https://doi.org/10.15388/vu.thesis.271
https://orcid.org/0000-0002-7204-3451

VILNIUS UNIVERSITY
INSTITUTE OF LITHUANIAN LITERATUTE AND FOLKLORE

Ringailė
TRAKYMAITĖ

Adjectival definiteness marking: Lithuanian in a typological context

DOCTORAL DISSERTATION

Humanities,
Philology (H 004)

VILNIUS 2021

This dissertation was written between 2013 and 2020 at Vilnius University.

Academic supervisor:

Prof. Habil. Dr. Axel Holvoet (Vilnius University, Humanities, Philology – H 004).

CORRIGENDUM

As was highlighted during the public defence procedure that took place on December 29, 20021 in the Vincas Krėvė Auditorium of the Faculty of Philology at Vilnius University, both sections 3 (page 33 in this dissertation) and 4 (page 78 in this dissertation) are based on the two below listed articles with minor updates and modifications (the second article was published after the dissertation had been printed):

Section 3	Trakymaitė R. (2018). "Determination and modification: Topology of prenominal attributes in iLithuanian", Kalbotyra, 710 (2018), 84–133. DOI: 10.15388/Kalbotyra.2018
Section 4	Trakymaitė R. (2021). "Adjectival definiteness marking in Lithuanian – one more puzzle piece: Qualitative adjectives that could but do not take definite forms", Baltistica, Vol. 56. No. 1 (2021), 19–79. DOI: 10.15388/Baltistica.56.1.2438

https://doi.org/10.15388/vu.thesis.271
https://orcid.org/0000-0002-7204-3451

VILNIAUS UNIVERSITETAS
LIETUVIŲ LITERATŪROS IR TAUTOSAKOS INSTITUTAS

Ringailė
TRAKYMAITĖ

Būdvardinė apibrėžtumo raiška: lietuvių kalba tipologiniame kontekste

DAKTARO DISERTACIJA

Humanitariniai mokslai,
Filologijos kryptis (H 004)

VILNIUS 2021

Disertacija rengta 2013–2020 metais Vilniaus universitete.

Mokslinis vadovas:

prof. habil. dr. Axel Holvoet (Vilniaus universitetas, humanitariniai mokslai, filologija – H 004).

ACKNOWLEDGEMENTS

I would like to express my gratitude to several people who made this dissertation possible. Firstly, I am grateful to Ērika Sausverde, who one nice and sunny afternoon in Vilnius mentioned my future supervisor to me and compellingly advised me to pursue a PhD under his guidance. Secondly, most sincerely and wholeheartedly, I am grateful to my supervisor Axel Holvoet for guiding me through this long and strenuous process and allowing me to discover the whole universe of (in)definiteness and other grammatical categories at my pace and to make my mistakes and my discoveries. Also, I am truly grateful to Axel for welcoming me to the *Academia Salensis* and for introducing me to this new era of summer schools, conferences, workshops, and a whole group of linguist friends, some of whom today I proudly call my closest allies. Thirdly, I am grateful to many remarkable women – both colleagues and good friends – who were beside me on this long journey: my very special thanks go to Solveiga Armoskaite, Birutė Spraunienė, Erika Jasionytė-Mikučionienė, Asta Laugalienė, Loreta Vaicekauskienė and Giedrė Junčytė.

I am truly obliged to my colleagues at the Centre for Scandinavian Studies, who helped and supported me through a not so trivial muddle of administrative procedures and requirements, especially Liana Šimanskienė, who was my go-to person with all sorts of questions and needs. Also, to Birutė Gudelienė of the university administration for her efficient and kind help, especially in the final stages of my studies.

I am also thankful to colleagues and friends who let me pick their brains regarding various multilingual examples, translations and interpretations: Jim Degrenius, Therese Rantakokko, Amber Newman, Vladimir Panov, and Jurgis Pakerys. I am truly grateful to Cara Quinton, Sarah Collins and Jacqueline Hogue, who proof-read my articles, as well as to Wayles Browne, who edited this dissertation.

My special thanks go to Ignas Rudaitis for his great help with statistics and logic insights.

Finally, I am grateful to my husband Dainius, my biggest supporter and champion, and to my boys Bartas and Sembas for enabling me to dedicate myself to academic endeavours and to accomplishing this much, and my brother Vaidotas for always being there for me.

Last, but not least I would like to express my gratitude to my parents Antanas Trakymas and Rita Trakymienė for introducing me to the marvellous world of knowledge and science.

Vilnius 2021

CONTENTS

LIST OF TABLES

LIST OF APPENDICES

ABBREVIATIONS AND DATA SOURCES

Adj	adjective/adjectival modifier
ADM	adjectival definiteness marking
ART	article
BNC	British National Corpus
CB	Circum-Baltic
CCLL	Corpus of the Contemporary Lithuanian Language (‘Dabartinis lietuvių kalbos tekstynas’) 2011
CG	Cognitive Grammar
CGEL	*The Cambridge Grammar of the English Language* 2002
COMM	common gender (in Swedish *utrum*)
Dan.	Danish
DEF	definite
Dem	demonstrative
Det	determiner
DP	determiner phrase
FrD	*Frequency Dictionary of the Written Lithuanian Language* (‘Dažninis rašytinės lietuvių kalbos žodynas’) 2009
GDS	*Grammatik over det Danske Sprog* 2010
Germ.	German
INDEF	indefinite
Lat.	Latin
LF	long adjectival form (definite/strong)
Lith.	Lithuanian
ltTenTen14	Lithuanian Web Corpus 2014
NEUTR	neuter gender (in Swedish *neutrum*)
NFS	*Norsk som fremmedspråk: Grammatikk* 2008
NONDEF	non-definite
NP	noun phrase
NRG	*Norsk referansegrammatikk* 1997
PDEF	preposed free definite article
PM	peripheral modifier
PP	prepositional phrase
Q	quantifier
Rus.	Russian
SDEF	suffixed (postposed) definite article
SAE	Standard Average European
SAG	*Svenska akademiens grammatik* 1999
SF	short adjectival form (non-definite/weak)
Språkbanken	The Swedish Language Bank
Swed.	Swedish

1. INTRODUCTION

1.1. What is adjectival definiteness marking?

The object of this dissertation is adjectival definiteness marking (ADM). It is a grammatical feature whereby the (in)definite status of a noun phrase (NP) is marked partly or solely on an adjective that is an internal part of the NP. Definiteness is a category of an NP. Hence, ADM, which sometimes implies that the sole locus of definiteness marking is the adjective, is a morphosyntactic feature that is attributed to an entire NP rather than an adjective as a lexical class:

(1) a. Arabic

al-*kitab*	***al***-*qadim*	vs	*kitab*	*qadim*
old.DEF	book.DEF		old	book
'the old book'			'an old book'	

b. Bulgarian

*stara**ta***	*kniga*	vs	*stara*	*kniga*
old.DEF	book		old	book
'the old book'			'an old book'	

c. Hebrew

hasefer	***ha***yašan	vs	*sefer*	*yašan*
old.DEF book.DEF			old	book
'the old book'			'an old book'	

d. Latvian

*vec**ā***	*grāmata*	vs	*veca*	*grāmata*
old.DEF	book		old	book
'the old book'			'an old book'	

e. Lithuanian

*sen**oji***	*knyga*	vs	*sena*	*knyga*
old.DEF	book		old	book
'the old book'			'a/the old book'	

f. Romanian

*veche**a***	*carte*	vs	*o*	*carte*	*veche*
old.DEF	book		an	book	old
‘the old book’			‘an old book’		

Frequently, ADM is combined with other morphological or lexical definiteness markers (determinatives, various pronouns, clitics, etc.), adjectives being just one of the components, as demonstrated in both (1) and (2) below. This typologically rare feature is shared by a number of languages, e.g., Arabic, Albanian, Bulgarian, Hebrew, Romanian, some Slavonic, e.g., Slovenian, Baltic, both Latvian and Lithuanian, and some Germanic languages, e.g., continental Scandinavian languages (Kibort 2008). Some of these languages have elaborate systems for marking both indefiniteness and definiteness, i.e., they have special markers (most often articles) that trigger (in)definiteness marking on all the constituents of an NP, including adjectives, e.g.:

