where are you my prince, 44 to 50 years old, smart, tallish, good looking, with a twinkle in your eye, you who like to jog and dance.’ (Basile and Ivaska 2021: 18)

(20) is the only instance in which the verb *löytyä* is found indexed for a person different than the third person (singular or plural). I will refer to third persons as *allophoric* forms and to first and second persons as *locuphoric* forms. Because locuphoric forms of the verb *löytyä* are not common, in QUANTI they were disregarded. WEB confirms the claim of marginality of such locuphoric forms by providing corpus evidence for it. There are, in fact, only 449 locuphoric instances of the verb *löytyä* in the corpus, compared to the 613,650 3SG-forms. However, as Figure 5 below shows, the most common locuphoric form is not the 2SG found in QUANTI, but the 1SG, in the Present tense. This form is common because a large part of the sentences found in the sample pertain to the internet environment, where people tend to advertise themselves (21), sometimes by posting their usernames (22).

- (21) *Löydyn myös Facebookista*
 find.MM.1SG also Facebook.ELA
 ‘I am also found on Facebook./You can also find me on Facebook.’

- (22) *löydyn skypestä edelleen samalla vanhalla nimellä*
 find.MM.1SG Skype.ELA still same.ADE old.ADE name.ADE
perneri, että sinne vaan kaikki vanhat ja uudet tyypit
 USER CONJ there.ILL EMPH all old.PL and new.PL guy.PL
 ‘I can be found on Skype still under the same old name perneri, (I would like) all old and new people (to add me) there.’

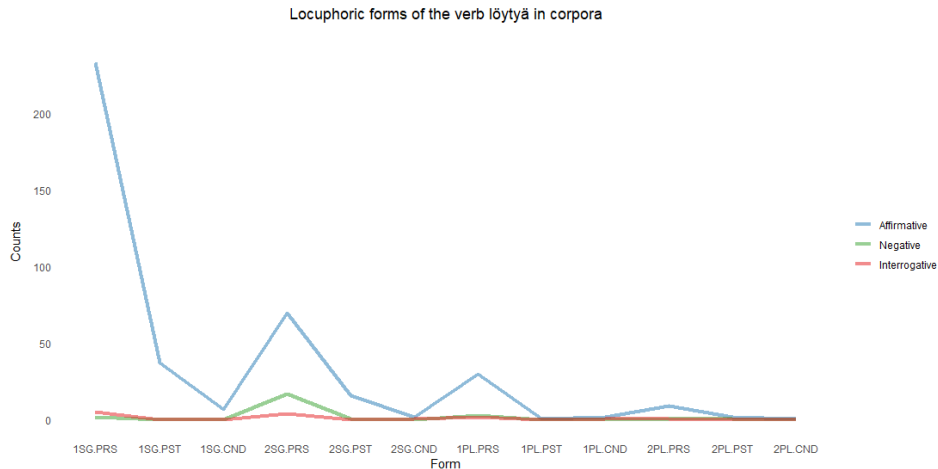


Figure 5. Locuphoric forms of the verb *löytyä* in the corpora (WEB)

While most of the examples pertain to an internet environment, the location expressed could also be a physical one. Moreover, (23) indicates the temporary location of the speaker, also favored by the choice of the verb *löytyä*, as well as by the adverbial phrase used, which provides a temporarily framed context (*iltakuuteen asti* ‘until six in the evening’).

- (23) *Löydyn Fastin pisteeltä kisojen jälkeen*
 find.MM.1SG Fast.GEN point.ABL competition.PL.GEN after
aina iltakuuteen asti.
 always evening.six.ILL until
 ‘After the competition, I am always to be found at Fast’s spot until six in the evening.’

Another interesting point brought up through the analysis of the sample used for WEB is the expression of mirativity. As we saw above, mirativity is frequently associated with FIND-based strategies (24).

- (24) *löydyn uudelleen ja uudelleen pakonomaisesti*
 find.MM.1SG again and again compulsively
tunkemassa ruokaa sisälleni
 shove.INF.INE food.PART inside.ALL.1PX
 ‘I find myself over and over again compulsively shoving food into myself.’

The speaker here finds herself doing something with low control. (24) almost seems like a copy of similar strategies in English, in which an analytical reflexive pronoun is involved in encoding a mirative event. The extent to which such a strategy can be the result of contact with a language like English, which dominates the internet environment, can only be guessed, as substantial data would be needed to corroborate such a claim.

3.5 Finnish Existential Partitive constructions and Expected Sample sizes (PARTEX)

This paper is a quantitative study of 10,000 examples from 10 different Finnish *existential* verbs. Such verbs include the invenitive verb *löytyä*, which behaves differently from the other verbs. It is worth mentioning that the article takes a language-specific point of view of Finnish Existential Partitive constructions, and that many instances of what would be considered an existential construction from the perspective of Fennistics, because it features a partitive-marked NP, could instead belong to what is commonly referred to as a presentational clause. In presentational clauses, the NP is indefinite like in existential constructions, but they are usually not considered to have a locational function.

Table 4. Results of collocation analysis applied to the Expected Sample size (ES) and to a reduced sample (adapted from PARTEX; column “English translation” added).

Verb	English translation	Expected Sample size	Occurrence Frequency	Fisher’s exact test (ES)	Fisher’s exact test (reduced sample)
löytyä	‘to be found’	787	490	4.470362e-11 *	3.531449e-53 *
sisältyä	‘to be included’	427	238	.06911321	.007108333 *
riittää	‘to suffice’	702	390	.02085609 *	5.097901e-17 *
valua	‘to flow’	350	173	.4391751	4.224982e-15 *
tulla	‘to come’	404	200	.4374663	1.53726e-08 *
ilmestyä	‘to appear’	272	130	.236718	2.20209e-30 *
pyöriä	‘to rotate/ circulate’	294	139	.149758	1.235105e-26 *
mahtua	‘to fit’	888	419	.005792884 *	2.156685e-25 *
sataa	‘to rain’	890	417	.002904667 *	7.447229e-25 *
jääädä	‘to remain/stay’	312	144	.05430832	1.090924e-24 *

Table 4 above shows the results of collocation analysis applied to the Expected Sample size and compares them to the results of collocation analysis applied to the original reduced sample. We can see that, while the results in the reduced sample vary in an unpredictable way and do not correlate with the Occurrence Frequencies of the verbs at all, by applying collocation analysis to an Expected Sample size we obtain p-values that can be interpreted more easily. The significant p-values can be ordered and can fulfill the real task of collocation analysis: determining which one of these verbs is more strongly associated with the construction under investigation.

The four verbs in boldface stand out as being the most strongly associated collexemes to the Finnish Existential Partitive collocation. If we exclude *sataa* ‘to rain’, all the other boldface verbs have, semantically speaking, something to

do with partitivity. In fact, *riittää* ‘to suffice’ and *mahtua* ‘to fit’ both refer to quantifying things in one way or another, hence it is no surprise that they are so strongly associated with this construction. What is surprising is that the verb *löytyä* ‘to be found’ is an outlier in both models, both the one based on the reduced sample and the one based on the Expected Sample size. It is, in fact, used almost 50% of the time (490/1000) with a Partitive-marked argument, and this is well reflected when testing for significance. Moreover, the inventive verb *löytyä* is also more likely to occur with the Finnish Existential Partitive construction compared to the copula *olla* ‘to be’. Let us now look at *mahtua* ‘to fit’ and *riittää* ‘to suffice’ (25–26).

- (25) a. *Maailma-an mahtu-u paljon itkeskeis-i-ä ego* “*surffare-i-ta*”.
 world-ILL fit-3SG many egotistical-PL-PART ego=surfer-PL-PART
 ‘There are a lot of egotistical ‘egosurfers’ in the world.’
- b. *Päivä-än mahtu-u suunnaton-ta suru-a, ahdistus-ta,*
 day-ILL fit-3SG enormous-PART sorrow-PART anxiety-PART
viha-a, kiintymys-tä, raivo-a, rakkau-tta.
 hate-PART affection-PART rage-PART love-PART
 ‘There is enormous sorrow, anxiety, hate, affection, rage, and love in every day.’
- (26) a. *katu=pöly-ä riittä-ä kaikkialle.*
 street=dust-PART suffice-3SG everywhere.ALL
 ‘There is street dust everywhere.’
- b. *Hiljaisu-tta riittä-ä täällä koti=maa-ssa.*
 silence-PART suffice-3SG here home=country-INE
 ‘There is (enough/so much) silence in our home country.’

These two verbs, significantly associated with the collocation under investigation, do not primarily have a locational function, but rather a presentational one. It is, however, interesting to observe that, while they somehow retain their semantics, they also seem to express an existential function. It is therefore no surprise that their p-values are statistically significant.

3.6 Discussion

One way of understanding how this dissertation contributes to pioneering research on inventive constructions in Finnish and beyond is through visualizing the interactions between the articles that compose it. Figure 6 shows how the articles relate to each other.

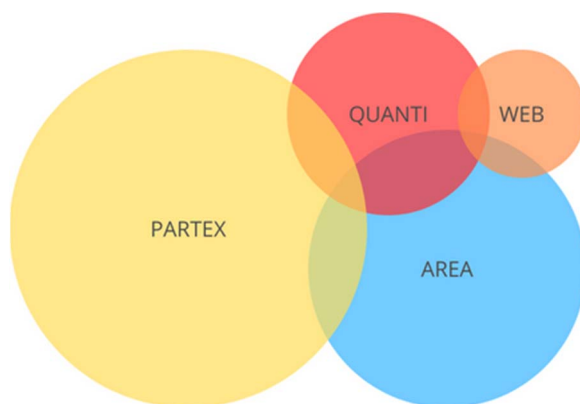


Figure 6. Interrelation of the four dissertation articles (diagram made on Canva by the author).

As can be seen in Figure 6, the articles intersect. The size of the spheres which represent the articles relate to the breadth of the research scope adopted by each article. All four articles contribute to answering the following research questions, albeit to different extents:

1. What are inventive-locational constructions and how do they differ from other locational constructions? (Q1)
2. Why are inventive-locational constructions so common in Finnish and what are their main features? (Q2)
3. What are the main features of the marginal locuphoric inventive strategies used in Finnish? (Q3)
4. What are the most typically used existential verbs in Finnish? (Q4)

Q1 is mainly answered by AREA, which presents an overview of a novel class of construction-strategies called inventive-locational constructions. Such constructions are shown to be common in Europe and presented from a synchronic point of view. Even though Article I also presents the different types of possible inventive-locational constructions, it is not a typological work in a classical sense. This is mainly because typological studies usually aim at surveying language samples that are balanced enough to include as much variety as possible, choosing from as many language families as possible from all the main geographic areas (Miestamo et al. 2016). Instead, my study lays the groundwork for future typological studies by analyzing a small convenience sample of European languages. Inventive-locational constructions have the peculiarity of expressing a locational function through a verb with the meaning FIND.

To answer Q2, the results from QUANTI, WEB, and PARTEX ought to be considered. Inventive-locational constructions appear to be widespread in the Finnish data and can therefore be regarded as productive copulas. This claim requires a couple of comments. First, it is not always easy to determine whether the verb

löytyä is used as a ‘pure’ locational copula or whether it retains its lexical semantic traits. It often triggers ambiguous readings that are open to interpretation and influenced by certain morphosyntactic traits (e.g., past tenses are more likely to trigger the FIND meaning instead of the locational copular meaning). Second, not all the criteria established for inventive-locational constructions (AREA) apply to such constructions in Finnish. The verb *löytyä* is in fact mostly used with allophoric (i.e., third person) forms, while locuphoric (i.e., first and second person) forms are marginal and limited to certain uses. Moreover, the mirative reading, which in other European languages appears to be typical of inventive-locational constructions, is present in Finnish only in a few instances of the verb’s locuphoric uses. This points towards the novelty and episodic nature of such uses, which in my opinion could well be copies of the corresponding English strategies (*to find oneself*). Perhaps this is an indicator of language change, and in the future, mirative-marked inventive strategies might become widespread because of the constant influence of English-speaking media. This has probably already been underway in the sister analytical inventive-locational construction *löytää itsensä* ‘to find oneself’, which marks a mirative event and is consistently used in other European languages (incl. English).

WEB and PARTEX further support these insights by answering Q3 and Q4. Locuphoric forms of the verb *löytyä* are indeed rare and presumably limited to some internet uses and other episodic uses, such as indicating one’s time-framed availability at a certain location. This thought contributes to answering Q1 as well, by offering an insight into what happens in other European languages. For example, in Italian, locuphoric forms of the inventive verb *trovarsi*, albeit more productive than in Finnish, may indicate the temporary presence of the located referent in a certain location (27). Of course, this also depends on other morphosyntactic and pragmatic factors since the verbs *essere* ‘to be’ and *stare* ‘to be/stay’ may also indicate temporary presence.

- (27) Italian
Mi trovo a casa
 MM.1SG find.1SG at home
 ‘I am at home (momentarily)’

By answering Q4, PARTEX also corroborates the results obtained in WEB. The verb *löytyä* is found to be strongly associated with the Finnish Existential Partitive construction, alongside a few other verbs that mostly have presentational features. The corpus search points toward the ordinariness of the verb *löytyä* occurring with the Existential Partitive construction, but it should be kept in mind that the sample used in PARTEX was obtained by narrowing down the corpus search to all occurrences of the verb in 3SG. This means that there is no way locuphoric forms of this verb could have appeared in this corpus search. However, considering the quantitative insight provided by both PARTEX and QUANTI, we can affirm that, because *löytyä* is so common in Existential constructions, its locuphoric forms must be rare even in larger samples.

Some of the interrelations between the papers presented above only arise by looking at this dissertation from a holistic perspective, taking advantage of mixed methods to investigate neighboring topics from different angles. First, the quantitative component, which is in direct communication with the qualitative observations made, helps one gain a broader view of what these inventive-locational constructions really are: In Finnish, among the long list of verbs that are traditionally considered existential, only one, *löytyä* ‘to be found’, stands out as being as strongly associated with the Finnish Existential construction as the verb *olla* ‘to be’. Second, the qualitative component of the dissertation points towards more research questions that can be answered by applying either quantitative methods or carrying out further qualitative research.

Before going to the concluding section of this thesis, it is necessary to discuss one aspect of the quantitative method used in PARTEX. Here, the p-values are not to be interpreted strictly as measures of statistical significance, but they rather mirror the relative strength of the association between each verb and the construction studied, in relation to the other verbs. This means that p-values are operationalized in an unconventional way and should be handled with care. Because multiple tests are conducted, the article should have presented significance thresholds separately corrected for each individual test (through the Holm-Bonferroni method or the Šidák correction¹⁰). Nonetheless, because of the balanced sample size between the different verbs, the relative order of the strength of association can be considered reliable. Hence, PARTEX indicates the existence of differing strengths of association between the verbs investigated and the existential construction. The frequencies presented in the paper also confirm that verb semantics intertwine with construction-level meaning, hence giving a quantitative basis to the general understanding of existential constructions, as well as to the justification for the concept ‘existential verb’.

¹⁰ I thank Unni Leino for sharing expertise over this topic.

4 CONCLUSION

The traits that distinguish inventive-locational constructions from other possible locational constructions allow for typologically-oriented research aimed at showing how such constructions work in a small areal sample of European languages. However, if we were to expand the sample and include languages from outside of Europe, it would still be possible to verify whether these constructions are found in the larger sample and how they behave.

The main difficulty a future researcher is faced with when choosing to carry out a typological survey of inventive-locational constructions in the world's languages is the availability of data. Because inventive-locational constructions are not yet well-established in linguistic research, and because they are not a linguistic category defined purely on functional grounds, they are hard to find, especially in scarcely documented languages. It could be that one or a few examples of inventive-locational constructions do appear in some reference grammars, but writing a grammatical sketch often does not leave space for marginal strategies expressing locational meaning, especially if such strategies are not established and have been neglected for most of the history of linguistic research. Moreover, inventive-locational strategies might not be regarded as conveying locational meaning in the first place. As already discussed, they are not fully grammaticalized and there is a fundamental ambiguity concerning the pragmatic function of FIND-based strategies.

It is also less likely that inventive-locational constructions appear in spoken-language corpora consisting of elicited data. Elicited sentences are usually quite simple and may focus on more “core” strategies of locational constructions as established in the grammatical tradition, such as canonical copulas, zero-copulas, posture verbs, specialized existential copulas, or negative existential markers (Veselinova & Hamari 2022). These macrocategories leave little space for marginal constructions such as inventive-locational constructions.

While inventive strategies seem to convey a locational function, in future research it would be interesting to investigate what are commonly considered locational constructions (i.e., constructions that involve a copula or some other linking element) as pragmatically expressing the meaning FIND. That is, the primary function of a locational copula is that of locating a referent in space, but what does it really mean to locate a referent in space? Could a copula like ‘to be’ pragmatically express the availability of something or someone in space, and that that something or someone is then *findable* or available (at all times or for a limited period) to the interlocutor?

To sum up, these are the main results of this dissertation. *Löytyä* ‘to be found’ behaves similarly to the copula *olla* ‘to be’ and can often substitute it. Quantitative methods (QUANTI) indicate that *löytyä* often appears in Existential constructions and becomes semantically bleached, thus, behaving like a copula and providing further evidence for the previous claim. *Löytyä* is used more often than other so-called existential verbs (except *olla* ‘to be’) in Existential constructions,

as quantitative methods (PARTEX) confirm. The association between *löytyä* and the Finnish Existential Partitive construction is strong. Furthermore, this must mean that locuphoric forms are rare, as also confirmed by QUANTI, and that there are differences in the indexing of the verb *löytyä*. In fact, it is mostly used in third persons, and WEB shows that locuphoric forms are indeed marginal. Finally, verbs like *löytyä* are not uncommon among languages spoken in Europe, and similarly have a locational function.

REFERENCES

- Abbott, Barbara 1992. Definiteness, Existentials, and the “List” Interpretation. In Chris Barker and David Dowty (eds.), *Proceedings of Semantics and Linguistic Theory II*. Columbus, OH: The Ohio State University. 1–16.
- Abbott, Barbara 1993. A Pragmatic Account of the Definiteness Effect in Existential Sentences. *Journal of Pragmatics* 19. 39–55.
- Aikhenvald, Alexandra Y. and Robert M.W. Dixon. 1998. Evidentials and areal typology: a case study from Amazonia. *Language Sciences*, 20(3), 241–57.
- Arkadiev, Peter. 2014. Towards an areal typology of prefixal perfectivization. *Scando-Slavica*, 60(2), 384–405.
- Borin, Lars, Markus Forsberg, and Johan Roxendal. 2012. Korp – the corpus infrastructure of Språkbanken. *Proceedings of LREC 2012*. Istanbul: ELRA, pages 474–478.
- Chomsky, Noam. 1965. *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.
- Comrie, Bernard. 1988. Linguistic typology. *Annual Review of Anthropology*, 17(1), 145–159.
- Creissels, Denis. 2014. Existential Predication in Typological Perspective. Available at <http://www.deniscreissels.fr/public/Creissels-Exist.Pred.pdf> (Last accessed 7 Apr 2024).
- Creissels, Denis. 2019. Inverse-locational predication in typological perspective. *Italian Journal of Linguistics*, 31(2), 37–106.
- Croft, William. 2003. Typology. *The handbook of linguistics*, 337–368.
- Croft, William. 2022. *Morphosyntax: constructions of the world’s languages*. Cambridge University Press.
- Dąbrowska, Ewa. 2015. What exactly is Universal Grammar, and has anyone seen it? *Frontiers in psychology*, 6.
- Dahl, Östen and Maria Koptjevskaja-Tamm. 2001. Circum-Baltic Languages. *Circum-Baltic Languages*, 355.
- Daniel, Michael. 2010. Linguistic typology and the study of language. In: Song, Jae Jung (ed.). 2010. *The Oxford handbook of linguistic typology*. OUP Oxford.
- Declerck, Renaat. 1988. *Studies on copular sentences, clefts and pseudo-clefts*. Leuven: Leuven University Press & Foris.
- De Vries, Mark. 2008. The representation of language within language: A syntactico-pragmatic typology of direct speech. *Studia Linguistica*, 62(1), 39–77.
- Diessel, Holger. 2017. Usage-based linguistics. In *Oxford research encyclopedia of linguistics*.
- Dryer, Matthew. S. 1992. The Greenbergian word order correlations. *Language*, 68(1), 81–138.
- Gaeta, Livio. 2023. Pure existentials vs. pure presentationals. In Laure Sarda and Ludovica Lena (eds.), *Existential Constructions across Languages: Forms, meanings and functions*, 102–138. John Benjamins. <https://doi.org/10.1075/hcp.76.04gae>
- Geeraerts, Dirk. 2010. *Theories of Lexical Semantics*. Oxford: Oxford University Press.
- Goldberg, Adele E. 1995. *Constructions: A construction grammar approach to argument structure* (Cognitive Theory of Language and Culture). Chicago: University of Chicago Press.
- Goldberg, Adele E. 2006. *Constructions at work: The nature of generalization in language*. Oxford: Oxford University Press.
- Gries, Stefan Th. 2015. The most under-used statistical method in corpus linguistics: Multi-level (and mixed-effects) models. *Corpora*, 10(1), 95–125.

- Gries, Stefan Th. 2019. 15 years of collocations: some long overdue additions/corrections (to/of actually all sorts of corpus-linguistics measures). *International Journal of Corpus Linguistics* 24 (3), 385–412.
- Gries, Stefan Th. 2021. *Statistics for linguistics with R: A practical introduction*. Walter de Gruyter GmbH & Co KG.
- Gries, Stefan Th. and Anatol Stefanowitsch. 2004. Extending collocation analysis: A corpus-based perspective on ‘alternations’. *International Journal of Corpus Linguistics* 9 (1), 97–129. <https://doi.org/10.1075/ijcl.9.1.06gri>.
- Hakanen, Aimo. 1972. Normaalilause ja eksistentiaalilause. *Sananjalka*, 14(1), 3–76.
- Haspelmath, Martin. 2010. Comparative concepts and descriptive categories in cross-linguistic studies. *Language*, 86(3), 663–687.
- Haspelmath, Martin. 2011. On S, A, P, T, and R as comparative concepts for alignment typology. *Linguistic Typology* 15, 535–567.
- Haspelmath, Martin. 2016. The serial verb construction: Comparative concept and cross-linguistic generalizations. *Language and linguistics*, 17(3), 291–319.
- Haspelmath, Martin. 2018. How comparative concepts and descriptive linguistic categories are different. In Daniël Van Olmen, Tanja Mortelmans, and Frank Brisard (eds.). *Aspects of linguistic variation*. De Gruyter Mouton, 83–113.
- Haspelmath, Martin. 2021. <https://dlc.hypotheses.org/2516>
- Haspelmath, Martin. 2022. Nonverbal clause constructions. Submitted manuscript. [Lingbuzz/006673](https://lingbuzz/006673). <https://ling.auf.net/lingbuzz/006673>, last accessed 7 Apr 2024.
- Heine, Bernd. 2017. Grammaticalization. In Brian D. Joseph and Richard D. Janda (eds.) *The handbook of historical linguistics*, 573–601.
- Heine, Bernd. 1993. *Auxiliaries: Cognitive forces and grammaticalization*. Oxford University Press.
- Helasvuo, Marja-Liisa. 1996. Ollako vai eikö olla – Eksistentiaalilauseen subjektin kohtalonkysymys. *Virittäjä* 100(3), 340–356.
- Hengeveld, Kees. 1992. *Non-verbal predication: Theory, typology, diachrony*. Berlin: Mouton de Gruyter.
- Huumo, Tuomas. 2003. Incremental existence: The world according to the Finnish existential sentence, *Linguistics* 41–3, 461–493.
- Huumo, Tuomas and Marja-Liisa Helasvuo. 2015. On the subject of subject in Finnish. In Marja-Liisa Helasvuo and Tuomas Huumo (eds.), *Subjects in constructions—canonical and non-canonical* (Constructional Approaches to Language 16), 13–41. Amsterdam: John Benjamins. <https://doi.org/10.1075/cal.16.02huu>.
- Huumo, Tuomas and Liina Lindström. 2014. Partitives across constructions: On the range of uses of the Finnish and Estonian ‘partitive subjects’. In Silvia Luraghi and Tuomas Huumo (eds.), *Partitive cases and related categories*. Berlin, Boston: De Gruyter. <https://doi.org/10.1515/9783110346060.153>.
- Idiatov, Dmitry. 2018. An areal typology of clause-final negation in Africa: Language dynamics in space and time. In Daniël Van Olmen, Tanja Mortelmans, and Frank Brisard (eds.). *Aspects of linguistic variation*. De Gruyter Mouton, 115–163.
- Ivaska, Ilmari. 2010. Eksistentiaalilauseen subjektijäsen edistyneiden suomenoppijoiden kirjoituksessa. *Sananjalka* 52. 67–88. <https://doi.org/10.30673/sja.86697>.
- Ivaska, Ilmari. 2011. Lausetyyppien sekoittuminen edistyneessä oppijansuomessa – näkökulmana eksistentiaalilause. *Lähivördlusi. Lähivertailuja* 21. 65–85. <https://doi.org/10.5128/lv21.03>.

- Ivaska, Ilmari. 2022. Regression analysis. In Jan-Ola Östman & Jef Verschueren (eds.), *Handbook of Pragmatics*, vol. 24th annual installment, 3–32. Amsterdam: John Benjamins. <https://doi.org/10.1075/hop.24.reg2>.
- Ivaska, Laura. 2020. *A mixed-methods approach to indirect translation*. Turku: University of Turku. [https://www.utupub.fi/handle/100,24\(15075\),5](https://www.utupub.fi/handle/100,24(15075),5).
- Kilgarriff, Adam, Vít Baisa, Jan Bušta, Miloš Jakubíček, Vojtěch Kovář, Jan Michelfeit, Pavel Rychlý, and Vít Suchomel. 2014. The Sketch Engine: ten years on. *Lexicography*, 1: 7–36.
- Kittilä, Seppo, Jussi Ylikoski, and Katja Västi. 2011. *Case, animacy and semantic roles. Case, Animacy and Semantic Roles*. John Benjamins.
- Klavan, Jane. 2020. Pitting corpus-based classification models against each other: a case study for predicting constructional choice in written Estonian. *Corpus Linguistics and Linguistic Theory* 16(2), 363–391. <https://doi.org/10.1515/cllt-2016-0010>.
- Klavan, Jane, Maarja-Liisa Pilvik, and Kristel Uibo. 2015. The Use of Multivariate Statistical Classification Models for Predicting Constructional Choice in Spoken, Non-Standard Varieties of Estonian. *SKY Journal of Linguistics*, 28.
- Lakoff, George. 1987. *Women, fire, and dangerous things: What categories reveal about the mind*. Chicago: University of Chicago Press.
- Lander, Yury and Peter Arkadiev. 2016. On the right of being a comparative concept. *Linguistic Typology*, 20(2), 403–416.
- Langacker, Ronald W. 1987. *Foundations of cognitive grammar. Vol. II. Theoretical prerequisites*. Stanford: Stanford University Press.
- Langacker, Ronald W. 2008. *Cognitive grammar*. Berlin: De Gruyter Mouton. <https://doi.org/10.1515/9783110199901.29>.
- Levshina, Natalia. 2015. *How to do linguistics with R. Data exploration and statistical analysis*. Amsterdam: Benjamins.
- Lindström, Liina, Maarja-Liisa Pilvik, and Helen Plado. 2021. Variation in negation in Seto. *Studies in Language*, 45(3), 557–597.
- McNally, Louise. 2016. Existential sentences crosslinguistically: variations in form and meaning. *Annual Review of Linguistics* 2: 211–31.
- Miestamo, Matti. 2005. Iso suomen kielioppi typologin silmin. *Virittäjä*, 109(4), 570–576.
- Miestamo, Matti, Dik Bakker, and Antti Arppe. 2016. Sampling for variety. *Linguistic Typology*, 20(2), 233–296.
- Mikkelsen, Line. 2011. Copular clauses. In Claudia Maienborn, Klaus von Heusinger and Paul Portner (eds.), *Semantics: An international handbook of natural language meaning, volume 2*, 1805–1829. Berlin: De Gruyter Mouton.
- Pook, Hanna. 2021. Object case variation of the pronoun mis ‘what’ in spontaneous spoken Estonian and Estonian dialects. *Eesti ja soome-ugri keeleteaduse ajakiri. Journal of Estonian and Finno-Ugric Linguistics*, 12(1), 259–301.
- R Core Team. 2018. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>.
- Rossi, Giovanni. 2020. Other-repetition in conversation across languages: Bringing prosody into pragmatic typology. *Language in Society*, 49(4), 495–520.
- Roy, Isabelle. 2013. *Nonverbal predication: Copular sentences at the syntax-semantics interface*. Oxford: Oxford University Press.
- Seržant, Ilja. A. 2021. Typology of partitives. *Linguistics*, 59(4), 881–947.
- Song, Jae Jung. 2010. Setting the stage. In: Song, Jae Jung (Ed.). 2010. *The Oxford handbook of linguistic typology*. OUP Oxford.

- Speelman, Dirk. 2014. Logistic regression. *Corpus methods for semantics: Quantitative studies in polysemy and synonymy*, 43, 487–533.
- Stefanowitsch, Anatol. 2020. *Corpus linguistics: A guide to the methodology*. (Textbooks in Language Sciences 7). Berlin: Language Science Press. DOI: 10.5281/zenodo.3735822
- Stefanowitsch, Anatol and Stefan Th Gries. 2003. Collocations: Investigating the interaction of words and constructions. *International Journal of Corpus Linguistics* 8 (2), 209–243.
- Talmy, Leonard. 2000. *Toward a cognitive semantics (Vol. 2)*. MIT press.
- Varjo, Mikael. 2022. Greater than zero? A study of referentially specific and open necessity constructions in Finnish everyday conversation. *Eesti ja soome-ugri keeleteaduse ajakiri. Journal of Estonian and Finno-Ugric Linguistics*, 13(2), 5–46.
- Varjo, Mikael and Karita Suomalainen. 2018. From zero to ‘you’ and back: A mixed methods study comparing the use of two open personal constructions in Finnish. *Nordic Journal of Linguistics*, 41(3), 333–366.
- VISK = Auli Hakulinen, Maria Vilkuna, Riitta Korhonen, Vesa Koivisto, Tarja Riitta Heinonen, and Irja Alho 2004: *Iso suomen kielioppi*. Helsinki: Suomalaisen Kirjallisuuden Seura. Web version. Available at: <https://kaino.kotus.fi/visk/etusivu.php> (last accessed on 7 Apr 2024).
- Wilbur, Joshua. 2014. *A grammar of Pite Saami*. (Studies in Diversity Linguistics 5). Berlin: Language Science Press. DOI: 10.17169/FUDOCs_document_00000020749
- Winter, Bodo. 2019. *Statistics for linguists: An introduction using R*. Routledge.

EESTIKEELNE KOKKUVÕTE

Inventiiv-lokatsioonid soome keeles: Segametoodiline lähenemine

Sissejuhatus

Erinevates keeltes on olemas eri viise väljendamaks seda, kuidas asjad on ruumis paigutatud. Selline asukoha väljendamine toimub Euroopa keeltes peamiselt nii-öelda koopulakonstruktsioonide kaudu (nt eesti keele *õlu on külmkapis*), kus *olema*-koopulal on teatud lokationaalne funktsioon. Samasugust funktsiooni väljendatakse ka ilma koopulata (nt sama tähendusega venekeelses lauses *pivo v kholodil'nike*) või asendit väljendava tegusõna kaudu (nt eesti *poiss magab voodis*). On aga võimalik ka kasutada muid grammatilisi strateegiaid koopula asemel, millele keeletüpoloogiline traditsioon on siiani jätnud tähelepanu pööramata. Kutsun neid strateegiaid *inventiiv-lokationaalseteks konstruktsiooniteks* ja selles töös uurin nende käitumist soome keeles, võrrelduna teiste Euroopa keeltega. Soome keeles on selles konstruktsioonis kasutusel tegusõna *löytyä* 'leiduma', mis käitub samamoodi kui muud koopulad. Termin *inventiivne* tuleb ladinakeelsest sõnast *invenire* 'leidma'. Väitekiri koosneb sissejuhatavast osast ja neljast artiklist, millele viitan lühenditega AREA, QUANTI, WEB ja PARTEX. Siinne väitekiri on segametoodiline uurimus, mis keskendub soome keelele. See püüab vastata neljale suurele uurimisküsimusele:

1. Mis on inventiiv-lokationaalset konstruktsioonid ja kuidas need erinevad muudest lokationaalsetest konstruktsioonidest? (K1)
2. Miks on inventiiv-lokationaalset konstruktsioonid soome keeles nii tavalised ja mis on nende peamised omadused? (K2)
3. Mis on soome keele marginaalsete lokufooriliste inventiivsete strateegiade peamised omadused? (K3)
4. Mis on kõige tavalisemad kasutatud eksistentsiaalsed tegusõnad soome keeles? (K4)

K1 vastab artikkel AREA, mis annab ülevaate inventiiv-lokationaalsetest konstruktsioonidest Euroopa keeltes. Kuigi artikkel ei ole laiaulatuslik keeletüpoloogiline uurimus, esitab see siiski inventiiv-lokationaalset konstruktsioone kui komparatiivseid mõisteid (*comparative concept*), mis võimaldab selliste konstruktsioonide tuvastamise ja uurimise muudes maailma keeltes.

K2 vastavad artiklid QUANTI, WEB ja PARTEX. Inventiiv-lokationaalset konstruktsioonid soome keeles on produktiivsed ning neid kasutatakse nii erinevates morfosüntaktilistes kui ka pragmatilistes kontekstites.

K3 vastavad artiklid QUANTI ja WEB. QUANTI esitab lokufoorilised *löytyä* – tegusõnaga strateegiaid ja QUANTI-artiklis esitatud tõlgendused saavad osaliselt kinnitust erinevast vaatepunktist lähtuvas artiklis WEB. Samuti leiab WEB, et

lokufoorilisi inventiivseid strateegiaid kasutatakse veebitekstides peamiselt selleks, et väljendada keelekasutaja olemasolu mingil internetisaidil või veebiplatvormil.

K4 vastab artikkel PARTEX, mis uurib soome keele muid võimalikke eksistentsiaalseid tegusõnu. Artikkel leiab, et *löytyä* -teigusõna on statistiliselt palju tihedamini seotud soome osastavas käändes subjekti sisaldava eksistentsiaalse konstruktsiooniga kui muud tegusõnad. Artikkel kinnitab seega töö teistes artiklites olevaid väiteid, et *löytyä* võib mõnikord asendada koopulat *olla* 'olema'.

Taust

Siinses eestikeelses kokkuvõttes kasutan uut katusterminid *lokatsioonaalne konstruktsioon* (inglise keeles *locational construction*). See hulk konstruktsioone sisaldab seega nii lokationaalseid konstruktsioone (*locative constructions*) kui ka eksistentsiaalseid konstruktsioone (*existential constructions*). Haspelmathi (2022) tüpoloogiat järgides on lokationaalsed konstruktsioonid tüüpi *õlu on külmkapis*, kus *õlu* viitab definiitsele referendile, mis tavaliselt paikneb lause alguses. See, et referent paikneb külmkapis, on uus informatsiooniosa. Eksistentsiaalset konstruktsiooni väljendavas lauses *külmkapis on õlu* lause alguses aga definiitne koht. Lauses öeldakse, mis selles kohas paikneb. *Õlu* on selles lauses seega indefiniitne referent. Esimene osa väitekirjast paigutab inventiiv-lokationaalsed konstruktsioonid muude lokationaalsete konstruktsioonide hulka ja uurib neid areaalse keeleteaduse perspektiivist.

Teine osa keskendub inventiiv-lokationaalsete konstruktsioonide kasutamisele soome keeles, analüüsib neid pigem keelespetsiifilisest vaatenurgast ega võrdle neid muude keeltega. Samamoodi kui eesti keeles, kasutatakse ka soome keele eksistentsiaalsetes lausetes osastava käändega markeerituid subjekte, mida kutsutakse ka eksistentsiaalseteks NP-deks (e-NP; vt Huumo ja Helasvuo 2015). Soome keeleteaduse traditsioonis on nimetavas ja osastavas käändes subjektide vaheldus olnud pikka aega tähelepanu keskpunktis.

Materjal ja meetod

See töö on segamethodiline uurimus soome keele inventiiv-lokationaalsetest strateegiatest. See tähendab, et nii kvalitatiivsed kui ka kvantitatiivsed meetodid on ühendatud, et püüda seletada inventiiv-lokationaalseid strateegiaid erinevatest vaatenurgadest.

Väitekirja keeletüpoloogiline osa (AREA) analüüsib inventiivseid strateegiaid 13 Euroopa keeles, mida uuritakse kvalitatiivsest vaatenurgast. Kogusin erinevate keelte kohta materjali veebist ja kontrollisin saadud infot vastavaid keeli emakeelena rääkijate käest küsides, mida nad arvavad inventiiv-lokationaalsetest strateegiatest oma keeles, või kogusin näitelauseid otse rääkijatelt. Keelte valim on kirjas Tabelis 1.

Tabel 1. Analüüsiks kasutatud keelte valim (AREA)

Keel	Keelkond, genus
Albaania	Indoeuroopa, isolaat
Baski	Isolaat
Inglise	Indoeuroopa, germaani
Eesti	Uurali, läänemeresoome
Soome	Uurali, läänemeresoome
Saksa	Indoeuroopa, germaani
Kreeka	Indoeuroopa, helleeni
Ungari	Uurali, ugri
Itaalia	Indoeuroopa, romaani
Läti	Indoeuroopa, balti
Malta	Afroaasia, semi
Vene	Indoeuroopa, slaavi
Sardi	Indoeuroopa, romaani

Soome keelele keskenduv osas (QUANTI, WEB ja PARTEX) kogusin materjali veebikorpustest. Analüüsin materjali, kasutades kvantitatiivseid meetodeid.

Artiklis QUANTI kasutasin korpust *Kansalliskirjaston lehtikokoelman (KLK) suomenkieliset lehdet*, millest võtsin 779 juhuslikku näitelauseid, mis sisaldavad soome verbi *löytyä* 'leiduma'. Analüüsisin valimit statistiliste meetoditega, nimelt logistilise regressiooni segamudeliteabil. See meetod võimaldab nime-tavas ja osastavas käändes subjektide uurimist soome *löytyä* -teguõnaga lausetes sõltumatute muutujate põhjal (nt sõnajärg, aeg, subjekti lemma jne.).

Artiklis WEB kasutasin platvormist Sketch Engine võetud korpust Finnish Web 2014 (fiTenTen2014). Materjal koosneb marginaalsest valimist, mis sisaldab 449 *löytyä* -teguõnaga lauset, kus teguõna on lokufoorilises vormis ehk 1. ja 2. grammatilises isikus.