(2) a. Danish

den	*gaml**e***	*bog*	vs	*en*	*gammel*	*bog*
the	old.DEF	book		an	old	book
‘the old book’				‘an old book’		

b. Dutch

het	*oud**e***	*boek*	vs	*een*	*oud*	*boek*
the	old.DEF	book		a	old	book
‘the old book’				‘an old book’		

c. Swedish

den	*gaml**a***	*bok**en***	vs	*en*	*gammal*	*bok*
the	old.DEF	book.DEF		a	old	book
‘the old book’				‘an old book’		

Much has been written on the (in)definiteness marking in the world’s languages with a primary focus on determiners. Yet, it is seldom that focus has stayed on the role adjectives have to play in rendering an NP (in)definite. The most comprehensive description of the ADM as a cross-linguistic occurrence is to be found in the canonical Cambridge textbook *Definiteness*

by Lyons, where a chapter entitled "Definite adjectives" is devoted to it (Lyons 2003, 82-85). Definiteness marking in continental Scandinavian languages has in particular received much attention due to the so-called double determination (or even triple marking, as we see in (2c)) of definiteness in an NP (Perridon 1989, Delsing 1993, Börjars 1994, Julien 2005, Lohndal 2007, Lohrmann 2011). Yet, the focus, e.g., in Swedish, has been on the interplay between the free and the bound definiteness marker (article), which in certain cases can be absent while the adjective always remains marked as definite. Occasionally, the absence of these dedicated markers provides even more important insights into the interplay between the ADM and the definite status of an NP. Such cases prove that multiple definiteness marking in Swedish NPs do not necessarily need to be interpreted as obligatory agreement, as is argued by Lyons (2003, 85).

The two Baltic languages, Latvian and Lithuanian, have only one way to morphologically mark the definite status of an NP, viz., by usage of a paradigm of long adjectival forms[1] (Valeckienė 1957, Ulvydas 1965, Rosinas 1976, Valeckienė 1986, Paulauskienė 1994, Rosinas 1996, Ambrazas et al. 2006, Holvoet, Tamulionienė 2006, Mikulskas 2006, Sprauniene 2008b, 2011). While Latvian definiteness marking is much more grammaticalised, Lithuanian is somewhat more flexible, allowing for grammatically definite contexts to still be able to accommodate short forms, e.g., in possessive NPs:

(3) a. Latvian

mans	*jaun**ais** /*jauns*	*mētelis*
my	new.DEF/new	coat

'my new coat'

b. Lithuanian

mano	*nauj**asis**/naujas*	*paltas*
my	new.DEF/new	coat

'my new coat'

Writing about the world's languages in general, Dahl notes: "NPs with adjectival modifiers tend to have at least as much and often more definiteness

[1] In Germanic linguistics the non-definite adjectival forms are traditionally referred to as strong, while in studies of the Slavic and Baltic languages they are traditionally referred to as short. Correspondingly, the definite adjectival forms are referred to as weak and as long. To sum up, short = strong = non-definite; long = weak = definite. In this dissertation, I will be referring to them as long (LF) *versus* short (SF).

marking than simple NPs. Furthermore, there is a clear tendency for any definiteness marking, excessive or otherwise, to show up next to or on the adjective" (Dahl 2004, 151-152). It is precisely this phenomenon that will be closely examined in the dissertation with a particular emphasis on Lithuanian, placing it in a typological context.

1.2. Aim and tasks of the dissertation

The aim of the dissertation is to analyse and describe the Lithuanian ADM, to place it in a typological context and hence to deepen the understanding of it, and to contribute to cross-linguistic research into this phenomenon. In this consistently synchronic study, I will, through several cross-sections of the ADM in contemporary Lithuanian (introduced in sections 3, 4 and 5), present its impact on the structure of the NP, as well as characterise its actual distribution (the use of long versus short paradigms) in the contemporary language.

The tasks of the dissertation are as follows:

1. to describe the linear structure of the definite Lithuanian NP,
2. to establish the function(s) of internal adjectival constituents in a Lithuanian NP,
3. to analyse the interplay between adjectival and other constituents in a Lithuanian NP and the word-order phenomena in an NP resulting from it,
4. to describe the semantics and functions of adjectival constituents in an NP and their impact on the ADM in Lithuanian;
5. to discuss the usage and distribution of long *versus* short adjective paradigms and to provide tangible quantified data underlying the discussion;
6. to demonstrate the relationship between the ADM and other major and minor cross-linguistic categories.

1.3. Data, research methods and challenges

The dissertation contains more than 200 examples of multilingual data, exemplifying various aspects of the ADM. Most of the data is from the following written contemporary languages: Lithuanian (the dominant data source), Danish, Dutch, English, Russian, and Swedish (and a few random examples from other languages, as e.g., in (1)). Spoken language, even though

it might be mentioned in the periphery of the dissertation, is not of interest here. The Lithuanian data exclusively presents samples of the written language accessed in two corpora, the Corpus of Contemporary Lithuanian (CCLL)[2] with instances from spoken language filtered out; and the Lithuanian Web Corpus (ltTenTen14)[3], which has mostly been used to cross-check the results and findings of the CCLL. Examples of all the other languages were also extracted from their respective corpora (see Abbreviations and Data Sources). The research presented in the dissertation is data-driven, based on empirical evidence, both qualitative (e.g., when defining the linear, topological structure of the definite Lithuanian NPs) and quantitative. The core of the data used for quantitative analysis is a comprehensive list of 111 most frequently used Lithuanian adjectives that can form a paradigm of long forms, based on the *Frequency Dictionary of the Written Lithuanian Language*[4] (see Appendix A), compiled by Utka (2009), henceforth referred to as FrD. The exact methods used in various sections of the dissertation are described in detail under the respective section headings.

The novelty in research methods presented in the dissertation lies in an attempt to quantify the actual usage of certain long and short adjectival paradigms. Even though a relatively small number of adjectives has been investigated in the quantitative analysis (30 qualitative adjectives that are predominantly used in their SFs; and 18 qualitative adjectives that are frequently used in their LFs), it, however, yields quantified results enabling a discussion on the contemporary status of the two adjectival paradigms in modern Lithuanian.

The typological aspect of the dissertation entails certain theoretical limitations and notions that are addressed in subsections 2.1 and 2.2 on the comparative concepts in cross-linguistic studies. Amongst the challenges while assessing and structuring data, the following three need to be highlighted:

1. The difficulties of working with the CCLL due to the lack of annotation; one has to manually extract, sort and treat data in the numbers of tens of thousands due to the unresolved issues of homonymy, lack of diacritical marks

[2] Dabartinės lietuvių kalbos tekstynas (CCLL), accessible online: http://tekstynas.vdu.lt/tekstynas/menu?page=about

[3] The Lithuanian Web Corpus, also called ltTenTen14, accessible online: https://www.sketchengine.eu/lttenten-lithuanian-corpus/

[4] It is available online at http://donelaitis.vdu.lt/publikacijos/Dazninis_zodynas.pdf. Further also referred to as FrD.

and systemic failures causing, e.g., repeat instances of the same sample. The FrD caused a few problems too because of mistaken blending of paradigms of different adjectives under one lexical entry, filing adjectives under verb headings, etc.
2. The absence of modern descriptions of linguistic contexts, e.g., in the labelling of nominal attributes, semantic and/or functional classification of adjectives etc. in Lithuanian grammars.
3. The lack of clarity in descriptions and classifications of certain word classes, e.g., pronouns, in the Lithuanian Grammar, where, e.g., universal quantifiers (definite pronouns) are classified with "Indefinite pronouns".

Much determination was required to resolve the problems arising from the four above-mentioned challenges. Sometimes I felt that gaps in the description of contemporary Lithuanian grammar had to be filled, e.g., by addressing the structure of the definite Lithuanian NP prior to focusing on the ADM *per se*. Yet, I must say that these challenges made me discover important aspects of ADM and contributed in a positive way to shaping my dissertation.

1.4. Novelty of the dissertation

Even though numerous works have been published, mostly focusing on the historical developments of adjectival definiteness markers in the Baltic languages, e.g., two recent comprehensive studies on the historical morphology of definiteness in Baltic have appeared, Gelumbeckaitė (2020) and Sommer (2018), the ADM as a typological phenomenon has been insufficiently analysed. As mentioned, as a cross-linguistic feature it is dealt with in a brief chapter in a canonical textbook on definiteness (Lyons 2003, 82-85). My analysis of the Lithuanian ADM and my attempt at situating it in a typological context through a comparison with mostly Slavonic and Germanic (customarily Scandinavian, occasionally Dutch) languages will contribute to the recognition of ADM as an important linguistic phenomenon in the field of nominal determination alongside determiners, demonstratives, phrasal clitics and other means of marking the category of (in)definiteness, giving it the attention it rightly deserves.

It has to be mentioned that some work has been done on the ADM of Lithuanian NPs from a cross-linguistic comparative perspective. Already in 1957, in her study of the use of contemporary Lithuanian long adjectival forms, Valeckienė provides a short comparative exploration of Lithuanian and German (Valeckienė 1957). In 2004, Mickūnaitė-Griškevičienė in her

doctoral dissertation on the functional sentence perspective in Lithuanian and Norwegian, concisely describes linguistic strategies used to convey the definite status of a Lithuanian NP, with the ADM being one of them (Mickūnaitė-Griškevičienė 2004). In 2006, in the new series of *Studies in Lithuanian Grammar*, a special volume dedicated to the Lithuanian NP was published (Holvoet, Mikulskas 2006) that contains three novel articles entirely or partially concerning the Lithuanian ADM. Holvoet and Tamulionienė present a fresh and modern take on the category of definiteness and its marking in Lithuanian providing some typological insights (Holvoet, Tamulionienė 2006, 11-32); Mikulskas in a symbolically entitled paper *A new perspective on the description of definite adjectives* outlines the novelty in thinking about this category in Lithuanian, basing his approach on the concept of grounding within the framework of cognitive linguistics (Mikulskas 2006, 33–65); and finally, Vaičiulytė-Semėnienė in her study of the category of possession and its morphosyntactic marking briefly discusses the relationship between possessives and the category of definiteness, its implications on the word-order pattern of an NP, founding her paper on cross-linguistic studies and typological hierarchies (Vaičiulytė-Semėnienė 2006, 151–178). In 2011, Spraunienė in her doctoral dissertation focuses on definiteness marking in Lithuanian against the background of Danish and other article languages (English and, to some extent, Hebrew) and presents a novel approach to the opposition of long and short adjectival forms in contemporary Lithuanian (Spraunienė 2011). Subsequently, these findings and a new way of approaching definiteness are reflected in Judžentis' *Grammatical Categories of the Lithuanian Language* where it is deservedly recognised as a category and rewarded with its own chapter (Judžentis 2012, 99–110). In 2015, Gillon and Armoskaite debate the NP/DP divide based on the Lithuanian evidence, focusing mostly on bare nouns, but briefly mentioning pronominal adjectives too (Gillon, Armoskaite 2015). Recently, two relevant papers on subjects concerning the ADM were written by Vaitkutė. The first one analyses strategies of conveying the definite status of Swedish nominals in Lithuanian translations (Vaitkutė 2017); the second one focuses on Lithuanian *toks* 'such' as a member of a cross-linguistic category of similatives (Vaitkutė 2019). Also, an article presenting a novel take on the two adjectival paradigms as cases of strong (LF) *versus* weak (SF) definiteness was written by Šereikaitė (Šereikaitė 2019).

Yet, many important aspects associated with ADM, such as an interplay between the adjectives and other (in)definiteness markers (determiners), the contexts disabling the uses of LF, the contexts disabling the uses of SF, the

impact of word order on the referential value of an NP, the relationship between the semantics of adjectives and their ability to serve as (in)definiteness markers, the relationship between ADM and other minor categories such as *similatives* (van der Auwera, Sahoo 2018), and many more need to be explored and described in a coherent and comprehensive way, highlighting, where possible, the cross-linguistic characteristics of ADM.

Despite studies on the Lithuanian language (and on other languages that exhibit the feature of ADM), ADM it is not acknowledged as a full-fledged cross-linguistic category, which is reflected in its absence from the *World Atlas of Language Structures Online*[5] (Dryer, Haspelmath 2013). I hope that one of the side-effects of this dissertation will be an increased global interest in the ADM as a cross-linguistic phenomenon circling in the orbit of (in)definiteness as a category related to many other major and minor cross-linguistic categories.

1.5. Structure of the dissertation

This dissertation is structured as follows. In section 1 the object of the study, the background, data and method, the aims and tasks, and the key notions used in the dissertation are introduced. Also, other aspects, such as the novelty and the added value, and the current *status quaestionis* are briefly presented. Section 2 discusses a few theoretical prerequisites, such as descriptive categories *versus* comparative concepts in cross-linguistic studies, and the possibility of a data-driven framework-free approach; it includes a discussion of whether ADM should be seen as an areal typological feature of Standard Average European (SAE) and of the Circum-Baltic area. Also, key concepts and definitions are introduced here along with a more detailed account of the Lithuanian definiteness marking system and the role of ADM in it. One of the subsections is devoted to adjectives as a universal comparative part-of-speech. The main data source for this dissertation, the list of the most frequently used Lithuanian adjectives that form both LF and SF paradigms is looked at from a typological perspective of semantic types of adjectives attested in the world's languages. Section 3 presents a topological structure of the definite Lithuanian NP and highlights the role ADM plays in it, including interaction between several adjectival constituents in an NP. In section 4, the instances of the absence of LF in Lithuanian NPs are discussed on the basis of empirical data; the relationship between the semantics of adjectives and determiners is analysed. Section 5, which is also empirically driven, deals with those

[5] The World Atlas of Language Structures Online (2013): https://wals.info/.

Lithuanian adjectives that are predominantly or frequently used in their LFs. Finally, section 6 contains some concluding remarks and outlines prospects for future research in the field.

1.6. Theses to be defended

The ADM is represented in Lithuanian by a non-trivial cluster of distinct features:

1. ADM directly correlates with the syntactic structure of the NP whereby the feature [+DEF] may and often is encoded in several loci with different degrees of impact. ADM is highly sensitive to the syntactic structure of the NP as a whole.
2. The different loci of [+DEF] correlate with the relative scopes of the categories established by the NP.
3. In the NP, constituents in the left periphery encode discourse-bound definiteness while constituents in the right (closer to the head) denote identifiability-based definiteness.
4. ADM correlates with adjectival classes and their inherent abilities to establish *ad hoc* or more permanent categories (generic nominals).
5. There exists a group of adjectives that are predominantly used in their SF; this use is because of their inability to establish a category (either *ad hoc* or taxonomic) a) due to semantic-pragmatic reasons; b) because rather than denoting a property, they denote something else, i.e., sentence-level modification, quantification, similarity, etc.
6. There exists a group of adjectives that are predominantly used in their LF; they form a small group whose definiteness mostly reflects socially and/or culturally determined identifiability.

2. SOME THEORETICAL PREREQUISITES

As the title of the dissertation implies, this is a study with aspirations to typological research. Therefore, some aspects associated with typological linguistics need to be addressed prior to proceeding to the research question. Lithuanian ADM will be placed in a typological context, requiring firstly a definition of what it entails as a language-particular category, and secondly a broader agreement, a cross-linguistic definition of what ADM is and how it can be described, assessed, compared and generalised over across several languages and language groups. Also, ADM in neighbouring languages, i.e., the Baltic, some Slavonic, German and continental Scandinavian, might be an areal linguistic phenomenon, which also poses several theoretical implications.