Artiklis PARTEX kasutasin veebikorpust Suomi24:2017, mis koosneb aastal 2017 kirjutatud Suomi24-foorumi postitustest. Sellest korpusest kogusin 10 erineva soomekeelse eksistentsiaalse teguõnaga näitelauseid (kokku 10 000 lauset, 1000 per teguõna), mida uurisin kollostruktsioonilist analüüsi kasutades. Kollostruktsiooniline analüüs (Stefanowitsch ja Gries 2003; Gries ja Stefanowitsch 2004) on meetod, mis võimaldab keelekonstruktsioonide (siin: eksistentsiaal-lauseid) ja nendega seotud sõnade (siin: eksistentsiaalsete teguõnade) omavaheliste suhete uurimist. Meetod uurib, kui tihedalt on konstruktsioonid ja sõnad omavahel seotud võrreldes muude sõnadega. Artiklis esitlen ka uut valimi moodustamise meetodit, mida kutsun *Expected Sample size* (ES). See meetod püüab arvesse võtta üksikute verbide esinemissagedused kogu korpuse ulatuses ja püüab võrrelda neid 1000 lausega valimist saadud esinemissagedustega.

Tulemused

Inventiiv-lokationaalsed konstruktsioonid on laialt levinud Euroopa keeltes, nagu näitab AREA. Inventiiv-lokationaalseid konstruktsioone iseloomustab inventiivne tegusõna, millel on järgmised omadused:

- I. selle sõnatüve põhitähendus on LEIDMA ja see on *semantiliselt pleekunud* (ing.k. *semantic bleaching*);
- II. see sisaldab valentsi või tegumoodi muutvat morfoloogilist või perifrastilist markerit;
- III. see väljendab lokationaalset funktsiooni ilma spetsiifilist asendit semantiliselt markeerimata.

Inventiiv-lokationaalseid konstruktsioone on kaht tüüpi, mis muutuvad sõltuvalt sellest, kas kohas paiknev element (ehk *locatum*, vt Haspelmith 2022) on elus või elutu (Tabel 2).

Tabel 2 – Inventiiv-lokationaalsete konstruktsioonide tüübid (AREA)

	Tüüp 1: elus locatum	Tüüp 2: elutu locatum
Kriteerium 1: lokatiivse fraasi konkreetsus	konkreetne/abstraktne	konkreetne
Kriteerium 2: tegusõna indekseerimine	lokufoorised ja allofoorised vormid (keelespetsiifilised erinevused)	allofoorised vormid
Keelespetsiifilised alltüübid: miratiivsus	miratiivne tähendus võimalik (oleneb konstruktsioonist)	NA

Soome keele inventiivseid konstruktsioone, mis põhinevad tegusõnal *löytyä* 'leiduma', kasutatakse tihti lokationaalsete koopulatena. Tavaliselt ei kasutata sellistes soomekeelsetes konstruktsioonides allofoorilisi vorme ja elus referentide kohta kasutatakse neid ainult seoses inim- või loomarühmadega, kuid mitte üksikute referentidega. Üksikute inim-referentide või loom-referentide kasutamine korreleerub miratiivse tähendusega ning väljendub teise analüütilise konstruktsiooni kaudu (*löytää itsensä* 'ennast leidma'). *Löytyä*-tegusõnaga konstruktsioonides võib *locatum* esineda nimetavas või osastavas käändes markeeritud vormis, millest on erinevus lokatiivse (1a) ja eksistentsiaalse (1b) konstruktsiooni vahel.

(1) Soome

- a. *Olut löytyy jääkaapista*
õlu leidma.REFL.3SG külmkapp.ELA
'Õlu leidub külmkapist.'
- b. *Jääkaapista löytyy olutta*
külmkapp.ELA leidma.REFL.3SG õlu.PART
'Külmkapist leidub õlut.'

QUANTI selgitab nimetava ja osastava vaheldust *löytyä* -tegusõnaga konstruktsioonides statistiliste meetoditega. Artikkel näitab, et *löytyä* on laialt kasutatud nii lokatiivsetes kui ka eksistentsiaalsetes konstruktsioonides ning seda kasutatakse samamoodi kui koopulat *olla*. Erinevad muutujad mõjutavad nimetava ja osastava käände valimist. Näiteks subjekti loendatavus korreleerub grammatilise arvuga ning nad mõjutavad koos sõltuva muutuja ennustamist. Samuti mõjutab ennustamist kongruentsuse ja sõnade järjekorra interaktsioon, aga sõnaliigi ja sõnade järjekorra interaktsiooni mõju ei ole statistiliselt oluline. Tegusõna grammatilise aja mõju on statistiliselt oluline ja seda väidet toetab see, et *löytyä* -verbil on tavaliselt koopula tähendus olevikus, kuid minevikus säilitatakse selle leksi-kaalne tähendus ja verb määrab subjektile nimetava käände. Samuti mõjutab oluliselt subjekti lemma: teatud lemmad, näiteks *ruumis* 'laip', korreleeruvad nimetava käändega, sest nimetav kirjeldab tavaliselt finiiitseid sündmusi, näiteks laiba leidumine kusagilt.

WEB kinnitab QUANTI tulemust, et *löytyä*-tegusõna lokufoorilised vormid on marginaalsed. Tegelikult on kogu korpus ainult 449 lokufoorilist esinemist, mida kasutatakse ainult teatud kontekstides. Sellistel esinemistel on enamasti spetsiifiline pragmaatiline funktsioon, sest lokufooriliste vormide kaudu reklaamivad kõnelejad oma kohalolekut internetis. Lokufooriliste vormide kasutus korreleerub miratiivsete kasutustega, millega väljendatakse kõnelejate tahtmatut osalemist teatud tegevuses või üllatuslikku/ootamatu olemist teatud kohas.

PARTEX uurib soome keele eksistentsiaalseid konstruktsioone, kus subjekt on markeeritud osastava käändega. Artikkel kinnitab, et kõikidest soome eksistentsiaalsetest tegusõnadest on *löytyä* 'leiduma' kõige rohkem seotud osastava käändega eksistentsiaalse konstruktsiooniga. Muud sageli kasutatud eksistentsiaalsed tegusõnad on *mahtua* 'mahtuma', *riittää* 'piisama' ja *sataa* 'sadama'. Tulemused kinnitavad oletust, et tegusõna *löytyä* 'leiduma' kasutatakse lokationaalse koopulana, samamoodi kui koopulat *olla* 'olema'.

Kokkuvõtteks

Siinne väitekiri esitab uue komparatiivse mõiste, mida võib kasutada tulevastes keeleteaduslikes uurimistöödes. Invenitiiv-lokationaalsed konstruktsioonid on nii-öelda konstruktsiooni-strateegiad, sest nendel on küll funktsionaalne mõõde, aga neid defineeritakse peamiselt formaalsete omaduste põhjal. 'Leiduma'-verbid lokatiivses ja eksistentsiaalses funktsioonis on Euroopa keeltes laialt levinud ning neid võib näha areaalse nähtusena. On oluline mõista, et invenitiiv-lokationaalsete konstruktsioonide defineerimine annab keeleteadlastele võimaluse väljuda areaalse lingvistika lähenemisest ning läheneda *leiduma*-konstruktsioonide keeletüpoloogilisest vaatenurgast. See tähendab seda, et invenitiiv-lokationaalseid konstruktsioone defineerivate formaalsete kriteeriumite põhjal on võimalik välja selgitada, kas sellised konstruktsioonid on olemas näiteks Austraalia või Lõuna-Ameerika keeltes.

Esmased tulemused juba näitavad, et invenitiiv-lokationaalsed konstruktsioonid on olemas mitmetes Aafrika keeltes, nimelt mande keeltes, ning Filipiinidel

tagalogi keeles, milles inventiiv-lokationaalsete konstruktsioonide kasutus võib olla hispaania ja inglise keelega pikaajase kontakti tulemus. Mande keeltes võib nende kasutus olla ka prantsuse keelega kontakti tulemus.

Edaspidistes uurimustes võiks olla huvitav küsida, kas inventiiv-lokationaalset konstruktsioonid grammatiseeruvad ja kuidas nad seda teevad. Rootsi keel annab hea näite sellest võimalusest, sest eksistentsiaalne konstruktsioon *det finns* on grammatiseerunud *finna* 'leidma'-teguõnast. Kuna rootsi keel on mõjutanud soome keelt pikka aega, on võimalik et *löytyä*-teguõna sage kasutus on seotud just sellise eksistentsiaalse kasutusega rootsi keeles ja et tulevikus samasugune kasutus grammatiseerub spetsiaalseks lokatiivseks ning eksistentsiaalseks koo-pulaks.

SUOMENKIELINEN TIIVISTELMÄ

Inventiivis-lokationaaliset konstruktiot suomen kielessä: Monimenetelmätutkimus

Johdanto

Eri kielissä on erilaisia tapoja ilmaista, missä asiat sijaitsevat. Tällainen paikan ilmaiseminen tapahtuu Euroopan kielissä enimmäkseen niin sanottujen kopularakenteiden kautta (esim. *olut on jääkaapissa*), joissa *olla*-kopulalla on sijaintia ilmaiseva funktio. Samanlaista funktiota voidaan ilmaista myös ilman kopulaa (esim. samamerkityksellisessä venäjänkielisessä lauseessa *pivo v kholodil'nike*) tai asemaa ilmaisevan verbin kautta (esim. suomen *poika nukkuu sängyssä*). Kopulan sijaan on kuitenkin myös mahdollista käyttää muita kieliopillisia strategioita, joihin kieliopillinen perinne ei ole tähän mennessä kiinnittänyt huomiota. Kutsun näitä strategioita *inventiivis-lokationaalisiksi konstruktioiksi* ja tässä työssä tutkin niiden käyttäytymistä suomen kielessä muihin Euroopan kieliin verraten. Tällaisessa konstruktiossa suomen kielessä käytetään *löytyä*-verbiä, joka käyttäytyy samalla tavalla kuin muut kopulat. Termi *inventiivinen* tulee latinankielisestä sanasta *invenire* eli 'löytää'. Väitöskirja koostuu johdantokappaleesta sekä neljästä osatutkimuksesta, joihin viitataan lyhenteillä AREA, QUANTI, WEB ja PARTEX. Käsillä oleva väitöskirja on suomen kieleen keskittyvä monimenetelmätutkimus. Se pyrkii vastaamaan neljään suureen tutkimuskysymykseen:

1. Mitkä ovat inventiivis-lokationaaliset konstruktiot ja millä tavalla ne poikkeavat muista paikkaa ilmaisevista konstruktioista? (K1)
2. Miksi ovat inventiivis-lokationaaliset konstruktiot niin tavallisia suomen kielessä ja mitkä ovat niiden keskeiset ominaisuudet? (K2)
3. Mitkä ovat suomen kielen marginaalisten lokuforisten inventiivisten strategioiden keskeiset ominaisuudet? (K3)
4. Mitkä ovat yleisimmät eksistentiaaliset verbit suomen kielessä? (K4)

K1:een vastaa osatutkimus AREA, joka antaa yleiskuvan inventiivis-lokationaalisista konstruktioista Euroopan kielissä. Vaikka artikkeli ei ole laaja kieli-tytologinen tutkimus, se käsittelee kuitenkin inventiivis-lokationaalisia konstruktioita vertailtavissa olevina käsitteinä (*comparative concept*), mikä mahdollistaa niiden konstruktioiden tunnistamisen ja tutkimisen myös muissa maailman kielissä.

K2:een vastaavat osatutkimukset QUANTI, WEB ja PARTEX. Tulokset osoittavat, että inventiivis-lokationaaliset konstruktiot ovat produktiivisia suomen kielessä, sekä sen, että niitä käytetään erilaisissa morfosyntaktisissa ja pragmaattisissa konteksteissa.

K3:een vastaavat osatutkimukset QUANTI ja WEB. QUANTI käsittelee *löytyä*-verbin lokuforisia strategioita ja QUANTI-artikkelissa esitetyt tulkinnat saavat osin vahvistusta näkökulmaltaan eriävässä WEB-artikkelissa. WEB-artikkelissa havaitaan lisäksi, että lokuforisia inventiivisiä strategioita käytetään verkkoteksteissä enimmäkseen ilmaisemaan kielenkäyttäjän olemassaoloa jollakin internetin sivustolla tai verkkoalustalla.

K4:een vastaa osatutkimus PARTEX, joka tutkii niitä verbejä, joita voidaan *olla*-verbin ohella käyttää ilmaisemaan olemassaoloa osana eksistentiaalista konstruktioita. Artikkelit tuo esille, että *löytyä*-verbi on tilastollisesti assosioitu suomen partitiivisubjekteja sisältävien eksistentiaalisten konstruktioiden kanssa paljon vahvemmin kuin muut verbit. Artikkelit vahvistaa siten väitöskirjan toisissa osatutkimuksissa esitetyn tulkinnan, että *löytyä* voi ajoittain korvata *olla*-kopulan.

Tausta

Tässä suomenkielisessä tiivistelmässä käytän uutta kattotermiä *lokationaalinen konstruktio* (englanniksi *locational construction*). Niinpä tällainen ryhmä konstruktioita sisältää sekä lokatiivisia konstruktioita (*locative constructions*) että eksistentiaalisia konstruktioita (*existential constructions*). Haspelmathin (2022) typologiaa noudattaen lokatiiviset konstruktioit ovat tyyppiä *olut on jääkaapissa*, jossa *olut* viittaa määräiseen tarkoitteeseen, joka suomessa yleensä sijaitsee lauseen alussa. Se, että tarkoite sijaitsee jääkaapissa, esitetään uutena tietona. Eksistentiaalista konstruktioita ilmaisevan lauseen *jääkaapissa on olut* alkuosassa on toisaalta paikanilmaus, joka ilmaisee tyyppillisesti jo tunnetun sijainnin. Lause ilmaisee sitä, mikä kyseisessä paikassa sijaitsee, ja *olut* esittelee uuden tiedon jostakin kyseisessä paikassa sijaitsevasta. Ensimmäinen osa väitöskirjaa asemoi inventiivis-lokationaaliset konstruktioit suhteessa muihin lokationaalisiin konstruktioihin ja tutkii niitä alueellis-typologisesta näkökulmasta.

Toinen osa keskittyy inventiivis-lokationaalisten konstruktioiden käyttämiseen suomen kielessä, analysoiden niitä pikemminkin kielispesifistä näkökulmasta ja vertailematta niitä muihin kieliin. Suomenkielisissä eksistentiaalisissa lauseissa voidaan käyttää partitiivilla merkittyjä subjekteja, joita kutsutaan myös eksistentiaalisiksi substantiivilausekkeiksi (e-NP; ks. Huumo ja Helasvuo 2015). Suomalaisessa kielitieteellisessä perinteessä nominatiivi- ja partitiivisubjektien vaihtelu on ollut jo pitkään huomion keskipisteenä.

Materiaali ja menetelmät

Tämä työ on monimenetelmätutkimus suomen kielen inventiivis-lokationaalista strategioista. Se tarkoittaa, että työssä yhdistetään laadullisia ja määrällisiä menetelmiä inventiivis-lokationaalisten strategioiden selittämiseksi erilaisista näkökulmista.

Väitöskirjan kielitypologinen osa (AREA) analysoi inventiivisiä strategioita 13:ssa Euroopan kielessä ja niitä tutkitaan laadullisesta näkökulmasta. Keräsin

verkosta materiaaleja erilaisista kielistä ja tarkistin saadut tiedot kysymällä kielten äidinkielisiltä puhujilta, mitä mieltä he ovat inventiivis-lokationaalisista strategioista omilla kielillään. Joissakin tapauksissa keräsin esimerkkilauseita suoraan puhujilta. Kielten otos löytyy Taulukosta 1.

Taulukko 1. Analyysissa käytetty kielten otos (Area)

Kieli	Kielikunta, suku
Albania	Indoeurooppalainen, isolaatti
Baski	Isolaatti
Englanti	Indoeurooppalainen, germaaninen
Viro	Uralilainen, itämerensuomalainen
Suomi	Uralilainen, itämerensuomalainen
Saksa	Indoeurooppalainen, germaaninen
Kreikka	Indoeurooppalainen, helleeninen
Unkari	Uralilainen, ugrilainen
Italia	Indoeurooppalainen, romaaninen
Latvia	Indoeurooppalainen, balttilainen
Malta	Afroaasialainen, semiittinen
Venäjä	Indoeurooppalainen, slaavilainen
Sardi	Indoeurooppalainen, romaaninen

Suomen kieleen keskittyvissä osatutkimuksissa (QUANTI, WEB ja PARTEX) keräsin materiaalia verkkokorpuksista. Materiaalin analysoinnissa käytin määrällisiä menetelmiä.

Osatutkimuksessa QUANTI käytin korpusta *Kansalliskirjaston lehtikokoelman (KLK) suomenkieliset lehdet*, josta otin 779 satunnaista esimerkkilauseetta, jotka sisältävät *löytyä*-verbin. Analysoin otosta tilastollisin menetelmin, logistisen regressiosekamallin avulla. Menetelmä mahdollistaa nominatiivi- ja partitiivi-subjektien vaihtelun tarkastelun *löytyä*-verbillisissä lauseissa riippumattomien muuttujien suhteen niin, että muuttujien keskinäiset vaikutukset voidaan vakioda (esim. sanajärjestys, tempus, subjektin lemma jne.).

Osatutkimuksessa WEB käytin Sketch Engine -alustalta otettua korpusta Finnish Web 2014 (fiTenTen2014). Materiaali koostuu pienestä otoksesta, joka sisältää 449 *löytyä*-verbillistä lausetta, joissa verbi on lokuforisessa muodossa eli 1. ja 2. kieliopillisessa persoonassa.

Osatutkimuksessa PARTEX käytin verkkokorpusta Suomi24:2017, joka koostuu vuonna 2017 Suomi24-foorumille kirjoitetuista viesteistä. Tästä korpuksesta keräsin esimerkkilauseita 10:llä eri eksistentiaalisella verbillä (yhteensä 10 000 lausetta, 1000 per verbi), joita tarkastelin kollostruktioanalyysin avulla. Kollostruktioanalyysi (Stefanowitsch ja Gries 2003; Gries ja Stefanowitsch 2004) on menetelmä, joka mahdollistaa kielen konstruktioiden (tässä: eksistentiaali-lauseiden) ja niiden osana esiintyvien sanojen (tässä: eksistentiaalisten verbien)

suhteiden tarkastelun. Menetelmällä tutkitaan, kuinka tiiviisti konstruktiot ja sanat ovat yhteydessä toisiinsa vertaamalla niiden yhteisesiintymistä konstruktion ja kunkin sanan kokonaisesiintymiseen. Artikkelissa esittelen myös uuden otantamenetelmän, jota kutsun nimellä *Expected Sample size* (ES). Menetelmä pyrkii ottamaan huomioon yksittäisten verbien esiintymistiheydet koko korpuseseen verraten ja vertailemaan niitä 1000 lauseotoksesta saatuihin esiintymistiheksiin.

Tulokset

Invenitiivis-lokationaaliset konstruktiot ovat yleisiä Euroopan kielissä, kuten AREA näyttää. Invenitiivis-lokationaalisia konstruktioita luonnehtii invenitiivinen verbi, jolla on seuraavat ominaisuudet:

I. sen sanavartalon perusmerkitys on löytää ja se on *semanttisesti haalistunut* (engl. *semantically bleached*);

II. se sisältää valenssia tai pääluokkaa muuttavan morfologisen tai perifrastisen elementin;

III. se ilmaisee lokationaalista funktiota ilman erityisen asennon semanttista merkitsemistä.

Invenitiivis-lokationaalisia konstruktioita on kahta tyyppiä, jotka eroavat toisistaan sen mukaan, onko paikassa sijaitseva elementti (eli *locatum*, ks. Haspelmath 2022) elollinen vai eloton (Taulukko 2).

Taulukko 2 – Invenitiivis-lokationaalisten konstruktioiden tyypit (AREA)

	Tyyppi 1: elollinen locatum	Tyyppi 2: eloton locatum
Kriteeri 1: lokatiivisen lausekkeen konkreettisuus	konkreettinen/abstrakti	konkreettinen
Kriteeri 2: verbin indeksi	lokuforiset ja alloforiset muodot (kielispesifiset eroavuudet)	alloforiset muodot
Kielispesifiset alatyypit: miratiivisuus	miratiivinen merkitys mahdollinen (riippuu konstruktiosta)	N/A

Suomen kielen invenitiivisiä konstruktioita, jotka perustuvat *löytyä*-verbiin, käytetään monesti paikkaa ilmaisevina kopulina. Yleensä näissä suomenkielisissä konstruktioissa ei käytetä alloforisia muotoja, ja kun tarkoite on elollinen, niitä käytetään ainoastaan ihmis- ja eläinryhmien eikä yksittäisten tarkoitteiden kanssa. Yksittäisten ihmis- tai eläintarkoitteiden käyttäminen korreloi miratiivisen merkityksen kanssa ja ilmenee toisen analyttisen konstruktion kautta (*löytää itsensä*). *Löytyä*-verbillisissä konstruktioissa *locatum* voi esiintyä nominatiivissa tai

partitiivissa, josta riippuu eroavuus lokatiivisen (1a) ja eksistentiaalisen (1b) konstruktion välillä.

(1a) *Olut löytyy jääkaapista.*

(1b) *Jääkaapista löytyy olutta.*

QUANTI selvittää *löytyä*-verbillisten konstruktioiden eksistentiaalisten subjektien nominatiivi- ja partitiivivaihtelua tilastollisin menetelmin. Artikkelissa käy ilmi, että *löytyä* on laajasti käytetty sekä lokatiivisissa että eksistentiaalisissa konstruktioissa ja että sitä käytetään samalla tavalla kuin kopulaa *olla*. Eri muuttajat vaikuttavat nominatiivin ja partitiivin valintaan. Esimerkiksi subjektin laskettavuus korreloi kieliopillisen luvun kanssa ja ne yhdessä vaikuttavat eksistentiaalisen subjektin sijamerkintään esimerkiksi niin, että ei-laskettavat monikolliset subjektit lisäävät partitiivin todennäköisyyttä. Myös kongruenssi ja sanajärjestys korreloivat subjektin sijamerkinnän kanssa niin, että partitiivin todennäköisyys lisääntyy silloin, kun kongruenssia ei ole. Sanaluokan ja sanajärjestyksen välinen suhde ei sen sijaan nouse esiin subjektin sijamerkintään vaikuttavana tekijänä. Verbin tempuksen vaikutus nousee niin ikään esiin, mitä väitettä tukee se huomio, että *löytyä*-verbillä on yleensä kopulan merkitys preesensissä, kun taas imperfektissä leksikaalinen merkitys säilyy ja subjekti esiintyy tyypillisesti nominatiivimuotoisena. Subjektin lemmalla on silläkin merkitystä: tietyt lemmat (kuten *ruumis*) korreloivat nominatiivin kanssa, sillä nominatiivi kuvailee yleensä punktuaalisia tapahtumia, niin kuin ruumiin löytyminen jostakin paikasta.

WEB tukee QUANTI-artikkelin havaintoa siitä, että *löytyä*-verbin lokuforiset muodot ovat marginaalisia. Nimenomaan koko korpuksessa on ainoastaan 449 lokufoorista esiintymistä, joita käytetään vain tietyissä konteksteissa. Näillä esiintymisillä on useimmiten erityinen pragmaattinen funktio, sillä lokuforisten muotojen kautta puhujat mainostavat omaa läsnäoloaan internetissä tai jollakin tietyllä sivustolla. Lokuforisten muotojen käyttö korreloi miratiivisten käyttöjen kanssa, joilla ilmaistaan puhujien tahatonta osallistumista tiettyyn toimintaan tai yllättävää/odottamatonta sijaintia tietyssä paikassa.

PARTEX tutkii suomen kielen eksistentiaalisia konstruktioita, joissa subjekti esiintyy partitiivissa. Artikkelit vahvistaa sen havainnon, että *löytyä*-verbi esiintyy partitiivisubjektillisissa eksistentiaalisissa konstruktioissa enemmän kuin kaikki muut eksistentiaaliset verbit. Muita usein käytettyjä eksistentiaalisia verbejä ovat *mahtua*, *riittää* ja *sataa*. Tulokset vahvistavat oletuksen, että *löytyä*-verbiä käytetään paikkaa ilmaisevana kopulana samalla tavalla kuin *olla*-kopulaa.

Yhteenveto

Käsillä oleva väitöskirja esittää uuden vertailtavissa olevan käsitteen, jota voidaan käyttää tulevaisuudessa kielitieteellisissä tutkimuksissa. Invenitiivis-lokationaaliset konstruktiot ovat ns. konstruktio-strategioita, sillä niillä on kyllä funktionaalinen ulottuvuus, mutta ne määritellään enimmäkseen muodollisten ominaisuuksiensa

pohjalta. *löytyä*-tyyppiset verbit lokatiivisessa ja eksistentiaalisessa funktiossa ovat Euroopan kielissä laajasti käytettyjä, ja niiden tällaista käyttöä voidaan pitää alueellisena ilmiönä. On tärkeä ymmärtää, että inventiivis-lokationaalisten konstruktioiden määrittely antaa kielitieteilijöille mahdollisuuden poistua alueellisen kielitieteen lähestymistavoista ja tutkia *löytyä*-konstruktioita pikemminkin kielitypologisesta näkökulmasta. Tämä tarkoittaa sitä, että inventiivis-lokationaalisia konstruktioita määrittelevien kriteerien pohjalta on mahdollista selittää, onko sellaisia konstruktioita olemassa vaikkapa Australian tai Etelä-Amerikan kielissä.

Alustavat tulokset osoittavat, että inventiivis-lokationaalisia konstruktioita on olemassa useissa Afrikan kielissä, kuten mandelaisissa kielissä, ja Filippiineillä puhutussa tagalogin kielessä, jossa inventiivis-lokationaalisten konstruktioiden käyttö saattaa olla espanjan ja englannin kielen kanssa tapahtuneen pitkäaikaisen kontaktin tulosta. Myöskin mandelaisissa kielissä niiden käyttö saattaa olla ranskan kielen kanssa tapahtuneen kontaktin tulosta.

Tulevissa tutkimuksissa voisi olla kiinnostava kysyä, kieliopillistuvatko inventiivis-lokationaaliset konstruktiot ja miten ne tekevät sen. Ruotsin kieli antaa hyvän esimerkin tällaisesta mahdollisuudesta, sillä eksistentiaalinen konstruktio *det finns* on kieliopillistunut verbistä *finna* 'löytää'. Koska ruotsi on vaikuttanut suomen kieleen pitkään, on mahdollista, että *löytyä*-verbin polyseeminen käyttö johtuu juuri siitä eksistentiaalisesta käytöstä ruotsin kielessä ja että tulevaisuudessa samankaltainen käyttö kieliopillistuu erityiseksi lokatiiviseksi ja eksistentiaalisiksi kopulaksi.

PUBLICATIONS

Basile, Rodolfo (Accepted for publication).
Invenitive-locational constructions in the languages of Europe.

Inventive-locational constructions in the languages of Europe

ABSTRACT

This paper presents a novel class of construction-strategies called inventive-locational constructions. These constructions make use of a verb root with meaning FIND, which becomes semantically bleached and behaves like a locational copula. Inventive-locational constructions functionally express predlocative and existential meaning without semantically marking a specific posture. The constructions at hand are defined at the beginning of the article and then discussed with a focus on a subset of the languages of Europe that includes not only Indo-European languages, but also Maltese, Basque, and three Uralic languages (Finnish, Estonian, and Hungarian). The paper also discusses the semantic and morphosyntactic features of inventive constructions. The main result of this paper is that inventive-locational constructions seem to be grouped into two types, according to whether the located element (here: locatum) is animate or inanimate.

Keywords: locational constructions, existential constructions, predlocative constructions, typology, inventive

1 INTRODUCTION

Much of the work done on grammar during the last decades refers to linguistic structures as CONSTRUCTIONS. Each construction has a specific function and/or form that, according to frameworks such as (Cognitive) Construction Grammar¹ (Lakoff 1987; Goldberg 1995), can be deduced from analyzing the construction as a whole. In other words, the meaning or the form of each construction is not inferable from the meanings or forms of its compositional parts. This understanding of constructions has become mainstream, and it has been adopted in typological studies. Works like Haspelmath (2022), while not explicitly defining the concept of construction or referring to Construction Grammar, use the term *construction* profusely as a synonym of CLAUSE CONSTRUCTION or CLAUSE TYPE, referring to syntactic configurations that have a specific function.

The aim of this paper is to present a class of locational clause constructions that I call INVENTIVE-LOCATIONAL CONSTRUCTIONS. This term is divided into two parts: the first part indicates the form or the strategy used (INVENTIVE), and

¹ Variants of Construction Grammar include Radical Construction Grammar (Croft 2001; Barðdal 2006) and Diachronic Construction Grammar (Traugott and Trousdale 2013).

the second part refers to the construction's function (LOCATIONAL). Inventive-locational constructions can also be called construction-strategies, which means that they are formally defined but functionally express locational meaning (see Haspelmath 2021; Croft 2022). The term *inventive* comes from the Latin verb *invenire* 'to find', because these strategies are all based on verb roots that primarily have the meaning FIND. For the verbs used in these constructions, I will use the term *inventive verbs*. Inventive-locational constructions are common in European languages (1).

- (1) Italian (personal knowledge)²
Il gatto si trova sull' albero
 DEF cat MID.3SG find.3SG on.DEF tree
 'The cat is in the tree.'

The example above expresses a PREDLOCATIVE CONSTRUCTION (Haspelmath 2022). Predlocative constructions have also been known as locative constructions or locative predications, terms that according to Haspelmath do not suffice for designating these clause types, since existential constructions also express location. Because these constructions feature a definite locatum (i.e., the located element), it becomes apparent that the other part of the clause construction is a predicate, more specifically a predicative locative phrase (hence, *predicative locative* or *predlocative*). For the sake of clarity, (1) can be compared with the examples below, which instead respectively express a predlocative construction with a locati copula (2a), a predlocative construction with a posture verb phrase (2b), and an EXISTENTIAL CONSTRUCTION (2c). Existential constructions differ from predlocative constructions in that they introduce a discourse-new, indefinite locatum, and hence they do not involve a predicate (Haspelmath 2022).

- (2) Italian
 a. *Il gatto è sull' albero*
 DEF cat be.3SG on.DEF tree
 'The cat is in the tree.'
 b. *Il gatto sta sdraiato sull' albero*
 DEF cat stay.3SG lie.PST.PTCP on.DEF tree
 'The cat is lying in the tree.'
 c. *C'è un gatto sull' albero*
 EXV INDEF cat on.DEF tree
 'There is a cat in the tree.'

Just like in (2a), the verb in (1) does not add any semantic information to the cat being in the tree, but simply states that the cat is there (cf. 2b, where the cat has

² All the Italian examples are the author's.

a specific posture). The verb in (1) becomes semantically bleached, to some extent, of its original lexical meaning FIND, consequently expressing a locational function as its primary meaning. This locational function can hence be said to be coexpressed by a verb with meaning FIND, that elsewhere would express its original meaning (for *coexpression patterns*, see Haspelmath 2019; Geniušienė 1987). For these reasons, such strategies may be considered locational copulas, i.e., specialized copulas featured in either a predlocative or an existential construction. Inventive-locational constructions can be defined as being characterized by what I call an *inventive verb*, which satisfies two main criteria:

- I. it has a root with meaning FIND which undergoes semantic bleaching;
- II. it features a morphological or periphrastic valency- or voice-changing marker.

In addition, in an inventive-locational construction, an inventive verb:

- III. expresses a locational function without semantically marking a specific posture.

The third criterion specifies one of the possible functions inventive verbs can have, which is that of conveying a relation between two syntactic arguments, one being the located element and the other a locative phrase, in a nonverbal predlocative or existential clause construction. However, it is worth noting that the third criterion can change, based on what kind of clause constructions these inventive verbs appear in. An example is (3).

- (3) Italian
- | | | | | |
|----------|------------|-----------|-----------------|---------------|
| <i>I</i> | <i>due</i> | <i>si</i> | <i>trov-ano</i> | <i>(bene)</i> |
| DEF.PL | two | MM.3PL | find-3PL | (well) |
- ‘The two get along.’

The verb used in the Italian example above qualifies as inventive: it is both semantically bleached (i.e., no actual finding is happening as another meaning takes over, criterion I) and features a middle marker that here expresses reciprocal meaning (criterion II). However, it does not express a locational function; Hence, criterion III becomes instead:

- III. expresses the meaning GET_ALONG.

In their current formulation, the definitional criteria of inventive verbs can be divided in two parts: A formal part, represented by the first two criteria, and a functional part, represented by the third criterion. The definition takes into account the possibility of expressing a variety of meanings by singling out the third criterion, while the first two remain unchanged. Other examples are given in (4).

- (4) a. Italian
Mi **trovo** *male* (*a lavoro*)
 MM.1SG find.1SG badly at work
 ‘I am having a bad/difficult time (at work).’
- b. Spanish (personal knowledge)
Me **encuentro** *mal*
 MM.1SG find.1SG badly
 ‘I don’t feel well.’
- c. English (Jay Zameska, p.c.)
They finally found each other.

The Spanish inventive verb in (4b) expresses a different, yet related, meaning compared to the Italian inventive verb in (4a) even if they are both used with the same adverb. Furthermore, it is not uncommon to hear utterances such as (4c) for expressing the meaning FALL IN LOVE. This latter meaning may be directly connected to the idea of concretely finding a romantic partner and may work only in certain contexts, but it contributes to showing why it could be worth looking into the polysemy of FIND-based strategies more in detail, especially when they present some kind of morphosyntactic marking, even though the cases above may only be marginal. Within the paper at hand, I will focus only on those inventives that express a locational function. I will show that inventive-locational constructions are common in European languages, and that they are arranged according to two main types, based on the animacy of the locatum (i.e., the located element). The sample used for this investigation consists of elicited utterances collected for 13 different European languages.

In Section 2, I will provide background information about nonverbal and locational predication, as well as some other key concepts. Following this, in Section 3, I will comment on the definition proposed above and present some borderline cases of inventive constructions. Section 4 will be devoted to data and methodology, and Section 5 will report the main results. The paper ends with a discussion in Section 6.

2 KEY CONCEPTS

This section provides an overview of some key concepts that are relevant for the analysis of inventive-locational constructions. First, I will introduce the general theoretical framework and the work done on locational constructions, since inventive-locational constructions represent a subset of them. Second, I will elaborate on three important factors that are used to analyze inventive-locational constructions: indexing, middle markers, and mirativity.

2.1 Locational constructions

Locational constructions can be identified as subtypes of what Hengeveld (1992) calls *non-verbal predication*, which refers to all constructions that have a non-verbal main predicate (see also Roy 2013). Because of the difference between morphosyntactic and semantic analyses, Hengeveld specifies that non-verbal is not a synonym of nominal or verbless, hence instances of non-verbal predication can also correspond to verbal sentences, i.e. copula constructions (Hengeveld 1992: 26). Non-verbal predication is also known under the labels *nonprototypical predication* (Croft 2022: 289), or *copular clauses* (Declerck 1988; Mikkelsen 2011). In Stassen's (1997: 55–61) terminology, it is called *intransitive predication*, which includes what he calls *locational strategy* (see also Heine 1997). In Haspelmath (2022) *nonverbal clause constructions*, in line with Hengeveld's (1992) definition, are said to lack a typical verb, i.e. they can feature a copula. Within the seven types of nonverbal clause constructions proposed by Haspelmath, two of them, i.e. predlocative constructions and existential constructions, can be subsumed by what has also been called *locational predication* (Creissels 2019), *locative paradigm* (Freeze 1992: 554), or *location clauses* (Croft 2022: 304). Predlocative constructions have also been referred to as *predicational location* (Croft 2022: 304), while existential constructions are also called *presentational constructions* (Croft 2022: 304). For some types of what are traditionally considered instances of existential predication, Creissels (2019) uses the term *inverse-locational predication*.

The problem with Croft's term presentational construction is that it can be confused with what Gast and Haas (2011) call *presentationals* (see also Lambrecht 1994), which represent a formal structure where a discourse-new subject is introduced. This type of configuration does not, however, usually have a locational function. Creissel's term inverse-locational predication is contrasted with *plain-locational predication* and assumes the unmarkedness of the latter, which is encoded by what he calls the prototypical *figure-ground relationship*, i.e. the perspectivization 'figure>ground'. Inverse-locational predication expresses instead a marked perspectivization 'ground>figure' (see also Koch 2012; McNally 2016; Creissels 2014; Huumo 2003). These semantic concepts date back to Talmy (1983; 2000: 311): the figure, also called *pivot* (Milsark 1977; Bentley et al. 2013), is a rather movable entity and the ground (in some traditions called *coda*, see Bentley et al. 2013) is typically fixed and unmovable. Haspelmath (2022) does not speak of perspectivization but refers to the pivot/figure as the *locatum* in predlocative constructions. The locatum is a definite argument and is said to be in a place expressed by a locative phrase. In existential constructions, the locatum is also called the *existent* and represents a discourse-new, indefinite nominal phrase. According to Haspelmath (2022), existential constructions obligatorily include a

locative phrase³. Existential constructions can also feature a specialized *existive* (Haspelmath 2022), i.e., a restricted copula like Tagalog *may* or the Italian construction *c'è*, which contains the expletive locational element *ci* (sometimes referred to as a *dummy*, like English *there* in *there is/are* or Swedish *det* in *det finns* ‘there is/are’) and a copula. In the paper at hand, I will use the terms *predlocative construction* and *existential construction* as two separate functional instances of locational constructions, similarly to Haspelmath (2022).

2.2 Indexing, middle markers, and mirativity

As I will show in Section 5, types of inventive constructions are also characterized based on the indexing form of the inventive verb, namely whether it can appear indexed for all persons or not. For this reason, I will adopt the terminology in Haspelmath’s (2013) article about argument indexing, which distinguishes two types of person forms: speech-role forms, or *locuphoric* forms, that refer to both speaker (grammatically encoded as first persons) and addressee (second persons), and *allophoric* forms (Dahl 2000), that refer to non-speech-role referents (i.e., third persons). Because of the intrinsically animate nature of first and second persons, we can establish a direct causal relation between animacy and locuphoric forms. Allophoric forms, on the contrary, do not correlate directly with animacy because they do not index speech-role forms. These distinctions will become relevant in the typology of inventive constructions since several languages show different morphosyntactic strategies according to the animacy of the referent and the indexing form of the inventive verb.

Inventive verbs are often morphologically or analytically marked via a middle marker. I will adopt the definition of middle markers (MMs) proposed by Inglese (2022), whose first point is that a MM “occurs with bivalent (or more) verbs to encode one or more of the following valency changing operations: passive, anticausative, reflexive, reciprocal, antipassive”. These are all valency-reducing operations (Inglese 2021, 2022; see also Kemmer 1993; Kulikov 2013), which confirms the fact that MMs are hence “inherently polyfunctional constructions”. The underlying idea is that valency-reducing markers are polyfunctional, unlike valency increasing markers (Nichols et al. 2004: 175; Bahrt 2021: 161; Inglese 2021). The manipulation of verb valency by MMs takes place in two ways, as brought to our attention by Inglese (2022): on the one hand, reflexives and reciprocals operate on the semantic structure of a two-participant event (see also Givón 2001: 95–116), on the other hand, passives and antipassives operate only

³ Existential constructions without a locative phrase are here called *hyparctic* (see also Section 3).

on syntactic valency (Zúniga and Kittilä 2019). The same middle marker can co-express several voices, as pointed out by studies focused on passive syncretism⁴ (Haspelmath 1990) or reflexive syncretism (Geniusiene 1987). These two types of voice syncretism are central in what has also been called middle syncretism (Bahrt 2021: 49). For example, the Russian suffix *-sya* encodes reflexive, passive, reciprocal and anticausative functions, plus an antipassive function (Bahrt 2021: 140) that does not belong to the meanings expressed by middle syncretism. The Italian middle-voice marking particle *si* can coexpress, among others, reflexive, anticausative, reciprocal, and antipassive functions, but it is also used in specific medio-transitive constructions (Masini 2012) or in impersonal constructions. As I will discuss further, determining whether the particle *si* expresses a middle/reflexive function or an impersonal one is usually problematic and highly context-dependent. For simplicity, I will use the term *middle marker* (glossed MM) as a convenience term for all the coexpressed functions.

To understand some of the subtypes of inventive constructions presented below, it is worth introducing the concept of *mirativity*. Mirativity is a linguistic category that has been first defined by DeLancey (1997; 2001) as conveying information that is new or unexpected to the speaker. Despite its similarity with another category, evidentiality, mirativity is considered as a separate linguistic category. In fact, even though sometimes mirative markers are coexpressed by evidential markers, it has been shown that in some languages the two categories can co-occur, occupying different slots in the verb structure, and even relate differently to other categories (Aikhenvald 2012: 436). Moreover, Hengeveld and Olbertz (2012: 488) add that the information expressed by mirative markers can be newsworthy, unexpected or surprising not only for the speaker, but also for the addressee. Some subtypes of inventive constructions show mirative readings, meaning that they present information that is unexpected or surprising for both participants of the speech act. These readings also seem to correlate with the degree of control that an animate speaker has over the situation and can be grammatically expressed by specialized strategies.

2.3 On the grammaticalization of inventives

Grammaticalization is a process that involves the transformation of lexical material into grammatical material, and of grammatical material into more grammatical material. An overview is given by Heine (2017). The overlap model (Heine 1993: 48–53) involves three stages: 1. An expression is recruited for grammaticalization; 2. Such expression acquires a second use pattern, which results in ambiguity between the two; 3. The first expression is lost, resulting in

⁴ The terms *coexpression* and *voice syncretism* can be considered synonymous in this context.

the second expression taking over. It is important to keep in mind that not all instances of grammaticalization go fully through all three stages but may stop at stage 2. This is also the case of inventive-locational constructions, which have not completely replaced other uses of FIND-based strategies.

An important mechanism of grammaticalization to be kept in mind for the definition of inventives is desemantization, or semantic bleaching. During this phase, the expression under investigation loses its original meaning. In our case, a verb with meaning FIND loses its lexical meaning and starts to be used in new ways. This mechanism is called extension, or context generalization. Such context generalization does not necessarily imply that the use of the same linguistic material in other contexts be lost (see stage 2 above). Grammaticalization has also been argued to require appropriate contexts to take place, which contributes to both expanding the contexts in which it can take place and to increase the frequency of use of the grammaticalized item. An example of what appears to be a fully grammaticalized inventive strategy is given by the Swedish Existential construction *det finns*. In this construction, the verb *att finna* ‘to find’ presents a middle marking morpheme *-s*, which is used productively in Swedish for several functions (e.g., passive, reflexive). The presence of the expletive *det* is typical of existential constructions (see, e.g., English *there is*).

3 FEATURES OF INVENTIVE-LOCATIONAL CONSTRUCTIONS AND BORDERLINE CASES

This section introduces first the pragmatic and usage features on inventives based on Italian data, then the typical use of inventives in predlocative and existential constructions, comparing them to copulas. Furthermore, it elaborates on one borderline case: inventive-hyparctic⁵ constructions.

3.1 Some pragmatic and usage remarks on inventives in Italian

To understand how inventive-locational constructions behave, I will here present the case of Italian, where the verb *trovare* ‘to find’ expresses a variety of functions depending on its morphosyntactic marking, the number of its arguments, the semantic features of these and, ultimately, context.

- (5a) *Trovate tutto l'occorrente in questa scatola.*
 find.2PL all DEF-necessary in this box
 ‘You will find all the necessary in this box.’

⁵ Following Haspelmath (2022), hyparctic construction are constructions that express “pure” existence, and do not have a locational element.

- (5b) *Tutto l'occorrente è in questa scatola.*
 all DEF-necessary be.3SG in this box
 'All the necessary is in this box.'
- (5c) *Tutto l'occorrente si trova in questa scatola.*
 all DEF-necessary MM find.3SG in this box
 'All the necessary is found in this box.'
- (6) *Pierpaolo si trova sull'albero.*
 P. MM find.3SG on.DEF-tree
 'Pierpaolo is on the tree.'
- (7) *I due si trovano.*
 DEF.PL two MM find.3PL
 'Those two get along.'
- (8) *Mi trovo bene qui.*
 MM.1SG find.1SG well here
 'I am comfortable here.'

In (5a), the verb 'find' does not present any special marking, and it could be argued that it is not interpretable in any other way than in the one provided by the verb's original meaning. However, from a pragmatic perspective, it is clear that there exist more neutral ways of conveying the same piece of information, e.g. by using a copula (5b). Instead, (5a) specifically addresses a generic addressee, who is then given instructions about the whereabouts of something. Moreover, it encodes pragmatic presupposition, in that what is assumed to be part of the common ground is that the addressee is supposed to be looking for the contents of the box, whereas by employing a regular BE-copula this feature is not overt. Similarly, (5c) may also encode pragmatic presupposition, albeit through a different grammatical strategy.

The above-mentioned literature on MMs calls for a reanalysis of constructions like the one expressed in (6) since what appears to be a full-fledged verb phrase is in fact a locational copula. The Italian verb *trovare* 'to find' is bivalent, in that it encodes both an agent (commonly referred to as A) and a patient (P). When associated with the MM *si*⁶, it changes its argument structure and mainly expresses anticausative, reflexive (including grooming events, such as 'to shave'), or reciprocal meaning, but also what Inglese (2022) calls impersonal agentless passive and what Masini (2012: 11) calls medio-transitive⁷. I refer to Inglese (2021: § 2.1) for the whole range of functions encoded by the Italian MM *si*. However, there is one specific function that the literature has so far not accounted

⁶ Also called Italian Reflexive *si*. Language-specific concepts are indicated with capital letters and are to be distinguished from comparative concepts (see e.g. Haspelmath X).

⁷ It. "costruzione media transitiva".

for, namely the expression of locational meaning through the perhaps particular case provided by the combination of the verb *trovare* ‘to find’ and the MM *si*, which yields the intransitive *trovarsi* ‘find.MM’. This strategy conveys a range of possible readings, represented by the following constructions:

- a. A finds P in X, where A and P are coreferential (REFLEXIVE);
- b. A’ find P in X, where A’ refers to a generic referent (IMPERSONAL);
- c. S finds.MM in X, where the verb undergoes semantic bleaching (LOCATIONAL).

In these constructions, native speaker intuition is often misleading. Because the MM *si* is so multifunctional, it is difficult to determine whether we can talk about one or the other construction. Possibly because educated Italian speakers often refer to the MM *si* as a Reflexive marker, this interpretation is usually considered the most plausible, even though this term only refers to the formal characteristics of the marker, not necessarily to its function and meaning. This is not to say that Italian *trovarsi* can never be interpreted in a reflexive sense: It can inasmuch as it is representing specific constructions in which it is clear that (i) the meaning FIND is maintained, and (ii) the event encoded is a proper reflexive event, where there is coreferentiality between A and P (Zúñiga and Kittilä 2019: 154; see also Haspelmath 2023). However, most of the Italian examples are up to debate because it is unclear whether the meaning FIND is maintained. Moreover, such middle mark forms in (6), if interpreted transitively, could more likely convey impersonal meaning rather than reflexive meaning (people are looking for Pierpaolo > he can be found on the tree). Perhaps what comes closer to a reflexive construction is a construction involving the intransitive verb *trovare* ‘to find’ and an analytical reflexive nominal (9, cf. 10 – for *reflexive nominal* see Haspelmath 2023).

- (9) *Pierpaolo trova se stesso (sull'albero).*
 P. find.3SG REFL self (on.DEF-tree)
 ‘Pierpaolo finds himself (on the tree).’

- (10) *I don't know why I find myself tearing up nearly every day now.* (CoCA)

In the English example above there are three clear clues as to why we should consider it as a reflexive construction rather than something else. First, English is fundamentally different from Italian in that it uses the analytical pronoun *myself* as a reflexive marker, and it is undoubted that here the action of finding is being performed on the same referent that is performing it (coreferentiality). Second, the verb ‘find’ marks an inherently mirative event. Finding something (and, by extension, oneself) involves a lower degree of control and intention compared to e.g., eating, sending, or looking for something. Thus, mirative may play an

inherently central role in utterances where the inventive verb is not completely semantically bleached, i.e., utterances where it somehow retains its original meaning. Third, the context helps confirming the low control of the action, as the speaker specifies ‘*I don’t know why*’. In (6) it could be the case that the referent was not in full control of climbing the tree either, but this interpretation must be left aside as there are no context clues or pragmatic presupposition devices that mark whether this is true or not. Two unambiguously mirative examples that contrast (6) are given below (11).

- (11a) *Pierpaolo si trova sull'albero a sua insaputa.*
 P. MM find.3SG on.DEF-tree [unbeknownst to him]
 ‘Pierpaolo finds himself on the tree, unbeknownst to himself.’
- (11b) *Pierpaolo si ritrova sull'albero.*
 P. MM MIR.find.3SG on.DEF-tree
 ‘Pierpaolo finds himself on the tree.’

While (11a) uses an overt periphrastic device to specify the unwillingness of Pierpaolo being on the tree, (11b) uses a morphological one. I argue here that the prefix *ri-* adds unambiguous mirative meaning to this verb based on my intuition, but further research is needed to assess whether this is an intrinsic characteristic of this polyfunctional morpheme or just a lexical feature of the verb *ritrovarsi* ‘to end up/find oneself at/doing’ (an analysis of Italian *ri-* is Cardinaletti 2003).

3.2 Inventives as copulas

Inventive strategies can equally express both predlocative and existential constructions, as shown respectively by (12a) and (12b), although they more commonly express predlocative constructions.

- (12) Russian (Anna Branets, p.c.)
- a. *Еда находит-ся на столе*
Eda nachod-it-sya na stole
 food find-3SG-MID on table.PREP
 ‘The food is on the table.’
- b. *На столе находит-ся еда*
Na stole nachod-it-sya eda
 on table.PREP find-3SG-MID food
 ‘There is food on the table.’

Because inventive-locational constructions are coexpressed by a verb root that usually conveys its lexical meaning, semantic bleaching is a central criterion in their definition. The verb with meaning FIND becomes a sort of copula. According

to Haspelmath (2022), common copulas indicate “a stative link between the two argument positions of an equational, ascriptive or locational clause (i.e., a predlocative or existential clause)”. Inventives do indicate this stative link but do so in a morphologically or periphrastically marked way, for example by employing a grammatical strategy that falls within the middle-voice spectrum. We can then say that inventives are special copulas that have a specific morpho-syntactic contour. Because they are copulas that express a locational function, instances of verbal predication, in which the verb retains its original meaning FIND, do not belong to inventive-locational constructions. In some cases, there is a blurry demarcation line between verbal and non-verbal predication, which makes some instances of inventive-locational constructions difficult to identify. This problem arises from the fact that middle forms, in languages like Italian, can coexpress several functions, such as impersonals (13a; see also § 3.1).

(13) Italian

- a. *Le patate si trovano al mercato?*
 DEF.PL potato.PL MM find.3PL at.DEF market
 ‘Can one find potatoes at the market?’
- b. *Le patate ci sono al mercato?*
 DEF.PL potato.PL EXPL be.3PL at. DEF market
 ‘Are there potatoes at the market?’
- c. *Le patate si trovano sottoterra*
 DEF.PL potato.PL MM find.3PL underground
 ‘Potatoes are located underground.’
- d. **Le patate si ritrovano sottoterra*
 DEF.PL potato.PL MM MIR.find.3PL underground
 * ‘Potatoes end up underground.’

The verb in (13a) does not unambiguously express a middle voice and seems to presuppose an actual search for the referent indicated by the subject, possibly also because the construction is in the form of a question. The example in (13a) would also be problematic in the affirmative form if the utterance had a prosodic emphasis on *si trovano* (personal knowledge). This may indicate that the verb should be interpreted as expressing its original meaning, and the sentence as an instance of verbal predication. However, (13a) is semantically very similar to (13b), which is instead considered a locational construction, hence nonverbal. Conversely, in (13c) the semantic bleaching is more evident, and the sentence can be unambiguously characterized as an inventive-locational construction⁸. The

⁸ It is worth pointing out that, like in (13a), here suprasegmental features such as prosodic marking may come into play as well. A prosodic emphasis on *si trovano* in (13c) would also

main reason is pragmatic: While it may not be obvious to the interlocutor that potatoes can be found at the market (because maybe they are not available at a certain point in time), it is common knowledge that potatoes grow underground. Hence, it is more likely that the interlocutor be not implying an actual search for the potatoes underground, and rather that the speaker be stating the obvious. This reasoning is in line with semantic nuances like the permanent (against the temporary) presence of the referent in a specific location. This difference in usage might be relevant in other languages, and it requires further investigation. (13d) shows that the verb *ritrovarsi* ‘to end up (in/doing sth)’ only works with animate referents, since it encodes a mirative/non volitional reading.

3.3 Usage of inventives in hyparctic and possessive constructions in Finnish

The term *hyparctic* has first been suggested by Haspelmath (2022) to indicate all those nonverbal clause constructions that express “pure existence”, without expressing a locational function. Hyparctic constructions are often coexpressed by morphosyntactic devices like the existives that are often also employed in existential sentences (14).

- (14) Italian
Dio c’è
 God EXV
 ‘God exists.’

Similarly, inventive verbs may appear, at least in Finnish, in hyparctic constructions instead of the copula *olla* ‘to be’. They hence fulfill all the formal criteria that define inventive constructions, except for conveying a locational function (15a). Further evidence for why inventives in Finnish are to be considered copulas is given by their usage in possessive constructions (15b).

- (15) Finnish
- | | | | |
|----|--|----------------|-----------------|
| a. | <i>Ero-j-a-kin</i> | <i>toki</i> | <i>löyt-y-y</i> |
| | difference-PL-PTV-ENCL | certainly | find-REFL-3SG |
| | ‘There are certainly also differences.’ (Basile and Ivaska 2021) | | |
| b. | <i>Liikuntataustaa</i> | <i>minulla</i> | <i>löyt-y-y</i> |
| | exercise=background.PTV | 1SG.ADE | find-REFL-3SG |
| | ‘I have experience in exercising.’ | | |

entail a change in the utterance’s meaning towards an impersonal interpretation. Similarly, an unmarked interrogative *Le patate si trovano sottoterra?* would retain a locational meaning, while a prosodically marked version with an emphasis on the verb would likely not.

Using quantitative methods, Basile and Ivaska (2021) argue that in Finnish FIND-based strategies often replace the copula *olla* ‘to be’, typically also used in hyparctic constructions and possessive constructions. Possessive and existential constructions are often investigated in tandem (see, e.g., Haspelmath 2022; Creissels 2023), since in many languages the possessor can be morphosyntactically considered a locative phrase (in Finnish marked for adessive, one of the locative cases). While this is not the most common strategy of expressing possession in Finnish its usage is spread enough to be worthy being investigated as a construction of its own. Based on the Finnish data, it can be said that inventive verbs can be used in three different functional types of constructions: inventive-locational constructions (further divided into inventive-predlocative and inventive-existential constructions), inventive-hyparctic constructions, and inventive-possessive constructions.

4 DATA AND METHODOLOGY

The aim of this paper is to show the usage of inventive-locational constructions by analyzing first-hand data from a sample of European languages. With European languages I mean languages spoken in Europe, hence I refer not only to Indo-European languages, but also to Afro-Asiatic (Maltese) or Uralic (Finnish, Estonian, and Hungarian) languages, and to the isolate Basque. The languages analyzed are listed in **Table 1**.

Table 1 – Language sample

Language	Language family, group
Albanian	Indo-European, isolate
Basque	Isolate
English	Indo-European, Germanic
Estonian	Uralic, Finnic
Finnish	Uralic, Finnic
Greek	Indo-European, Hellenic
German	Indo-European, Germanic
Hungarian	Uralic, Ugric
Italian	Indo-European, Romance
Latvian	Indo-European, Baltic
Maltese	Afroasiatic, Semitic
Russian	Indo-European, Slavic
Sardinian	Indo-European, Romance

I will give an overview of the types of inventive-locational constructions in the languages of the sample and comment on what characterizes them. The data were either elicited from native speakers or collected on the web, and in this latter case the grammaticality of the examples was checked with native speakers and language experts. The language sample is based on availability of native speakers or experts for consultation. Moreover, it is smaller than I originally intended because in many languages inventive constructions behave the same as in the other phylogenetically related ones (e.g., Spanish and French use inventive verbs similarly to Italian and Sardinian).

5 TYPES OF INVENTIVE-LOCATIONAL CONSTRUCTIONS

In this section, I will present and analyze the two main types of inventive constructions in my sample. For each language, I have found at least one example confirming that inventive constructions are used. Although marginal in Basque, inventive constructions seem to be spread across the whole European linguistic area, which points to the possibility of an areal phenomenon. Inventive constructions are arranged according to the semantic criterion of the animacy of the referent expressed by the locatum, which establishes two types. The main features of these two types are shown in **Table 2**.

Table 2 – Inventive construction types

	Type 1: animate locatum	Type 2: inanimate locatum
Criterion 1 (semantic): concreteness, locative phrase	concrete/abstract	concrete
Criterion 2 (morphosyntactic): indexing, verb	both locuphoric and allophoric forms (language-specific differences)	allophoric forms
Language-specific subtypes: mirativity	mirative reading possible (construction specialization)	N/A

In my sample, the general tendency is that all languages feature Type 2 constructions with an inanimate locatum, a concrete locative phrase and allophoric forms, while Type 1 can also be either absent (e.g. Hungarian) or confined only to some subtypes (e.g. only with allophoric forms, see Finnish and Estonian below). As shown in the table, Type 1 has an animate locatum that can be found in a concrete or abstract/metaphorical location expressed by a locative phrase. Because the locatum is animate, the inventive verb can appear in both locuphoric (16a-b) and allophoric (9c) forms, as shown by Campidanese Sardinian.

- (16) Campidanese Sardinian (personal documentation)
- a. *M'agattu* *in* *Casteddu*
 MM.1SG-find.1SG in Cagliari
 'I am in Cagliari.'
- b. *M'agattu* *in* *una* *situazioni* *malla*
 MM.1SG-find.1SG in INDEF situation bad
 'I find myself in a bad situation.'
- c. *Su* *pisci* *si* *agattara* *in* *s'acqua*
 DEF fish MM.3SG find.3SG in DEF-water
 'The fish is in the water.'

Type 2, instead, has an inanimate locatum that can be found mainly in a concrete location expressed by a locative phrase, and the inventive verb appears only with allophoric forms (17a). Abstract locations are rare and appear to be directly correlated with locati that encode abstract referents⁹ (17b).

- (17) Italian
- a. *L'Italia* *si* *trov-a* *nel* *Mediterraneo*
 DEF-Italy MM.3SG find-3SG in.DEF Mediterranean.sea
 'Italy is situated in the Mediterranean.'
- b. *La* *felicità* *si* *trov-a* *nelle* *piccole* *cose*
 DEF happiness MM.3SG find-3SG in.DEF.PL little.PL thing.PL
 'Happiness is found in little things.'

Italian represents one of the most consistent languages with regards to inventive constructions, in that it uses the middle form not only for both types (i.e., with animate and inanimate locati), but also indiscriminately with all grammatical persons, and with both concrete and abstract locations. The same forms are also allowed, as already introduced in § 3, for mirative readings – compare (18a) with (18b) and (18c).

- (18) Italian
- a. *Ti* *trovi* *a* *casa*
 MM.2SG find.2SG at home
 'You are at home.' – non-mirative reading
- b. *Ti* *ri-trovi* *a* *casa*
 MM.2SG MIR-find.2SG at home
 'You find yourself/end up at home.' – mirative reading
- c. *Ti* *trovi* *in* *una* *situazione* *spiacevole*
 MM.2SG find.2SG in INDEF situation unpleasant
 'You find yourself in an unpleasant situation.' – mirative reading

⁹ See Basile and Ivaska (2021: 16, 33) for a few examples on Finnish.

The Italian prefix *ri-/re-* is productive and usually expresses the meaning AGAIN. If used with an inventive verb, however, it does not have any iterative meaning, at least not in the prototypical sense. Instead, it marks a mirative event. When it is absent, the only way to determine whether the utterance has a mirative reading is by looking at the pragmatic context (cf. 18a and 18c). Table 3 below summarizes the features of inventive-locational constructions in the languages under investigation.

Table 3 – Features of inventive-locational constructions in the sample

Language	Animate locatum ¹⁰	Inanimate locatum	Concrete location	Abstract location	Locuphoric forms	Allophoric forms	Dedicated mirative strategy
Albanian	(+)	(+)	(+)	(+)	(+)	(+)	(-)
Basque	(-) only mirative	(+)	(+)	(+)	(-) only mirative	(+)	(+)
English	(-) only mirative	(+)	(+)	(+)	(-) only mirative	(+)	(+)
Estonian	(+) except sg	(+)	(+)	(-)	(-)	(+)	(+)
Finnish	(+) except sg	(+)	(+)	(-)	(+) marginal	(+)	(+)
Greek	(+)	(+)	(+)	(+)	(+)	(+)	(-)
German	(+)	(+)	(+)	(+)	(+)	(+)	(-)
Hungarian	(-)	(+)	(+)	(-)	(-)	(+)	(-)
Italian	(+)	(+)	(+)	(+)	(+)	(+)	(-) ¹¹
Latvian	(+)	(+)	(+)	(+)	(+)	(+)	(-)
Maltese	(+)	(+)	(+)	(-)	(+)	(+)	(+)
Russian	(+)	(+)	(+)	(+)	(+)	(+)	(-)
Sardinian	(+)	(+)	(+)	(+)	(+)	(+)	(-)

¹⁰ In this column, Basque and English are marked (-) because they allow for an animate locatum only in their dedicated mirative strategies (e.g., *I found myself at the hospital*, **I am found at the hospital*). Similarly, Estonian and Finnish only allow animate locati if these are marked for (partitive) plural.

¹¹ Even though the verb *ritrovarsi* ‘to end up (in/doing sth)’ always has a mirative or non-volitional reading, and hence could be thought of as a dedicated strategy, Italian inventives do not in principle need the additional morphemic marking *ri-* to express mirative or non-volitional meaning. For this reason, I do not consider Italian as having a(n obligatory) dedicated mirative strategy.

5.1 Differences within Type 1: locuphoric and allophoric forms

Some languages like Finnish or Estonian employ specialized grammatical strategies depending on whether the inventive verb presents a locuphoric or an allophoric form. In fact, while the Finnish inventive verb *löytyä* (19) and the Estonian inventive verb *leiduma* (20) can be used both with inanimate (19a; 20a) and animate (19b; 20b) referents, they are usually not used with locuphoric forms (first or second persons), except for a few rare cases (21). This means that the verb is usually in the third person. Animate referents are also often limited to groups of animals or people rather than single individuals. The suffix *-U-* is usually referred to as a reflexive suffix, hence I gloss it accordingly, rather than as MM (the same goes for Estonian).

(19) Finnish (personal knowledge)

- a. *Jääkaapista löyt-y-y omenoita*
 fridge.ELA find-REFL-3SG apple.PL.PTV
 ‘There are apples in the fridge.’
- b. *Metsästä löyt-y-y erilaisia ötököitä*
 forest.ELA find-REFL-3SG different.PL.PTV bug.PL.PTV
 ‘There are all sorts of bugs in the forest.’

(20) Estonian (etTenTen – Web 2019¹²)

- a. *Metsa all leid-u-b kukeseeni ja*
 forest.GEN under find-REFL-3SG chanterelle.PL.PTV and
metsamaasikaid
 wild.strawberry.PL.PTV
 ‘There are chanterelles and wild strawberries on the forest floor.’
- b. *maailmas leid-u-b veel häid inimesi*
 world.INE find-REFL-3SG still good.PL.PTV person.PL.PTV
 ‘There are still good people in the world.’

(21) Finnish (Basile and Ivaska [2021: 18])

- Mistä löyd-y-t prinssini, 44–50-v. fiksu, pitkähäkö,*
 where.ELA find-REFL-2SG prince.1PX 44-50-y.o. smart tallish
ulkonäkö ok, pilke silmäkulmassa, lenkkeilet ja tanssit
 appearance ok twinkle eye.corner.INE jog.2SG and dance.2SG
 ‘Where are you my prince, 44 to 50 years old, smart, tallish, good looking, with a twinkle in your eye, who likes to jog and dance.’

The specialized allophoric form and its abundant use with partitive-marked NPs (Basile and Ivaska 2021), which are traditionally considered existential NPs

¹² Information about the corpus can be found at https://app.sketchengine.eu/#dashboard?corpname=preloaded%2Fettentent19_fil2&corp_info=1 (accessed on November 9th, 2022)

(Huumo and Helasvuo 2015, Huumo & Lindström 2014, Metslang 2012), could be the sign of the grammaticalization of this inventive form into a dedicated existitive (like Spanish *hay*), although this claim must be substantiated with further research. For the locuphoric forms, hence exclusively with animate referents, Estonian and Finnish use a dedicated analytical reflexive strategy (22a). This means that the verb does not have a reflexive marking, instead the transitive verb *löytää* ‘to find’ is used with an accusative-marked reflexive pronoun (in Estonian, the verb *leidma* ‘to find’ is used with the partitive-marked reflexive pronoun *end/ennast*).

(22) Finnish

- a. *Löysin* *itseni* *keskeltä* *metsää*
 find.PRT.1SG self.ACC.1PX middle.ABL forest.PTV
 ‘I found myself in the middle of the forest.’
- b. *Olin* *yhtäkkiä* *keskellä* *metsää*
 be.PST.1SG suddenly middle.ADE forest.PTV
 ‘Suddenly, I was in the middle of the forest.’

This analytical strategy seems, in addition, to mark the mirativity of the action, which could be directly motivated by the animacy of the referent, morpho-syntactically derived from the locuphoric form used. However, the reason why this reading arises could also lie in the higher transitivity of this type of analytical constructions, where the verb could also be perceived as conveying its original meaning. On the other hand, semantically analogous mirative copula constructions like (22b) can usually be considered locational constructions.

5.2 Differences between types: the animacy criterion

According to the animacy criterion, inventive verbs can have different dedicated grammatical forms. Consider English (23).

(23) American English (Jay Zameska, p.c.)

- a. *Bogs and marshes are found throughout Estonia.*
- b. *Wolves are found throughout the Northern hemisphere.*
- c. *Ukraine found itself at war.*
- d. **Russia finds itself in Eurasia.*
- e. *Jay found himself in the woods.*
- f. *I found myself at the hospital.*

In English, an animate referent represented by the locatum is the criterion for grammaticality of the reflexive strategy in inventive constructions (23e-f), while the passive strategy can be used with both inanimate (23a) and some animate (23b) referents. When an inanimate referent does not metaphorically refer to a collectivity or a metonymic plurality of people or animate referents (23c), the reflexive strategy is perceived as ungrammatical (23d).

Construction specialization also applies to Maltese, which uses distinct grammatical strategies to differentiate between non-mirative and mirative readings. Compare the examples below (24).

- (24) Maltese (elicited, Ray Fabri, p.c.)
- a. *Il-lukanda t-in-sab il-Belt*
 DEF-hotel 3FSG.PRS-DERV7-FIND DEF-City
 ‘The hotel is situated/found in Valletta (the City).’
- b. *N-in-sab Ruma*
 1SG.PRS-DERV7-find Rome
 ‘I am in Rome.’ – non-mirative reading
- c. *Sib-t ruh-i f'Malta*
 find-1SG.PST soul-1SG.POSS in-M.
 ‘I found myself in Malta.’ – mirative reading
- d. *Sib-t ruh-i f'sitwazzjoni diffiċli*
 find-1SG.PST soul-1SG.POSS in-situation difficult
 ‘I found myself in a difficult situation.’

In (24a-b), the derivational form 7 (DERV7) can coexpress several functions of the middle voice (in this case passive, Ray Fabri, p.c.). In (24a), the construction has a concrete location and an inanimate locatum, like in English, while in (24b) the locatum is animate. (24b) contrasts with the periphrastic reflexive strategy used in (24c), which encodes a mirative reading. The same strategy, with the same reading, can also be used with abstract locative phrases (24d).

Basque presents two different morphosyntactic strategies that accurately align with the two types described above. One of them (25a) is a present participle strategy and allows for both allophoric and locuphoric forms of the verb *izan* ‘to be’ but does not allow for inanimate locati (Type 1). The other one (25b) only allows the 3SG potential form of *izan* ‘to be’ to be used with the infinitive short stem of the verb *aurki* ‘to find’ (Irantzu Epelde, p.c.), and is used only with inanimate locati (Type 2).

- (25) Basque (Urtzi Etxeberria, p.c.)
- a. *Egoera arraro batean aurki-tzen naiz*
 situation strange one.INE find-PRP be.1SG
 ‘I find myself in a strange situation.’

- b. *Webgunean hainbat informazio osagarri aurki*
 website.INE several information.ABS additional find.INF
daiteke
 be.PRS.POT.3SG
 ‘More information can be found on the website.’

As for (25b), it could be an innovation due to contact with Spanish (Irantzu Epelde, p.c.), which could confirm that (25b) is to be analyzed as an inventive-locational construction rather than an instance of verbal predication. According to native speakers, the most natural way of expressing a locational function in Basque is, however, the locative copula *egon* (Urtzi Etxeberria, p.c.). This fact questions the productivity of the Basque inventive-locational construction, which should be seen as a rather marginal phenomenon that is developing thanks to language contact with its neighboring languages.

Even for a morphosyntactically distant language like Basque, the same criteria seem to apply. To distinguish between the two types of inventive constructions, or within subtypes of the first type, specialized morphosyntactic strategies often play a role. They vary according to the animacy of the referent represented by the locatum, and to the morphological characteristics of the verb phrase.

5.3 Marginal cases

In Greek, middle-marked forms of the verb *vrisko-* ‘to find’ are used to indicate locational meaning. However, this meaning is not available when the verb is marked for perspective aspect (26).

- (26) Greek (Anna Kampanarou, p.c.)
*I Eladha vriskete/*vrethike stin*
 the Greece.NOM find.NACT.PRES.3SG/find.NACT.PAST.PERF.3SG at.the
Evropi.
 Europe
 ‘Greece is/*was located in Europe.’

The non-active form, when marked for perfective aspect, can only be read as the literal non-active form of ‘find’, i.e., ‘be found by someone’ (27) or receive an inchoative (change-of-state) use (28; Anna Kampanarou, p.c.). The latter can be also interpreted as expressing mirative reading.

- (27) *I Maria vrethike apo tin astinomia.*
 the Mary.NOM find.NACT.PERF.3SG by the police
 ‘Mary was found by the police.’

- (28) *Vrethika* *sti* *thesi* *na* *apologume.*
 find.NACT.PERF.1SG in.the position to apologize
 ‘I found myself in the position to apologize.’

For indicating locational functions in Albanian, in addition to inventive-locational constructions, I have also found another strategy, based on a verb with meaning HAPPEN. I propose that this type of construction be called *eventive-locational* (Lat. *evenire* ‘to happen’). Just like in inventive constructions, the verb *ndodh* ‘to happen’ undergoes semantic bleaching and behaves like a copula. However, it does not satisfy any other morphosyntactic or semantic criteria. Example (29) shows the Albanian inventive-locational construction alongside the eventive-locational construction.

- (29) Albanian (Nensi Islami, p.c.)
 a. *Italia* *gje-nd-et/ndodh-et* *në* *Mesdhe*
 Italy find-REFL-3SG/happen-3SG in Mediterranean.ACC
 ‘Italy is situated in the Mediterranean.’
 b. *Ndodh-em/gje-nd-em* *në* *një* *situatë*
 happen-1SG/find-REFL-1SG in INDEF.ACC situation.ACC
të *keqe*
 ADJART.ACC bad.ACC
 ‘I find myself in a bad situation.’

This construction has not been found in any other language of the sample. Because it is so marginal and not morphologically marked, it could be a simple case of polysemy of the verb with meaning HAPPEN. Further investigation is needed.

6 DISCUSSION AND FUTURE RESEARCH

This paper has proposed a novel definition for certain locational constructions I call inventive-locational constructions. These constructions feature a verb with the meaning FIND which is called an inventive verb and satisfies certain criteria. This verb, in fact, has a root whose prototypical meaning is FIND and becomes semantically bleached. This verb is also morphologically marked (e.g., via a reflexive/middle voice strategy), and only expresses a locational function without semantically marking a specific posture. Semantic bleaching can be regarded as a central criterion because it makes it possible to identify inventive constructions as having a purely locational meaning. This means that inventive-locational constructions are instances of what is traditionally considered nonverbal predication.

Inventive-locational constructions can be classified in two main types, which are distinct on the basis of the animacy of the referent represented by the locatum.

The first type, which has an animate referent, can have language-specific subtypes that correlate not only with the animacy criterion, but also with the concreteness of the location expressed by the locative phrase and with whether the inventive verb is found in a locuphoric or an allophoric form. It is still debatable whether some of these subtypes (as in Finnish or Maltese) can be distinguished on the sole basis of the animacy of the referent, or whether there is another variable that comes into play – mirativity. Mirativity could in fact be directly correlated with animacy, but Maltese, for example, uses a dedicated analytical strategy exclusively for mirative utterances. Analytical strategies like the one used by Maltese could also be interpreted as instances of verbal predication. In this case, there would be no semantic bleaching and, instead of a mirative reading, we would be dealing with a prototypical meaning FIND, which entails *per se* a higher degree of unintentionality and unexpectedness of the action.

While I had initially thought that inventive constructions might be peculiar to European languages, and that it is possible that they constitute an areal phenomenon, I have recently been finding them in other languages as well. Tagalog (Austronesian < Malayo-Polynesian) uses forms of the verbs *kita* ‘to see/find’ and *hanap/tagpo* ‘to find’ to express locational functions (personal documentation). While this phenomenon could be the result of prolonged contact with European languages like Spanish and English, other non-European languages also present inventive-locational constructions. One clear example is provided by Mandinka (Mande, Creissels and Sambou 2013: 145). These recent findings call for further research on a broader spectrum of languages, that can be facilitated by the definition criteria presented above (Section 1). These criteria make it possible to use inventive constructions as a comparative concept. It is worth considering that there might be not only other language families in the world that use FIND-based strategies to express a locational function, but also construction-strategies based on other verbs that become semantically bleached and coexpress a locational function alongside their original meaning, in a similar way to posture verbs or to the Albanian eventive-locational construction introduced above. I suggest that verbs like these be part of a class of construction-strategies called *situative-locational constructions*. Situative-locational constructions would ideally be a larger term that groups inventive-locational constructions with other similar constructions, where semantic bleaching plays a role in the definition of locational strategies.

One possible research question for the future is, of course, the grammaticalization of inventive constructions, especially because semantic bleaching is a central component of both these constructions and grammaticalization itself (Hopper and Traugott 2003: 76). Perhaps there are traceable historical reasons behind why exactly this kind of verb has become popular to express locational meaning. This development could be directly linked to its lexical aspect, which points, in addition, toward the suitability of a cognitive approach.

APPENDIX

This appendix illustrates the data used in this article. Each inventive verb form is associated with the type it is used in (in parentheses). *Type 1* can be followed by a comment, which indicates that the data collected shows the usage of Type 1 only in certain conditions (e.g. “Type 1, abstract LOC” means that the form appears with an animate referent located in an abstract location only).

Language	Inventive verb form	Example	Source
Albanian	<i>gjend-</i>	in text	consultation with native speaker (Nensi Islami)
Basque	<i>aurkitzen naiz</i> (Type 1, abstract LOC, mirative) <i>aurki daiteke</i> (Type 2)	in text	Irantzu Epelde, Urtzi Etxeberria, p.c.
English	<i>to find oneself</i> (Type 1) <i>to be found</i> (Type 2)	in text	consultation with native speaker (Jay Zameska)
Estonian	<i>ennast leidma</i> (mirative) <i>leiduma</i>	<i>Leidsin ennast</i> find.PRT.1SG self.PTV <i>haiglast</i> hospital.ELA 'I found myself at the hospital.'	Liina Lindström, p.c.
Finnish	<i>löytää itsensä</i> (mirative) <i>löytyä</i>	in text	Ilmari Ivaska, p.c.
German	<i>sich befinden</i>	Der Bahnhof be-finde-t DEF station DER-find-3SG sich zwischen zwei Städten REFL between two cities 'The station is located between two cities.'	Gerson Klumpp, p.c.
Greek	<i>vriskome-</i>	in text	Anna Kampanarou, p.c.
Hungarian	<i>található</i> (Type 2, concrete LOC)	Magyarország Európában Hungary Europe.INE talál-hat-ó find-POT-PRP 'Hungary is located in Europe.'	Bogáta Timár, Kata Kubínyi, p.c.
Italian	<i>trovarsi</i>	in text	personal knowledge

Language	Invenitive verb form	Example	Source
Latvian	<i>atrasties</i>	Alus atrodas ledusskapī beer find.3SG.REFL fridge.LOC 'The beer is found in the fridge.' Jūs atrodaties centrā 2PL find.2PL.REFL center.LOC 'You are (located) in the center.'	Daiki Horiguchi, p.c.
Maltese	<i>-sab</i> (Type 1, non-mirative; Type 2) <i>sab- ruħ-</i> (Type 1, mirative)	in text	Ray Fabri, p.c.
Russian	<i>nachodit'sya</i>	in text	Denys Teptiuk, p.c.
Sardinian	<i>s'agattare</i>	in text	consultation with native speaker

Glossing abbreviations

1	first person	INF	infinitive
2	second person	MIR	mirative
3	third person	NACT	non active
ABL	ablative	PERF	perfective
ABS	absolute	PL	plural
ACC	accusative	POSS	possessive suffix
ADJART	adjectival article	POT	potential
APPL	applicative	PREP	prepositional case
DEF	definite	PRT	preterite
DER	derivational form	PRP	present participle
DERV7	derivational form 7	PRS	present
ELA	elative	PST	past
ENCL	enclitic	PTCP	participle
EXPL	expletive	PTV	partitive
EXV	existive	PX	personal suffix
F	feminine	REFL	reflexive
GEN	genitive	SG	singular
INDEF	indefinite	TEL	telic
INE	inessive		

ACKNOWLEDGMENTS

I am especially thankful to Gerson Klumpp and Mariann Bernhardt for the fruitful in-class discussions that, I think, ultimately led me to come up with this topic. Gerson was also the first person to ever encourage me to investigate inventives more in detail.

Martin Haspelmath also deserves my most sincere thanks for providing valuable feedback at various stages of this work. He suggested that I use the term *inventive* (instead of my less transparent term *situative*) after a brief online meeting we had in 2021.

Moreover, I thank John Beavers for his thoughts on the initial stages of this work, and for providing a different perspective over this topic.