2.1. Comparative concepts in cross-linguistic studies

Several linguists, especially typologists, have discussed the challenges of cross-linguistic comparison. Their thinking and argumentation for and against is well summarised in an article published by Haspelmath in 2010 putting forward some key questions concerning cross-linguistic work (Haspelmath 2010). He argues that descriptive formal categories cannot be compared across languages because they are language-particular, assigned to a particular language based on a particular set of criteria that might differ from language to language. He claims that linguists need to create comparative concepts that enable them to identify relevantly equivalent phenomena in languages rather than comparing language-particular categories. It is against these comparative concepts that "descriptive categories of particular languages can be matched" (ibid., 681). In 2018, he wrote another article explaining how comparative concepts and descriptive linguistic categories are different (Haspelmath 2018). Later, he revisited the subject (Haspelmath 2020), reformulating the question as to where the difference between general linguistics (also known as theoretical linguistics) and particular linguistics (research on individual languages) lies. Trying to resolve what Haspelmath calls *the general linguistics paradox*[6], he claims that: 1) "language description is true to categories of each language, but is inspired by the accumulated knowledge of comparative linguistics" (Haspelmath 2020, 14); 2) "comparison of

[6] **The general linguistics paradox**: We want to explore and understand the nature of Human Language, but what we can observe directly is particular languages. (Haspelmath, 2020, 7).

languages is not expected to be based on the categories of the particular languages, because there is no uniform set of building blocks of which all grammatical systems are composed" (ibid.). Therefore, linguists must focus on and use what he calls *independent yardsticks for comparative measurement*. Furthermore, he claims that grammatical descriptions contribute to general linguistics (studies of Human Language in the broadest sense) by providing crucial data for large-scale cross linguistic comparisons, potentially leading to discovery of cross-linguistic universals (ibid.) The latter being an aspiration for this work, I hope that this dissertation will fill in a gap in larger-scale typological research on ADM as a cross-linguistic category.

Reflecting on ADM in a typological context, I would like to remark that knowledge of other languages, Scandinavian and Slavic in particular, has helped me greatly in advancing my thinking on the Lithuanian ADM, i.e., manifold and detailed studies of first and furthermost Scandinavian definiteness marking systems (but also, studies of the long and short adjectival paradigms in Slavic languages) made me look at Lithuanian data through the glasses of previously acquired knowledge of the linear placements of prenominal attributes in a definite NP, and helped me discover and better describe the topology of Lithuanian prenominal attributes in a tradition that has been alien to the Baltic descriptive grammars.

2.2. ADM as a Standard Average European /Circum-Baltic feature

Besides the general *versus* language-particular linguistic analysis, the concept of generalising on the basis of common linguistic areas and language-contact induced similarities, viz. areal linguistics, needs to be addressed here. The existence of Standard Average European as a linguistic area (or *Sprachbund*) with specific features (so-called *Europeanisms*) seems to be a well-established fact (Haspelmath 2001b). A linguistic feature is viewed as a *Europeanism* if it meets certain requirements, as outlined by Haspelmath (ibid., 1493), whereby most European languages exhibit this feature; the geographic neighbours lack this feature; the eastern Indo-European languages lack it too (Armenian, Iranian, Indic), and it is not a feature that is shared by a majority of the world's languages. Consequently, even though Haspelmath does not include the ADM amongst the 12 SAE features, all of which are syntactic or "concern the existence of certain morphosyntactic categories" (ibid.), it could be a strong candidate, alongside another adjectival feature, i.e., comparative marking of adjectives. Moreover, it should be considered as a strong candidate for the inventory of the Circum-Baltic (CB) features (Koptjevskaja-Tamm, Wälchli 2001), as it seems to be shared by all

the languages except for Finnish in the Baltic area today[7]. It needs to be mentioned that opposed to the SAE, which is considered a prototypical *Sprachbund*, the Circum-Baltic area is *a contact superposition zone* between SAE and South Eurasia (Koptjevskaja-Tamm, Wälchli 2001, 624). In his article on the Circum-Baltic linguistic area, Wälchli, discussing the acquiring of definite articles by some CB languages, refers briefly to "an ancient parallel" of development of the paradigm of long adjectival forms in Germanic and Slavic/Baltic, which has different origins and different ranges of use, and yet could potentially be seen as a result of an areal linguistic interference (Wälchli 2011, 331).

2.3. Adjectives as a subject for typological research

Since the subject of this dissertation is ADM, a few words must be said about adjectives as a part-of-speech concept and its place in typological research. In 1982, Dixon published a study entitled *Where Have All the Adjectives Gone?*, a first attempt at a comprehensive typological analysis of adjectives in world's languages based on their *universal semantic types* (Dixon 1982). In this study he claimed that adjectives are not a universal part-of-speech (Dixon 1982, 2): some languages lack it all together, in others it is a small and non-productive one. In 2004, he revisited the topic and changed his claim, saying that adjectives should be acknowledged as a distinct word class in every human language. Moreover, he identifies 4 core semantic types typically associated with adjective classes regardless of their size in a given language: 1) dimension; 2) age; 3) value, and 4) colour. In addition to these, 3 peripheral semantic types, typical for medium and large adjective classes (which is the case of Lithuanian) can be named, viz., 5) physical property; 6) human propensity, and 7) speed (Dixon 2004, 3-4). In his study of 1982, Dixon looked at languages that had small adjective classes (he did a survey of 17 languages with small adjective classes[8]) and noticed that they seemed to exhibit similarities suggestive of "some type of syntactico-semantic

[7] German has 3 paradigms of adjectival forms, referred to as 'strong', 'mixed' and 'weak'. The choice of the paradigm is decided by the type of element in the determiner position. Whether this is a case of 'pure' manifestation of the category of definiteness or simply a form of agreement can be debated (Kibort 2008). Zwicky argues that German ADM is a case of government (Zwicky 1986). Polish has a finite list of SF, mostly used as constituents of a predicate and analysed as fossilised lexical entries. The situation in Slovak, Ukrainian and Belarussian is similar (Hansen 2004, 70).

[8] Igbo – 8 adjectives, organised in antonym pairs representing the 4 core semantic types, viz. large – small, new – old, good – bad, black/dark – white/light; Hausa – 12 adjectives; Venda, Malak Malak, Southern Paiute, etc., just to mention a few.

universals" (Dixon 1982, 7), which was further explored and resulted in the cross-linguistic typology of adjective classes of 2004 (Dixon, Aikhenvald 2004). In addition to the semantics of the adjective class, an important notion explored was overlapping between the classes of nouns, verbs and adjectives (see also Haspelmath 2001b). Most of the European languages that have large adjectival classes are characterised by an overlap amongst the three classes, implying that there are noun-like, verb-like and adjective-like adjectives. Also, adjectives in predicates and the modifier function of an NP were examined.

I would like to make two remarks here with regard to the data of Lithuanian adjectives analysed in this dissertation in the light of what was said above. Firstly, the list of 111 Lithuanian adjectives extracted from the FrD exhibited a strong correspondence with the 7 identified semantic types[9] with members appearing in antonym pairs, e.g., *didelis – mažas* 'big – small', *ilgas – trumpas* 'long – short', *aukštas – žemas* 'high – low', *naujas – senas* 'old – new', *geras – blogas* 'good – bad', *juodas – baltas* 'black – white', *raudonas* 'red', *karštas – šiltas – šaltas* 'hot – warm – cold', *stiprus – silpnas* 'strong – weak', *didis* 'great', *garsus* 'famous', etc., except for type (7) describing speed (see Appendix A for the comprehensive list). This shows that Lithuanian behaves typically with regard to the typological expectations.

Secondly, the overlap between the three word classes (noun, verb, and adjective) is of importance for explaining the distribution of LF and SF paradigms in predicative function (for detailed discussion see Section 4.3.2 on Attribution and predication). Also, as demonstrated by Corbett in the study on Russian adjectives (Corbett 2004, 209), adjectives that take a second argument exhibit a specific syntactic pattern (this will be discussed in Section 4.4.3 on verb-like adjectives), which is also attested in Lithuanian.

Concluding, I would say that a comprehensive typological study of Lithuanian adjectives on a par with the ones mentioned above would add to our understanding of this rich and interesting part-of-speech attested in Lithuanian.