REFERENCES

- Aikhenvald, Alexandra Y. 2012. The essence of mirativity. *Linguistic Typology*, 16(3), 435–485.
- Bahrt, Nicklas N. 2021. *Voice syncretism* (Research on Comparative Grammar 1). Berlin: Language Science Press.
- Barðdal, Jóhanna. 2006. Construction-specific properties of syntactic subjects in Icelandic and German. *Cognitive Linguistics* 17(1). 39–106.
- Basile, Rodolfo and Ilmari Ivaska. 2021. Löytyä-verbin konstruktioiden yhteydessä esiintyvä subjektin sijanvaihtelu. *Eesti ja soome-ugri keeleteaduse ajakiri. Journal of Estonian and Finno-Ugric Linguistics*, 12(1), 11–39.
- Bentley, Delia, Francesco Maria Cicone and Silvio Cruschina. 2013. Existential constructions in crosslinguistic perspective. *Italian Journal of Linguistics*, 25(1), 15–43.
- Creissels, Denis. 2014. Existential Predication in Typological Perspective. *Ms., Université Lyon*.
- Creissels, Denis. 2019. Inverse-Locational Predication in Typological Perspective. *Italian Journal of Linguistics* 31 (2): 38–106.
- Creissels, Denis. 2023. Existential predication and have-possessive constructions in the languages of the world. In Laure Sarda & Ludovica Lena (Eds.), *Existential constructions across languages: Forms, meanings and functions*. Amsterdam/Philadelphia: John Benjamins. 34-67.
- Creissels, Denis and Pierre Sambou. 2013. *Le mandinka: phonologie, grammaire, textes*. Paris: Karthala.
- Croft, William. 2001. *Radical Construction Grammar: syntactic theory in typological perspective*. Oxford: Oxford University Press.
- Croft, William 2022. *Morphosyntax: Constructions of the World's Languages*. Cambridge: CUP.
- Dahl, Östen. 2000. Egophoricity in discourse and syntax. *Functions of Language* 7(1). 37–77.
- Declerck, Renaat. 1988. *Studies on copular sentences, clefts and pseudo-clefts*. Leuven: Leuven University Press & Foris.

- DeLancey, Scott. 1997. *Mirativity: The grammatical marking of unexpected information*.
DeLancey, Scott. 2001. The mirative and evidentiality. *Journal of pragmatics*, 33(3), 369–382.
- Freeze, Ray. 1992. Existentials and other locatives. *Language*, 553–595.
- Gast, Volker and Florian Haas. 2011. *On the distribution of subject properties in formulaic presentational of Germanic and Romance*, 127–166. Amsterdam: John Benjamins Publishing.
- Geniušienė, Emma. 1987. *The Typology of Reflexives*. Berlin: Mouton de Gruyter.
- Goldberg, Adele E. 1995. *Constructions: a Construction Grammar approach to argument structure*. Chicago: University of Chicago Press.
- Haspelmath, Martin. 1998. How young is standard average European. *Language sciences*, 20(3), 271–287.
- Haspelmath, Martin. 2013. Argument indexing: A conceptual framework for the syntactic status of bound person forms. *Languages across boundaries: Studies in memory of Anna Siewierska*, 197–226.
- Haspelmath, Martin. 2019. Comparing reflexive constructions in the world's languages. Ms. MPI-SHH Jena & Leipzig University, Germany.
- Haspelmath, Martin. 2021. Existential constructions and two types of comparative concepts: Construction-functions and construction-strategies. *Descriptive Grammar and Typology 2, Paris 2021 December 1–3*. (https://grammar2021.sciencesconf.org/data/pages/Handout_Paris.pdf)
- Haspelmath, Martin. 2022. Nonverbal clause constructions. (https://www.academia.edu/71539642/Nonverbal_clause_constructions)
- Heine, Bernd. 1997. *Possession: cognitive sources, forces, and grammaticalization*. Cambridge: Cambridge University Press.
- Hengeveld, Kees. 1992. *Non-verbal predication: Theory, typology, diachrony*. Berlin: Mouton de Gruyter.
- Hengeveld, Kees, and Hella Olbertz. 2012. Didn't you know? Mirativity does exist!. *Linguistic Typology*, 16(3), 487–503.
- Hopper, Paul J. and Elizabeth C. Traugott. 2003. *Grammaticalization (2nd edition)*. Cambridge: Cambridge University Press.
- Huumo, Tuomas. 2003. Incremental existence: the world according to the Finnish existential sentence. *Linguistics* 41 (3): 461–493.
- Huumo, Tuomas and Marja-Liisa Helasvuo. 2015. On the subject of subject in Finnish. In Tuomas Huumo and Marja-Liisa Helasvuo (eds.), *Subjects in constructions—canonical and non-canonical*, 13–41. Amsterdam: John Benjamins.
- Huumo, Tuomas and Liina Lindström. 2014. Partitives across constructions: on the range of uses of the Finnish and Estonian ‘partitive subjects’. In Silvia Luraghi and Tuomas Huumo (eds.), *Partitive Cases and Related Categories* 54: 153. Berlin: De Gruyter Mouton.
- Inglese, Guglielmo. 2022. Towards a typology of middle voice systems. *Linguistic Typology* 26(3). 489–531.
- Inglese, Guglielmo. 2023. The rise of middle voice systems: a study in diachronic typology. *Diachronica*.
- Kemmer, Suzanne. 1993. *The middle voice*. Amsterdam/Philadelphia: John Benjamins Publishing Company.

- Koch, Peter. 2012. Location, existence and possession: a constructional-typological exploration. *Linguistics* 44: 569–612.
- Lakoff, George. 1987. *Women, fire, and dangerous things: what categories reveal about the mind*. Chicago: University of Chicago Press.
- Lambrecht, Knud. 1994. *Information structure and sentence form*. Cambridge: CUP.
- Masini, Francesca. 2012. Costruzioni verbo-pronominali ‘intensive’ in italiano. *Language and the Brain–Semantics*.
- McNally, Louise. 2016. Existential sentences crosslinguistically: variations in form and meaning. *Annual Review of Linguistics* 2: 211–31.
- Metslang, Helena. 2012. On the case-marking of existential subjects in Estonian. *SKY Journal of Linguistics* 25: 151–204.
- Mikkelsen, Line. 2011. Copular clauses. In Claudia Maienborn, Klaus von Stechow and Paul Portner (eds.), *Semantics: An international handbook of natural language meaning, volume 2*, 1805–1829. Berlin: De Gruyter Mouton.
- Milsark, Gary L. 1977. Toward an explanation of certain peculiarities of the existential construction in English. *Linguistic Analysis* 3(2). 1–29.
- Partee, Barbara H. and Vladimir Borschev. 2007. Existential Sentences, BE, and the genitive of negation in Russian. *Existence: Semantics and Syntax*, 147–190. Springer, Dordrecht.
- Roy, Isabelle. 2013. *Nonverbal predication: Copular sentences at the syntax-semantics interface*. Oxford: Oxford University Press.
- Stassen, Leon. 1997. *Intransitive predication*. Oxford University Press.
- Talmy, Leonard. 1983. How language structures space. In *Spatial orientation*, 225–282. Boston: Springer.
- Talmy, Leonard. 2000. *Toward a Cognitive Semantics. Vol 1*. The MIT press.
- Zúñiga, Fernando, and Seppo Kittilä. 2019. *Grammatical voice*. Cambridge University Press.

Basile, Rodolfo, and Ilmari Ivaska. 2021.
Löytyä-verbin konstruktioiden yhteydessä esiintyvä subjektin
sijanjaihtelu. *Eesti ja soome-ugri keeleteaduse ajakiri*.
Journal of Estonian and Finno-Ugric Linguistics, 12(1), 11–39.

LÖYTYÄ-VERBIN KONSTRUKTIOIDEN YHTEYDESSÄ ESIINTYVÄ SUBJEKTIN SIJANVAIHTELU

Rodolfo Basile^{a,b}, Ilmari Ivaska^b

^aTarton yliopisto, EE

^bTurun yliopisto, FI

rodolfo.basile@ut.ee, ilmari.ivaska@utu.fi

Abstrakti. Artikkelitarkastelee *löytyä*-verbin konstruktioiden nominatiivi- ja partitiivi-subjektin vaihtelua. Aineistona on korpuksista poimittu 779 havainnon satunnaisotos, jota tarkastellaan sekä kvantitatiivisesti tilastollisin menetelmin että kvalitatiivisesta näkökulmasta. Tutkimus pyrkii selvittämään, mitkä muuttujat vaikuttavat *löytyä*-verbin sisältävien lauseiden subjektien sijanvalintaan. Valikoidut muuttujat ovat subjektin luku, subjektin jaollisuus, subjektin sanaluokka, sanajärjestys, aikamuoto, subjektin ja verbin välinen kongruenssi sekä subjektin lemma, joka toimii satunnaismuuttujana. Regressioanalyysin keinoin subjektin sijanvalintaa ennustetaan mainittujen muuttujien ja niiden välisten vuorovaikutussuhteiden avulla. Laadullisessa analyysissä käsitellään myös näiden morfosyntaktisten ja semanttisten seikkojen vaikutusta lauseen eksistentiaalisuuden sekä subjektin kvantiteetin ja definiittisyyden tulkintaan.

Avainsanat: subjektin sijanvalinta, partitiivi, konstruktiot, morfosyntaksi, regressioanalyysi, määrällinen tutkimus, korpusmenetelmät, suomen kieli

DOI: <https://doi.org/10.12697/jeful.2021.12.1.01>

1. Johdanto

Viime aikoina kielitieteessä on keskusteltu usein nominatiivi- ja partitiivisubjektin välisestä vaihtelusta. Sitä on tutkittu esimerkiksi typologian (Luraghi & Kittilä 2014) ja murteiden (Lindström 2017) näkökulmista. Slaavilaisissa kielissä, kuten puolassa (Blaszczak 2009) tai venäjässä, kyseinen vaihtelu tapahtuu nominatiivin ja genetiivin välillä, erityisesti kun verbi on kielteisessä muodossa (Blaszczak 2012, Seržant 2014, Partee & Borshev 2008): tämä ilmiö on suhteessa konstruktioiden eksistentiaalisuuteen. Eksistentiaalilauseita (Creissels 2014, McNally 2016) on myös tutkittu kieltenvälisesti (Sawyer 1973, Bentley ym. 2013) ja ylipäätään monissa muissa kielissä, kuten englannissa

(Breivik 1999, Milsark 2014), italiassa (Cruschina 2012) tai kiinassa (Hu & Pan 2008). Koska eksistentiaalilauseita on joskus ongelmallista määritellä, niitä on myös tutkittu verbin perusteella: englanniksi voi esimerkiksi olla yhtäältä *existential 'there'*, joka esiintyy yhdessä *be*-verbin ('olla') kanssa, toisaalta *presentational 'there'*, jonka yhteydessä käytetään muita verbejä (Ward 1999: 5).

Tässä artikkelissa tarkastelemme *löytyä*-verbin käyttäytymistä erilaisten muuttujien analyysin perusteella. Pyrimme selvittämään, miten *löytyä*-verbin sisältävät lauseet käyttäytyvät yhtäältä nominatiivi- ja toisaalta partitiivisubjektin yhteydessä, sekä sitä, minkälaisessa vuorovaikutuksessa tämä vaihtelu on suhteessa muihin eksistentiaalilauseeseen liitettyihin kielenpiirteisiin. Esimerkit (1) ja (2) kuvaavat vastaavasti nominatiivi- ja partitiivisubjektin käyttöä.

(1) Savo löytyy Savosta! (KLK_FIN_2000)

(2) Taitoa löytyy kun tarvitaan. (KLK_FIN_1998)

Tutkimuksen idea syntyi havaitessamme *löytyä*-verbin käyttövalikoiman laajuuden. Huomasimme, että kyseinen verbi on erittäin mielenkiintoinen semanttiselta kannalta erityisesti ns. eksistentiaalisissa konstruktioissa, mutta myös tavanomaisten nominatiivisubjektien kanssa. Toisin kuin monet muut kyseisissä konstruktioissa esiintyvät prototyyppiset ”leksikaaliset eksistentiaalit” (eli muut kuin *olla*, ks. Larjavaara 2019: 149), kuten esimerkiksi *lähteä*, *tulla*, *tippua* jne., jotka yleensä kuvailevat jotakin toimintaa eivätkä pelkästään (e-)subjektin olemassaoloa, *löytyä* toimii monesti *olla*-verbin korvikkeena eli korvaa eksistentiaalisen verbin *par excellence*. Sen lisäksi kukaan ei ole aiemmin keskittänyt huomiota nimenomaan tähän verbiin.

Tutkimuksen pääkysymys on: mitkä muuttajat vaikuttavat *löytyä*-verbin sisältävien lauseiden subjektien sijamuotovalintaan? Käytetty aineisto on *Kansalliskirjaston lehtikokoelman (KLK)* suomenkieliset lehdet 1990- ja 2000-luvulla. Valitsimme tämän aineistoksemme, sillä sen voidaan katsoa sisältävän tyypillistä nykypäivän kirjoitettua suomea ja korpus kattaa laajan valikoiman journalistisia tekstejä. Aineiston laajuus mahdollistaa sekä ilmiön tarkastelun määrällisesti tilastotieteellisiä menetelmiä käyttäen että esiinnousevien ilmiöiden tarkemman laadullisen tarkastelun. Artikkelin rakenne on seuraava: luvussa 2 esittelemme *löytyä*-verbillisiä lauseita aiemman tutkimuksen valossa.

Luvussa 3 kerromme, miten olemme toteuttaneet nyt käsillä olevan tutkimuksen. Luvussa 4 siirrymme analysoimaan subjektivaihtelua sekä määrällisin että laadullisin menetelmin. Luvussa 5 teemme tutkimuksen yhteenvedon.

2. Löytyä-verbilliset lauseet aiemman tutkimuksen valossa

Tässä luvussa esittelemme artikkelin taustateoriaa ja terminologiaa. Alaluvussa 2.1. määrittelemme subjektin ja esittelemme eksistentiaalilauseiden kategoriaa ja sen yhteyttä subjektin sijavaihteluun. Alaluvussa 2.2. puhumme jaollisuudesta ja kvantiteetista sekä siitä, miten nämä käsitteet liittyvät lausesemantiikkaan. Lopuksi alaluvussa 2.3. kerromme lähemmin tutkimuksen lähestymistavasta sekä siitä, miksi valitsemamme muuttujat ovat tärkeitä.

2.1. Subjektin määritelmä ja eksistentiaalilauseet

Huumo ja Helasvuo (2015) tekevät yhteenvedon keskustelusta, jota on käyty fennistiikassa kautta aikojen: voivatko partitiivissa olevat substantiivilausekkeet oikeastaan toimia subjekteina? Heidän artikkelissaan antamaansa subjektin roolin määritelmää voidaan pitää suppeana: semanttisesta ja diskurssianalyttisesta näkökulmasta sekä kognitiivisen kieliopin pohjalta subjekti on lauseen lähtöpiste (eng. *starting point*, ks. esim. Langacker 2008). Subjektia määrittelevät suomen kielessä kieliopilliset ominaisuudet kuten nominatiivimuoto, predikaatin kongruenssi sekä yleensä preverbaalinen eli temaattinen positio lauseessa (Huumo & Helasvuo 2015: 37). Huumo ja Helasvuo suhtautuvat siis ns. partitiivisubjektiin epäilevästi ja ehdottavat erillistä ‘eksistentiaalisen substantiivilausekkeen’ (eng. *e-NP* eli *existential Noun Phrase*) kategoriaa. Tämän näkemyksen mukaan kaikki partitiivitunnukselliset subjektin kaltaiset substantiivilausekkeet kuuluvat siis eksistentiaalilauseeseen. Tätä näkökulmaa edustaa esimerkiksi Ikola (1954: 224–226), jonka mukaan partitiivisubjektit esiintyvät ainoastaan eksistentiaalilauseissa, jotka sallivat myös nominatiivisubjekteja. Toinen semanttista roolia luonnehtiva ja kaikki eksistentiaaliset subjektit kattava käsite on Huumon (2003: 462) *e-theme*.

Vakiintuneet eksistentiaalilauseet ovat sanajärjestykseltään konstruktioita, joissa adverbiaali sijoittuu yleensä temaattiseen, verbiä

edeltävään positioon. Sanajärjestys voi kuitenkin olla käänteinen, missä tapauksessa e-subjekti esiintyy ennen verbiä ja paikkaa ilmaiseva adverbiaali sen jälkeen, eikä kongruenssia ole. Jos lause sen sijaan kongruoi eli subjekti on esimerkiksi monikon nominatiivissa ja verbi monikon 3. persoonassa, kyseessä on sijaintia ilmaiseva normaalilause (vrt. Hakanen 1972, Larjavaara 2019: 118; sijainnin ilmaisemisesta [engl. locational predication], ks. myös Creissels 2014): lause ei siis tuo esiin uutta referenttiä eikä näkökulma ole eksistentiaalilauseista poiketen holistinen (ks. myös Lees 2015: 7). Ivaska (2011: 81) osoittaa, että oppijankielen kontekstissa lauseissa käytetyt epäprototyypiset piirteet johtavat usein lausetyyppien sekoittumiseen: toisin sanoen eksistentiaalisen ja ei-eksistentiaalisen lauseen välinen raja hämärtyy (termeistä *eksistentiaalilause* ja *eksistentiaalinen lause* ks. Ivaska 2010, 2011) eikä välttämättä ole mahdollista erottaa eksistentiaalilauseita muista lausetyypeistä kategorisesti ainoastaan subjektin sijamuodon, kongruenssin tai sanajärjestyksen perusteella. Alaluvussa 2.3. käsittelemme yksityiskohtaisemmin tällaisten muuttujien vaikutusta lausetyyppien määrittelyyn.

2.2. Subjektin semantiikan vaikutus lausesemantiikkaan

Löytyä-verbillisten lauseiden semanttiset piirteet kiertyvät osaksi keskustelua lauseiden eksistentiaalisuudesta, ja esimerkiksi subjektin jaollisuus vaikuttaa lauseiden semanttiseen tulkintaan. ISK:ssa käytetään termejä *jaollinen* ja *jaoton* (VISK §554–555) viitaten subjektin tarkoitteeseen. Larjavaara (1988) käyttää termiä *avoin* (vrt. *sulkeinen*) kuvatessaan jaolliseen subjektin tarkoitteeseen viittaavia ns. *ainesanoja* tai jaottomia *kappalesanoja* monikon partitiivissa (ks. myös Huumo 2003); *sulkeisia* ovat toisaalta kaikki jaottomiin subjektin tarkoitteisiin viittaavat *kappalesanat*. Sijanvalinta vaikuttaa kuitenkin monesti jaollisten ainesanojen kvantiteettiin: lauseessa *Viini on lasissa* subjektin kvantiteettia käsitellään siis sulkeisena (vrt. avoimeen vaihtoehtoon *Viiniä on lasissa*). Jos prototyypisesti jaoton subjekti on monikkumuotoisena partitiivissa, muuttaa sijamerkintä subjektin kvantiteetin avoimeksi. Tämä vaikuttaa myös verbin lukuun, sillä tällaisessa tapauksessa verbi ei sulkeisesta kvantiteetista poiketen kongruoi subjektin kanssa. Monikollisia verbitapauksia on tästä syystä eksistentiaalilauseiden joukossa huomattavasti vähemmän kuin yksiköllisiä.

Partitiivisubjekti ja inkongruenssi ovat siis tärkeitä elementtejä eksistentiaalilauseen määrittämisessä esimerkiksi Hakasen (1972) mukaan. Prototyypisellä eksistentiaalilauseella on myös tietty rakenne, jolla luonnehditaan alussa olevan lokatiiviadverbiaalin tarkoitetta kertomalla mitä tämä sisältää sekä esitetään diskurssiin uusia epämääräisiä substantiivilausekkeita, jotka voivat olla nominatiivi- tai partitiivisijaisia (Huumo & Lindström 2014: 154). Tässä rakenteessa yllä mainitut substantiivilausekkeet eivät kuitenkaan ole prototyypisiä subjekteja vaan e-subjekteja, joilla on eksistentiaalinen funktio yhtä hyvin myös lauseen alussa, mikäli ne ovat partitiivissa.

2.3. Tutkimuksen lähestymistapa sekä muuttujien esittely

Lähestymme *löytyä*-verbillisiä lauseita käyttöpohjaisesta näkökulmasta konstruktiokieliopin avulla (Goldberg 1995, 2006): lähdemme liikkeelle siitä oletuksesta, että nominatiivi–partitiivi -vaihteluun vaikuttaa samanaikaisesti useita toisiinsa kietoutuneita muuttujia, jotka kaikki osaltaan vaikuttavat sijamerkintään. Näiden eri muuttujien välinen vuorovaikutteinen kokonaisuus ohjaa konstruktion kokonaismerkityksen hahmottumista joko eksistentiaaliseksi tai sijaintia ilmaisevaksi normaalilauseeksi. Seuraavaksi esittelemme tarkastelemamme muuttujat. Kaikki muuttujat on valikoitu aiemman tutkimuksen perusteella: riippumattoman eli selitettävän muuttujamme eli subjektin sijamuodon lisäksi myös riippuvat eli selittävät muuttujamme kuten kongruenssi, sanajärjestys, sekä subjektin luku ja jaollisuus ovat lauseen eksistentiaalisuuteen tyypillisesti yhdistettyjä kielenpiirteitä. Aikamuoto on syytä ottaa huomioon, jotta voidaan huomioida sen mahdollinen suhde lauseen aspektiin ja teonlaatuun. Verbin aikamuodon suhde muihin lauseen kieliopillisiin ominaisuuksiin (kuten tässä tutkimuksessa e-NP:n sijamuoto) määrittää lauseen aspektia, kun taas lauseen teonlaatu on sidottu verbi-
lekseemien aspektiominaisuuksiin (VISK § 1500). Subjektin sanaluokka – ennen kaikkea pronomien ja substantiivien välinen suhde – puolestaan liittyy siihen Helasvuon (1996: 352) esittämään huomioon, että “[eksistentiaalilauseiden subjektit] eivät ole diskurssin kannalta keskeisiä vaan jäävät useimmiten kertamaininnan varaan.” Pronomien tarkoitteet määrittyvät tyypillisesti kontekstissaan joko deiktisesti tai anaforisesti (VISK §714), mistä syystä ne eivät lähtökohtaisesti vastaa Helasvuon määritelmää tyypillisestä eksistentiaalilauseen subjektista.

Subjektin lemman sisällyttäminen tarkastelunalaisiin muuttujiin sen sijaan mahdollistaa kollostruktionaalisten preferenssien eli konstruktion ja siinä esiintyvien sanojen myötäesiintymä-suhteiden huomioimisen (Stefanowitsch & Gries 2003, Gries & Stefanowitsch 2004).

Aiemman tutkimuksen perusteella tärkeimmät muuttujat subjektin sijamuotovaihtelun määrittämisessä ovat subjektin luku (3) suhteessa verbin lukuun (muuttujana kongruenssi kyllä 3a ja 3b/ei 3c) sekä subjektin jaollisuuteen (jaollinen 3a/jaoton 3b) ja sanajärjestykseen (SV 3c, VS 3d). Näiden kaikkien muuttujien lisäksi käytetään subjektin sanaluokkaa eli substantiivia (NP, 3a) tai muita (3e), niin kuin adjektiiveja, pronomineja, partisiippeja (-*vA* tai -*nUt*) ja 4. infinitiivejä sekä aikamuotoja (preesens 3a, preteriti 3c, perfekti 3f). Satunnaismuuttujana eli muuttujana, jonka kaikkia mahdollisia arvoja ei ole ennalta määritetty, on subjektin lemman.

- (3) a. Luomu löytyy yhdeltä kartalta. (KLK_FI_1999)
 b. Tyypilliset esimerkit löytyvät pääkaupunkiseudulta. (KLK_FI_1999)
 c. Kumiteriäkin löytyi – tosin vain hirvittäviä kokoja. (KLK_FI_1998)
 d. Näyttelemistyyleistä löytyi myös eroja. (KLK_FI_1997)
 e. Silti pelattavaa löytyy. (KLK_FI_1996)
 f. Jonkunlainen malli on löytynytkin. (KLK_FI_1996)

Larjavaara (2019: 12) esimerkiksi selittää, että ”partitiivin lauseopillinen eli syntaktinen luonne määräytyy lähes kokonaan sen eri merkityksistä (= semantiikasta). Sijan merkitys ja syntaksi sulautuvat yhteen”. Konstruktiokieliopin peruseriaate selittää satunnaismuuttujamme (eli subjektin lemman) vaikutuksen lauseiden merkityksiin. Toisin sanoen, koska konstruktiomerkitykset kumpuavat prototyypisistä käytöstä ja prototyypisessä käytössä useimmin esiintyvien sanojen merkityksestä (Goldberg, Casenheiser & Sethuraman 2004), eri käyttökontekstien sanojen voidaan olettaa preferoivan eri konstruktioita.

3. Tutkimuksen toteutus

Tämä luku koostuu kahdesta alaluvusta, joissa kerromme, miten toteutamme tutkimuksen. Aloitamme esittelemällä käyttämämme aineiston, minkä jälkeen kuvaamme ne menetelmät, joilla analysoimme aineistoa.

3.1. Aineisto

Aineistona käytetään *Kansalliskirjaston lehtikokoelman (KLK)* suomenkielisiä lehtiä. Tarkastelumme on luonteeltaan synkronista, mistä syystä olemme rajanneet tarkastelumme vuoden 1990 alun ja vuoden 2000 lopun väliseen aineistoon. Tutkimuksen toteutettavuuden vuoksi emme ole analysoineet korpuksen kaikkia 48969 *löytyä*-verbin sisältävää lausekontekstia yksityiskohtaisesti, vaan olemme nostaneet tarkastelun kohteeksi 500 affirmaatiivisesta SV-lauseesta ja 500 affirmaatiivisesta VS-lauseesta koostuvan tasapainotetun satunnaisotoksen. Ratkaisu on konstruktoiden sisäistä ja lähikonstruktoiden välistä vaihtelua tarkastelevassa tutkimuksessa tyyppillinen (ks. esim. Gries & Wulff 2013, Klavan 2020), sillä näin voidaan varmistaa, että tilastollisessa analyysissä voidaan mallintaa tarkasteltavan ilmiön (tässä: subjektin sijamerkintä) vaihtelua tasapainoisesti (ks. tarkemmin 3.2). Kielto-lauseet on jätetty tarkastelun ulkopuolelle, sillä niissä eksistentiaali-lauseen subjekti esiintyy odotuksenmukaisesti partiiivimuotoisena riippumatta muista sijamerkintään vaikuttavista tekijöistä. Lopullinen aineisto koostuu 779 lauseesta, joista SV-lauseita on 387 ja VS-lauseita 392. Tämä eroavuus johtuu ensinnäkin siitä, että molemmissa ryhmissä (SV ja VS) esiintyi vastakkaisen ryhmän esimerkkilauseita johtuen korpuksen lukuvirheistä mutta myös siitä, että **Taulukossa 1** olevat lauseet poistettiin.

Taulukko 1. Poistetut lauseet.

Kieltolauseet	39
Kvanttori- ja numeraalilauseet	95
<i>Löytyä</i> ei pääverbinä	49
Tapaluokka potentiaali	6
Aikamuoto pluskvamperfekti	2
Muut ongelmalliset	29
<i>Löytyä</i> -verbin käyttö 2. persoonassa	1
Yhteensä	221

39 kieltolauseetta jätettiin tarkastelun ulkopuolelle, sillä ne olivat virheellisesti jääneet tunnistamatta negatiivisiksi. VISK:n mukaan (§ 902) intransitiiviset kvanttorilauseet esiintyvät myös *löytyä*-verbin yhteydessä

ja niillä on eksistentiaalilauseen piirteitä. Jätimme kvanttorilauseet pois tarkastelusta, sillä niissä ei yleensä ole subjektin sijavaihtelua vaan (e-)subjektit esiintyvät partitiivissa. Intransitiivisen kvanttorilauseen tuntomerkkejä näkyy erityisesti SV-lauseissa (4), joissa partitiivimuotoista subjektia (e-subjektia) seuraavat predikaatti ja määränilmaus:

- (4) [...] ruumiita oli torstaiamuun mennessä löytynyt 41. (KLLK_FI_1994)

VS-vastineiden (5) teemana on yleensä elatiivi- tai ablatiivimuotoinen adverbiaali myös kvanttorilausekkeiden kanssa:

- (5) Aasiasta löytyy paljon esimerkkejä. (KLLK_FI_1999)

Monessa esimerkkilauseessa *löytyä*-verbi esiintyy 1. infinitiivinä (6) modaaliverbin kanssa tai 1. partisiippina (7), monesti verbien *saattaa* (8), *arvella*, *näyttää* tai *uskoa* kanssa:

- (6) Muutamille voi löytyä virka myös yläasteelta. (KLLK_FI_2000)
- (7) Paras oivallus valmistajalta ovat kuitenkin vakiona autosta löytyvät kuormakiskot. (KLLK_FI_1999)
- (8) Myöhemmin rakennuksesta saattaa kuitenkin löytyä merkittävä homeongelma. (KLLK_FI_1999)

Yllä mainituissa lauseissa *löytyä* ei ole pääverbinä. Kategoriaan ”Muut ongelmalliset” kuuluvat kaikki ne lauseet, jotka eivät syystä tai toisesta ole ymmärrettäviä: monesta puuttuu olennainen osa tai sanat ovat automaattisen tekstien digitoimisprosessin virheiden vuoksi fragmentoituneita. Tuhannesta esimerkkilauseesta vain yhdessä esiintyy *löytyä*-verbin käyttö 2. persoonassa (9).

- (9) Mistä löydyt prinssini, 44–50-v. fiksu, pitkäkö, ulkonäkö ok, pilke silmäkulmassa, lenkkeilet ja tanssit. (KLLK_FI_1997)

Ongelmallisia tapauksia oli myös sen tyyppisissä lauseissa (10), joissa koko lauseen näkökulmasta tarkoitteeltaan monikolliseksi käsitettävä, subjektina toimiva substantiivilauseke koostuu kahdesta yksikön nominatiivimuodossa olevasta substantiivista. Analyysin perusteella

tiedämme kuitenkin, että verbin monikkomuotoisuus edellyttää aina tarkoitteeltaan monikolliseksi käsitettävää subjektia.

- (10) Äiti ja lapsi löytyivät surmattuina torstaina iltapäivällä Nivalassa. (KLLK_FI_1995)

Lopullisessa aineistossa nominatiiviesiintymiä on 455 kappaletta ja partitiiviesiintymiä 324 kappaletta.

3.2. Menetelmät

Tarkastelemme *löytyä*-verbillisten lauseiden subjektien sijanvalintaa sekä määrällisin että laadullisin menetelmin. Määrällisessä osassa tarkastelemme monimuuttujaisen logistisen sekamallin (ks. esim. Gries 2015) avulla sitä, mitkä edellä käsitellyt, aiemman tutkimuskirjallisuuden perusteella identifioidut lauseen eksistentiaalisuuden tulkintaa ohjaavat muuttujat vaikuttavat sijanvalintaan joko yksin tai vuorovaikutuksessa keskenään. Tavoitteena on siis kuvata ja mallintaa *löytyä*-verbillisten lauseiden subjektien sijavaihteluun vaikuttavia osatekijöitä, ei sitä, mikä on *löytyä*-verbille ylipäätään tyypillistä ja mikä ei. Kaikki muuttujat esitellään **Taulukossa 2** (tämän alaluvun lopussa).

Logistinen sekamalli tarkoittaa regressioanalyysia, jossa kaksi-arvoista riippuvaa muuttujaa – tässä tutkimuksessa subjektin sijanvalintaa – ennustetaan sekä kiinteillä muuttujilla että satunnaismuuttujilla. Kiinteiden muuttujien mahdolliset arvot tulevat ennalta tiedetystä joukosta (esim. subjektin ja verbin välinen sanajärjestys) kun taas satunnaismuuttujien arvot valikoituvat suuremmasta, ennalta määrittämättömästä joukosta mahdollisia arvoja. Tämän tutkimuksen muuttujista subjektin lemma on ainoa satunnaismuuttuja, sillä muiden muuttujien kaikki mahdolliset arvot on määritelty etukäteen. Regressioanalyysissa on kaksi selvää etua perinteiseen, yksittäisiä muuttujia ja niiden merkitystä kutakin erikseen tarkastelemaan lähestymiseen. Yhtäältä se vakioi muuttujat suhteessa toisiinsa niin, että kunkin muuttujan vaikutusta ja merkitystä voidaan tarkastella suhteessa muihin muuttujiin: miten muuttuja X vaikuttaa tarkastelunalaiseen ilmiöön, kun kaikki muut muuttujat pysyvät samoina. Toisaalta se mahdollistaa myös muuttujien välisen vuorovaikutuksen tarkastelun: vaikuttaako muuttuja X tarkastelunalaiseen ilmiöön eri tavoin silloin, kun muuttujan Y arvo muuttuu? On syytä korostaa,

että monien tarkastelunalaisten muuttujien – kuten subjektin jaollisuuden ja luvun suhde – vaikutus sijavaihteluun on kuvattu aiemmassa tutkimuksessa huolellisesti ja tunnetaan hyvin. Näiden huomioiminen osana tilastollista analyysia on kuitenkin ensiarvoisen tärkeää, jotta eri muuttujien ja niiden välisen vuorovaikutuksen vaikutus sijamerkintään voidaan mallintaa luotettavasti. Tähän perustuu myös aineiston edellä (ks. 3.1) kuvattu, VS- ja SV-sanajärjestyksellisiä *löytyä*-lauseita tasapainoisesti sisältävä satunnaisotos. Satunnaisotos kaikista *löytyä*-verbillisistä lauseista kuvaisi näiden lauseiden käyttöä ylipäätään, mutta se ei kuitenkaan mahdollistaisi harvinaisempien käyttökontekstien vaikutusten luotettavaa mallintamista.

Tutkimusasetelmamme on luonteeltaan konfirmatorinen eikä eksploratorinen, ja kaikki malliin sisällytetyt muuttujat samoin kuin niiden mahdolliset keskinäiset vuorovaikutussuhteet ovat aiemman tutkimuksen motivoimia, mistä syystä emme pyri optimoimaan mallia muuttujanvalintamenetelmien avulla (ks. Winter 2020: 277–279). Emme myöskään tarkastele mallin tilastollista merkitsevyyttä kokonaisuudessaan vaan kuvaamme sen ennustavuutta niin sanotun R-neliön avulla. R-neliö kuvaa sitä, kuinka hyvin riippuvan muuttujan vaihtelu – tässä tutkimuksessa subjektin sijamerkintä – voidaan ennustaa riippumattomien muuttujien avulla. R-neliön arvot vaihtelevat 0:n ja 1:n välillä niin, että 1 indikoi mallin riippumattomien muuttujien ennustavan vaihtelun täydellisesti ja 0 ei lainkaan. Tarkastelemme kuitenkin kunkin muuttujan tilastollista merkitsevyyttä vertaamalla ANOVA-testin avulla käyttämäämme mallia sellaiseen, josta kulloinkin kyseessä oleva muuttuja on jätetty pois (ks. Winter 2020: 260–263). Vakiintuneen käytännön mukaisesti tilastollisen merkitsevyyden kynnyksarvo on $p=0.05$. Kaikki tilastolliset analyysit on tehty R-ohjelmointiympäristössä (R Core Team 2018), ja sekamalleissa olemme käyttäneet `lme4`-pakettia (Bates ym. 2015), R-neliön laskemiseen `MuMIn`-pakettia (Bartoń 2018) ja mallia havainnollistavien kuvaajien tekoon `effects`-pakettia (Fox & Hong 2009).

Lopullinen mallimme on seuraava:

```
SUBCASE ~ SUBQUAN +
          SUBNUM +
          SUBQUAN:SUBNUM +
          AGREEMENT +
```

WO +
 SUBPOS +
 AGREEMENT:WO +
 SUBPOS:WO +
 TENSE +
 (1|SUBLEMMA)

Edellä kuvattujen muuttujien yksittäistarkastelun lisäksi tarkastelemme siis myös subjektin jaollisuuden ja subjektin luvun välistä vuorovaikutusta, verbin ja subjektin välisen kongruenssin ja subjektin ja verbin välisen sanajärjestyksen välistä vuorovaikutusta sekä subjektin sanaluokan ja subjektin ja verbin välisen sanajärjestyksen välistä vuorovaikutusta. Nämä vuorovaikutukset ovat kieliopillisesti perusteltuja: subjektin semanttinen jaollisuus on sidoksissa lukuun niin, että prototyypillisesti jaolliset eivät tyypillisesti esiinny monikkomuotoisina (VISK §555). Vastaavasti niin kongruenssi kuin sanajärjestyksen kuvaavat nimenomaan verbin ja subjektin luvun välistä suhdetta, minkä lisäksi niin kongruoimattomuutta kuin VS-sanajärjestystäkin on tyypillisesti pidetty partitiivisijaisen subjektin kanssa eksistentiaalilauseen ominaispiirteinä (Hakanen 1972). Myös subjektin sanaluokan osalta etenkin pronominiin voidaan olettaa deiktisen tulkintansa vuoksi vaikuttavan lauseiden informaatorakenteeseen ja sen myötä sanajärjestykseen. Lisäksi nämä ilmiöt saattavat tahoillaan vaikuttaa lause-tyyppien sekoittumiseen vuorovaikutuksessa sanajärjestyksen kanssa (Ivaska 2011), mistä syystä niiden väliset mahdolliset vaikutussuhteet on syytä ottaa huomioon myös tilastollisessa analyysissä.

Tämän jälkeen keskitymme kuvaamaan laadullisesti näitä havaittuja ilmiöitä suhteessa aiempaan tutkimukseen. Pyrimme tällä tavoin vastaamaan tutkimuskysymykseen mahdollisimman kattavasti menetelmällisen triangulaation keinoin, käyttämällä peräkkäistä monimenetelmäisyyttä (engl. *sequential mixed methods*), jossa määrällisen analyysin tulokset ohjaavat laadullista analyysia (määrällisten ja laadullisten menetelmien yhdistämisestä ks. esim. Hashemi 2012).

Taulukko 2. Tilastollisessa analyysissä käytetyt muuttujat ja niiden esiintymämäärät aineistossa.

MUUTTUJA	ARVOT	LÄHDE
SUBCASE: subjektin sija	nom (nominatiivi): 455 part (partitiivi): 324	Korpuksen annotaatio (tarkastettu)
SUBQUAN: subjektin jaollisuus	count (jaoton): 603 uncount (jaollinen): 176	Manuaalinen annotaatio
SUBNUM: subjektin luku	sg (yksikkö): 495 pl (monikko): 284	Korpuksen annotaatio (tarkastettu)
VERBNUMBER: verbin luku	sg (yksikkö): 687 pl (monikko): 92	Korpuksen annotaatio (tarkastettu)
AGREEMENT: kongruenssi	yes (subjekti ja verbi samassa luvussa): 575 no (subjekti ja verbi eri luvussa): 204	Korpuksen annotaatio (tarkastettu)
SUBPOS: subjektin sanaluokka	N (substantiivi): 724 A (adjektiivi): 15 Pron (pronomini): 30 V (verbi): 10	Korpuksen annotaatio (tarkastettu)
WO: sanajärjestys	SV (subjekti ennen verbiä): 387 VS (verbi ennen subjektia): 392	Korpuksen annotaatio (tarkastettu)
TENSE: aikamuoto	PRES (preesens): 554 IMPERF (preteriti): 187 PERF (perfekti): 38	Korpuksen annotaatio (tarkastettu)
SUBLEMMA: subjektin lemma	satunnaismuuttuja, jolla aineistossa 568 eri arvoa	Korpuksen annotaatio (tarkastettu)

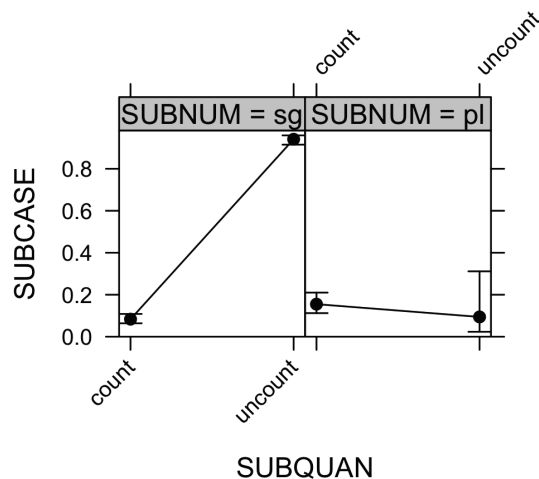
4. Analyysi

Tilastollinen mallimme onnistuu kuvaamaan subjektin sijavalinnassa esiintyvän vaihtelun melko hyvin ($R^2_{\text{marginaalinen}}=0.65$, $R^2_{\text{ehdollinen}}=0.73$). Marginaalinen R-neliö kuvaa mallin kiinteiden muuttujien ennustavuutta ja ehdollinen R-neliö kiinteiden ja satunnaismuuttujan yhteistä ennustavuutta. Muuttujien keskinäisen riippuvuuden mittari (engl.

Variance Inflation Factor) ei myöskään indikoi muuttujien liiallista multikollinearisuutta, sillä kaikki arvot ovat selvästi alle yleisenä kynnysarvona pidetyn 10:n. Näin ollen voimmekin oikeutetusti siirtää tarkastelemaan tuloksia kunkin muuttujan osalta.

4.1. Subjektin jaollisuus ja luku

Täyden mallin vertaaminen malliin, jossa subjektin jaollisuudella ja luvulla ei ole vuorovaikutussuhdetta, osoittaa, että vuorovaikutus on tilastollisesti erittäin merkitsevä (LRT $X(1)=17.056$, $p<0.0001$). Edelleen, kun tätä mallia verrataan malleihin joista nämä muuttujat on vuorollaan jätetty pois, nähdään, että jaollisuudella (LRT $X(1)=282.14$, $p<0.0001$) on myös itsessään tilastollisesti erittäin merkitsevä vaikutus subjektin sijanvalintaan, kun taas luvulla ei ole (LRT $X(1)=0.1624$, $p=0.6869$).



Kuvio 1. Subjektin jaollisuuden ja luvun suhde subjektin sijamerkinäisyys¹.

¹ Kuviot 1–4 kuvaavat tarkasteltujen muuttujien vaikutusta *löytyä*-verbillisten lauseiden subjektin sijanvalintaan. Pystyakselin arvot vaihtelevat välillä 0–1 niin, että pienemmät arvot tarkoittavat todennäköistä nominatiivisubjektia ja suuremmat arvot todennäköistä partitiivisubjektia. Kuvion 1 vasemmanpuoleisesta paneelista näkyy, että yksikkömuotoisten subjektien (SUBNUM = sg) kohdalla subjektin jaollisuus (uncount) lisää partitiivisijaisuuden todennäköisyyttä kun taas monikkomuotoisten subjektien (SUBNUM = pl) kohdalla näin ei ole.

Kuten **Kuvio 1** havainnollistaa, prototyypisesti jaottomat (tai *laskettavat*, ‘count’) subjektit ovat yksikössä lähes poikkeuksetta nominatiivi-muotoisia (y-akselin arvo on lähellä 0:aa), ja subjektin prototyypinen jaollisuus (*ei-laskettavuus*, ‘uncount’) lisää partitiivin todennäköisyyttä. Subjektin monikollisuus lisää partitiivin todennäköisyyttä prototyypisesti jaottomien osalta (esim. *sanat-sanoja*), ja tällaisia monikomuotoja onkin perinteisesti pidetty jaollisina. Kun monikollinen jaoton subjekti on partitiivissa, tilanne voidaan tulkita joko kollektiivisena tai distributiivisena (Huumo 2006, Vilkuna 1992). Lause (11) voi mahdollisesti olla kumpaakin: kollektiivinen, sillä parkkipaikat ovat kaikki samanaikaisesti olemassa; distributiivinen, sillä niitä vapautuu ja täyttyy jatkuvasti, eikä koskaan ole sama parkkipaikka vapaana verrattuna esimerkiksi aikaisempaan tuntiin. Kollektiivisessa tulkinnassa viitataan pääasiassa parkkipaikkojen olemassaoloon eikä siihen, että ne ovat vapaina toiminnan hetkellä. Lause (12) luetaan taas pääosin distributiivisena, sillä lauseen lopussa oleva adverbiaali antaa ymmärtää, että osallistujia on monenlaisiin erilaisiin kohteisiin sekä mahdollisesti peräkkäisinä hetkinä.

(11) Mutta myös parkkipaikkoja löytyy konserttipaikan vierestä. (KLK_FI_1997)

(12) Dharman kampanja alkaa hyvin, avustajia löytyy kaikkiin tehtäviin. (KLK_FI_1999)

Jaollisten subjektien kohdalla (esim. *rahat-rahoja*) subjektin monikollisuus ei lisää partitiivin todennäköisyyttä. Tällaisia tapauksia on aineistossa ainoastaan kymmenen, ja nähdäksemme nämä tapaukset ovat esimerkkejä tilanteista, joissa prototyypisesti jaollisten tarkoitteiden jaottomuutta ilmaistaan nimenomaan lukumerkinnän avulla – niin, että ne kuvaavat esimerkiksi laskettavien alkioiden joukkoa (13).

(13) Poliisi on vihdoin päässyt murtomiehen jäljille, ja Ukon rahatkin löytyvät. (KLK_FI_1995)

Sanan *raha* kaltaiset tapaukset voidaan ymmärtää myös puhtaan jaottomaksi (esim. *kolikko*), mutta olemme päättäneet käsitellä aineiston *raha*-tyyppisiä sanoja jaollisina, jotta soveltamamme tilastollinen menetelmä heijastaisi tätä edellä kuvattua vaihtelua: nähdäksemme

kyseessä on prototyypillisesti jaollisten tarkoitteiden joukko, jonka semanttista tulkintaa ohjataan numeruksen avulla tunnusmerkkiseen jaottomaan (laskettavaan) suuntaan – ja näin sijamerkintä, jaollisuus ja luku muodostavat vuorovaikutteisen kokonaisuuden. Sama koskee myös sanoja *työ* (14), *ruoka* (15) tai *vesi* (16).

(14) Hänen tavoitteenaan on olla Kaliforniassa ainakin puoli vuotta ja jos töitä löytyy, pitempäänkin. (KLLK_FI_1993)

(15) Sadan erilaisen ohjeen joukosta löytyy kylmiä ruokia, salaatteja, keittoja, patoja ja risottoja, kastikkeita, laatikoita, paistoksia, murekeruokia, fileitä ja annospaloja. (KLLK_FI_1999)

(16) Mistä löytyy sulimmat uimavedet? (KLLK_FI_1996)

Vastaavasti prototyypillisesti laskettava, jaoton yksiköllinen subjekti kuvaa monikossa jaollista paljoutta, joka voi olla indefiniittinen ja ilmaista partitiivissa avokvantiteettia (17) tai definiittinen ja ilmaista nominatiivissa sulkeiskvantiteettia (18). Katsoimme silti tarpeelliseksi säilyttää määrällisen analyysin muuttujien koodauksessa eri paljouksien yksittäisten tarkoitteiden prototyypillisen jaottomuuden (count), jotta näkisimme, millaisessa vuorovaikutuksessa se on subjektin luvun kanssa. Monikollinen nominatiiviviittaus voi joka tapauksessa kohdistua myös jaottomaan kohteeseen (ks. Larjavaara 2019: 184) tai määrittää jaotontarkoitteista *paria*, *settiä* tai *nippua* (ks. myös 4.2, jossa näiden roolia analysoidaan sanajärjestyksen kannalta): käsittäksemme paljous- ja settitulkinnan välinen raja ei kuitenkaan ole aina aivan selvä.

(17) Muita kuvauspaikkoja on löytynyt muun muassa Suomenlinnasta (...). (KLLK_FI_1999)

(18) Puhelinnumerot löytyvät jäsenkirjeestä. (KLLK_FI_1996)

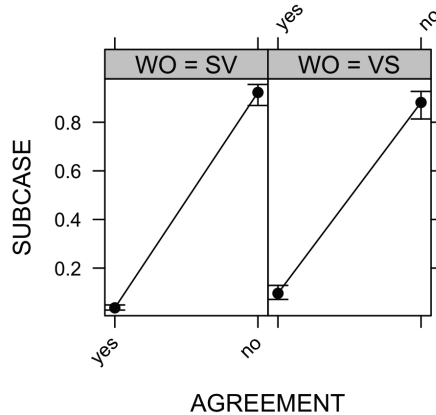
Tarkasteltaessa laadullisesti niitä tapauksia, joissa prototyypillisesti jaoton subjekti esiintyy yksikössä partitiivimuotoisena, huomio kiinnittyy lauseiden interrogatiivisuuteen ja siihen liittyvään tarkoitteiden avoimuuteen: esimerkin (19) *hienoa herraa* ja esimerkin (20) *somaa, muodokasta naista* ovat kumpikin tarkoitteiltaan sekä epäspesifejä että indefiniittisiä, mitä tulkintaa partitiivisijaisuus ilmaisee.

- (19) Löytyisikö varakasta, hienoa herraa, joka olisi tarpeessa edustuskel-
poisesta, sivistyneestä naisesta, ehkä kutsuilla, matkoilla, työasioista,
apukättä. (KLLK_FI_1999)
- (20) Löytyykö rehell., omlll. toim. tulevaa, 50–57-v., somaa, muodokasta
naista, et tupakoi etkä liioin kapakoi. (KLLK_FI_1999)

Tulokset ovat monin osin odotuksenmukaiset ja seuraavat aiempaa tutkimusta, sillä jaollisuus ja luku ovat luonnollisesti toisiinsa kietoutuneita syntaktis-semanttisia ilmiöitä. On kuitenkin huomionarvoista, että subjektin prototyypinen jaollisuus lisää partitiivisubjektin todennäköisyyttä vain yksiköllisenä. Tämän voidaan katsoa tarkoittavan sitä, että monikollisuus muuttaa automaattisesti subjektin semanttista kategoriaa. Vastaavasti yksikkömuotoisten prototyyppisesti jaottomien subjektien partitiivisijaisuus näyttäytyy keinona, jolla ohjataan tarkoitteiden semanttista tulkintaa epäspesifiksi ja indefiniittiseksi.

4.2. Kongruenssi ja sanajärjestys

Kongruenssin ja sanajärjestyksen välinen vuorovaikutus on tilastollisesti merkitsevä (LRT $X(1)=6.9673$, $p=0.008301$). Tämän mallin vertaaminen edelleen malliin ilman kongruenssia osoittaa lisäksi, että kongruenssin vaikutus subjektin sijanvalintaan on tilastollisesti erittäin merkitsevä myös itsessään (LRT $X(1)=195.61$, $p<0.0001$). Myös sanajärjestys itsessään vaikuttaa tilastollisesti merkitsevästi subjektin sijanvalintaan (LRT $X(5)=11.671$, $p=0.03959$), mutta tulos ei ole täysin verrannollinen, sillä sanajärjestyksen poistaminen mallista vaikuttaa myös sen vuorovaikutukseen subjektin sanaluokan kanssa (ks. 4.3). Kuten **Kuviosta 2** nähdään, partitiivisubjektit ovat huomattavasti todennäköisempiä silloin, kun subjekti ja verbi eivät kongruoi (y-akselin arvo on lähellä 1:ä, mikä tarkoittaa, että se on useimmiten partitiivissa). Kuten odottaa saattaa, myös VS-sanajärjestys lisää partitiivisubjektin todennäköisyyttä edelleen. Nämä molemmat seikat osoittavat sen, miten molemmat muuttujat osaltaan ohjaavat kohti eksistentiaalista tulkintaa. Tämä heijastanee lauseen eksistentiaalisuuden ilmaisemiseen liittyvää rajankäyntiä: kun yksi tyyppi-irteistä – kuten inkongruenssi eli subjektin ja verbin erilukuisuus – täyttyy, muilla tyyppi-irteillä – kuten sanajärjestyksellä – on vähemmän merkitystä.



Kuvio 2. Kongruenssin ja sanajärjestyksen suhde subjektin sijamerkintään.

Verbin kanssa kongruoimaton monikollinen nominatiivisubjekti esiintyy yhteensä 20 esimerkkilauseessa: kyse on nähdäksemme ns. *pareista* (21), *seteistä* (22) tai *nipuista* (ks. Larjavaara 2019: 184), minkä takia inkongruenssi on mahdollinen, mutta VS-järjestyksellisissä lauseissa luontevampi kuin SV-järjestyksellisissä.

(21) *Tanssikengät löytyy.* (KLK_FI_1998)

(22) *Tuotteet löytyy myös rannasta.* (KLK_FI_1995)

Kongruoimattomia VS-lauseita, joissa verbinjälkeinen nominatiivimuotoinen e-subjekti on monikollinen, voidaan pitää sijamerkinnän osalta epäprototyypisinä eksistentiaalilauseina (vrt. VISK §893): tämän tulkinnan mahdollistaa lauseenalkuinen adverbiaali, jonka jälkeen predikaatti esiintyy melkein aina kolmannessa persoonassa. Vaikka ne ilmaisevat reeman paikalla olevaa uutta referenttiä, nominatiivimuotoinen e-subjekti indikoi tunnistettavaa tarkoitetta sekä sulkeista määrää (tunnistettavuudesta ks. Vilkuna 1992). Esimerkissä (23) tätä tunnistettavuuden määrittämistä auttaa myös relatiivilauseen antama konteksti (ks. tästä myös 4.3). Lauseen (24) adverbiaalilla *ratkaisuun* on sama rajoittava funktio: kaikista mahdollisista *perusteista* otetaan huomioon ainoastaan sulkeinen setti juuri siihen ratkaisuun tarkoitettuja perusteita.

(23) Sieltä löytyy myös sosiaalitulat, jonne kaveriporukka tyttöystävineen kerääntyy viikonlopun viettoon. (KLK_FI_1998)

(24) [...] ratkaisuun löytyi perusteet ulkomaalaislaista. (KLK_FI_1995)

Lause (25) sisältää sekä nominatiivi- että partitiivisubjektin. Nominatiivisubjektin tarkoitteen sulkeismääräisyys on vielä ilmeisempi tässä, jos sitä verrataan partitiivisubjektin ilmaisemaan avomäärään.

(25) Siitä löytyy myös urheiluseurojen ja kylätoimikuntien yhteistiedot sekä joitakin maalaiskunnan nähtävyyksiä. (KLK_FI_1995)

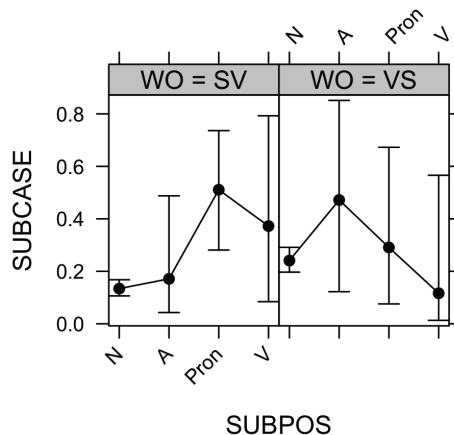
Yhteenvetona sanajärjestys yhdessä (in)kongruenssin kanssa ohjaa monesti lauseen eksistentiaalisuuden tulkintaa. On kiinnostavaa, että definiittisiä ja tarkoitteeltaan tunnistettavia, monikollisia nominatiivisubjekteja esiintyy myös eksistentiaalilauseissa (ks. myös Larjavaara 2019: 185), vaikka kyseisissä konstruktioissa reeman paikka on yleensä varattu yksikön nominatiivissa tai yksikön ja monikon partitiivissa esiintyvälle uusille, indefiniittisille tarkoitteille.

4.3. Subjektin sanaluokka ja sanajärjestys

Subjektin sijamerkintä korreloi subjektin sanaluokan ja sanajärjestyksen kanssa niin, että etenkin pronominisubjektit ovat muita useammin partitiivimuotoisia silloin, kun ne edeltävät verbiä (ks. **Kuvion 3** vasen paneeli, y-akseli kuvaa partitiivisubjektin todennäköisyyttä). Muiden kuin substantiiviedussanaisten subjektien pienen määrän vuoksi sen kummemmin sanaluokan ja sanajärjestyksen vuorovaikutus ($LRT X(3)=2.0077, p=0.5708$) kuin sanaluokka itsessään ($LRT X(6)=4.4454, p=0.6166$) ei kuitenkaan ole tilastollisesti merkitsevä. On silti huomionarvoista, että etenkin pronominisubjektien (nom=12, part=18) kohdalla eksistentiaalilauseelle ominaisia partitiivisubjekteja on suhteellisesti enemmän kuin substantiiviedussanaisissa subjekttilausekkeissa (nom=431, part=297). Toisaalta pronominisubjektit (SV=22, VS=8) esiintyvät suhteellisesti substantiiviedussanaisia subjekttilausekkeitä (SV=350, VS=374) useammin normaalilauseelle tyypillisessä verbinetisessä asemassa. Esimerkit (26)–(29) havainnollistavat aineiston pronominisubjektillisiä lauseita niin nominatiivissa kuin partitiivissakin.

- (26) Se löytyy tämän hallituksen kautta. (KLIK_FI_1993)
- (27) Niitä löytyy myös Sri Lankasta. (KLIK_FI_1996)
- (28) Varmasti löytyy joku, joka lähtee sitä viljelemään. (KLIK_FI_1995)
- (29) Löytyykö sitä temppuun ilmoitetulta Santeri Ampujalta? (KLIK_FI_1998)

Havainto on mielenkiintoinen sen seikan valossa, että eksistentiaalisten lauseiden subjekteille on havaittu tyypilliseksi se, että ne esittelevät uusia tarkoituksia, joihin ei yleensä viitata toistamiseen. Nämä ominaisuudet ovat tyypillisiä nimenomaan substantiiviedussanaisille subjekteille (Helasvuo 1996, 2001), jotka kuitenkin esiintyvät tässä aineistossa pronominusubjekteja yleisemmin perussubjektille ominaisessa nominatiivissa. Voidaankin kysyä, halutaanko *löytyä*-lauseiden tulkintaa ohjata eksistentiaaliseksi subjektin sijamerkinän avulla silloin, kun subjekti on eksistentiaalilauseelle epätyypillinen pronomini. Tätä tulkintaa puoltaa se, että ero näkyy nimenomaan SV-sanajärjestyksellisissä lauseissa, jotka siis niin ikään ovat lähtökohtaisesti epäprototyypisiä eksistentiaalilauseelle. SV-sanajärjestyksen yleisyys pronominisubjektillisissa *löytyä*-lauseissa liittyy eksistentiaalilauseen informaatorakenteeseen: konstruktion teemapaikassa esiintyy todennäköisesti edeltävään kontekstiin viittaava ja siksi tunnettu elementti. Näin ollen lauseen eksistentiaalista tulkintaa ilmaistaan useammin subjektin partitiivisijaisuudella.



Kuvio 3. Subjektin sanaluokan ja sanajärjestyksen suhde subjektin sijamerkintään (N = substantiivi; A = adjektiivi; Pron = pronomini; V = verbi).

Samalla on kiinnostavaa huomata, kuinka partitiivi vaikuttaa kvantiteettiin pronomien yhteydessä. Esimerkissä (30) relatiivilause määrittää edeltävää pronomia *niitä*, joka toimii e-subjektina: partitiivi ilmaisee avokvantiteettia, mutta kokonaisuus, johon pronomini viittaa, on sulkeinen joukko.

- (30) Aina löytyy niitä, joiden mielestä ruoho on vihreämpää aidan takana.
(KLK_FI_1993)

Toisin sanoen kuvitellaan, että on olemassa tietty sulkeinen ryhmä, joka koostuu kaikista ihmisistä, jotka luulevat ruohon olevan vihreämpää aidan takana – partitiivisijaisella pronominilla viitataan moniin kyseiseen ryhmään kuuluviin ihmisiin, jotka koostavat puolestaan toisen indefiniittisen, avoimen joukon. Lauseenalkuinen aika-adverbiaali mahdollistaa lisäksi tilanteen iteratiivisen tulkinnan, joka on tyypillinen niin partitiivisubjekteille kuin eksistentiaalilauseillekin ylipäättään. Seuraavassa alaluvussa pohdimme lisää iteratiivisuudesta sekä muista aspektiin liittyvistä ominaispiirteistä.

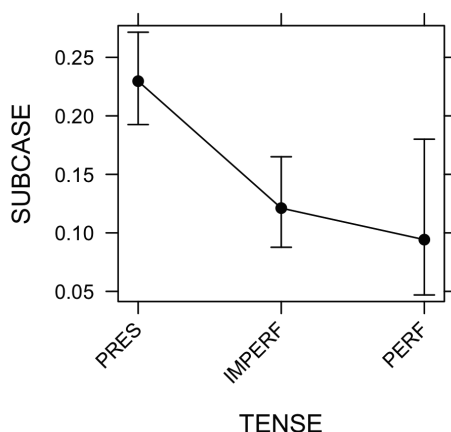
4.4. Aikamuoto

Myös aikamuodolla on tilastollisesti merkitsevä rooli *löytyä*-lauseiden subjektin sijavalinnassa (LRT $X(2)=7.5495$, $p=0.02294$). Kuten **Kuvio 4** havainnollistaa, preesensmuotoisten lauseiden subjektit ovat muita aikamuotoja useammin partitiivimuotoisia (y-akseli kuvaa partitiivisubjektin todennäköisyyttä). Tämä on osin odotuksenmukaista: tilanteiden aikarakenne ja subjektin jaollisuus kietoutuvat yhteen (VISK §1522), preteriti on tyypillinen nimenomaan narratiivisessa käytössä, jossa se kuvaa peräkkäisiä, puhehetken kannalta menneitä, määräisiä tapahtumia (VISK §1531; §1538), kun taas preesens kuvaa usein pikemminkin pelkkää olemassaoloa. Aikamuodon voidaan siis katsoa vaikuttavan *löytyä*-verbin semantiikkaan erityisesti aspektin ja teonlaadun kannalta: preteritillä kuvataan jaotonta tilannetta. *Löytyä*-verbi on aspektiltaan teelinen ja resultatiivinen sekä yleensä implikoi ennen löytymistä tapahtunutta etsintää, etenkin preteritissä (31) tai perfektissä (32). Sitä voidaan semanttisesti siis pitää koostuvana kahdesta osasta, joista toinen on duratiivinen ja toinen punktuaalinen.

(31) Etsinnöissä hänen pyöränsä löytyi läheisen koulun pihasta. (KLLK_FI_1995)

(32) Jonkunlainen malli on löytynytkin. (KLLK_FI_1996)

On kuitenkin yllättävää, että perfekti näyttää lisäävän nominatiivi-muotoisen subjektin todennäköisyyttä – sillä sitä käytetään tyypillisesti ilmaisemaan aspektuaalista imperfektiivisyyttä (VISK §1537). Perfektin osalta ei kuitenkaan ole syytä tehdä kovin pitkälle meneviä johtopäätöksiä, sillä tapausten kokonaismäärä on melko pieni (n=38).



Kuvio 4. Aikamuodon suhde subjektin sijamerkintään.

Aikamuodolla on kiinnostavia vaikutuksia verbi- ja lausesemantiikkaan. Huumon (2003: 477) mukaan partitiivisijainen *e-theme* esiintyy teelisten verbien yhteydessä myös ilman lokatiivista adverbialia; toisaalta atelisten verbien yhteydessä pitää olla lokatiivi (vrt. **Vieraita tanssii*; *Vieraita tanssii lattialla*, *ibid.*). Hän selittää tämän ilmiön seuraavasti: partitiivisijaiset *e*-subjektit ohjaavat punktuaalis-iteratiiviseen tulkintaan, jossa yksittäisten referenttien jatkuvaa osallistumista ei edellytetä. Ateelisten verbien kohdalla referenttien on osallistuttava toimintaan jatkuvasti, joten punktuaalis-iteratiivinen tulkinta ei ole mahdollinen; mahdollinen lokatiivinen adverbialia tarjoaa tulkintaan jonkinlaisen näkökulmapaikan, joka toimii lähtöpisteenä lauseen merkitykselle – mahdollistaen näin partitiivin referenttien ajoittaisen vaihdon. Kun siis otetaan huomioon *löytyä*-verbin aspektin punktuaalinen osa,

tämä voi esiintyä iteratiivisessa merkityksessä eikä referenttien osallistuminen ole kontinuaatiivista, vaikka lokatiivista adverbialia ei olisi (33).

(33) Erojakin toki löytyy. (KLK_FI_2000)

Tässä *erojakin*-subjektin ymmärretään osallistuvan toimintaan yksi kerrallaan, ja sitä määrää verbin punktuaalinen tulkinta. Toisaalta lause voi myös yksinkertaisesti tarkoittaa, että eroja on olemassa: tästä näkökulmasta katsoen iteratiivisuus ei enää vaikuta lauseen kokonais-semantiikkaan. Kaiken kaikkiaan aikamuoto vaikuttaa *löytyä*-verbillisten lauseiden tulkintaan suhteessa aspektuaalisuuden eri tahoihin.

4.5. Subjektin lemma

Kuten analyysin alussa toteamme, subjektin lemman sisällyttäminen malliin nostaa sen ennustavuutta selvästi ($R^2_{\text{marginaalinen}}=0.65$ vs. $R^2_{\text{ehdollinen}}=0.73$). Vertaamalla lopullista mallia vastaavaan monimuuttujamalliin ilman satunnaismuuttujaa nähdään niin ikään, että tämä ero on tilastollisesti merkitsevä ($\text{LRT } X(1)=5.9604, p=0.01463$)². Eri sanoilla on siis erilaisia preferenssejä sen suhteen, missä muodossa niitä käytetään *löytyä*-verbillisten lauseiden subjektina. **Taulukko 3** listaa aineiston viisi tyypillisintä nominatiivi- ja partitiivisubjektin lemmaa. Kaikki tyypilliseksi nominatiivisubjekteiksi listatut lemmat esiintyvät aineistossa nominatiivissa vähintään kolme kertaa useammin kuin partitiivissa – ja tyypilliseksi partitiivisubjekteiksi listatut vastaavasti vähintään kolme kertaa useammin partitiivissa kuin nominatiivissa.

Kuten alla olevan **Taulukon 3** esimerkit osoittavat, nominatiivisijalle tyypillisten lemموjen tarkoitteet ovat seikkoja, joista ilmaistaan tyypillisesti niiden konkreettista tai metaforista sijaintia, kun taas partitiivisijalle tyypillisten lemموjen tarkoitteet ovat sellaisia, joille on tyypillistä niiden olemassaolon tai -olemattomuuden ilmaiseminen. Tämä vastaa monin tavoin konstruktiokieliopin keskiössä olevaa vuorovaikutusta konstruktion kokonaismerkityksen ja sen osien välillä

2 On osin epäselvää, voiko eri menetelmin sovitettuja tilastollisia malleja verrata yksiselitteisesti keskenään. Käyttämämme glmer-funktion dokumentaatio kuitenkin toteaa seuraavasti: “glmer and glm log-likelihoods are consistent”, mistä syystä glmer- ja glm-funktioiden avulla sovitettuja malleja voi nähdäksemme verrata näihin arvoihin perustavalla uskottavuusosamäärättestillä (engl. *likelihood ratio test*, LRT).

(Goldberg, Casenheiser & Sethuraman 2004). Lisäksi havainto tukee aiempia tutkimuksia lähimerkityksisten konstruktioiden välisen vaihtelun yhteydestä näiden konstruktioiden toisistaan eroaviin leksikaalisiin preferensseihin (Gries & Stefanowitsch 2004). Nähdäksemme eksistentiaalinen tulkinta korreloi subjektin semantiikan kanssa siten, että tietyt tarkoitteet ohjaavat tulkintaa eksistentiaaliseen suuntaan ja tietyt normaalilauseen suuntaan.

Taulukko 3. Aineiston viisi tyypillisintä nominatiivi- ja partitiivisubjektina esiintyvää lemmaa.

NOM		PART	
SYY (10)	<i>Syyt kissojen vihaamiseen löytyvät korvien välistä</i>	TYÖ (9)	<i>Töitä löytyy keittiön puolelta.</i>
RUUMIS (8)	<i>isän ruumis löytyi ratin takaa</i>	RAHA (6)	<i>palloiluun kyllä löytyy vain lisää rahaa.</i>
MIES (6)	<i>Kuollut mies löytyy puistonpenkiltä</i>	TIETO (6)	<i>Pankista löytyy siis tietoja lähes [...]</i>
PAIKKA (6)	<i>lähin [...] paikka löytyy [...] kirkonkylästä</i>	TILA (5)	<i>Keskustataajamasta löytyy kirjastolle tilaa.</i>
RATKAISU (5)	<i>Ratkaisu löytyy toisaalta pienistä avustusten nousuista.</i>	POIKKEUS (4)	<i>Poikkeuksiakin ammattikunnasta löytyy.</i>

5. Päätelmät

Tarkastelimme tässä tutkimuksessa *löytyä*-verbin sisältäviä lauseita niiden subjektijäsenen sijamerkintävaihtelun näkökulmasta. Tutkimuksemme keskiössä oli sijamerkinnän rooli ja vuorovaikutus muiden kielellisten ominaisuuksien kanssa tutkittaessa tällaisten lauseiden asemaa eksistentiaalilauseiden ja normaalilauseiden rajapinnalla. Tuloksemme osoittavat, että erilaiset muuttujat vaikuttavat sijamuotovalintaan subjektin osalta, mutta myös, että ne ovat usein vuorovaikutuksessa keskenään. Voidaan vahvistaa, että subjektin luku on vuorovaikutuksessa subjektin jaollisuuden kanssa: subjektin prototyypinen jaollisuus lisää partitiivin todennäköisyyttä. Monikossa partitiivi esiintyy todennäköisemmin jaottomien kuin jaollisten subjektien osalta. Monikollisuus

muuttaa subjektin semanttista kategoriaa, mikä tarkoittaa, että tarkoitteeltaan jaottomasta subjektista tulee kvantiteetiltaan avoin ainoastaan, kun kyseessä oleva monikollinen subjekti on merkitty partitiivilla: tarkoitteeltaan jaollisen subjektin ei tarvitse olla monikossa, jotta avoin kvantiteetti olisi ilmaistu. Päinvastoin, kun jaollinen subjekti esiintyy monikossa, sanan semantiikka muuttuu ja kyseessä on avoin kvantiteetti – esimerkiksi joukko (Larjavaaran 2019: 29 termein ”setti”) *vesiä, mehuja* tai *suklaita* (vrt. sulkeisiin vaihtoehtoihin *vedet, mehut*, tai *suklaat*). Sanajärjestys ei ole itsessään tilastollisesti merkitsevä subjektin sijamuotovalinnan kannalta, kun taas kongruenssi on, mikä johtune siitä, että monikkokongruenssi on partitiivin yhteydessä suorastaan epäkieliopillinen. Niiden vuorovaikutus on myös merkitsevä: subjektin ja verbin välinen inkongruenssi vaikuttaa partitiivin valintaan todennäköisemmin SV-järjestyksellisissä lauseissa. Nämä ovat tärkeitä muuttujia ohjaamassa kohti lauseen eksistentiaalista tulkintaa. Vaikka subjektin sanaluokka ei vaikuta tilastollisesti merkitsevältä, kun se on vuorovaikutuksessa sanajärjestyksen kanssa, huomasimme, että nominatiivisubjekteja on enemmän substantiivien joukossa, sillä ne tyypillisesti esittävät uusia tarkoitteita, kun taas pronominit viittaavat aiemman mainittuun tai ylipäätään jo tiettyyn tarkoitteeseen. VS-järjestys lisää partitiivin todennäköisyyttä substantiivien kanssa. Myös aikamuodon ja subjektin sijamerkinän välillä on tilastollisesti merkitsevä yhteys: *pre-sens* esiintyy partitiivin kanssa useammin kuin *menneisyyteen* viittaavat aikamuodot, sillä niiden kuvailemat menneet tapahtumat voidaan katsoa määräisiksi, puhujalle/kirjoittajalle jo tietyiksi. Menneen ajan muodoista perfekti esiintyy nominatiivissa preteritiä useammin, mutta havaintojen suhteellisen pieni määrä ei oikeuta tilastollisesti merkitsevien tulkintojen tekoon. Subjektin lemma on myös tilastollisesti merkitsevä, sillä tietyt sanat esiintyvät erityisesti normaalilauseen subjekteina niin, että lause ilmaisee niiden tarkoitteiden sijaintia, kun taas toiset sanat suosivat konteksteja, joissa kuvataan niiden olemassaoloa ylipäätään.

On edelleen epäselvää, mikä normaali- ja eksistentiaalilauseen välinen raja on monissa tapauksissa, mutta tutkimuksemme osoittaa, että valitsemiemme eri muuttujien välillä on vuorovaikutteinen suhde. *Löytyä*-verbillisten lauseiden konstruktio ei siis ole määriteltävissä ainoastaan sanajärjestyksen tai verbin ja subjektin välisen kongruenssin pohjalta. Jokaisella muuttujalla on oma vaikutuksensa subjektin sijamuodon valitsemiseen, ja nämä tekijät yhdessä ohjaavat yksittäisten

lauseiden tulkintaa enemmän tai vähemmän eksistentiaalisuutta ilmaisevaan suuntaan. Nähdäksemme tämä vastaa konstruktiokieliopin näkemystä siitä, että kunkin konstruktion sisäisessä vaihtelussa sen prototyyppisimmät variantit ovat myös lähinnä konstruktiomerkityksen mukaista tulkintaa (Goldberg, Casenheiser & Sethuraman 2004). On kiinnostavaa havaita, että riippumatta analysoiduista muuttujista usein nominatiivi- ja partitiivisubjektin vaihtelu johtaa juuri sulkeis- tai avokvantiteetin tulkintaan. Subjektin sijamuoto muuttaa tarkoitteen kvantiteettia, jolla on vaikutus lauseen informaatorakenteeseen, minkä voi huomata subjektin jaollisuuden, sanajärjestyksen sekä subjektin sanaluokan vuorovaikutteisesta suhteesta.

Semantiikan näkökulmasta *löytyä*-verbi korvaa monesti *olla*-verbin ja on tästä syystä kiinnostava verrattuna muihin suomenkielisissä eksistentiaalilauseissa esiintyviin verbeihin. Tämä semanttinen haalistuminen näkyy erityisesti prototyyppisissä eksistentiaalilauseissa, joissa uusi referentti tuodaan ”puhtaasti” esiin: tässä käyttöyhteydessä verbi voi kongruoida subjektin kanssa ja subjektin sijamuotovalinnalla ei ole niin tärkeä rooli, kun taas on tarpeellista, että lokatiiviadverbiaali on lauseen alussa. Normaaliilauseissa, eli yleensä kun lauseen subjekti on nominatiivissa ja sanajärjestys on suora (SV), *löytyä*-verbin semantiikka voi tietysti sisältää aspektuaalisia sävyjä kuten iteratiivisuutta tai viitata edeltävään, duratiiviseen etsintäprosessiin. Yleinen tendenssi sekä normaalilauseissa että eksistentiaalilauseissa on joka tapauksessa se, että näitä aspektuaalisia sävyjä ei säilytetä. Tämä näkyy esimerkiksi sellaisissa partitiivitapauksissa, joissa tulkinta on kollektiivinen. On kiinnostavaa huomata, että lokatiiviadverbiaalin olemassaolo ei sanajärjestyksestä riippumatta monesti vähennä *löytyä*-verbin semanttista haalistumista, vaikka sen suuntaisuus saattaa viitata jotenkin kompleksempaan semantiikkaan.

Yhtenä tämän tutkimuksen kannalta kiinnostavista uusista tutkimussuunnista olisi kääntää tutkimusasetelma osin päinvastaiseksi: mitkä verbit ylipäättään mahdollistavat partitiivisubjektin käytön? Olisi niin ikään mielenkiintoista paneutua syvemmin partitiivissa esiintyvien subjektien tyypillisiin semanttisiin piirteisiin. Nähdäksemme tutkimusessamme käytetyt menetelmät soveltuvat hyvin näiden kysymysten tarkasteluun. Regressioanalyysin avulla tarkasteltavaan ilmiöön samanaikaisesti vaikuttavat useat muuttujat voidaan vakioda niin, että niistä kutakin voidaan tarkastella erikseen ja vuorovaikutuksessa toistensa

kanssa ilman, että eri muuttujien vaikutukset peittävät alleen tai muuttavat toinen toisiaan. Me toivommekin tämän tutkimuksen kannustavan partitiivisubjektin mahdollistavien verbien ja tällaisten subjektien tyyppillisten semanttisten piirteiden selvittämiseen jatkossa.

Lähteet

- Bartoń, Kamil. 2018. *MuMIn: Multi-Model Inference*. <https://cran.r-project.org/web/packages/MuMIn/>.
- Bates, Douglas, Martin Mächler, Ben Bolker & Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1). 1–48. <https://doi.org/10.18637/jss.v067.i01>.
- Bentley, Delia, Francesco Maria Ciconte & Silvio Cruschina. 2013. Existential constructions in crosslinguistic perspective. *Italian Journal of Linguistics* 25(1). 15–43.
- Błaszczak, Joanna. 2009. Differential subject marking in Polish: The case of genitive vs. nominative subjects in ‘X Was Not at Y’-constructions. Helen de Hoop & Peter de Swart (toim.), *Differential subject marking*, 113–149. Dordrecht: Springer. https://doi.org/10.1007/978-1-4020-6497-5_6.
- Błaszczak, Joanna. 2012. A spurious genitive puzzle in Polish. T. Hanneforth & G. Fanselow (toim.), *Language and logos* (Studia Grammatica 72), 17–47. Berlin: Akademie Verlag. <https://doi.org/10.1524/9783050062365.17>.
- Breivik, Leiv Egil. 1999. On the rhetorical function of existential *there*. *Nordlit*. 3–14. <https://doi.org/10.7557/13.2125>.
- Creissels, Denis. 2014. Existential predication in typological perspective. *Ms., Université Lyon*. <http://www.deniscreissels.fr/public/Creissels-Exist.Pred.pdf>. (14.6.2021.)
- Cruschina, Silvio. 2012. Focus in existential sentences. Valentina Bianchi & Cristiano Chesi (toim.), *Enjoy linguistics! Papers offered to Luigi Rizzi on the occasion of his 60th birthday*, 77–107. Siena: CISCL Press.
- Fox, John & Jangman Hong. 2009. Effect displays in R for multinomial and proportional-odds logit models: Extensions to the *effects* package. *Journal of Statistical Software* 32(1). 1–24. <https://doi.org/10.18637/jss.v032.i01>.
- Goldberg, Adele E. 1995. *Constructions: A construction grammar approach to argument structure* (Cognitive Theory of Language and Culture). Chicago: University of Chicago Press.
- Goldberg, Adele E. 2006. *Constructions at work: The nature of generalization in language*. Oxford: Oxford University Press.
- Goldberg, Adele, Devin Casenheiser & Nitya Sethuraman. 2004. Learning argument structure generalizations. *Cognitive Linguistics* 15(3). 289–316. <https://doi.org/10.1515/cogl.2004.011>.
- Gries, Stefan Th. 2015. The most under-used statistical method in corpus linguistics: Multi-level (and mixed-effects) models. *Corpora* 10(1). 95–125. <https://doi.org/10.3366/cor.2015.0068>.