2.4. Key concepts and definitions

The study of the ADM in this dissertation will not be presented from the perspective of a particular linguistic school or theoretical approach. This research is language-particular, observation-based and data-driven with an

[9] It would be interesting to perform a thorough semantic analysis of all the 111 adjectives. However, this does not fall within the scope of this study.

ambition to contribute to the general understanding of the ADM as a cross-linguistic comparative concept. It will become obvious that certain linguistic approaches will be used to describe certain linguistic manifestations, e.g., in section 3, the concept of grounding from Cognitive Linguistics will be invoked to illustrate the reference assignment process; typological *Prepositional Noun Modifier* and *Animacy* hierarchies will be invoked to describe the topology of Lithuanian prenominal modifiers, etc., simply because they seem to best capture the essence of their occurrence, and offer the most clear and understandable way to describe it.

Below are a few key concepts and their definitions as used in this dissertation:

I feel that since the subject of this dissertation concerns definiteness, my stance on **the NP/DP hypothesis** should be clarified. The NP/DP (determiner phrase) controversy opposes different views as to which element of the NP should be treated as the head (sometimes also called *nucleus*) of a nominal structure. In the grammatical theories embracing the DP analysis, the DP layer is stacked on top of the NP, with determiner D being the head of the DP. In more traditional grammars, it is nouns (this term is not self-explanatory and might be treated differently in different grammars)[10] that constitute the core of a noun phrase, which can be definite without acknowledging the existence of the DP layer. In this thesis, I will refer to the noun or a noun-like element (see below) as the head of an NP and will speak exclusively of NPs. Also, since the main focus of this study is ADM, bare noun phrases (NPs without modifiers) and their definiteness status will not be analysed.

In this dissertation, the term **noun phrase (NP)** refers to a complete discourse-oriented structure containing a common noun or proper noun (or another word with a noun-like function[11]) with or without preposed (prenominal) or postposed (postnominal) attributes, e.g., *žaidimas* 'a game', *įdomus žaidimas* 'an interesting game', *populiarusis žaidimas* 'the popular game', *pirmasis jos stalo žaidimas* 'her first table game', *žaidimas apie elektros kilmę* 'a game about the origins of electricity', *keletas jos bičiulių pamėgtų žaidimų* 'a few games liked by her friends', etc.

[10] CGEL, e.g., considers pronouns to be a subcategory of nouns rather than a word class of their own (CGEL, 2002, 327) due to their functions that are very similar to those of common and proper nouns.

[11] Adjectival or ordinal nominalisations, pronouns and participles, e.g.: *linksmasis* ***šviesiaplaukis*** *iš penkto aukšto* 'the cheerful **blonde** from the fifth floor'; ***tie*** *naujai* ***atsikraustę*** *į šį namą* '**those** (who have) recently **moved** into this house'.

Even though the pragmatic category of definiteness is universal, as a reference coordination method, it does not necessarily have specific grammatical markers in the world's languages. In fact, only a minority of world's languages have dedicated (in)definiteness markers in their grammatical repertoire[12]. Most often it is other elements, e.g., demonstratives, possessives, etc., that serve this function instead. Therefore, the grammars of different languages have different traditions when it comes to recognising the grammatical category of **determiners** (also referred to as **determinatives** when speaking of them as a word-class)[13]. In Lithuanian linguistics, there is no tradition of speaking about determiners and determinatives. In Scandinavian linguistics, it is only Norwegian grammars that consistently use the term *determinativer* 'determinatives' to single out this category as a word-class. In SAG, the term *definite attributes* is consistently used to refer to elements that give an NP a definite reading (see section 3.3.2 for a detailed overview). Since in Lithuanian, we only have one dedicated morphological marker to denote the [+DEF] status of an NP, which is located in an adjective, in some instances the determiner and modifier function (traditionally performed by an adjective) coincide in the same NP constituent. Yet, not all adjectives used in their LF get a definite reading, and *vice versa*, not all adjectives used in the short form get an indefinite reading, as will be shown.

The grammatical category of definiteness, and in particular **ADM**, which is one of its types, is a morphosyntactic feature (whether its *modus operandi* is agreement or government will not be discussed here), encoding the semantic/pragmatic concept of identifiability (based on familiarity, uniqueness, inclusiveness or taxonomy). If a language has grammaticalised the feature of definiteness (has at least one dedicated (in)definiteness marker), it can be said to have a grammatical category of definiteness. As Lyons says: "[…] there is a distinction to be made between grammatical definiteness and semantic/pragmatic definiteness, and on the concept of grammaticalization. The proposal is that definiteness *stricto sensu* is not a semantic or pragmatic notion as assumed by almost all writers on the subject, but rather a grammatical category on a par with tense, mood, number, gender etc." (Lyons 2003, 274–275). According to him, grammaticalisation of this category is key – it needs to have a representation by a grammatical form or

[12] For more details, see wals.info, chapters on definite and indefinite articles: https://wals.info/chapter/37 and https://wals.info/chapter/38 respectively.

[13] In CGEL, determiner is a function in the structure of the NP, while determinatives are a category of words (and larger constructions) (CGEL 2002, 355).

forms, which, in case of ADM is the two sets of long and short adjectival forms.

As will be explained in detail in 2.5, short adjectival forms in Lithuanian are neutral with regard to definiteness marking, which means that they can be interpreted as getting both a definite and an indefinite reading depending on the NP and its context. I will therefore consistently mark them as AjdNONDEF in this dissertation. However, in certain syntactic constructions, e.g., in sentences with rhematic NPs, only short forms can be used, and these get an exclusively indefinite reading, as they are introducing new referents. In these instances, to mark their **indefinite status**, I will use the mark of AdjINDEF to signal just this, as opposed to AdjNONDEF. The reader should expect very few of these instances.

Finally, the term ***ad hoc* category**, borrowed from pragmatics, will be instrumental in this dissertation. An *ad hoc* category is a pragmatic category with an overt linguistic encoding constructed instantaneously to achieve communication goals. It is not available as a structure in long-term memory; it is highly context-dependent and based on an exemplar. Yet, the category itself is more relevant in discourse than the mentioned exemplar. The *ad hoc* categories do not appear with ready-made linguistic labels, but rather by means of complex expressions, e.g., *things to do on a rainy Sunday afternoon* (Mauri 2014). It correlates with generic definiteness in that, like a generic/taxonomic reference that refers to the kind rather than individual objects, an *ad hoc* category refers to the category rather than an individual object(s).

There will, most certainly be other terms and concepts used in this dissertation, but they will be explained in the running text where relevant.

2.5. Definiteness marking in Lithuanian: a summary of findings up to date

Lithuanian has one overt morphological indicator coding definiteness, namely the set of suffixal adjectival endings added to the short forms of adjectives (agreeing with these in case, number and gender) to form the so-called ‘long’ or ‘definite’ adjectival forms:

(4) *baltas* *laukas* vs
white.NOM.SG.M.NONDEF field.NOM.SG.[M]
‘a white field’

*balt**asis***	*laukas*
white.NOM.SG.M.DEF	field.NOM.SG.[M]
'the white field'	

It has to be noted that in contemporary Lithuanian, we witness the fusion of the former pronouns *jis/ji* with adjectival endings. In some paradigmatic forms one can still see and possibly analyse them as pronouns added to the basic form of the adjective, e.g., the adjective 'white' as in the example above, *baltų* 'white.GEN.PL.NONDEF' vs *baltųjų* white.GEN.PL.DEF' where the plural form could still be disassembled as *baltų* + *jų* with *jų* being GEN.PL of pronoun *jis/ji*; but in the majority of other paradigmatic forms this is no longer possible. Today, the process of synthesis between the former pronouns and the adjectival endings is rather advanced. Therefore, we often refer to Lithuanian adjectives as having two sets of endings – definite and non-definite, rather than saying that a clitic or an affix is added to the inflected forms of an adjective; the definite ones are used to mark the definiteness of the NP (Holvoet, Sprauniené 2012, 72). Hence the consistent marking of the definite adjectival endings in bold, e.g., *balt**ųjų***, and not *balt**ųjų*** in this dissertation.