- Gries, Stefan Th. & Stefanie Wulff. 2013. The genitive alternation in Chinese and German ESL learners: Towards a multifactorial notion of context in learner corpus research. *International Journal of Corpus Linguistics* 18(3). 327–356. <https://doi.org/doi:10.1075/ijcl.18.3.04gri>.
- Gries, Stefan Th. & Anatol Stefanowitsch. 2004. Extending collocation analysis: A corpus-based perspective on ‘alternations’. *International Journal of Corpus Linguistics* 9(1). 97–129. <https://doi.org/10.1075/ijcl.9.1.06gri>.
- Hakanen, Aimo. 1972. Normaalilause ja eksistentiaalilause. *Sananjalka* 14(1). 36–76. <https://doi.org/10.30673/sja.86366>.
- Hashemi, Mohammad R. 2012. Reflections on mixing methods in applied linguistics research. *Applied Linguistics* 33(2). 206–212. <https://doi.org/10.1093/applin/ams008>.
- Helasvuo, Marja-Liisa. 1996. Ollako vai eikö olla – eksistentiaalilauseen subjektin kohtalonkysymys. *Virittäjä* 100(3). 340–356.
- Helasvuo, Marja-Liisa. 2001. *Syntax in the making: The emergence of syntactic units in Finnish conversation* (Studies in Discourse and Grammar 9). Amsterdam/Philadelphia: John Benjamins. <https://doi.org/10.1075/sidag.9>.
- Hu, Jianhua & Haihua Pan. 2008. Focus and the basic function of Chinese existential *you*-sentences. Ileana Comorovski & Klaus von Heusinger (toim.), *Existence: Semantics and syntax*, 133–145. Dordrecht: Springer. https://doi.org/10.1007/978-1-4020-6197-4_5.
- Huomo, Tuomas. 2003. Incremental existence: The world according to the Finnish existential sentence. *Linguistics* 41(3). 461–493. <https://doi.org/10.1515/ling.2003.016>.
- Huomo, Tuomas. 2006. Kvantiteetti ja aika I. Nominaalisen aspektin näkökulma suomen objektin ja subjektin sijamerkintään. *Virittäjä* 110(4). 504–538.
- Huomo, Tuomas & Marja-Liisa Helasvuo. 2015. On the subject of subject in Finnish. Marja-Liisa Helasvuo & Tuomas Huomo (toim.), *Subjects in constructions – canonical and non-canonical* (Constructional Approaches to Language 16), 13–41. Amsterdam: John Benjamins. <https://doi.org/10.1075/cal.16.02huu>.
- Huomo, Tuomas & Liina Lindström. 2014. Partitives across constructions: On the range of uses of the Finnish and Estonian ‘partitive subjects’. Silvia Luraghi & Tuomas Huomo (toim.), *Partitive cases and related categories*. Berlin, Boston: De Gruyter. <https://doi.org/10.1515/9783110346060.153>.
- Ikola, Osmo. 1954. Suomen lauseopin ongelmia [I-III]. *Virittäjä* 58(3). 209–244.
- Ivaska, Ilmari. 2010. Eksistentiaalilauseen subjektijäsen edistyneiden suomenoppijoiden kirjoituksessa. *Sananjalka* 52. 67–88. <https://doi.org/10.30673/sja.86697>.
- Ivaska, Ilmari. 2011. Lausetyyppien sekoittuminen edistyneessä oppijansuomessa – näkökulmana eksistentiaalilause. *Lähivõrdlusi. Lähivertailuja* 21. 65–85. <https://doi.org/10.5128/lv21.03>.
- Klavan, Jane. 2020. Pitting corpus-based classification models against each other: a case study for predicting constructional choice in written Estonian. *Corpus Linguistics and Linguistic Theory* 16(2), 363–391. <https://doi.org/10.1515/cllt-2016-0010>.
- Langacker, Ronald W. 2008. *Cognitive grammar*. Berlin: De Gruyter Mouton. <https://doi.org/10.1515/9783110199901.29>.
- Larjavaara, Matti. 1988. Suomen kvantitatiivinen spesies. *Virittäjä* 92(4). 469–503.

- Larjavaara, Matti. 2019. *Partitiivin valinta*. Helsinki: Suomalaisen Kirjallisuuden Seura.
- Lees, Aet. 2015. *Case alternations in five Finnic languages: Estonian, Finnish, Karelian, Livonian and Veps*. Leiden: Brill. <https://doi.org/10.1163/9789004296367>.
- Lindström, Liina. 2017. Partitive subjects in Estonian dialects. *Eesti ja soome-ugri keeleteaduse ajakiri. Journal of Estonian and Finno-Ugric Linguistics* 8(2). 191–231. <https://doi.org/10.12697/jeful.2017.8.2.07>.
- Luraghi, Silvia & Seppo Kittilä. 2014. Typology and diachrony of partitive case markers. Silvia Luraghi & Tuomas Huumo (toim.), *Partitive cases and related categories* (Empirical Approaches to Language Typology 54), 17–62. Berlin, Boston: De Gruyter Mouton. <https://doi.org/10.1515/9783110346060.17>.
- McNally, Louise. 2016. Existential sentences crosslinguistically: Variations in form and meaning. *Annual Review of Linguistics* 2. 211–231. <https://doi.org/10.1146/annurev-linguistics-011415-040837>.
- Milsark, Gary L. 2014. *Existential sentences in English* (RLE Linguistics D: English Linguistics). London: Routledge. <https://doi.org/10.4324/9781315856728>.
- Partee, Barbara H. & Vladimir Borschev. 2008. Existential sentences, BE, and the genitive of negation in Russian. Ileana Comorovski & Klaus von Heusinger (toim.), *Existence: Semantics and syntax*, 147–190. Dordrecht: Springer. https://doi.org/10.1007/978-1-4020-6197-4_6.
- R Core Team. 2018. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>.
- Sawyer, Janet. 1973. Existential sentences: A linguistic universal? *American Speech* 48(3/4). 239–245. <https://doi.org/10.2307/3087831>.
- Seržant, Ilja A. 2014. The independent partitive genitive in North Russian. Björn Wiemer & Ilja Seržant (toim.), *Contemporary approaches to dialectology: the area of North, Northwest Russian and Belarusian vernaculars* (Slavica Bergensia 12), 270–329. Bergen: University of Bergen.
- Stefanowitsch, Anatol & Stefan Th. Gries. 2003. Collocations: Investigating the interaction of words and constructions. *International Journal of Corpus Linguistics* 8(2). 209–43. <https://doi.org/10.1075/ijcl.8.2.03ste>.
- Vilkuna, Maria. 1992. *Referenssi ja määräisyys suomenkielisten tekstien tulkinnassa* (Suomi 163). Helsinki: Suomalaisen Kirjallisuuden Seura.
- VISK = Hakulinen, Auli, Maria Vilkuna, Riitta Korhonen, Vesa Koivisto, Tarja Riitta Heinonen & Irja Alho. 2004. *Iso suomen kielioppi*. Helsinki: Suomalaisen Kirjallisuuden Seura. Verkkoversio. <http://scripta.kotus.fi/visk> (4.3.2020).
- Ward, Gregory. 1999. A comparison of postposed subjects in English and Italian. Akio Kamio & Ken-ichi Takami (toim.), *Function and structure: Studies in honor of Susumu Kuno*, 3–21. Amsterdam: John Benjamins. <https://doi.org/10.1075/pbns.59.03war>.
- Winter, Bodo. 2020. *Statistics for linguists: An introduction using R*. London: Routledge. <https://doi.org/10.4324/9781315165547>.

Abstract. Rodolfo Basile, Ilmari Ivaska: Subject case alternation in constructions containing the Finnish verb *löytyä*. This article examines the nominative-partitive subject alternation occurring with constructions containing the Finnish verb *löytyä*. The material used is taken from corpora and consists of a random sample of 779 observations, analyzed both quantitatively by means of statistical methods, and from a qualitative point of view. The research aims at investigating which variables influence the case alternation of subjects of constructions containing the verb *löytyä*. The chosen variables are subject number, subject divisibility, subject part of speech, word order, tense, agreement and subject lemma, the only random variable. With the help of regression analysis, the subject case is predicted on the basis of said variables and of interactions between them. The qualitative analysis will also discuss the relationship these morphosyntactic and semantic variables have with the existential interpretations of the clause as well as with the subject quantity and definiteness.

Keywords: subject alternation, partitive, constructions, morphosyntax, regression analysis, quantitative research, corpus linguistics, Finnish language

Kokkuvõte. Rodolfo Basile, Ilmari Ivaska: Subjekti käändevaheldus *löytyä*-verbiga konstruktsioonides. Artiklis uuritakse nominatiivi- ja partitiivikujulise subjekti vaheldumist soome keele *löytyä*-verbi sisaldavates konstruktsioonides. 779 vaatlust sisaldavat korpustest pärinevat juhuvalimit analüüsitakse nii kvantitatiivsete kui ka kvalitatiivsete meetoditega. Uurimuse eesmärk on välja selgitada, millised tegurid *löytyä*-verbi sisaldavate konstruktsioonide subjekti käändevalikut mõjutavad. Käsitletavad tegurid on subjekti arv, loendatavus, sõnaliik, konstruktsiooni sõnajärg, ajavorm ning subjekti ja verbi ühildumine. Juhusliku muutujana kaasatakse ka subjekti lemma. Regressioonanalüüsi abil ennustatakse subjekti käändevalikut mainitud tegurite ja nende vaheliste koosmõjude kaudu. Kvalitatiivse analüüsi käigus arutletakse ka selle üle, milline on nimetatud morfosüntaktiliste ja semantiliste tegurite mõju lause eksistentsiaalsele tõlgendusele ning subjekti kvantiteedile ja definiitsusele.

Märksõnad: subjekti käändevaheldus, partitiiv, konstruktsioonid, morfosüntaks, regressioonanalüüs, kvantitatiivne uurimus, korpuslingvistika, soome keel

Basile, Rodolfo (Accepted for publication).
'I am also found on Facebook': Locuphoric 'find'-based strategies in
Finnish internet corpora. In Glaser, Sleeman, Strobel & Tamm (eds.).
Partitive constructions and partitive elements within and across
language borders in Europe Vol. – num. 5 (prov.)
Linguaggio e Variazione. Variation in Language.
Edizioni Ca'Foscari, Venice University Press.

**Partitive Constructions and Partitive Elements Within
and Across Language Borders in Europe**

edited by Elvira Glaser, Petra Sleeman,
Thomas Strobel, Anne Tamm

‘I Am Also Found on Facebook’ Locuphoric ‘Find’-Based Strategies in Finnish Internet Corpora

Rodolfo Basile

University of Tartu, Estonia; University of Turku, Finland

Abstract Allophoric forms of the Finnish verb *löytyä* ‘to be found’ are used in locational constructions similarly to copulas, while the verb’s locuphoric forms have been claimed to be marginal. This paper confirms this claim by looking into internet corpora and provides a qualitative overview of the types of utterances found in the sample, arguing that locuphoric forms may as well convey a locational function in certain contexts. The article also suggests a parallel between locuphoric ‘find’-based strategies and competing impersonal constructions containing the verb *löytää* ‘to find’, discussing the role of partitive-marked object arguments.

Keywords Locational. Existential. Partitivity. Corpus linguistics. Invenitive strategies.

Summary 1 Introduction. – 2 Background. – 2.1 Locational Constructions. – 2.2 The Use of *löytyä* as a Locational Copula. – 3 Material and Method. – 4 Results. – 4.1 Self-Advertisement in Internet Contexts. – 4.2 Permanent Versus Temporary Location. – 4.3 Mirative Marking. – 4.4 Other Uses. – 4.5 Competing Constructions. – 5 Discussion. – 6 Conclusion.



Edizioni
Ca' Foscari

LiVVal. Linguaggio e Variazione | Variation in Language 5

e-ISSN 2974-6574 | ISSN 2974-6981

ISBN [ebook] 978-88-6969-795-1 | ISBN [print]

Peer review | Open access

Submitted 2023-05-23 | Accepted 2023-10-23 | Published: forthcoming

© 2024 Basile | 4.0

DOI 10.30687/978-88-6969-795-1/000

1 Introduction

Basile and Ivaska (2021) have argued that the Finnish intransitive verb *löytyä* 'to be found' is used similarly to the locational copula *olla* 'to be'. Constructions containing this verb are highly specialised in that they often only convey a locational function, without involving the verb's original meaning FIND nor any kind of posture. For these reasons, they too can be considered copulas.¹

- (1a) *Puhelinnumerot löytyvät jäsenkirjeestä.*
phone.number.NOM.PL find.MM.3PL newsletter.ELA
'The phone numbers are (found) in the newsletter.' (Basile, Ivaska 2021, 25)

- (1b) *Poikkeuksiakin ammattikunnasta löytyy*
exception.PL.PART.ENCL profession.ELA find.MM.3SG
'There are also exceptions in the profession.' (Basile, Ivaska 2021, 33)

Like *olla* 'to be', *löytyä* triggers the nominative-partitive alternation in its first argument, also called the *locatum*, which is the subject or located element (Haspelmath 2022). Partitive subject-like arguments are typical of the Finnish Existential construction, which usually also features the lack of agreement between verb and NP (cf. 1a, 1b). Because of this lack of agreement, *löytyä* 'to be found' naturally appears almost always indexed for third persons. However, in their study, Basile and Ivaska also found one instance in which this verb is indexed for second person singular (2).

- (2) *Mistä löyd-y-t prinsini, 44-50-v. fiksu,*
where.ELA find-MM-2SG prince.1PX 44-50-y.o. smart
pitkähkö, ulkonäkö ok, pilke silmäkulmassa,
tallish appearance ok twinkle eye.corner.INE
lenkkeilet ja tanssit
jog.2SG and dance.2SG
'Where are you my prince, 44 to 50 years old, smart, tallish, good-looking, with a twinkle in your eye, you who like to jog and dance.' (Basile, Ivaska 2021, 18)

¹ I would like to thank Petra Sleeman and the two anonymous reviewers for their valuable comments and suggestions that substantially improved the quality of this paper. List of glossing abbreviations: 1 - first person; 2 - second person; 3 - third person; ABE - abessive; ABL - ablative; ACC - accusative; ADE - adessive; ALL - allative; COMP - comparative; COND - conditional; CONJ - conjunction; CONNEG - connegative; ELA - elative; EMOJI - emoji; EMPH - emphatic; ENCL - enclitic; ESS - essive; GEN - genitive; ILL - illative; INE - inessive; INF - infinitive; INTERJ - interjection; ITER - iterative; MM - middle marker; NEG - negation; NOM - nominative; PART - partitive; PASS - passive; PL - plural; PST - past; PTCP - participle; PX - personal suffix; RELP - relative pronoun; SG - singular; USER - username.

They then argued that the example above is marginal and used only in certain contexts. But how marginal is it? Within the scope of this paper, I will try to answer this question, as well as to determine the semantic and pragmatic constraints under which this verb appears indexed for locuphoric forms (i.e., speaker and addressee forms, see Haspelmath 2013). I will show that most uses of locuphoric forms of the verb *löytyä* pertain to an internet environment. These strategies are used to advertise the internet presence of users on various platforms. I will also discuss other uses that do not necessarily refer to the internet environment but are nevertheless not common in everyday speech (Juha-Matti Aronen, p.c., among others). After an overview of the relevant literature, including Basile and Ivaska (2021), in Section 2, I define the material and methods of the research at hand in Section 3. Section 4 presents the results and provides an analysis of the material. In Section 5, I discuss the results and, finally, Section 6 concludes the article and mentions some future research possibilities.

2 Background

The verb *löytyä* 'to be found' is an intransitive derived from the verb *löytää* 'to find' through the deverbal morpheme *-U-* (realising as *-u/ y-* following vowel harmony), which has been argued to have several functions, such as automative, passive, and reflexive (Kulonen-Korhonen 1985; VISK § 335). For this morpheme, I use the term "middle marker" (Kemmer 1993, 41; Zúñiga, Kittilä 2019, 168; Inglese 2022; 2023), which conveniently subsumes its different functions. It is also to be considered that locuphoric forms of the verb *löytyä* do indeed retain their meaning FIND more often than the verb's allophoric forms (third persons, see Dahl 2000; Haspelmath 2013; cf. "aliophoric" [Haspelmath 2020]), hence the middle markers they employ often express a passive function. This matter should however be further investigated by contrasting middle-marked forms of verbs with a root meaning FIND with unmarked ones (e.g., *löytää* 'to find'). The latter forms are arguably far more common in everyday speech but, for reasons of space and scope, I focus on the former and sketch a brief analysis of one of the possible competing constructions at the end of the article. Although the paper at hand does not analyse 'find'-based strategies from a cross-linguistic perspective, it is fundamental to bear in mind that these strategies are indeed used in other European languages for a variety of functions, especially to convey locational meaning (e.g., Italian *trovarsi* find.MM 'be located', Russian *nachodit'sya* find.MM 'be located'), and that they exhibit different usage patterns compared to Finnish when it comes to the

acceptability of allophoric versus locuphoric forms (Basile 2023).² The reasons why there exists this degree of variation are perhaps of a diachronic nature.

2.1 Locational Constructions

Locational constructions are forms of intransitive predication (Stassen 1997). They also fall within the concepts “non-verbal predication” (Hengeveld 1992; Roy 2013) and “copular clauses” (Declerck 1988; Mikkelsen 2011). An overview of these constructions is found in Haspelmath (2022), who distinguishes two main types: “predlocative constructions” and “existential constructions”. These two construction types involve two arguments, a located element and a location, which are linked by a stative linking element called a copula (Haspelmath 2022). Predlocative constructions (e.g., ‘The beer is in the fridge’) predicate about a *locatum* (also called “figure” [Talmy 2000; Koch 2012; Creissels 2014] or “pivot” [Milsark 1977; Bentley, Ciconte, Cruschina 2013]) which is represented by a usually definite referent (‘the beer’) that is said to be in a location expressed by a locative phrase (‘in the fridge’) by means of a copula (‘is’). In existential constructions (e.g., ‘There are beers in the fridge’), the locatum is instead also called the *existent*, and represents an indefinite and discourse-new referent (‘beers’). Existential constructions can be additionally marked for word order (see also Creissels 2019) or, in the case of English, expressed by an expletive followed by a copula (‘there are’). Both predlocative and existential constructions express a locational function and feature an overt locative phrase.

In the Finnish tradition, the prototypical Existential construction features a clause-initial locative adverbial, as well as no verbal agreement (such as in French *Il y a des hommes* ‘There are men’, where the existential copula is marked for singular and the existent *des hommes* is a partitive construction), and a discourse-new subject-like referent whose existence is being predicated (Hakanen 1972; VISK § 893; Huumo 2003). The subject-like referent has also been called an *e-NP* (existential Noun Phrase) because it does not satisfy the typical criteria assigned to subjecthood: for example, it is often marked for partitive case (Huumo, Helasvuo 2015). One problem with the Finnish traditional definition of Existential construction is that it is often similar to structures that have been called “presentationals” (Gast, Haas 2011), which can also feature partitive-marked arguments that

² Outside of Europe, ‘find’-based locational strategies have been found in, e.g., Tagalog and some Mande languages (Basile 2023). A cross-linguistic study is however needed to assess the productivity of such constructions.

do have the function of introducing new referents but typically do not express locational meaning. This happens because Finnish grammarians also allow for other (intransitive) 'existential verbs' to be used in the Finnish Existential construction, verbs that do not have a locational function. Such verbs are, e.g., *tulla* 'to come', *juosta* 'to run', *ilmestyä* 'to appear', and many others. Among these, however, the verb *löytyä* 'to be found' is a special case because not only its frequency in the Finnish Existential construction is much higher than the frequency of the rest of the existential verbs (Basile forthcoming), but also because it can express a purely locational function. It hence belongs to what Basile (2023) calls "inventive verbs" ('inventives'). Inventives are a class of verbs that have a root with meaning FIND (from Latin *invenire* 'to find') which becomes semantically bleached and fulfils other grammatical functions. Typically, inventives feature a valency-changing operation and are part of inventive-locational constructions, where they express a locational function without marking a specific posture (like posture verbs such as German *liegen* 'to lie' instead do). Inventives mostly occur in predlocative constructions in European languages, but in Finnish (and in Estonian) they frequently appear in existential constructions as well. For this reason, it is often the case that Finnish and Estonian inventive-locational constructions feature partitive-marked existents (or e-NPs, see above). When *löytyä* is instead indexed for locuphoric forms, the pronoun it relates to cannot be marked for partitive case. Given that this verb is an intransitive, the locuphoric pronominal form for which it is indexed, if overt, will be marked for nominative case and constitute the syntactic subject of the clause.

2.2 The Use of *löytyä* as a Locational Copula

Basile and Ivaska (2021) investigate the nominative-partitive alternation of subjects in sentences containing the Finnish verb *löytyä* 'to be found' from a quantitative point of view.³ For subjects it is meant both canonical subjects (nominative-marked, clause-initial NPs that trigger verb agreement) and e-NPs, which can be either nominative-marked clause-final NPs or partitive-marked NPs. In both cases, e-NPs do not trigger verb agreement. When they are partitive-marked, they can occur both in clause-initial and clause-final position and usually suggest an existential reading. The method used in their article

³ The sample used in Basile and Ivaska (2021) is a random sample of 779 sentences taken from the corpus *Kansalliskirjaston lehtikokoelma (KLK)*, made of newspapers and magazines written in contemporary Finnish. Given the synchronic nature of the research, the collection of examples was limited to papers written between the beginning of 1990 and the end of 2000.

is mixed-effect logistic regression (Gries 2015), which consists in a binary (NOM-PART) response variable (subject case marking) explained based on both fixed-effect and random-effect explanatory variables. They find that several variables have a statistically significant effect on nominative-partitive alternation, such as the NP's number in correlation with whether it is a count noun or not, and word order in correlation with verb agreement. Verbal tense also showed statistical significance – this is particularly interesting since past tenses, often encoding perfective aspects, increase the likelihood of the verb *löytyä* preserving its original meaning FIND. The examples analysed are almost all indexed for allophoric forms, and the only example indexed for a locuphoric form is justified as being marginal. The example in question is taken into consideration since it interestingly seemed to convey a locational function, although no nominative-partitive alternation occurred. Basile and Ivaska (2021, 35) generally conclude that allophoric forms of the verb *löytyä* indeed function as copulas, and that semantic bleaching plays a role in this. As we see from the analysis below, locuphoric forms of the verb *löytyä* can also be considered as copulas, in that they too undergo semantic bleaching to some extent. This feature comes into play especially when considering the uses of *löytyä* to mark permanent versus temporary location.

3 Material and Method

The material used in the research at hand is taken from the corpus Finnish Web 2014 (fiTenTen2014), a Finnish corpus consisting of various text types taken from the web. It is a large corpus that contains ca. 1.7 billion tokens and ca. 127 million sentences. I accessed the corpus through the platform Sketch Engine (Kilgarriff et al. 2014), where I performed a simple search of all the affirmative, negative, and interrogative forms of the verb *löytyä* 'to be found' when indexed for locuphoric forms (first and second persons, both singular and plural). While in Finnish interrogative forms are obtained by adding the morpheme *-ko/-kö*, negative forms are analytical constructions – e.g., *en löydy* 'I am not to be found' for the first person singular – made of a negation (*en* NEG.1SG) followed by a connegative form (*löydy*). Between the negation and the connegative form it is possible to find several other parts of speech, such as adverbs. However, I only searched for negative constructions that were written subsequently, with no additional language material in between the negative and the connegative form. Additionally, none of the negative-interrogative forms (such as *enkö löydy* 'am I not found') nor past conditional forms (such as *olisin löytynyt* 'I would have been found') yielded any result, so I did not include them in Table 1, which lists all the locuphoric forms used.

Table 1 Locuphoric forms of the verb *löytyä*

	Affirmative	Negative	Interrogative
1SG.PRS	löydyn	en löydy	löydyinkö
1SG.PST	löydyin	en löytynyt	löydyinkö
1SG.CND	löytyisin	en löytyisi	löytyisinkö
2SG.PRS	löydyt	et löydy	löydytkö
2SG.PST	löydyit	et löytynyt	löydyitkö
2SG.CND	löytyisit	et löytyisi	löytyisitkö
1PL.PRS	löydymme	emme löydy	löydymmekö
1PL.PST	löydymme	emme löytäneet	löydymmekö
1PL.CND	löytyisimme	emme löytyisi	löytyisimmekö
2PL.PRS	löydytte	ette löydy	löydyttekö
2PL.PST	löydytte	ette löytäneet	löydyttekö
2PL.CND	löytyisitte	ette löytyisi	löytyisittekö

I provide a descriptive statistical outlook of the search results in Section 4. The sentences in the sample are then analysed from a qualitative perspective.

4 Results

The corpus search yielded a total of 540 occurrences containing the verb *löytyä* in locuphoric forms. After manually going through all the occurrences, 91 were discarded for various reasons, such as:

- substantival uses of *löydyt* (a misspelled version of *löydöt* ‘findings, bargains’, and a homograph of *löytyä* when indexed for 2SG);
- transitive uses of *löytyä*, likely due to misspelling *löytää* ‘to find’;
- misspelled forms of *löytyy*, i.e., find.MM.3SG.

When we are dealing with internet texts, we must account for what has also been called “bad language” (Eisenstein 2013). Posts on social media and forums are not post-edited and users will often misspell words for several reasons (Drouin, Davis 2009). Additionally, users may be second language learners or unbalanced bilinguals (for research on L2 Finnish, see, e.g., Ivaska [2010; 2011], who found that learners of Finnish tend to mix sentence types when producing complex constructions such as the Existential construction in writing). The final sample amounts to 449 sentences [fig. 1].

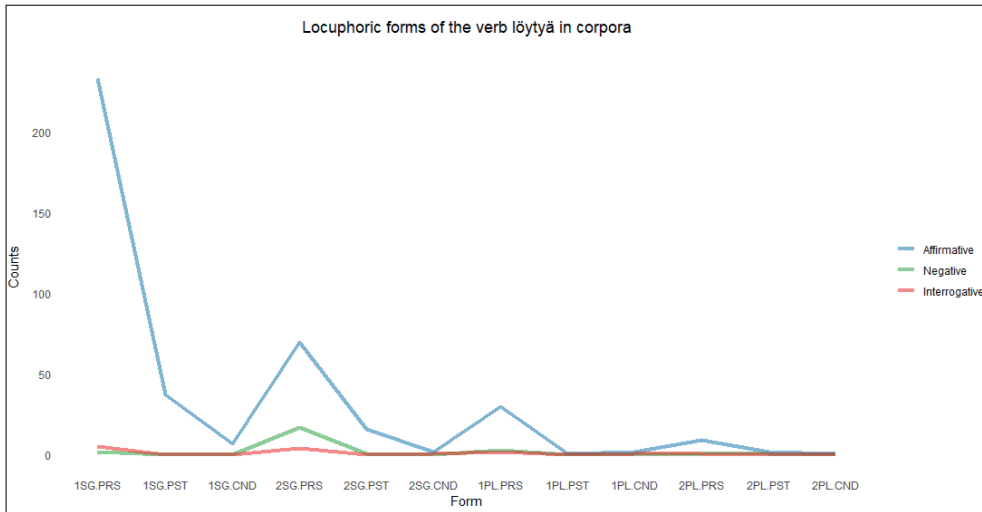


Figure 1 Results of the corpus search. Realised with RStudio by the author

Figure 1 confirms the claim by Basile and Ivaska (2021) according to which *löytyä* is rarely indexed for second person singular and, by extension, for all its locuphoric forms. In fact, locuphoric forms occur 449 times out of a total of ca. 1.2 million occurrences of the verb *löytyä* in the whole corpus. By comparison, performing a simple search of the allomorphic form *löytyy* within the same corpus yields a total of 613,650 occurrences. The results show that the most common locuphoric form is the first person singular in the present tense, indicative mood. In the following subsections, I analyse the main functional characteristics of the sentences found in my sample from a qualitative point of view.

4.1 Self-Advertisement in Internet Contexts

Most occurrences of the verb *löytyä* with locuphoric forms are indexed for first person singular (284 sentences, 63.25%). Some marginal instances with other locuphoric forms, such as second persons, are discussed below, but it is to be kept in mind that there is no big difference in use when it comes to grammatical number. In general, people tend to frequently talk about themselves and index most verbs for first person singular. However, the reasons why it is also the case with locuphoric forms of *löytyä* is that the first person singular is mainly used in forum environments and social media, where speakers advertise their presence on other websites or platforms (3).

- (3) **Löydyn** *myös* *Facebookista*
 find.MM.1SG also Facebook.ELA
 'I am also found on Facebook. / You can also find me on Facebook.'

The speakers also use this strategy to make their usernames available and clear to their interlocutors and, through the internet forum format, to the rest of the community. In these cases, the first singular form is accompanied by an adessive-marked NP such as *nimi* 'name', *nimimerkki* 'pseudonym', *tunnus* 'username' that specifies the users' internet identity. The main pragmatic intent of this construction is made explicit in examples such as (4), where a verbless relative construction (*että...*) has the clearly conative function of wanting the addressee to add the speaker to their contact list.

- (4) **Löydyn** *skypestä* *edelleen* *samalla* *vanhalla* *nimellä*
 find.MM.1SG Skype.ELA still same.ADE old.ADE name.ADE
perneri, *että* *sinne* *vaan* *kaikki* *vanhat*
 USER CONJ there.ILL EMPH all old.PL
ja *uudet*
 and new.PL
 'I am to be found on Skype still under the same old name perneri, (I would like) all old and new people (to add me) there.'

In some examples, the difference between the speakers' real and internet identity is made more evident. In (5), the NP containing the username is used with the postposition *taka-* 'behind' instead of being marked for adessive case like in (4). The same postposition can be used in presentational constructions such as (6), where the speaker's username is, however, not mentioned.

- (5) *Eli* *instagrammista* **löydyn** *niinkin* *tutun*
 that.is Instagram.ELA find.MM.1SG as.well familiar.GEN
nimimerkin *kuin* *partfour* *takaa.*
 pseudonym.GEN as user behind.ELA
 'So, I am also found on Instagram behind the username partfour.'

- (6) *Tämän* *blogin* *takaa* **löydyn** *minä*
 this.GEN blog.GEN behind.ELA find.MM.1SG 1SG
 'The person behind this blog is me.'

These constructions seem to pragmatically imply the permanent presence of the speakers on the internet. This can be a feature encouraged by the very internet environment, in which all usernames

can be considered to be available at all times, even when the person they refer to is not logged into the platform. This strategy is also employed with plural first persons, while second persons are mostly used to confirm to other users about their presence in a certain list or group, most likely on social media (7).

(7a) **Löydyt** jo suosikeistani!
find.MM.2SG already favourite.PL.ELA.1PX
'You are already among my favourites / in my favourites list!'

(7b) **Löydyt** kuitenkin edelleen sieltä tykkäjien joukosta.
find.MM.2SG anyway still there.ABL liker.PL.GEN group.ELA
'Anyway, you are still in the list of likers.'

4.2 Permanent Versus Temporary Location

According to Haspelmath (2022, 6), German uses two different strategies in existential constructions to mark whether a referent is permanently located at a certain location or only temporarily (*es gibt* lit. 'it gives' marks permanent location, while *stehen* 'to stand' refers to temporary presence). In my sample, different strategies are similarly used to mark these two locational strategies. The difference is, however, that the verb is not lexically differentiated like in German. The first strategy is the one we see in § 4.1, and marks permanent location through lexical devices that pertain to the speakers' internet presence. The second strategy is characterised by temporal adverbials (8).

(8) **Löydyn** Fastin pisteeltä kisojen jälkeen aina.
find.MM.1SG F.GEN point.ABL competition.PL.GEN after always
iltakuuteen asti
evening.six.ILL until
'After the competition, I am to be found at Fast's spot right until six in the evening.'

In this type, the locative phrase marks a real-world location, in which the speaker can be found for a limited time indicated by temporal adverbials like *iltakuuteen asti* 'until 6 in the evening'. The adverb *aina* 'always' also marks the event as recurrent, as does *viikonloppuisin* 'on weekends' in (9).

(9) *Viikonloppuisin löydyn hyvin poikkeuksetta pullon pohjalta.*
weekend.ITER find.MM.1SG very exception.ABE bottle.GEN BOTTOM.ABL
'On the weekends I am always drinking.' (lit. 'found at the bottom of the bottle')

The difference is that while (8) indicates the availability of the speaker, in (9) the location is fictive and metaphorical. The speaker here infers about her drinking habit. In another example, time framing is used as a strategy to communicate the appearance of the speaker in a video, inferring permanent presence within the video, however confined to a certain time interval (10).

- (10) *miehän löydyn TÄSTÄ pätkästä 50 sekunnin kohdilta*
1SG.EMPH find.MM.1SG this.ELA part.ELA 50 second.GEN place.PL.ABL
'I am to be found in THIS snippet at around 50 seconds.'

4.3 Mirative Marking

Mirativity is a linguistic category that expresses information that is surprising or unexpected to both the speaker and the addressee (DeLancey 1997; DeLancey 2001; Hengeveld, Olbertz 2012). Several European languages use 'find'-based strategies to mark mirative events (e.g., 'I found myself on the top of the hill'; 'I found myself thinking about you'), as the verb *löytyä* can also occasionally do (11).

- (11) *löydyn uudelleen ja uudelleen*
find.MM.1SG again and again
pakonomaisesti tunkemassa ruokaa sisälleni
compulsively shove.INF.INE food.PART inside.ALL.1PX
'I find myself over and over again compulsively shoving food in my mouth'

This way of marking one's involuntary involvement in a certain situation is similar to the more common analytical construction *löytää itsensä* 'find oneself', that also encodes mirative meaning (12).

- (12) *Löysin itseni sairaalasta.*
find.PST.1SG self.1PX hospital.ELA
'I found myself at the hospital.'

Like in English and in other European languages (Basile 2023), the mirative strategy can also be used with concrete locations instead of abstract situations and states of mind. It is still debatable whether (13a) can be interpreted as a mirative-marked strategy. The poly-functionality of the middle marker may also simply suggest a passive-resultative reading, but then again the question arises: why didn't the speaker choose to use a passive construction (e.g., *minut löydettiin* 1SG.ACC find.PASS.PST lit. 'I was found'). On the other hand, it is clear that (13b) does not express a mirative function.

- (13a) *Minulle kerrottiin, että löydyin kadulta*
 1SG.ALL tell.PASS.PST CONJ find.MM.PST.1SG street.ABL
verissäni, josta minut vietiin sairaalaan
 blood.PL.INE.1PX RELP.ELA 1SG.ACC carry.PASS.PST hospital.ILL
ja sen jälkeen mielisairaalaan.
 and it.GEN after mental.hospital.ILL
 'I was told that I was found on the street covered in blood, then carried to the hospital and then to the mental hospital.'
- (13b) *Viime tiistaina meillä oli meidän*
 last Tuesday.ESS 1pl.ADE be.PST.3SG 1SG.GEN
lukion musa-abien konsertti, jossa
 high.school.GEN music-major.PL.GEN concert rel
itse sitte löydyin toisinaan mikin,
 self then find.MM.PST.1SG occasionally microphone.GEN
pianon, kitaran tai basson takaa.
 piano.GEN guitar.GEN or bass.GEN behind.ELA
 'Last Tuesday we had our high school's music majors' concert, in occasion of which I could be found behind the microphone, the piano, the guitar, or the bass.'

While usually used with an elative-marked spatial argument, *löytyä* can occasionally also be used with illative-marked NPs (14a) or verbs in the third infinitive illative form (14b), which expresses movement towards rather than from. Both examples encode a mirative event which is strictly connected to the internet environment.

- (14a) *Löydyin tekstiin sattumalta ja ilahduin*
 find.MM.PST.1SG text.ILL accidentally and cheer.PST.1SG
että aiheesta käydään vilkasta keskustelua.
 CONJ topic.ELA run.PASS active.PART conversation.PART
 'I accidentally stumbled upon the text and was glad that the topic is being discussed actively.'
- (14b) *hohoo, vasta nyt löydyin*
 INTERJ only now find.MM.PST.1SG
lukemaan sun mahtavia ekoisi-huomioita!
 read.INF.ILL 2SG.GE amazing.PL.PART ecodaddy-remark.PL.PART
 'Oh, I am only now reading your amazing eco-daddy remarks!'

Mirative readings are also triggered when, pragmatically, speakers presuppose the high unlikelihood of the situation. While for many other examples it is debatable whether the verb retains its original meaning FIND, (15) seems unambiguously resultative and passive in meaning.

- (15) *Siksi olikin suoranainen ihme,*
 hence be.PST.EMPH absolute miracle
kun sinä löydýt.
 when 2SG find.MM.PST.2SG
 ‘For this reason it was an absolute miracle when you were found.’

4.4 Other Uses

Among the other sentence types in my sample, I found occurrences of *löytyä* that mark desiderative functions. The literature usually refers to desideratives as morphosyntactic devices that encode volitive modality. Notable is the case of the Japanese bound verbal suffix *-tai*, which conveys the meaning WANT (Izutani 2003). When the suffix *-tai* co-occurs with the nominative marking *-ga* on the NP, we have a desiderative construction. The Finnish multifunctional particle *-pa* is the closest relative to a morphological desiderative marker (VISK § 833-5; for other desiderative constructions in Finnish see VISK § 1659). Together with the conditional mood, it indicates wanting an unlikely event to take place. Similarly, the conditional mood in (16), together with the fundamentally mirative meaning encoded by the verb, marks a desiderative function.

- (16) *jos vaikka joku aamu yllättäisin itseni*
 if though some morning surprise.COND.1SG self.1PX
ja oikeesti löytyisin tuolta lenkkeilemästä
 and really find.MM.COND.1SG there.ABL jog.INF.ELA
klo 7 aamulla. :D
 hour 7 morning.ADE EMOJI
 ‘fonly one morning I could surprise myself and actually be found jogging at 7 am.’

The form *et löydy* ‘you are not found’ is almost always (15 out of 17 times) used referring to business contexts. This type of utterance communicates the importance of the internet presence for enterprises (17a,b).

- (17a) *Jos joku etsii vaikkapa autonhuoltoa*
 if someone search.3SG for.example car.maintenance.PART
Jyväskylästä, olet aikailla ulkona pelistä jos
 Jyväskylä.ELA be.2SG pretty.much out.ESS game.ELA if
et löydy *Googlessa*
 NEG.2SG find.MM.CONNEG Google.INE
 ‘If someone is looking for, say, car maintenance services in Jyväskylä, you are going to be pretty much out of the game if you are not to be found on the first page of Google’s search results.’

(17b)	<i>eri</i>	<i>yhteisöihin</i>	<i>liittyvästä</i>	<i>kokonaisuudesta</i>
	various	organisation.PL.ILL	related.ELA	whole.ELA
	<i>yrittäjien</i>	<i>hakukonenäkyvyydessä</i>	<i>on</i>	<i>kysymys.</i>
	company.GEN	search.engine.visibility.INE	BE.3SG	question
	<i>Jos et löydy,</i>	<i>et</i>	<i>ole</i>	
	if	NEG.2SG find.MM.CONNEG	NEG.2SG	be.CONNEG
	<i>olemassa.</i>			
	exist.INF.INE			
	‘...it is a matter of the company’s search engine visibility for what concerns various organisations. If you are not to be found, you don’t exist.’			

4.5 Competing Constructions

Since the locuphoric forms of the verb *löytyä* are expectedly marginal, speakers could prefer other constructions when talking about themselves and where they are located or to be found. Instead of using *löytyä*, which features only one argument (the subject of an intransitive clause, or S), competing constructions employ its transitive counterpart *löytää* ‘to find’, which features two arguments (the agent A and the patient P; for reference see, e.g., Haspelmath 2011). P, which corresponds to the object of a transitive verb, can be affected by case marking alternation (accusative-partitive in the case of personal pronouns, genitive-partitive in the case of other nouns), similarly to S (nominative-partitive) when *löytyä* is indexed for allophoric forms (Basile, Ivasaka 2021). Competing constructions using *löytää* ‘to find’ may include Impersonal and Impersonal Passive constructions, where A is not expressed and P is susceptible to nominative-partitive alternation. Here, I will only consider locuphoric forms of P. In the Impersonal construction, the first/second person pronoun is followed by the verb indexed for 3SG (e.g., *minut/minua löytää* 1SG.ACC/1SG.PART find.3SG ‘they find me, I am found’), while in the Impersonal Passive construction the only difference is the passive marking *-tAAn/-ttiin* on the verb, with basically no difference in meaning (e.g., *minut/minua löydetään/löydettiin* 1SG.ACC/1SG.PART find.PASS.PRS/find.PASS.PST ‘I am found’). While I do not analyse all the possible variations of these constructions, I will briefly elaborate on two of them, namely the Impersonal constructions *minut löytää* and *minua löytää*, where the first person singular pronoun is indexed respectively for accusative and partitive case.

By performing a simple phrase search within the same corpus used above, I found that the accusative-marked construction *minut löytää* is far more common than its partitive-marked counterpart *minua löytää* (raw frequencies 910/10). It seems like the accusative-marked construction is often used similarly to locuphoric forms of the intransitive *löytyä*, with contexts ranging from internet environments (18a) to expression of time-framed permanent presence (18b).

(18a) *Sieltä minut löytää nimellä @evehei.*
there.ABL 1SG.ACC find.3SG name.ADE USER
'You can find me / I am to be found by the name @evehei.'

(18b) *Vapaa-ajalla minut löytää usein koripallokentältä*
free-time.ADE 1SG.ACC find.3SG often basketball.court.ABL
'In my free time you can often find me / I am often to be found at the
basketball court.'

Of the 10 occurrences of *minua löytää*, only 3 are relevant because used impersonally (19).

(19a) *kyseisillä nimillä minua löytää pahraiten*
in.question.PL.ADE name.PL.ADE 1SG.PART find.3SG best
'Ideally, you can find me through the names in question.'

(19b) *Harvemmin minua löytää kuitenkaan valittamasta*
rarely.COMP 1SG.PART find.3SG anyway.NEG complain.INF.ELA
musiikista mikä huoneessa soi.
music.ELA RELP room.INE play.3SG
'It is even rarer that you can find me complaining about the music playing
in the room.'

(19c) *Minua löytää DC:stä ja*
1SG.part find.3SG DC:ELA and
yllä olevasta osoitteesta, Myrskylinnusta.
above.ADE be.PTCP.ELA address.ELA Myrskylintu.ELA
'I am to be found in DC and at the address above, Myrskylintu.'

It is interesting to notice that both accusative-marked and partitive-marked Impersonal constructions do not seem to encode mirative events. The Finnish Partitive case is often associated with changes in clause-level aspect, marking indefinite events or events with low control, as happens with certain verbs indicating feelings (e.g., *minua itkettää* 1SG.PART cry.CAUS.3SG 'I feel like crying') where the experiencer is marked for partitive case. In this type of construction there can also be a causer (A) marked for nominative case, rendering the experiencer a sort of P. The fact that low control can be associated with mirative events and that the Partitive can be used to mark this type of events could point toward the suitability of the Finnish Partitive case to mark mirativity in *minua löytää* constructions; however, this is not the case. It is also true that the *minua löytää* construction is too marginal in the sample to draw general conclusions.

5 Discussion

In a similar way to what Basile and Ivaska (2021) do for the allophoric forms of *löytyä*, it can be argued that locuphoric forms of this verb may also function as copulas, since they too can undergo semantic bleaching, at least to some degree. In fact, we saw that many of the sentences found in the corpus have the main function of conveying the permanent presence of the speaker/hearer, often in an internet environment. Alongside competing constructions that use the verb *löytää* 'to find', which would perhaps constitute a valid alternative, one could expect to find the copula *olla* 'to be' as the most frequent and unmarked way of expressing the same meaning. The additional meaning provided by *löytyä*, that of prompting the addressee to look for the located referent, certainly cannot be ignored, but it is often the case that the main function this verb has is a locational function, similarly to the copula *olla*.

The data indicates that, as Basile and Ivaska (2021) argued, locuphoric forms of *löytyä* are not as productive as allophoric forms, in that the latter are used overall more, more widely, and in a variety of contexts (Basile, Ivaska 2021). This means that the allophoric forms are already established as locational strategies, because they indicate a link between a located referent and a location, just as copulas do. It could be the case that by analogy with allophoric forms of *löytyä*, its locuphoric forms have started spreading to similar functional domains and have hence started conveying locational meanings. We could say, perhaps speculatively, that this is the first step towards an enhanced productivity of locuphoric forms of *löytyä*, thanks to the increased use of its allophoric forms.

We also should not forget that a language is not an isolated system. Finnish is surrounded by Indo-European languages with which it has been in contact for a long time. Of these, two of the arguably more influential languages, Swedish and Russian, similarly present 'find'-based (inventive) strategies to convey locational meaning (Swedish *att befinna sig* 'to find oneself/be located'; Russian *nachodit'sya* 'to be located'). These strategies are productive, and they might have played a role in helping their Finnish counterpart rise as a locational copula, both in its locuphoric and allophoric forms.

6 Conclusion

This paper discusses locuphoric forms of the Finnish locational copula *löytyä* 'to be found'. The corpus search confirms the claim by Basile and Ivaska (2021) about the marginality of use of these forms, as only about 500 examples were found in a corpus of more than 1 billion tokens. Most occurrences are indexed for first person singular,

as speakers tend to refer to themselves in internet forums and social media, in order to advertise their internet presence on other platforms. This finding points toward the specificity of use of these forms in certain contexts, while in everyday conversation they are arguably nearly absent and substituted by other 'find'-based strategies or the copula *olla* 'to be'. Locuphoric uses of *löytyä* can also mark mirative events, similarly to analytical 'find'-based constructions such as *löytää itsensä* 'to find oneself'. Moreover, they can mark permanent versus temporary presence of referents at a certain location, as well as desiderative functions. The sample is too small to make claims about the productivity of such verbal forms, especially because they are not common in everyday speech. It is, however, big enough to raise the question about the possible reasons why these forms developed only so marginally, and what the future of this line of research holds. One possible development is a study that contrasts locuphoric forms of *löytyä*, which features a middle marker, with unmarked intransitive constructions featuring the verb *löytää* 'to find'. The preliminary considerations about the spreading by analogy of locational functions from allophoric to locuphoric forms of *löytyä* call for a more detailed diachronic study of the development of *löytyä* as a locational copula in the first place. This development might also be supported by language-contact hypotheses, since the use of 'find'-based strategies in locational constructions seems to be widespread in the languages of Europe, including Swedish and Russian. Furthermore, cross-linguistic evaluations about the productivity of 'find'-based (inventive) strategies are needed. Using cognitive and usage-based frameworks to study such constructions could shed light on whether there exist general tendencies that could explain the development of inherently dynamic verbs with meaning FIND into stative copulas in different languages.

Bibliography

- Basile, R. (2023). "Inventive-Locational Constructions in the Languages of Europe". <https://doi.org/10.17605/OSF.IO/ES3FH>.
- Basile, R. (Forthcoming). "The Finnish Existential Partitive Construction Analyzed: Comparing Two Applications of Collostructional Analysis". Leuschner, T. et al. (eds), *How to Do Things with Corpora – Methodological Issues and Case Studies*. Berlin/Heidelberg: J.B. Metzler. Linguistik in Empirie und Theorie/Empirical and Theoretical Linguistics.
- Basile, R.; Ivaska, I. (2021). "Löytyä-verbin konstruktioiden yhteydessä esiintyvän subjektin sijanvaihtelu". *Eesti ja soome-ugri keeleteaduse ajakiri. Journal of Estonian and Finno-Ugric Linguistics*, 12(1), 11-39. <https://doi.org/10.12697/jefu.2021.12.1.01>.
- Bentley, B.; Ciconte, F.M.; Cruschina, S. (2013). "Existential Constructions in Cross-Linguistics Perspective". *Italian Journal of Linguistics*, 25, 15-43.

- Creissels, D. (2014). "Existential Predication in Typological Perspective". Ms., University of Lyon.
- Creissels, D. (2019). "Inverse-Locational Predication in Typological Perspective". *Italian Journal of Linguistics*, 31(2), 37-106.
- Dahl, Ö. (2000). "Egophoricity in Discourse and Syntax". *Functions of Language*, 7(1), 37-77. <https://doi.org/10.1075/foL.7.1.03dah>.
- Declerck, R. (1988). *Studies on Copular Sentences, Clefts and Pseudo-Clefts*. Leuven: Leuven University Press & Foris. <https://doi.org/10.1075/foL.7.1.03dah>.
- DeLancey, S. (1997). "Mirativity: The Grammatical Marking of Unexpected Information". *Linguistic Typology*, 1, 33-52. <https://doi.org/10.1515/lity.1997.1.1.33>.
- DeLancey, S. (2001). "The Mirative and Evidentiality". *Journal of Pragmatics*, 33(3), 369-82. <https://doi.org/10.1515/9783110572261-009>.
- Drouin, M.; Davis, C. (2009). "R U Txtng? Is the Use of Text Speak Hurting your Literacy?". *Journal of Literacy Research*, 41(1), 46-67. <https://doi.org/10.1080/10862960802695131>.
- Eisenstein, J. (2013). "What to Do about Bad Language on the Internet". *Proceedings of NAACL-HLT 2013. Atlanta, Georgia, 9-14 June 2013*. Association for Computational Linguistics, 359-69.
- Gast, V.; Haas, F. (2011). "On the Distribution of Subject Properties in Formulaic Presentationals of Germanic and Romance". Malchukov, A.; Siewierska, A. (eds), *Impersonal Constructions: A Cross-Linguistic Perspective*. Amsterdam: John Benjamins, 127-66. <https://doi.org/10.1075/slcs.124.05gas>.
- Gries, S.Th. (2015). "The Most Under-Used Statistical Method in Corpus Linguistics: Multi-Level (and Mixed-Effects) Models". *Corpora*, 10(1), 95-125. <https://doi.org/10.3366/cor.2015.0068>.
- Hakanen, A. (1972). "Normaalilause ja eksistentiaalilause". *Sananjalka*, 14(1), 36-76. <https://doi.org/10.30673/sja.86366>.
- Haspelmath, M. (2011). "On S, A, P, T, and R as Comparative Concepts for Alignment Typology". *Linguistic Typology*, 15, 535-67. <https://doi.org/10.1515/lity.2011.035>.
- Haspelmath, M. (2013). "Argument Indexing: A Conceptual Framework for the Syntactic Status of Bound Person Forms". Bakker, D.; Haspelmath, M. (eds), *Languages across Boundaries: Studies in Memory of Anna Siewierska*. Berlin: De Gruyter Mouton, 197-226. <https://doi.org/10.1515/9783110331127.197>.
- Haspelmath, M. (2020). "Locuphoric Person Forms and Speech Act Participants". <https://dlc.hypotheses.org/2430>.
- Haspelmath, M. (2022). "Nonverbal Clause Constructions". https://www.academia.edu/71539642/Nonverbal_clause_constructions.
- Hengeveld, K. (1992). *Non-Verbal Predication: Theory, Typology, Diachrony*. Berlin: Mouton de Gruyter. <https://doi.org/10.1515/9783110883282>.
- Hengeveld, K.; Olbertz, H. (2012). "Didn't you Know? Mirativity Does Exist!". *Linguistic Typology*, 16(3), 487-503. <https://doi.org/10.1515/lity-2012-0018>.
- Huumo, T. (2003). "Incremental Existence: The World According to the Finnish Existential Sentence". *Linguistics*, 41(3), 461-93. <https://doi.org/10.1515/ling.2003.016>.
- Huumo, T.; Helasvuo, M.L. (2015). "On the Subject of Subject in Finnish". Huumo, T.; Helasvuo, M.L. (eds), *Subjects in Construction: Canonical and*

- Non-Canonical*. Amsterdam: John Benjamins, 13-41. <https://doi.org/10.1075/cal.16.02huu>.
- Inglese, G. (2022). "Towards a Typology of Middle Voice Systems". *Linguistic Typology*, 26(3), 489-531. <https://doi.org/10.1515/Lingty-2020-0131>.
- Inglese, G. (2023). "The Rise of Middle Voice Systems: A Study in Diachronic Typology". *Diachronica*, 40(2), 195-237. <https://doi.org/10.1515/Lingty-2020-0131>.
- Ivaska, I. (2010). "Eksistentiaalilauseen subjektijäsen edistyneiden suomenoppijoiden kirjoituksessa". *Sananjalka*, 52(1), 67-88. <https://doi.org/10.30673/sja.86697>.
- Ivaska, I. (2011). "Lausetyyppien sekoittuminen edistyneessä oppijansuomesa-näkökulmana eksistentiaalilause". *Lähivördlusi. Lähivertailuja*, 21, 65-85. <https://doi.org/10.5128/lv21.03>.
- Izutani, M. (2003). "Desiderative Construction in Japanese Revisited". *Journal of The College of Liberal Arts and Sciences Tokyo Medical and Dental University*, 33, 19-27.
- Kemmer, S. (1993). *The Middle Voice*. Amsterdam: John Benjamins. <https://doi.org/10.1075/tsl.23>.
- Kilgarriff, A et al. (2014). "The Sketch Engine: Ten Years On". *Lexicography*, 1, 7-36. <https://doi.org/10.1007/s40607-014-0009-9>.
- Koch, P. (2012). "Location, Existence and Possession: A Constructional-Typological Exploration". *Linguistics*, 44, 569-612. <https://doi.org/10.1515/Ling-2012-0018>.
- Kulonen-Korhonen, U. (1985). "Deverbaalisten U-verbijohdosten semantiikkaa". *Virittäjä*, 89, 290-309.
- Mikkelsen, L. (2011). "Copular Clauses". Maienborn, C.; Von Heusinger, K.; Portner, P. (eds), *Semantics: An International Handbook of Natural Language Meaning*, vol. 2, 1805-29. Berlin: De Gruyter Mouton. <https://doi.org/10.1075/la.85>.
- Milsark, G.L. (1977). "Toward an Explanation of Certain Peculiarities of the Existential Construction in English". *Linguistic Analysis*, 3(2), 1-29.
- Roy, I. (2013). *Nonverbal Predication: Copular Sentences at the Syntax-Semantics Interface*. Oxford: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199543540.001.0001>.
- Stassen, L. (1997). *Intransitive Predication*. Oxford: Oxford University Press.
- Talmy, L. (2000). *Toward a Cognitive Semantics*. Vol. 1, *Concept Structuring Systems*. Cambridge (MA): The MIT Press. <https://doi.org/10.7551/mitpress/6847.001.0001>.
- VISK = Hakulinen, A.; Vilkuna, M.; Korhonen, R.; Koivisto, V.; Heinonen, T.R.; Alho, I. (2004). *Iso suomen kielioppi*. Helsinki: Suomalaisen Kirjallisuuden Seura. Web version 1.11.2008. <https://kaïno.kotus.fi/visk/etusivu.php>.
- Zúñiga, F.; Kittilä, S. (2019). *Grammatical Voice*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/9781316671399>.

Basile, Rodolfo (Accepted for publication).
The Finnish Existential Partitive Construction Analyzed:
Comparing Two Applications of Collostructional Analysis.
In Leuschner, Barðdal, Delaby & Vajnovszki (eds.). How to Do
Things with Corpora. Berlin/Heidelberg: J.B. Metzler. (Linguistik in
Empirie und Theorie/Empirical and Theoretical Linguistics).

The Finnish Existential Partitive Construction Analyzed: Comparing Two Applications of Collostructional Analysis

Rodolfo Basile

ABSTRACT

This article investigates the Finnish Existential Partitive Construction in web corpora from a quantitative point of view, analyzing the occurrence frequencies of existential verbs in the construction. The aim is to examine how often different verbs feature in the Finnish Existential Partitive Construction in corpora. Verb frequencies in this type of Finnish construction have not been studied before. The study proposes two different applications of collostructional analysis (Stefanowitsch & Gries 2003): one based on a reduced sample, and another based on an Expected Sample size. It is then shown that the former sampling leads to biased results, while the latter considers the absolute frequency of the verbs on a whole-corpus level. Specifically, the verb *löytyä* ‘to be found’ is more prone to be associated with the Finnish Existential Construction, and this corroborates the idea of a copular interpretation of said verb.

Keywords: collostructional analysis, Construction Grammar, copula, existential constructions, Finnish, partitive

1 INTRODUCTION

In Finnish, subject-like arguments marked for partitive case may often occur in what are broadly considered existential predications, or also existential clause constructions (Creissels 2014, 2019; Ikola 1954; Hakanen 1972; Haspelmath 2022). In such constructions, said subject-like arguments usually have the main function of introducing new referents and locating them in a location expressed by a locative Noun Phrase.

However, as it is explained below, the Finnish Existential Construction does not always have the locational function that is commonly attributed to the existential construction as a comparative concept (on “comparative concept” see Haspelmath 2010), but rather often just introduce new referents into discourse.¹

¹ As a general rule, I refer to language-specific categories and grammatical constructions by capitalizing the initials of the terms I use, while lower-case initials are reserved to comparative concepts, that are valid cross-linguistically. This is done in a similar fashion to e.g. Croft (2001: 51), Barðdal (2006), and Haspelmath (2010: 665).

It is also argued that the Finnish Existential Construction produces a holistic interpretation of the events happening within a certain location (Huumo 2003).

While most existential constructions feature the copula *olla* ‘to be’, in the Finnish tradition other intransitive verbs are also considered as *existential verbs*, i.e. verbs that can appear in existential constructions (see also Huumo 2018: 424 on the occurrences of transitive clauses with partitive subjects). The examples in (1) illustrate the use of both the copula *olla* ‘to be’, and another intransitive and less frequent verb, *mädäntyä* ‘to rot’, as well as the alternation between the nominative, as in (1a, 2a), and the partitive, as in (1b, 2b), marking on the argument: this alternation primarily determines whether the referent in question is seen as a whole (nominative) or as a part of itself (partitive). The partitive case is also used when the referent is a mass noun, when it is marked for plural, or in negative sentences.²

- (1) a. *Pöydä-llä on salaatti.*
 table-ADE be.3SG salad.NOM
 ‘There is a salad on the table.’
- b. *Pöydä-llä on salaatti-a.*
 table-ADE be.3SG salad-PTV
 ‘There is (some) salad on the table.’
- (2) a. *Pöydä-llä mädänty-y salaatti.*
 table-ADE rot-3SG salad.NOM
 ‘There is a salad rotting on the table.’
- b. *Pöydä-llä mädänty-y salaatti-a.*
 table-ADE rot-3SG salad-PTV
 ‘There is (some) salad rotting on the table.’

The aim of this paper is to fill a gap in quantitative studies about frequencies of partitive subject-like arguments in Finnish, which belong to what I here call the Finnish Existential Partitive Construction. The focus has often been on, e.g., the alternation between canonical (nominative or *totalitive*, see Larjavaara 2019) and partitive-marked subjects (Basile & Ivaska 2021), but never on with which verbs partitive subject-like NPs are more likely to occur.

I use collostructional analysis (Stefanowitsch & Gries 2003; Gries & Stefanowitsch 2004) to answer the following research questions: How strongly are different intransitive verbs attracted to the Finnish Existential Partitive Construction? What are the main differences between the analyzed verbs? The reason why I have chosen to use collostructional analysis lies in the preliminary

² List of glossing abbreviations: 1 – first person; 3 – third person; ABL – ablative; ADE – adessive; ALL – allative; ILL – illative; INE – inessive; INT – intensifier; MM – middle marker; NOM – nominative; PL – plural; PRT – preterite; PTV – partitive; PX – personal suffix; SG – singular.

insights provided by looking into different verbs: can we say that some existential verbs are more prone to appear as collexemes of the Finnish Existential Partitive construction? And if so, why would that be?

The article is structured as follows: Section 2 gives an overview of the partitive case and its degree of subjecthood in Finnish, as well as of Existential constructions in Finnish. Section 3 introduces materials and methods used, while Section 4 analyzes the data from a quantitative point of view. Finally, Section 5 draws conclusions and discusses future developments.

2 PARTITIVE CASE, SUBJECTHOOD, AND EXISTENTIAL CONSTRUCTIONS IN FINNISH

Partitive subject-like NPs constitute a typologically common feature of the Baltic-Finnic area (Koptjevskaja-Tamm & Wälchli 2001; Luraghi & Kittilä 2014; Seržant 2015; Järventausta 1991). In Finnic, the partitive case emerged from an old Uralic separative case (see Kiparsky 1998; Kittilä & Ylikoski 2011). When it came into use, the separative strategy began to be conveyed by the elative or the ablative cases. In Mordvinic, for example, the ablative/elative separative strategy is still used in partitive constructions, although mainly with partitive objects (Bernhardt 2021: 54; see also Larjavaara 2019; Bartens 1999). Partitive subject-like NPs have also been studied in Estonian (Metslang 2014; Lindström 2017; see also Huumo & Lindström 2014 for a comparison with Finnish), as well as in other neighboring languages, such as Livonian, Veps, and Karelian (Lees 2015).

The notion of partitive subject and the degree to which a partitive-marked NP can be a subject has been discussed widely: Huumo & Helasvuo (2015) have proposed the alternative definition of *e-NP* (continuing the tradition of the ‘*existential NP*’, see Helasvuo 1996; Helasvuo & Huumo 2010), since partitive Noun Phrases occur in what are traditionally considered Existential sentences, and since these NPs do not satisfy the criteria designed for prototypical subjects. Other proposed terms include *e-subject* (see e.g. VISK § 910; Metslang 2012: 152), *e-theme* (Huumo 2003) or *Existential S argument* (Huumo 2018), which does not contemplate subjecthood at all. For constructions that include *existential NPs*, I use the term FINNISH EXISTENTIAL Construction. For the NPs found in examples and contained in the Finnish Existential Construction, I use the term *e-NP*. Within the Finnish Existential Construction, *e-NPs* can alternate between nominative and partitive marking, giving birth to two filial constructions: the FINNISH EXISTENTIAL NOMINATIVE Construction and the FINNISH EXISTENTIAL PARTITIVE Construction.

The Finnish Existential Construction has also been studied from the point of view of cognitive semantics (Huumo 2003) and that of (second) language acquisition (Ivaska 2010, 2011; Kajander 2013).

Because the Finnish Existential Construction allows for the use of semantically full verbs, it differs from the cross-linguistically valid, functional definition of *existential clause construction* proposed by Haspelmath (2022), who considers it as an instance of nonverbal clause constructions, which can also be characterized by a copula (including copulas specialized to existential constructions such as Tagalog *may* or Spanish *hay*). According to Haspelmath (2022), copulas indicate a stative link between two arguments, excluding verbs like *become* which are dynamic in nature. From a typological point of view, the Finnish Existential Construction does not necessarily have the function of locating a newly introduced referent into a certain location, but rather of introducing a new referent and predicate and what it is doing in a certain location. This means that these constructions can also employ dynamic verbs. The Finnish Existential Construction is hence functionally much more similar to *presentational* constructions (Gast & Haas 2011), but since this discussion is beyond the scope of this paper, I refer to the Finnish Existential Construction as having the following formal traits:

- a clause-initial locational adverbial;
- an intransitive verb or a copula always in 3SG, i.e. absence of verb agreement;
- a clause-final (indefinite) nominative or partitive e-NP.

The choice of these traits establishes a prototype based on the Finnish traditional literature on existential sentences (VISK § 893; Hakanen 1972) and defines the Finnish Existential Construction as a language-specific construction, rather than a typologically valid category. The three formal traits also distinguish the Finnish Existential Construction from the so-called *normaalilause* ‘normal/canonical sentence’ (Hakanen 1972), which usually has a prototypical subject in nominative and an agreeing predicate, as seen in (3). In Finnish, word order is relatively free and used for information-structural purposes. When there is a partitive e-NP instead of a nominative e-NP in preverbal position, a clause can also be considered a Finnish Existential Construction, as shown in (4). If the word order were to be inverted, with the locative in preverbal position, it could be argued that such an example offers an enhanced presentational reading (4c).

- (3) a. *Salaatti mädäntyy pöydällä.*
 salad.NOM.SG rot-3SG table-ADE
 ‘The salad is rotting on the table.’
- b. *Salaatti-t mädäntyy-vät pöydällä.*
 salad-NOM.PL rot-3PL table-ADE
 ‘The salads are rotting on the table.’
- (4) a. *Salaatti-a mädäntyy pöydällä.*
 salad-PTV rot-3SG table-ADE
 ‘There is (some) salad rotting on the table.’

- b. *Salaatte-ja* *mädänty-y* *pöydä-llä.*
 salad-PTV.PL rot-3SG table-ADE
 ‘There are salads rotting on the table.’
- c. *Pöydä-llä* *mädänty-y* *salaatte-ja.*
 table-ADE rot-3SG salad-PTV.PL
 ‘There are salads rotting on the table.’

In other words, partitive NPs in preverbal position often also suggest an existential (/presentational) reading of the construction (Ikola 1954). Moreover, all e-NPs are marked for partitive case in negated sentences. One of the latest accounts on the Finnish Existential Construction (Larjavaara 2019) includes certain constructions that are usually considered canonical clauses with reversed word order (5 – here cited from Larjavaara [2019: 124]).

- (5) *Piha-lla* *ol-i-vat* *serkku-ni*
 yard-ADE be-PRT-3PL cousin-PL.1PX
Kaisa *ja* *Maarit.*
 K.NOM and M.NOM
 ‘In the yard there were my cousins Kaisa and Maarit.’

The focus of the analysis at hand is to show how different intransitive verbs behave with partitive-marked e-NPs. These verbs are also called *lexical existentials* (Larjavaara 2019, see also Section 3); a term which refers to all verbs other than the copula *olla* ‘to be’ that can feature in the Finnish Existential Construction.

All the verbs used in the present analysis can appear in said construction, but there is reason to believe that some of them are more suitable than others because of their semantic contour. One in particular, *löytyä* ‘to be found’, has been claimed to be used as a productive locational copula, since it can convey a purely locational function (Basile & Ivaska 2021; Basile forthcoming). This means that it is often semantically bleached of its original meaning FIND, resulting in what Basile (2023) calls an *inventive* strategy, i.e. a FIND-based strategy that conveys locational meaning, fairly common in European languages.

3 MATERIAL AND METHOD

The material for this research has been collected from the corpus Suomi24:2017, a collection of all the utterances found in Finland’s largest online social networking forum Suomi24 in the year 2017.³ This corpus consists of ca. 168 million words. It is morphosyntactically annotated, and the annotation is carried out

³ Accessible at <https://www.suomi24.fi/> (last accessed 31/01/2023)

automatically. The corpus search has been done via the corpus user interface Korp, provided by the Language Bank of Finland (City Digital Group 2021).

I investigate the occurrences and frequencies of verbs used with the Finnish Existential Partitive Construction in said corpus, using *collostructional analysis* (Stefanowitsch & Gries 2003; Gries & Stefanowitsch 2004). This method allows one to check how strongly the Finnish Existential Partitive Construction and different verbs are attracted to each other, i.e. how often different verbs can feature in that construction in corpora. Collostructional analysis is a quantitative method of research which may aid in supporting qualitative findings as well.

3.1 Sampling

The material used in this study is based on a list of lexical existentials collected by Larjavaara (2019) in his research monograph on the Finnish partitive case. A random subset of verbs have been collected from that list, verbs that have fairly simple meanings which may be considered more likely to be frequent in everyday conversations (e.g., *elää* ‘to live’ instead of *eläytyä* ‘to empathize’). Then, the POTENTIAL OCCURRENCE ratio (for simplicity, PO) was calculated, expressed in percentage. The ten verbs with highest POs (>45%) were then considered. PO represents, in this case, the probability of occurrence of the Finnish Existential Partitive Construction and is indicated by the formula in (6). As mentioned in Section 2, the verbs that occur as lexical existentials can co-occur with either a nominative or a partitive-marked NP, of which only the latter may invariably lead to an existential/presentational reading and can thus be used as an operationalization to measure the verb’s potential for an existential/presentational interpretation of the sentence.

$$(6) \quad PO = RF / AF$$

The Potential Occurrence (PO) of a verb is hence the ratio between its RELATIONAL FREQUENCY (RF) and its ABSOLUTE FREQUENCY (AF). I call RELATIONAL FREQUENCY (RF) what is obtained by constraining the corpus search so that it yields, in this case, all the affirmative occurrences of the verb in 3SG. This is a precondition for the occurrence of the Existential Partitive Construction because of the absence of agreement (Hakanen 1972). Negative sentences are also not considered, as negative polarity often triggers partitive marking. In contrast, the ABSOLUTE FREQUENCY (AF) is the number of occurrences of a given verb in the whole corpus, regardless of the construction in which it appears. I have extracted from the corpus all the occurrences of each of the ten verbs marked for 3SG (see RF above), then I have randomized each sample obtained and extracted the first 1,000 sentences resulting from the randomization. From these

1,000 sentences (per verb), I have extracted the ones that contained the Finnish Existential Partitive Construction (see OCCURRENCE FREQUENCY in Section 3.1).

The ten verbs in Table 1 constitute all the verbs in my analysis, alongside their frequencies and their Potential Occurrence, expressed in percentage. It should be brought to attention that the frequencies vary significantly across verbs, since some of them are used, both absolutely and relatively to the analyzed construction, considerably more than others.

Table 1 Verbs used in the analysis

Verb	Gloss	Absolute Frequency	Relational Frequency	Potential Occurrence
löytyä	‘to be found’	116,068	72,276	62.27%
sisältyä	‘to be included’	4,104	2,287	55.73%
riittää	‘to suffice’	78,275	43,475	55.54%
valua	‘to flow’	6,053	2,995	49.48%
tulla	‘to come’	708,448	350,415	49.46%
ilmestyä	‘to appear’	10,951	5,232	47.78%
pyöriä	‘to rotate/circulate’	18,871	8,925	47.29%
mahtua	‘to fit’	12,029	5,674	47.17%
sataa	‘to rain’	7,759	3,636	46.86%
jäädä	‘to remain/stay’	143,980	66,537	46.21%

We can observe, for example, that *tulla* ‘to come’ has the highest number of occurrences on a corpus level (708,448), while *sisältyä* ‘to be included’ appears only 4,104 times. The verb *tulla* also has the highest Relational Frequency, i.e., it occurs 350,415 times in 3SG, but it appears frequently indexed for other persons, too. For this reason, we could expect *sisältyä* to feature in the Existential Partitive Construction more than *tulla*, because its ratio between the two (yet smaller) frequencies is higher (see 3rd column, PO).

3.2 Collostructional Analysis

In corpus linguistics, *collocation* typically refers to lexical co-occurrence patterns, usually defined based on the items’ proximity (for an overview see, e.g., Ivaska 2015: 25–29, and therein indicated sources). The term *collostruction* (see above) is inspired by said patterns, but instead of the immediate collocations of a word in a corpus, it takes into account the specific construction (as defined in, e.g., Goldberg 1995) the word occurs in, as well as its internal variation. Words, i.e., lexemes, are within this theoretical framework called *collexemes*.

In this paper, the collocation analyzed is the Finnish Existential Partitive Construction, while the ten verbs contained in Table 1 constitute the collexemes. The collexemes are tested for their *collocational strength*, i.e., their ability to attract the collocation under analysis. The strength of a collocation can be measured by means of non-parametric statistical tests. Usually, a good fit for this purpose is Fisher's exact test, which is also the test used in this paper.

Fisher's exact test is used to analyze contingency tables, such as 2×2 crosstabulations that display the frequency distribution of different variables. In this case, the variables involved are the verbs and their frequencies within and out of the construction analyzed (see §4 for details). As established in this type of research, *p*-values will be reported, and their significance threshold is conventionally set to $p < .05$. Collexemes may be ordered according to how strongly they are associated with the construction analyzed. The list of verbs is hence accompanied by an ordered range of *p*-values, whereby the smaller ones indicate stronger attraction to the collocation compared to the larger ones. A *p*-value of 1 would indicate complete repulsion of the collocation, i.e., no association whatsoever.

3.3 Dispersion

Within the study at hand, I also measure dispersion in the corpus to exclude the eventuality that constructions be associated with a specific topic or corpus subsection. A well-dispersed construction appears evenly within the sample. It has been shown (Biber et al. 2016: 459) that the most effective way to obtain dispersion is Gries' (2008) Deviations of Proportions, which considers the difference between expected and observed percentages of occurrence of a word or linguistic phenomenon in different corpus parts,⁴ and always ranges from 0 to 1 (Gries 2008: 415–416), where values closer to 0 mean even dispersion. Gries' Deviations of Proportions is included as a criterion in the extended version of collocational analysis (Gries 2019), which considers not only frequencies but also, e.g., the direction of attraction and repulsion of the collexemes to the collocation and vice versa.

Intuitively speaking, verbs like *sataa* 'to rain' are more prone to be associated with partitive-marked e-NPs, compared to how prone partitive-marked e-NPs are to attract the verb *sataa*, essentially because of the verb's semantics (for the problems involving meteorological expressions see Eriksen, Kittilä & Kolehmainen

⁴ The corpus is divided by topics, which include the following (here translated from Finnish): Entertainment and Culture, Family, Finance, Groups, Health, Hobbies, Home and Construction, Leisure, Pets, Places, Relationships, Science and Technology, Sex, Society, Sports and Fitness, Traveling, Vehicles and Transportation, Work and Study, Youth, among others. Each main topic is further divided into numerous subsections.

2010). Other verbs' much larger Potential Occurrence ratios point towards a not very successful analysis of said direction, since verbs like *löytyä* 'to be found' counterintuitively have a higher probability to attract partitive e-NPs, as the analysis presented here shows. In other words, I believe it is sufficient to analyze these collocations based on the criterion of attraction/repulsion, disregarding the direction of this attraction/repulsion. Moreover, partitive-marked e-NPs only alternate with nominative (e-)NPs in the Finnish Existential Construction, so the number of extra columns in the co-occurrence table suggested by Gries (2019: 391) is also not relevant for the present purposes.

3.4 Two Different Ways of Applying Collostructional Analysis

The first way of applying collostructional analysis is based on a reduced sample size: The sample used here consists of 1,000 sentences for each verb with the restrictions described for their Relational Frequency, totaling 10,000, and have been coded manually for the nominative/partitive alternation of the NP. This reduction of the sample was necessary because the Finnish Existential Partitive Construction is not transparent, even in syntactically coded corpora, meaning that in many cases the automatic annotation fails at recognizing this construction, and can include other constructions in which the partitive case expresses something else, e.g., a time adverbial. The present analysis shows that this approach can lead to bias, as it is not possible to analyze all the occurrences of the construction: The samples are all equal in size and the information about the total occurrences of the word in the whole corpus inevitably slips away.

For this reason, the second time I apply collostructional analysis to a different sample, which I call EXPECTED SAMPLE SIZE (ES). It is calculated based on the OCCURRENCE FREQUENCY (OF), i.e., the number of occurrences of a collexeme in the previously introduced reduced sample of 1,000 sentences, according to the following proportion (7a), where ES represents the incognita (7b).

- (7) a. $AF : RF = ES : OF$
 b. $ES = AF \times OF / RF$

The other constituents of the proportion are the already introduced Absolute Frequency (AF), Relational Frequency (RF) and Occurrence Frequency (OF). If we postulate that RF gives us the exact number of partitive e-NPs in the whole corpus, we understand how the two sides of the proportion above relate to each other. RF is the ideal occurrence frequency of the Existential Partitive Construction on a corpus level, while OF is the observed (i.e., the actual) occurrence frequency of the same construction. ES can hence only be an approximation (because the elements in RF do not all satisfy the criteria of the construction at hand)

but is a more reliable sample size than a reduced sample of 1,000 occurrences, which does not provide any information about the context in which these occurrences are found. The proportion is based on the idea of quantitatively including the whole corpus of sentences and has the advantage to observe large-scale effects even with a small sample at hand, providing an estimate, on a corpus level, of the tendencies of the construction under investigation.

For example, let us consider the verb *ilmestyä* ‘to appear’. As shown in **Table 1**, it has an Absolute Frequency of 10,951 and a Relational Frequency (i.e., 3SG indexing) of 5,232. Its Occurrence Frequency, as shown in Section 4, is 130: This number represents the occurrence of the Existential Partitive Construction in the reduced sample of 1,000 sentences. By applying the proportion (7a), we get to calculate the (rounded) ES, i.e., the Expected Sample size for the verb *ilmestyä*, in the following way (8).

$$(8) \quad ES = 10,951 \times 130 / 5,232 \approx 272$$

By calculating ES, my aim is to address the following question: How many sentences should we analyze to observe the frequency of the Existential Partitive Construction we observed in the first, reduced sample? For (8) above, the analyzed sample should ideally consist of at least 272 sentences to possibly observe the maximum Occurrence Frequency of 130 Existential Partitive sentences. This quantity can of course vary both in an absolute way, because performing random sampling will yield different results each time, and relatively to the size of the initial reduced sample.

4 RESULTS

In this section, I demonstrate the results of the different applications of collocation analysis. First, I apply the method to a reduced sample, then to the Expected Sample size. The ten verbs remain unaltered.