Another important remark is that not all adjectives have a paradigm of definite forms. It is only **qualitative adjectives**, as well as ordinal numbers, participles and some pronouns that can assume definite markers (Valeckienė 1957, 257-299, 299-301; Paulauskienė 1994, 220; Ambrazas et al. 2006, 185-187, 245, 260, 367-369, Sprauniené 2008b, 117). See Section 4.2 for a detailed discussion on the different types of Lithuanian adjectives in relation to their ability to take long adjectival forms. A few words yet need to be added to say that some sub-categories of qualitative adjectives do not take long forms either. Multisyllabic compound adjectives, e.g., *kilniaširdis* 'noble-hearted', *daugiapakopis* 'multilevel', *aštriabriaunis* 'sharp-edged', do not take long forms. Similarly to relational adjectives (see 4.2.1 for discussion on relation versus property), these adjectives function as productive replacements for non-determiner genitives[14], e.g., *kilniaširdis* → *kilnios*.GEN.SG.F *širdies*.GEN.SG.[F] 'of noble heart'; but rather than denoting a relation like relational adjectives, these still denote a property. Another group of qualitative adjectives that do not take long forms are adjectives formed with suffixes or prefixes indicating the degree of a quality, e.g., *didokas* 'rather big' from *didis/didelis*, *mažiukas* 'tiny' from *mažas.* Also belonging here are the so-called diminutive adjectives, e.g., *gražutis* 'the cute one', *greitukas* 'the fast one', which in Southern dialects

[14] See footnote 20 on non-determiner genitives.

replace the long adjectival forms (Ambrazas et al. 2006, 146). That these subgroups of adjectives do not take LFs can be explained by the fact that long forms of Lithuanian adjectives establish a category (either *ad hoc* or taxonomic), which is incompatible with degree modification. In case of the diminutive adjectives, they themselves are interpreted as definite and hence require no additional definiteness markers (for more on the role of adjectives in establishing categories, see 4.2.2).

In Lithuanian grammars[15] the tradition has been to attribute the feature of definiteness to adjectives and not to the domain of the NP by stating explicitly that the notion of definiteness arises in the opposition of AdjINDEF and AdjDEF. However, in more recent publications it has become clear that in Lithuanian the scope of the category of definiteness and the locus of its marking do not coincide (Holvoet, Tamulionienė 2006, 13, 22). While morphologically the marker is located in an adjective, it has scope over a whole NP. Moreover, the opposition of short and long forms does not directly map onto a noun phrase as [-DEF NP] versus [+DEF NP], but as Spraunienė (Spraunienė 2008b, 119; 2011, 73) points out, the AdjDEF marks the definiteness of an NP [+DEF NP] while the AdjNONDEF is neutral with regard to definiteness and can be described as [±DEF NP]. Moreover, even NPs containing long adjectival forms may not get their definite reading because of the presence of a definite adjectival form but rather because of some other factors, as illustrated in (5), explicitly in the reading (d):

(5) *Jis* *įsimylėjo* *gražiąją*
he fall-in-love.3PST beautiful.ACC.SG.F.DEF
karaliaus *dukterį.*
king.GEN.SG.[M] daughter.ACC.SG.[F]
'He fell in love with the beautiful daughter of the king.'[16]

This example might have 4 different interpretations with regards to the AdjDEF:
a) definite, if the context that the king had several daughters and only one of them was beautiful is provided;

[15] It was in the English version of the Lithuanian Grammar of 2006 that the relation between the category of definiteness and the noun rather than the adjective itself was first noted: "The category of definiteness in the adjective is based on the opposition of definite adjectival forms, which in addition to their lexical meaning of quality contribute definite status to the noun they determine, and simple, or indefinite, adjectival forms, which lack the meaning of definiteness." (Ambrazas et al. 2006,142).

[16] This example is cited from Holvoet, Tamulionienė (2006, 26).

b) anaphoric definite, if it was mentioned in the preceding context that a king had a daughter and that she was beautiful;
c) traditional grammars would argue that this example could also be the case of an "emphatic" use of the long adjectival forms, manifesting itself in folklore and fiction as standard traditional epithets, e.g., *senie**ji** tėveliai* 'the old.NOM.PL.M.DEF parents'. The more recent papers on the subject (Mikulskas 2006, 54-55, footnote 16; Spraunienė 2011, 111-112) agree that these instances should rather be treated as cases of generic reference or role reference. The adjectives here denote a property assumed to be inherent to the nominal referent, e.g., the parents are always old. In line with this, one could say that princesses are always beautiful.
d) non-contributing to the definite interpretation, as the referent is identifiable on its own, but rather adding to the meaning of the NP by disabling an interpretation whereby if a short form of the adjective was used, the description *king's daughter* would have an indefinite meaning *king's daughter = princess*[17], thus the presence of a long adjectival form here is not redundant, according to Holvoet and Tamulionienė. Moreover, it could be argued that it is in a sort of definiteness agreement with a defined referent. In Lithuanian, this agreement is not obligatory, but in a corresponding Latvian example, it would be the case:

(6) *Viņš* *iemīlēja* *skaist**o***
he.NOM.SG.M fall-in-love.3PST beautiful.ACC.SG.F.DEF
ķēniņa *meitu*.
king.GEN.SG.[M] daughter.ACC.SG.[F]
'He fell in love with the beautiful daughter of the king.'[18]

To sum up, in Lithuanian, as opposed to Swedish, there is no automatic agreement in definiteness (the NP may be definite without the adjective getting the definite ending). Moreover, often the adjective gets a definite ending because it makes a specific contribution to definiteness. Finally, while there is no general agreement in definiteness, there are certain types of definiteness which require the adjective to be definite even though it is not the only contributor to definiteness, e.g., the usage of adjectives with proper names, e.g., *gars**usis** Stephenas Hawkingas* 'the famous.DEF Stephen Hawking'.

[17] Other similar examples of the kind would be *maldos namai* 'house of prayer' = a church (or a mosque, or a shrine) or *žodžio žmogus* 'a man of his word = a man of integrity, a man who keeps his promises'.

[18] This is an exact translation of the example (5) into Latvian, verified by mother-tongue speakers.

Sprauniené argues that in contemporary Lithuanian the long adjectival forms always encode definiteness (Sprauniené 2011, 74-76), both on the level of individual and generic reference. As an argument for this interpretation, she claims that definite adjectival forms cannot be used with indefinite modifiers; neither can they be used when a noun phrase is rhematic (comment) rather than thematic (topic). The definite forms in these contexts are available only in cases where they represent NPs of generic reference, e.g., *balt**asis**/juod**asis** gandras* (the white.NOM.SG.M.DEF/black.NOM.SG.M.DEF stork):

7)	*Jam*	*paskambino*	*vienas*	
	he.DAT.SG.M.	telephone.3PST	one.NOM.SG.M	
	*[*gars**usis**]*		*garsus*	*žurnalistas*
	[famous.NOM.SG.M.DEF]		famous.NOM.SG.M.INDEF	journalist.NOM.SG.[M]

'He was telephoned by a famous journalist.' [19]

(8)	*Ant*	*palangės*		*tupėjo*	*[*balt**oji**]*
	on	windowsill.GEN.SG.[F]		sit.3PST	[white.NOM.SG.F.DEF]
	balta		*katė.*		
	white.NOM.SG.F.INDEF		cat.NOM.SG.[F]		

'On the windowsill there sat a white cat.'

A Lithuanian NP can be marked for definiteness in the following ways:

1. In an NP with an adjectival modifier, it is marked by the presence of the long rather than short adjectival form (LF).
2. Otherwise, it is conveyed through the use of definite attributes with or without an attributive adjective, incl. demonstratives, possessives and determiner-genitives[20], as well as universal quantifiers, including fractions.
3. Sometimes, the so-called definiteness effects (Lyons 2003, 227-251) come into play, e.g., mass nouns and plurals as objects of perfective verbs are interpreted as definite (Holvoet, Tamulionienė 2006, 30-32); certain word order patterns in which the thematic (topicalised) NP gets a definite reading due to the functional sentence perspective. Also, according to Lyons, property predication and superlatives, as well as a number of other syntactical constructions, are to be treated as definiteness effects.

[19] This example, as well as (8), is cited from Sprauniené (2011, 74).

[20] In Lithuanian, there are two types of genitive constructions, viz. determiner-genitives and non-determiner-genitives. A detailed account of possessives, determiner-genitives and non-determiner genitives is offered in Trakymaitė (2018, 117-122). See also Koptjevskaja-Tamm (2003).

The above can be summarised in the following table:

Table 1. Definiteness marking in Lithuanian

Prenominal attributes	**NP$_{DEF}$**	
Adj	**Adj$_{NONDEF}$ + N**	**Adj$_{DEF}$ +N**
	balta katė 'the white cat'	*balt**oji** katė* 'the white cat'
Other definite attributes[21]	**Det (+Adj$_{NONDEF}$)+ N**	**Det (+Adj$_{DEF}$) +N**
	ta balta katė 'the/that white cat'	*ta balt**oji** katė* 'the/that white cat'

[21] The types of definite attributes are definite adjectival forms, demonstratives, possessives and determiner-genitives, and universal quantifiers (including fractions).

3. DETERMINATION AND MODIFICATION: TOPOLOGY OF PRENOMINAL ATTRIBUTES IN LITHUANIAN*

3.1. Introduction

3.1.1. Background and aim

The section studies the topology of the definite Lithuanian noun phrase (NP). Hence the analysis of the structure of the definite Lithuanian NP will be seen through the interplay of morphology, syntax, and (pragmatic) definiteness. It has to be noted that limited research is available on the structure and syntax of the Lithuanian NP, thus this section will have a twofold aim, namely, to describe and propose a rationale for the linear structure of the definite Lithuanian NP; and do so with a particular emphasis on the morphological and syntactic encoding of definiteness, which, as we will argue, affects the line-up of prenominal attributes. The marking and expression of indefiniteness will not be explicitly analysed here.

3.1.2. Data and method

Lithuanian data will be compared to its Scandinavian analogues, mainly Swedish (Swed.), which exhibits elaborate and fully grammaticalised definiteness marking. Moreover, a feature shared by the Scandinavian languages and Lithuanian (Lith.) is the adjectival encoding of definiteness that will figure prominently in this thesis. Additionally, Swedish, similarly to Lithuanian, displays several loci of definiteness in an NP, which enables us to make an instructive comparison of the topology of the NP. According to Mikulskas (2006, 33), when researching the grammatical challenges surrounding definiteness in any language, especially in a language with weakly grammaticalised marking of definiteness, viz. lacking formal definiteness markers like articles, it is useful to compare the data to that of languages that represent a well-developed model of definiteness encoding, namely where marking of (in)definiteness is fully grammaticalised and obligatory. Therefore the comparison of the two opposite poles, Swedish on the one hand and Lithuanian with its minimal formal definiteness marking on the other, would enable a more detailed and refined analysis of the Lithuanian definite NP structure, as cross-linguistic typological research has a reversed perspective too: not only does it enable generalisations, but also working in the opposite direction, it informs the description of linguistic phenomenon in a particular language (Croft 2004, 9; Haspelmath 2010, 663-687). Also,

*This section, with minor updates and modifications, is based on this article: Trakymaitė R. (2018). "Determination and modification: Topology of prenominal attributes in Lithuanian", Kalbotyra, 710 (2018), 84–133.

for the purpose of a detailed and theory-neutral distributional grammar description the English data (authentic examples) and the *Cambridge Grammar of the English Language* (CGEL[22]) will be extensively used. For the Scandinavian languages, mostly the concepts and terms used in the *Swedish Academy Grammar* (SAG[23]) will be employed; a few references will be made to the *Grammar of the Danish Language* (GDS[24]) and the two grammars of the Norwegian language, viz. *Norsk referansegrammatik* (NRG[25]) and *Norsk som fremmedspråk: Grammatikk* (NFS[26]).

All the examples presented in this thesis, unless a specific source is quoted, are collected from the Danish, English, Lithuanian, and Swedish corpora of written language. No quantitative analysis was carried out for the purpose of this section. Instead, certain syntactic structures, e.g., *demonstrative + noun*, or *demonstrative + adjective+ noun* or *quantifier + demonstrative+ noun* just to mention a few, were searched in databases of the languages mentioned to reflect the unique syntactic usages in respective languages. The point of departure was the Swedish NP structure with clearly identified positions for various functions on a phrasal level (SAG, Vol 3, 13), e.g., determiner (Det) (in SAG *definita attribut*), demonstrative (Dem), quantifier (Q), adjectival modifier (Adj). A few Lithuanian examples (based on the similar structures attested for by the corpus) were constructed for illustrative purposes and reflect the author's native competence. Some were found online (not in corpora). Both of the latter are indicated in respective footnotes.

3.1.3. Key concepts and terms

Since the research object of the section is the phrase, the focus will be on the syntactic functions within the phrase, namely those of head and dependent(s). The head is a noun (or another word with noun-like function), e.g., *knyga* 'a book'. It can have dependents that are syntactically subordinate elements (CGEL, 24), viz. complements, e.g., *knyga* ***apie Kauno apskrities žmones*** 'a

[22] CGEL = Huddleston, Rodney & Geoffrey K. Pullum, eds. 2002. *The Cambridge Grammar of the English Language*. Cambridge: Cambridge University Press.

[23] SAG = SAG = Teleman, Ulf, Staffan Helleberg & Erik Andersson, eds. 1999. *Svenska Akademiens grammatik*. Stockholm: Svenska akademien, Nordstedts ordbok.

24 GDS = Hansen, Erik & Lars Heltoft. 2010. *Grammatik over det Danske Sprog*. Århus: Syddansk Universitetsforlag.

[25] NRG= Faarlund, Jan, Svein Lie & Kjell Ivar Vannebo, eds. 1997. *Norsk referansegrammatikk.* Oslo: Universitetsforlaget.

[26] NFS= Anne Golden, Kirsti Mac Donald & Else Ryen, eds. 2008. *Norsk som fremmedspråk: Grammatik.*(3. utgave). Oslo: Universitetsforlaget.

book about people from Kaunas County'; modifiers, e.g., ***garsi*** *knyga* 'a famous book'; and determiners, e.g., ***ta*** *knyga* 'the/that book'. Only the grammatical elements preceding the noun, called prenominal attributes, will be analysed. Postnominal elements and structures modifying the noun[27], e.g., relative clauses, prepositional phrases (PP), appositions, etc. will not be analysed here. To sum up, mainly modifiers and determiners will be examined. Most complements, such as postnominal PPs, e.g. *atsakymai į mano klausimus* 'the answers to my questions', are excluded because of their position; in the case of prenominal genitives, e.g. *Respublikos prezidentė* 'the President of the Republic', it is often difficult to establish whether they are complements or modifiers, and while we will of course deal with such genitives in this section, the important question for our purpose will be whether they perform a determiner function as well or not. Henceforth I will oppose only modifiers and determiners, distinguished on the basis of their having descriptive content or not.

In this chapter, I will describe an NP with prenominal attributes as a sequence of elements lining up from left to right with the head being the rightmost element; that is, I will examine the linear structure of the NP. It is evident that prenominal attributes form inner structures, phrases of their own,

e.g., *visi trys* ***mano vaikystės draugės*** *vaikai*[28] 'all three children of my childhood friend', where *my childhood friend* is an NP in its own right. As Perridon notes, each different field[29] in a Swedish NP "may contain syntagms of an internal structure of their own" (Perridon 1989, 201). The same can be said about Lithuanian prenominal attributes. However, these imbedded relations (hierarchical) between the inner phrases will not be analysed here; all the elements will be treated as co-existing on the same linear level.