4.1 Reduced Sample

Collocation analysis is here applied on a reduced sample of 10,000 sentences in total, 1,000 per verb. It is good to restate that the column labeled Occurrence Frequency in Table 3 and later indicates how many instances of the Finnish Existential Partitive Construction were found for each verb, within the 1,000 total examples that contain each verb. The Potential Occurrence (PO) is maintained as a reminder of the fact that these verbs were not chosen randomly, but rather based on the ratio between their previously defined Relational Frequency and overall Absolute Frequency. Table 2 gives an example of the crosstabulation used for carrying out Fisher’s exact test, while Table 3 exemplifies the results.

Table 2 Example of crosstabulation with the verb *löytyä* (reduced sample)

	(+) Existential Partitive	(-) Existential Partitive
(+) <i>löytyä</i>	490	510
(-) <i>löytyä</i>	2,250	6,750

By looking at the crosstabulation above, we can notice that:

- the first number of the first column (490) is the Occurrence Frequency of the verb *löytyä*, i.e. the number of occurrences of the Existential Partitive Construction out of 1,000 examples with said verb;
- the second number of the first column (2,250) is the sum of the Occurrence Frequencies of all the other verbs (i.e. the occurrences of the verbs within the construction);
- the first number of the second column (510) is the number of constructions that are different from the Existential Partitive Construction, and in which the verb *löytyä* appears (= [1,000 sentences] – 490);
- the second number of the second column (6,750) is the number of constructions that are different from the Existential Partitive construction, and in which all the other verbs appear (= [9,000 total sentences] – 2,250).

Table 3 Occurrence Frequencies of the Existential Partitive Construction in the corpus, Gries' DP values and Fisher's exact test (reduced sample)

Verb	Gloss	Potential Occurrence	Occurrence Frequency	Fisher's exact test ^a	Gries' DP
<i>löytyä</i>	'to be found'	62.27%	490	3.531449e-53 *	.54
<i>sisältyä</i>	'to be included'	55.73%	238	.007108333 *	.39
<i>riittää</i>	'to suffice'	55.54%	390	5.097901e-17 *	.49
<i>valua</i>	'to flow'	49.48%	173	4.224982e-15 *	.47
<i>tulla</i>	'to come'	49.46%	200	1.53726e-08 *	.45
<i>ilmestyä</i>	'to appear'	47.78%	130	2.20209e-30 *	.23
<i>pyöriä</i>	'to rotate/circulate'	47.29%	139	1.235105e-26 *	.47
<i>mahtua</i>	'to fit'	47.17%	419	2.156685e-25 *	.52
<i>sataa</i>	'to rain'	46.86%	417	7.447229e-25 *	.54
<i>jäää</i>	'to remain/stay'	46.21%	144	1.090924e-24 *	.41

a) Statistically significant *p*-values are marked with an asterisk.

As for the results, at first glance, we may notice that all the verbs are well dispersed in the corpus, since Gries' DP never approaches the value 1, but remains instead in the middle-to-low range. There is also a difference between the verbs'

PO and OF, meaning that a high value of Potential Occurrence does not effectively guarantee the appearance (Occurrence Frequency) of the construction. This is partly due to verb semantics since some verbs may naturally favor the introduction of new referents into discourse, which is the main function of the Finnish Existential Construction, more than others. For example, the verb *löytyä* ‘to be found’ appears 490 times, which means that the frequency of the Finnish Existential Partitive Construction is higher, by 71 sentences, than the frequency of the second most frequent verb, *mahtua* ‘to fit’, which appears 419 times. The verb *löytyä* also has the highest value of Potential Occurrence: 62.27%. This means that its PO is by 6.54% higher than the PO of the second most likely verb in the list, *sisältyä* ‘to be included’ (PO = 55.73%).

As shown by Basile & Ivaska (2021), *löytyä* often has a meaning that is equivalent to the copula *olla* ‘to be’, the most prototypical existential verb, and can be considered a sort of special locative copula (as defined by Haspelmath 2022). The study at hand corroborates this finding by providing quantitative evidence for it. Out of the observed verbs, *löytyä* is the most strongly attracted to the Finnish Existential Construction, both as Existential Nominative and Partitive Construction. In fact, the existential function of *löytyä* seems to have grammaticalized through coexpression from an original verb with meaning FIND, and to have undergone semantic bleaching, hence expressing a merely locational function without indicating a specific posture, exactly like the locative copula *olla* ‘to be’ (6a). Moreover, it is often used to only state the existence of a referent, without specifying its location, as in (9).⁵

- (9) a. *Varmaan Tamperee-lta löyt-y-y*
 surely Tampere-ABL find-MM-3SG
vastaav-i-a paikk-oj-a
 corresponding-PL-PTV place-PL-PTV
 ‘For sure there are similar places in Tampere.’
- b. *Tekosy-i-tä kyllä löyt-y-y.*
 excuse-PL-PTV yes.INT find-MM-3SG
 ‘There are always excuses.’

The copula *olla* ‘to be’ even has a lower value of Potential Occurrence compared to *löytyä* (54.75%, from the same corpus). This is supposedly due to its versatility, since it is used not only in locative or existential constructions, but also as a predicative copula and as an auxiliary (Relational Frequency 4,796,047; Absolute Frequency 8,759,759). This copula is also frequently used in the 1st and 2nd

⁵ These constructions are considered existential constructions in the Finnish linguistic tradition. From a cross-linguistic perspective, they are more similar to what Haspelmath (2022) calls *hyparctic* constructions. However, no consensus has been reached as to whether we need such a construction type.

persons (10) which, conversely, rarely happens with *löytyä* (Basile & Ivaska 2021: 18; Basile forthcoming).

- (10) *Jos ole-t tosi-ssa-si, oo-t*
 if be-2SG true-INE.2SG be-2SG
joko läski tai muuten ruma.
 either fat or otherwise ugly
 ‘If you are being serious, then you’re either fat or ugly.’

Apart from the case of *löytyä*, not much else can be said about the numbers in Table 2. The *p*-values resulting from the analysis indicate that there is an association between the Finnish Existential Partitive Construction and each one of those verbs, but it is unclear why they change in that way. Except for the first verb, they do not seem to correlate at all with their Occurrence Frequency, which would be logical if we consider the differences in Occurrence Frequencies; this flaw may be attributed to the bias provided by the limited number of example sentences analyzed.

4.2 Expected Sample Size

In this subsection I apply collostructional analysis to the verbs’ Expected Sample size (see 3.1.). ES is an estimate of the ideal sample, calculated proportionally to the overall occurrences of each verb in the corpus (Absolute Frequency, AF), as well as to their Relational Frequency (RF). The Occurrence Frequency is represented by the same numbers used in the previous analysis, i.e., the number of occurrences of the Finnish Existential Partitive Construction in a sample of 1,000 sentences. This number would of course be subject to change in case a new random sampling would occur, hence it is not a constant. The reason why I chose a new way of sampling is to eliminate the bias previously resulting from the reduced, non-proportional sample size of 10,000 sentences in total. An example of crosstabulation is given in Table 4. The results of the analysis are exemplified in Table 5.

Table 4 Example of crosstabulation with the verb *löytyä*

	(+) Existential Partitive	(-) Existential Partitive
(+) <i>löytyä</i>	490	297
(-) <i>löytyä</i>	2,250	2,289

In this case, while the first column of the crosstabulation remains the same as in the previous crosstabulation (see Table 2), the second one significantly changes, as it is based on the Expected Sample size. The sum of the two numbers contained

in the first row of the crosstabulation is the value of the Expected Sample size for the verb *löytyä* (787).

Table 5 Results of colostruational analysis applied on ES, compared with the results from the previous sampling (see Table 3)

Verb	Expected Sample size	Occurrence Frequency	Fisher's exact test (ES)	Fisher's exact test (reduced sample)
löytyä	787	490	4.470362e-11 *	3.531449e-53 *
sisältyä	427	238	.06911321	.007108333 *
riittää	702	390	.02085609 *	5.097901e-17 *
valua	350	173	.4391751	4.224982e-15 *
tulla	404	200	.4374663	1.53726e-08 *
ilmestyä	272	130	.236718	2.20209e-30 *
pyöriä	294	139	.149758	1.235105e-26 *
mahtua	888	419	.005792884 *	2.156685e-25 *
sataa	890	417	.002904667 *	7.447229e-25 *
jäädä	312	144	.05430832	1.090924e-24 *

As for the results in Table 5, compared to the results in Table 3, *p*-values are not all significant, because we have a sample size that reflects the overall occurrence (Absolute Frequency) of the verb on a corpus level. If we condition the *p*-value on a null hypothesis, according to which the association between the studied collexemes and colostruaction is coincidental, we can see that not all verbs are significantly associated with the construction at hand. Perhaps this means that the verbs with higher *p*-values, while still being associated with the relevant colostruaction, will be more likely to be associated with other construction types, such as the Existential Nominative Construction (where the e-NP is unmarked but always clause-final), or constructions in which an unmarked NP functions as subject.

In contrast, in the analysis carried out on the reduced sample (§ 4.1), we would expect significant association between the verbs and the Finnish Existential Partitive Construction, while excluding the possibility of having no association between collexemes and colostruaction: The aim is simply to detect an ordered set of *p*-values that would mirror and describe the Occurrence Frequency. Of course, since the sample is biased, this did not happen.

The Expected Sample sizes provide important information. For example, in the case of the verb *sataa* 'to rain', which has an Absolute Frequency of 7,759, its Expected Sample size tells us that of every 890 sentences, we might expect to find up to 417 occurrences (OF), almost half of the sample, of the Finnish Existential Partitive Construction. The non-significant values in Table 5 tell us that, given a certain sample size, a certain number of occurrences, and the relationship

between these and the other verbs' occurrences, the null hypothesis that the association between the studied collexemes and collocation is coincidental cannot be fully rejected. For this reason, there is a higher probability that the occurrences of the Finnish Existential Partitive Construction found in the smaller ESs be due to chance. Using the .05 threshold level of statistical significance, we can say that collexemes having an OF of 390 or more, and an ES above 700, are associated to the Finnish Existential Partitive Construction, and that this association is strong and not coincidental. Even though here the sample is also limited, it relates to the overall occurrence (Absolute Frequency, AF) of each verb on a corpus level and allows for possible resampling that would likely yield similar results. Again, the Expected Sample size of the verb *löytyä* is smaller than the ES of the second most frequent verb *mahtua* by 101 units (clauses), yet its Occurrence Frequency is significantly bigger.

With this method, not only can we determine an approximative ranking of association of the collexemes with the collocation, but we also have statistical significance to corroborate such claims. On a reduced sample of 1,000 sentences, at least more than 1/3 of them should be occurrences of the Existential Partitive Construction to say that a specific collexeme can be strongly associated with said collocation. This does not, however, mean that the appearance of the Existential Partitive Construction with verbs without a significant *p*-value is completely random, only that it is less likely to happen.

The highest-associated verb is *löytyä* 'to be found', followed by both *mahtua* 'to fit' and *sataa* 'to rain' equally on the second place, whose slight difference between ES, OF and *p*-value can be ignored, as it does not add any significant piece of information. The verb *riittää* 'to suffice' is the third most featured collexeme. We may observe that all the significantly associated collexemes differ by one decimal digit, except for the outlier *löytyä*.

The *p*-value of *löytyä* is, also in this sampling, so small compared to the others that its association with the Finnish Existential Partitive Construction is way beyond the threshold to be considered a mere coincidence. Even within this second sampling, we can substantiate with quantitative evidence the claim that *löytyä* constitutes a special copula that behaves similarly to the copula *olla* 'to be'.

The verbs *mahtua* 'to fit' and *riittää* 'to suffice', on the contrary, do not primarily have a locational function. Why the Finnish Existential Partitive Construction is significantly associated with them could also be explained by the fact that both these verbs are semantically associated with (un)bound quantities, which are often expressed by the partitive case in Finnish. They can, in fact, also have the function of introducing new referents into discourse, hence having a presentational configuration (Gast & Haas 2011). Despite this, I still found examples that feature these two verbs in what resembles a locational use (11–12).

- (11) a. *Maailma-an* ***mahtu-u*** *paljon*
 world-ILL fit-3SG many
itsekeskeis-i-ä *ego"surffare-i-ta"*.
 egotistical-PL-PTV ego=surfer-PL-PTV
 ‘There are a lot of egotistical ‘egosurfers’ in the world.’
- b. *Päivä-än* ***mahtu-u*** *suunnaton-ta*
 day-ILL fit-3SG enormous-PTV
suru-a, *ahdistus-ta,* *viha-a,*
 sorrow-PTV anxiety-PTV hate-PTV
kiintymys-tä, *raivo-a,* *rakkau-tta.*
 affection-PTV rage-PTV love-PTV
 ‘There is enormous sorrow, anxiety, hate, affection, rage and love in every day.’
- (12) a. *katu=pöly-ä* ***riittä-ä*** *kaikkialle.*
 street=dust-PTV suffice-3SG everywhere.ALL
 ‘There is street dust everywhere.’
- b. *Hiljaisuus-tta* ***riittä-ä*** *täällä*
 silence-PTV suffice-3SG here
koti=maa-ssa.
 home=country-INE
 ‘There is (enough/so much) silence in our home country.’

It is worth noting that the examples above not only seem to convey the main function of existential constructions, which is that of placing a discourse-new referent in a location, but also seem to add meaning related to the abundance and indefiniteness of the referent being located, in line with their semantics. Contrarily, the other verbs in the sample, such as *ilmestyä* ‘to appear’, seem to prefer definite quantity referents that are rarely marked for partitive.⁶

The case of *sataa* ‘to rain’ is *sui generis*, because in Finnish this verb triggers almost always obligatory case government in partitive (cf. 13).

- (13) *sataa* *lunta/*lumi*
 rain.3SG snow.PTV/*NOM
 ‘It is snowing’ [lit. ‘It is raining snow’].

⁶ I thank an anonymous reviewer for pointing out that in the closely related language Estonian both *leiduma* ‘to be found’ and *ilmuma* ‘to appear’ can be found in several existential clause types, with the difference that the former prefers partitive-marked e-NPs, while the latter nominative-marked e-NPs (Rätsep 1978).

5 DISCUSSION AND CONCLUSIONS

This research shows how collocation analysis can be applied to the Finnish Existential Partitive Construction. The main problem with this construction is that partitive-marked e-NPs, which constitute the key morphosyntactic element in this type of locational construction, are often not recognized by automatically annotated corpora. Morphologically they are, in fact, similar to other parts of speech also coded in the partitive case (e.g. temporal and other types of adverbials), but they specifically appear as the main argument of an Existential Construction, and hence have the main function of introducing a new referent in the discourse.

In a first attempt to address the question whether some existential verbs are more prone to appear in the Finnish Existential Partitive Construction, I have applied collocation analysis to a reduced sample. The reduction was necessary since it is not convenient to select a very large sample, as this would require manually polishing the whole corpus because of ambiguity in the coding of partitive e-NPs. I have hence selected a random sample of 10,000 clauses in total, 1,000 per verb. Then, I have counted the Occurrence Frequency (OF) of the Finnish Existential Partitive Construction for each verb and applied the method in question. The resulting values are not very revealing about the statistical significance of association between the verbs and the construction and I have attributed this to the bias provided by the reduced sampling.

Because of these first results, I have decided to reapply the method considering an Expected Sample size (ES), proportionally calculated by maintaining the verbs' Occurrence Frequencies (OF) obtained in the first sampling, and comparing them to their Absolute Frequencies (AF), i.e. the total occurrences of the verbs in the whole corpus, and their Relational Frequencies (RF), which I have obtained by limiting the corpus search to occurrences in 3SG and by eliminating occurrences with negative polarity, in order to favor the appearance of partitive e-NPs that are not triggered by negation. This new application has yielded unbiased results, which consist of four significant *p*-values out of ten. These findings may be interpreted as follows: the association between these four verbs and the Finnish Existential Partitive Construction is strong and not coincidental.

The most important contribution of this paper is corroborating the finding that the verb *löytyä* 'to be found' is used as a special existential copula. In fact, not only is it more likely to appear with the Finnish Existential Partitive Construction compared to the copula *olla* 'to be', but it also has the features of a prototypical locative copula, since it can undergo semantic bleaching and express a purely locational function, without semantically marking a specific posture. Two other verbs, *mahtua* 'to fit' and *riittää* 'to suffice', are also significantly associated with the Finnish Existential Partitive Construction. These two verbs may express typical locational traits of the Existential Partitive Construction, but they also mark semantic traits related to the indefiniteness of the referent being located.

They may also indicate that the referent being located is to be found in an excessive quantity.

The research conducted here leaves some important questions unanswered: Which other constructions provide the same challenges as the Finnish Existential Partitive Construction? How effective would it be to use an Expected Sample size on other constructions, possibly drawing from different languages? Are there any advantages in averaging smaller samplings and applying the method recursively? Despite these uncertainties, using Expected Sample sizes seems to fit well the purpose of this paper.

The statistically significant association between some of the collexemes and the collocation at hand is certainly interesting and points toward relevant qualitative outcomes, not necessarily excluding that the appearance of the rest of the collexemes within the collocation be coincidental. However, since the interpretation of the p -values obtained from the analysis on the Expected Sample size is conditioned on a random sample that can potentially yield very different results, the strength of association of the collexemes could vary significantly. Further research is required to develop a method or a set of methods that aim at tackling this variability.

COMPETING INTERESTS

The author has no conflicts of interest to declare that are relevant to the content of this chapter.

REFERENCES

- Barðdal, Jóhanna. 2006. Construction-specific properties of syntactic subjects in Icelandic and German. *Cognitive Linguistics* 17(1): 39–106.
- Bartens, Raija. 1999. *Mordvalaiskielten rakenne ja kehitys*. [The Structure and Development of the Mordvinic Languages] Helsinki: Suomalais-ugrilaisen seuran toimituksia.
- Basile, Rodolfo. 2023. *Inventive-locational constructions in the languages of Europe*. Preprint on Open Science Framework: <https://doi.org/10.17605/OSF.IO/ES3FH>
- Basile, Rodolfo. Forthcoming. ‘I am also found on Facebook’: Locuphoric find-based strategies in Finnish internet corpora. In *Partitive constructions and partitive elements within and across language borders in Europe*, ed. Elvira Glaser, Petra Sleeman, Thomas Strobel and Anne Tamm. Edizioni Ca’Foscari, Venice University Press.
- Basile, Rodolfo, and Ilmari Ivaska. 2021. Löytyä-verbin konstruktioiden yhteydessä esiintyvä subjektin sijanvaihtelu. [Subject Case Alternation in Constructions Containing the Finnish Verb *Löytyä*] *Eesti ja soome-ugri keeleteaduse ajakiri*. *Journal of Estonian and Finno-Ugric linguistics* 12(1): 11–39.

- Bernhardt, Mariann. 2021. *Määräisyys ja objektin merkitseminen mordvalaiskielissä*. [Definiteness and Object Marking in the Mordvinic Languages] Doctoral dissertation, University of Turku.
- Biber, Douglas, Randi Reppen, Erin Schnur, and Romy Ghanem. 2016. On the (non) utility of Juilland's D to measure lexical dispersion in large corpora. *International journal of corpus linguistics* 21 (4): 439–464.
- City Digital Group (2021). Suomi24 virkkeet – korpus 2001–2020, Korp-versio [Suomi24-Corpus 2001-2020]. Kielipankki. <http://urn.fi/urn:nbn:fi:lb-2021101525>
- Creissels, Denis. 2014. *Existential predication in typological perspective*. <http://www.deniscreissels.fr/public/Creissels-Exist.Pred.pdf>. Accessed 18 December 2021.
- Creissels, Denis. 2019. Inverse-locational predication in typological perspective. *Italian journal of linguistics* 31 (2): 38–106.
- Croft, William. 2001. *Radical construction grammar: Syntactic theory in typological perspective*. Oxford: Oxford University Press.
- Eriksen, Pål Kristian, Seppo Kittilä, and Leena Kolehmainen. 2010. The linguistics of weather: cross-linguistic patterns of meteorological expressions. *Studies in language. International journal sponsored by the foundation "Foundations of Language"* 34 (3): 565–601.
- Gast, Volker, and Florian Haas. 2011. On the distribution of subject properties in formulaic presentationals of Germanic and Romance. In *Impersonal constructions: A cross-linguistic perspective*, ed. Andrej L. Malchukov and Anna Siewierska, 127–166. Amsterdam/Philadelphia: John Benjamins.
- Goldberg, Adele E. 1995. *Constructions: A construction grammar approach to argument structure*. Chicago: University of Chicago Press.
- Gries, Stefan Th. 2008. Dispersions and adjusted frequencies in corpora. *International journal of corpus linguistics* 13 (4): 403–437.
- Gries, Stefan Th. 2019. 15 years of collocations: some long overdue additions/corrections (to/of actually all sorts of corpus-linguistics measures). *International journal of corpus linguistics* 24 (3): 385–412.
- Gries, Stefan Th., and Anatol Stefanowitsch. 2004. Extending collocation analysis: A corpus-based perspective on 'alternations'. *International journal of corpus linguistics* 9 (1): 97–129.
- Hakanen, Aimo. 1972. Normaali- ja eksistentiaalilause. [Normal Clause and Existential Clause] *Sananjalka* 14 (1): 36–76.
- Haspelmath, Martin. 2010. Comparative concepts and descriptive categories in cross-linguistic studies. *Language* 86(3): 663–687.
- Haspelmath, Martin. 2022. *Nonverbal clause constructions*. <https://ling.auf.net/lingbuzz/006673>. Accessed 5 September 2023.
- Helasvuo, Marja-Liisa. 1996. Ollako vai eikö olla – Eksistentiaalilauseen subjektin kohtalonkysymys. [To Be or not to Be – The Fate of the Subject of the Existential Clause] *Virittäjä* 100 (3): 340–356.
- Helasvuo, Marja-Liisa, and Tuomas Huomo. 2010. Mikä subjekti on? [On the subject of subject in Finnish]. *Virittäjä* 114 (2): 165–195.
- Huomo, Tuomas. 2003. Incremental existence: The world according to the Finnish existential sentence. *Linguistics* 41 (3): 461–493.

- Huumo, Tuomas. 2018. The partitive A: On uses of the Finnish partitive subject in transitive clauses. In *Diachrony of differential argument marking*, ed. Ilja A. Seržant and Alena Witzlack-Makarevich, 423–453. Berlin: Language Science Press.
- Huumo, Tuomas, and Marja-Liisa Helasvuo. 2015. On the subject of subject in Finnish. In *Subjects in constructions – canonical and non-canonical*, ed. Marja-Liisa Helasvuo and Tuomas Huumo, 13–41. Amsterdam/Philadelphia: John Benjamins.
- Huumo, Tuomas, and Liina Lindström. 2014. Partitives across constructions: On the range of uses of the Finnish and Estonian ‘partitive subjects’. In *Partitive cases and related categories*, ed. Silvia Luraghi and Tuomas Huumo, 153–176. Berlin, Boston: De Gruyter.
- Ikola, Osmo. 1954. Suomen lauseopin ongelmia [I–III]. [Problems in Finnish Syntax] *Virittäjä* 58 (3): 209–245.
- Ivaska, Ilmari. 2010. Eksistentiaalilauseen subjektijäsen edistyneiden suomenoppijoiden kirjoituksessa. [Subject Constituent of the Existential Sentence in the Writing of Advanced Learners of Finnish] *Sananjalka* 52: 67–88.
- Ivaska, Ilmari. 2011. Lausetyyppien sekoittuminen edistyneessä oppijansuomessa – näkökulmana eksistentiaalilause. [Clause Type Confusion in Advanced Learner Finnish – Existential Sentence as Our Perspective] *Lähivõrdlusi. Lähivertailuja* 21: 65–85.
- Ivaska, Ilmari. 2015. *Edistyneen oppijansuomen korpuspiirteitä korpusvetoisesti: Avainrakenteanalyysi*. [Corpus-driven Approach towards Constructional Features of Advanced Learner Finnish: Key Structure Analysis] Doctoral dissertation, University of Turku .
- Järventausta, Marja. 1991. *Das Subjekt im Deutschen und im Finnischen. Seine Formen und semantischen Rollen*. [The Subject in German and in Finnish. Its Forms and Semantic Roles.] Frankfurt am Main: Peter Lang.
- Kajander, Mikko. 2013. *Suomen eksistentiaalilause toisen kielen oppimisen polulla*. [Paths of Learning Finnish Existential Sentences] Doctoral dissertation, University of Jyväskylä .
- Kiparsky, Paul. 1998. Partitive case and aspect. In *The projection of arguments: Lexical and compositional factors*, ed. Miriam Butt and Wilhelm Geuder, 265–307. Stanford: CSLI Publications.
- Kittilä, Seppo, and Jussi Ylikoski. 2011. Remarks on the coding of goal, recipient and vicinal goal in European Uralic. In *Case, animacy and semantic roles*, ed. Seppo Kittilä, Katja Västi and Jussi Ylikoski, 29–64. Amsterdam/Philadelphia: John Benjamins.
- Koptjevskaja-Tamm, Maria, and Bernhard Wälchli. 2001. The Circum-Baltic languages: an areal-typological approach. In *Circum-Baltic languages*, ed. Östen Dahl and Maria Koptjevskaja-Tamm, 615–750. Amsterdam / Philadelphia: John Benjamins.
- Larjavaara, Matti. 2019. *Partitiivin valinta*. [Choosing of Partitive] Helsinki: Suomalaisen Kirjallisuuden Seura.
- Lees, Aet. 2015. *Case alternations in five Finnic languages: Estonian, Finnish, Karelian, Livonian and Veps*. Leiden: Brill.
- Lindström, Liina. 2017. Partitive subjects in Estonian dialects. *Eesti Ja Soome-Ugri Keeleteaduse Ajakiri. Journal of Estonian and Finno-Ugric Linguistics* 8 (2): 191–231.
- Luraghi, Silvia, and Seppo Kittilä. 2014. Typology and diachrony of partitive case markers. In *Partitive cases and related categories*, ed. Silvia Luraghi and Tuomas Huumo, 17–62. Berlin, Boston: De Gruyter.

- Metslang, Helena. 2012. On the case-marking of existential subjects in Estonian. *SKY Journal of Linguistics* 25: 151–204.
- Metslang, Helena. 2014. Partitive noun phrases in the Estonian core argument system. In *Partitive cases and related categories*, ed. Silvia Luraghi and Tuomas Huumo, 177–256. Berlin, Boston: De Gruyter.
- Rätsep, Huno. 1978. *Eesti keele lihtlausetete tüübid*. [Types of Simple Sentences in Estonian] Tallinn: Valgus. Tallinn.
- Seržant, Ilja A. 2015. The independent partitive as an Eastern Circum-Baltic isogloss. *Journal of language contact* 8(2): 341–418.
- Stefanowitsch, Anatol, and Stefan Th Gries. 2003. Collostructions: Investigating the interaction of words and constructions. *International journal of corpus linguistics* 8(2): 209–243.
- VISK = Auli Hakulinen, Maria Vilkuna, Riitta Korhonen, Vesa Koivisto, Tarja Riitta Heinonen, and Irja Alho. 2004. *Iso suomen kielioppi*. [The Large Grammar of Finnish] Helsinki: Suomalaisen Kirjallisuuden Seura.

CURRICULUM VITAE

Name: Rodolfo Basile
Place and date of birth: Caserta (Italy), 2 April 1993
Email: rodolfobasile993@gmail.com

Education and degrees awarded

Master's Degree Conservatorio Statale di Musica "Nicola Sala", Benevento (Italy) Classical Guitar, 20 July 2018

Master's Degree Università degli Studi di Napoli "L'Orientale", Naples (Italy) Comparative Literature and Culture Studies, Finnish and German languages, 13 April 2018

Bachelor's Degree Università degli Studi di Napoli "L'Orientale", Naples (Italy) Comparative Language, Literature and Culture Studies, Finnish and German languages, 18 February 2015

Work experience

Jan 2018 – Freelance translator

1.10.2021 – Junior Research Fellow in Estonian and Finno-Ugric Linguistics (workload 0.5), University of Tartu

Feb 2021 – Instructor at Communicating Science team, University of Tartu

Feb 2020 – Jun 2020 Teacher of Finnish language for foreigners (0 → A1), University of Tartu

2017 – 31.3.2020 In-house translator at Inter-LMP, d.o.o.

Sep 2018 – Feb 2019 Literary translation traineeship program at FILI (Finnish Literature Exchange), Helsinki

Sep 2017 – Dec 2017 Teaching assistant (Finnish class, teacher Emmi Pollari), University of Helsinki

Apr 2017 – May 2017 Teacher of Finnish language, Liceo Mazzini (Naples, Italy)

Sep 2013 – May 2014 Teaching assistant (Italian class, teacher Marja Härmänmaa), University of Helsinki

List of Publications

- Basile, Rodolfo & Ilmari Ivaska. 2021. Löytyä-verbin konstruktoiden yhteydessä esiintyvä subjektin sijanvaihtelu. *Eesti ja soome-ugri keeleteaduse ajakiri. Journal of Estonian and Finno-Ugric Linguistics*, 12(1), 11–39. <https://doi.org/10.12697/jeful.2021.12.1.01>
- Basile, Rodolfo. 2021. Situative constructions in European languages. *SAE2021 Book of Abstracts*.
- Basile, Rodolfo. 2020. Partitiivin mysteerejä ratkomassa. *Sananjalka*, 62(62), 307–309.
- Basile, Rodolfo. 2019. Alternation between nominative and partitive subjects in constructions containing the Finnish verb löytyä. In: *In Memoriam of Anne Vainikka. Conference Contributions*. November 22–23, 2019. Budapest. (21–26). Budapest: issuu.
- Basile, Rodolfo. 2018. Löytyä-verbin semantiikasta. Tartu: *XXXIV Rahvusvahelise soome-ugri tudengite konverentsi (XXXIV IFUSCO) tessikogumik* ISSN 2613–4985.

ELULOOKIRJELDUS

Nimi: Rodolfo Basile
Sünnikoht ja -aeg: Caserta (Itaalia), 2. aprill 1993
E-post: rodolfobasile993@gmail.com

Haridus ja kraadid

Magistrikraad, Conservatorio Statale di Musica “Nicola Sala”, Benevento (Itaalia)
Klassikaline kitarr, 20. juuli 2018

Magistrikraad, Università degli Studi di Napoli “L’Orientale”, Napoli (Itaalia)
Võrdlev kirjandus- ja kultuuriteadus, soome ja saksa keeled, 13. aprill 2018

Bakalaureusekraad, Università degli Studi di Napoli “L’Orientale”, Napoli (Itaalia)
Võrdlev keele-, kirjandus- ja kultuuriteadus, soome ja saksa keeled, 18. veebruar 2015

Töökogemus

Jan 2018 – Vabakutseline tõlkija

1.10.2021 – Nooremteadur eesti ja soome-ugri keeleteaduses (koormus 0,5), Tartu Ülikool

Veeb 2021 – Instruktor Communicating Science meeskonnas, Tartu Ülikool

Veeb 2020 – Juuni 2020 Soome keele õpetaja välismaalastele (tase 0 → A1), Tartu Ülikool

2017 – 31.3.2020 Tõlkija Inter-LMP, d.o.o. juures

Sept 2018 – Veeb 2019 Kirjandustõlke praktikaprogramm FILI-s (Finnish Literature Exchange), Helsinki

Sept 2017 – Dets 2017 Õppejõu assistent (soome keele kursus, õpetaja Emmi Pollari), Helsingi Ülikool

Apr 2017 – Mai 2017 Soome keele õpetaja, Liceo Mazzini (Napoli, Itaalia)

Sept 2013 – Mai 2014 Õppejõu assistent (itaalia keele kursus, õpetaja Marja Härmänmaa), Helsingi Ülikool

Publikatsioonide loetelu

- Basile, Rodolfo & Ilmari Ivaska. 2021. Löytyä-verbin konstruktioiden yhteydessä esiintyvä subjektin sijanvaihtelu. *Eesti ja soome-ugri keeleteaduse ajakiri. Journal of Estonian and Finno-Ugric Linguistics*, 12(1), 11–39. <https://doi.org/10.12697/jeful.2021.12.1.01>
- Basile, Rodolfo. 2021. Situative constructions in European languages. *SAE2021 Book of Abstracts*.
- Basile, Rodolfo. 2020. Partitiivin mysteerejä ratkomassa. *Sananjalka*, 62(62), 307–309.
- Basile, Rodolfo. 2019. Alternation between nominative and partitive subjects in constructions containing the Finnish verb löytyä. In: *In Memoriam of Anne Vainikka. Conference Contributions*. November 22–23, 2019. Budapest. (21–26). Budapest: issuu.
- Basile, Rodolfo. 2018. Löytyä-verbin semantiikasta. Tartu: *XXXIV Rahvusvahelise soome-ugri tudengite konverentsi (XXXIV IFUSCO) tessikogumik* ISSN 2613–4985.

DISSERTATIONES LINGUISTICAE UNIVERSITATIS TARTUENSIS

1. **Anna Verschik.** Estonian Yiddish and its contacts with coterritorial languages. Tartu, 2000, 196 p.
2. **Silvi Tenjes.** Nonverbal means as regulators in communication: socio-cultural perspectives. Tartu, 2001, 214 p.
3. **Iloa Tragel.** Eesti keele tuumverbid. Tartu, 2003, 196 lk.
4. **Einar Meister.** Promoting Estonian speech technology: from resources to prototypes. Tartu, 2003, 217 p.
5. **Ene Vainik.** Lexical knowledge of emotions: the structure, variability and semantics of the Estonian emotion vocabulary. Tartu, 2004, 166 p.
6. **Heili Orav.** Isiksuseomaduste sõnavara semantika eesti keeles. Tartu, 2006, 175 lk.
7. **Larissa Degel.** Intellektuaalsfäär intellektuaalseid võimeid tähistavate sõnade kasutuse põhjal eesti ja vene keeles. Tartu, 2007, 225 lk.
8. **Meelis Mihkla.** Kõne ajalise struktuuri modelleerimine eestikeelsele tekst-kõne sünteesile. Modelling the temporal structure of speech for the Estonian text-to-speech synthesis. Tartu, 2007, 176 lk.
9. **Mari Uusküla.** Basic colour terms in Finno-Ugric and Slavonic languages: myths and facts. Tartu, 2008, 207 p.
10. **Petar Kehayov.** An Areal-Typological Perspective to Evidentiality: the Cases of the Balkan and Baltic Linguistic Areas. Tartu, 2008, 201 p.
11. **Ann Veismann.** Eesti keele kaas- ja mäarsõnade semantika võimalusi. Tartu, 2009, 145 lk.
12. **Erki Luuk.** The noun/verb and predicate/argument structures. Tartu, 2009, 99 p.
13. **Andriela Rääbis.** Eesti telefonivestluste sissejuhatus: struktuur ja suhtlusfunktsioonid. Tartu, 2009, 196 lk.
14. **Liivi Hollman.** Basic color terms in Estonian Sign Language. Tartu, 2010, 144 p.
15. **Jane Klavan.** Evidence in linguistics: corpus-linguistic and experimental methods for studying grammatical synonymy. Tartu, 2012, 285 p.
16. **Krista Mihkels.** Keel, keha ja kaardikepp: õpetaja algatatud parandussekventside multimodaalne analüüs. Tartu, 2013, 242 lk.
17. **Sirli Parm.** Eesti keele ajasõnade omandamine. Tartu, 2013, 190 lk.
18. **Rene Altrov.** The Creation of the Estonian Emotional Speech Corpus and the Perception of Emotions. Tartu, 2014, 145 p.
19. **Jingyi Gao.** Basic Color Terms in Chinese: Studies after the Evolutionary Theory of Basic Color Terms. Tartu, 2014, 248 p.
20. **Diana Maisla.** Eesti keele mineviku ajavormid vene emakeelega üliõpilaste kasutuses. Tartu, 2014, 149 lk.
21. **Kersten Lehismets.** Suomen kielen väylää ilmaisevien adpositioiden *yli, läpi, kautta* ja *pitkin* kognitiivista semantiikkaa. Tartu, 2014, 200 lk.

22. **Ingrid Rummo.** A Case Study of the Communicative Abilities of a Subject with Mosaic Patau Syndrome. Tartu, 2015, 270 p.
23. **Liisi Piits.** Sagedamate inimest tähistavate sõnade kollokatsioonid eesti keeles. Tartu, 2015, 164 lk.
24. **Marri Amon.** Initial and final detachments in spoken Estonian: a study in the framework of Information Structuring. Tartu, 2015, 216 p.
25. **Miina Norvik.** Future time reference devices in Livonian in a Finnic context. Tartu, 2015, 228 p.
26. **Reeli Torn-Leesik.** An investigation of voice constructions in Estonian. Tartu, 2015, 240 p.
27. **Siiri Pärkson.** Dialoogist dialoogsüsteemini: partneri algatatud parandused. Tartu, 2016, 314 lk.
28. **Djuddah A. J. Leijen.** Advancing writing research: an investigation of the effects of web-based peer review on second language writing. Tartu, 2016, 172 p.
29. **Piia Taremaa.** Attention meets language: a corpus study on the expression of motion in Estonian. Tartu, 2017, 333 p.
30. **Liina Tammekänd.** Narratological analysis of Võru-Estonian bilingualism. Tartu, 2017, 217 p.
31. **Eva Ingerpuu-Rümmel.** Teachers and learners constructing meaning in the foreign language classrooms: A study of multimodal communication in Estonian and French classes. Tartu, 2018, 218 p.
32. **Kaidi Rätsep.** Colour terms in Turkish, Estonian and Russian: How many basic blue terms are there? Tartu, 2018, 181 p.
33. **Kirsi Laanesoo.** Polüfunktsionaalsed küsilauseid eesti argivestluses. Tartu, 2018, 176 lk.
34. **Maria Reile.** Estonian demonstratives in exophoric use: an experimental approach. Tartu, 2019, 240 p.
35. **Helen Türk.** Consonantal quantity systems in Estonian and Inari Saami. Tartu, 2019, 149 p.
36. **Andra Rumm.** Avatud küsimused ja nende vastused eesti suulises argivestluses. Tartu, 2019, 217 lk.
37. **Eleri Aedmaa.** Detecting Compositionality of Estonian Particle Verbs with Statistical and Linguistic Methods. Tartu, 2019, 271 p.
38. **Kristina Koppel.** Näitelausete korpuspõhine automaattuvastus eesti keele õppesõnastikele. Tartu, 2020, 249 lk.
39. **Ilze Tälberga.** On the equivalents of the Latvian verbal prefixes in Estonian. Tartu, 2020, 210 p.
40. **Roger M. A. Yallop.** The affect and effect of asynchronous written artefacts (cover letters, drafts, and feedback letters) within L2 English doctorate writing groups. Tartu, 2020, 312 p.
41. **Mariann Proos.** Meaning and usage of Estonian experience perception verbs. Tartu, 2021, 175 p.
42. **Helen Hint.** From full phrase to zero: a multifactorial, form-specific and crosslinguistic analysis of Estonian referential system. Tartu, 2021, 271 p.

43. **Anton Malmi.** The production of Estonian palatalization by Estonian and Russian speakers. Tartu, 2022, 180 p.
44. **Mari Aigro.** In any case? Estonian spatial cases as argument markers. Tartu, 2022, 223 p.
45. **Maria Tuulik.** Adjektiivide süstemaatiline polüseemia eesti keeles tajusadjektiivide näitel. Tartu, 2022, 182 p.