Two questions should be posed now: 1) What lexical classes perform as determiners, and what as modifiers? 2) What is the difference between determiners and modifiers; does it affect the placement of dependents within a positional linear structure of an NP? The answers to these questions are closely linked with the concepts of definiteness and referentiality. Both determiners and modifiers are crucial to reference identification (narrowing

[27] In the function-oriented grammar of the Lithuanian language, these are called *kvalifikatoriai* 'qualifiers' as opposed to the preposed *modifikatoriai* 'modifiers' (Valeckienė 1998, 118–130).

[28] There is an ambiguity in this example, the sub-phrase *my childhood friend* could be interpreted as *[[my childhood] friend]* → 'a friend from my childhood times' or as *[my [childhood friend]]* → 'my childhood friend' the intended reading being the latter.

[29] An overview of these fields as defined by Perridon is given in 3.7.1.

down the referential mass). However, it is only determiners that serve in assigning definiteness value to an NP, viz. "the determiner serves to mark the NP as definite or indefinite" (CGEL, 54). The elements that can trigger a definite reading of an NP will be called definiteness carriers to differentiate them from a lexical class of determinatives that do not necessarily mark the NP as definite, e.g., cardinal numbers do not assign definiteness, yet they do contribute to determination and in CGEL and some Scandinavian (Norwegian) grammars would be classified as determinatives.

This is of particular interest when examining adjectival modifiers, as definite (also referred to as long and/or weak) adjectival forms are the sole carriers of grammaticalised definiteness in Lithuanian; hence they could be treated as determiners. Should certain prenominal attributes in Lithuanian be treated as determinatives (a lexical class)? What is the syntactic function of genitives? What is the interaction between definite and indefinite adjectival forms in the same NP?

3.1.4. Structure of section 3

The section is structured in the following way: first, the definiteness-marking systems in Swedish and Lithuanian (in detail presented in 2.5) will be briefly introduced, focusing on the adjectival modifiers within an NP; then the distinction between determiner and modifier function will be discussed; this will be followed by detailed descriptions of various lexical classes of attributes functioning as either determiners or modifiers within an NP. A special section will be devoted to Lithuanian NPs with adjectival attributes – both long and short, and combining both types in the same NP. Finally, a topology of the definite Lithuanian NP will be proposed.

3.2. Definiteness marking systems

3.2.1. Swedish

This section will provide an essential summary of the definiteness marking system of the Swedish language.

Swedish nouns have an inflectional category called *species* (SAG, Vol 2, 96), opposing a form without a definiteness suffix, an indefinite form, to a form containing a definiteness suffix, a definite form. That is, they have a definiteness category expressed by a nominal suffix, which varies according to the gender and number of the noun, e.g.:

(9)	*en*	*film-Ø*	vs	*film**en***
	a.ART.COMM.INDEF	movie.SG.[COMM]		movie.SG.[COMM].DEF
	'a movie'			'the movie'

In the linguistic literature, this definiteness morpheme is called the suffixed definite article or marker (suffixal article) or the bound definite article, or the postposed definite article (here referred to as SDEF) (Perridon 1989, Delsing 1993, Börjars 1994, SAG, Lyons 2003, Dahl 2004, 2010, Julien 2005, Lohrmann 2011). Moreover, Swedish also has a free-form definite article or a syntactic determiner, also called the preposed definite article *den*, which SAG describes as a special kind of definite pronoun (SAG, Vol 2, 301). I will refer to the preposed definite article as PDEF. It agrees with the noun in gender and number. This article is normally obligatory in a definite NP containing a descriptive prenominal attribute (modifier), if other definite attributes are absent (SAG, Vol 2, 301) or, in terms of CGEL, if determiner position is not filled by any other determiner and an NP contains an internal adjectival modifier (CGEL 330), e.g.:

(10)	***den***	*långa*	*filmen*
	the.ART.COMM.DEF	long.SG.DEF	movie.SG.[COMM].DEF
	'the long movie'		

That is, this article is only used if a noun has an adjectival attribute. That is why it is also sometimes called the adjectival definite article (SAG, Vol 2, 301)). Other definite attributes include demonstratives[30] and determiner-possessives[31]. Among the adjectival attributes, the so-called relational pronouns and their sub-groups need to be mentioned, as their definiteness value sometimes clashes with that of other determiners, leading to what looks like irregularities (sometimes also called definiteness mismatches). These include comparative pronouns (e.g., *samma* 'same'), "ordinative"[32] pronouns (e.g., *första* 'first', *sista* 'last', *förra* 'previous'), perspectival pronouns (e.g., *nedersta* 'lowest', *vänstra* 'left') and focusing pronouns (e.g., *själv* 'self', *enda* 'the one', *blotta* 'only'). Also, quantifiers, especially universal quantifiers, play a special role in the structure of the Swedish NP, as well as

[30] The Swedish pronoun *sådan* 'such', which according to SAG is classified as a comparative pronoun, may in some uses function as a demonstrative in definite NPs (van der Auwera, Coussé 2016). See also footnote 67.

[31] There are two types of possessives, determiner-possessives or determiner-genitives, and non-determiner genitives (Koptjevskaja-Tamm 2003, 516).

[32] The term "ordinative" pronouns is used by Dahl (2004, 153), which is borrowed from SAG (Vol 2, 233, 236, 243, 246 & 436). This term is created in parallel to the term *ordinatives* or ordinal numbers.

in the reference identification process, as will be shown in the following paragraphs.

This preposed article is also used as a definite attribute in an NP if it is followed by a restrictive relative or narrative clause. In the case of a relative clause, the suffixed article may be omitted, e.g.:

(11)	***den***	*långa*	*film*	*som*	*jag*
	the.ART.COMM.DEF	long.SG.DEF	movie.SG.[COMM]	that	I
	såg	*igår*			
	see.PST	yesterday			
	'the long movie that I saw yesterday'[33]				

Besides the preposed and/or suffixed articles, one more feature characterises a definite NP, namely the obligatory definite (weak) adjectival form (here referred to as AdjDEF). It is probably one of the most consistent indicators of whether a singular NP is definite or not. This, however, applies to NPs where the head of an NP is in the singular, since the weak form of an adjective is identical with the plural form of that adjective, e.g.:

(12)	*en*	*lång-Ø*	*film*	vs
	a.ART.COMM.INDEF	long.SG.COMM.INDEF	movie.SG.[COMM]	
	'a long movie'			
	den	*lång**a***	*filmen*	
	the.ART.COMM.DEF	long.SG.DEF	movie.SG.[COMM].DEF	
	'the long movie'			

To sum up, in a simple Swedish definite NP containing an adjectival modifier[34], there can maximally be three different types of definiteness markers, namely, a preposed free definite article (PDEF), a definite (weak

[33] This example is my own. It builds on attested examples, like *den röda klänning(en) som jag mest älskade* 'the red dress that I loved the most' (SAG, Vol 2, 301).

[34] By the term "simple" I here refer to the NP containing one syntactic determiner being a free preposed article, one adjectival modifier, and the head marked for definiteness by the postposed morphological article. This is done in order to distinguish between this type and other more complex types of NPs where several determiners and modifiers can be present, as well as some of them omitted, as shown in Table 2. It has to be noted that a distinction is also made by SAG (Vol 3, 15) between simple and complex definite NPs, complex NPs differing from simple ones in that that they start with a pronominal attribute, e.g., a totality pronoun, e.g., *alla hans nya cyklar* 'all his new bicycles'.

form, AdjDEF) of an adjective, and a suffixed definite article (SDEF). The representation of a definite Swedish NP looks like this:

NPDEF = PDEF + AdjDEF + SDEF

Sometimes, for various reasons, one (or more) of the three markers above listed is omitted – it can be either PDEF or SDEF, or both PDEF and SDEF. AdjDEF, on the other hand, cannot be omitted under any circumstances[35], e.g.:

Table 2. Patterns of definiteness marking in the Swedish definite NP

NPDEF = PDEF + AdjDEF + SDEF[36]			**Translation**
den	galn-**e**	despot-**en**	*the crazy despot*
-	norr-**a**	sida-**n**	*northern side*
den	gråast-**e**	höstdag -	*the greyest autumn day*
-	först-**a**	pris-	*first prize*

To conclude, it must be said that in the well-developed Swedish system of determiners (definite attributes), containing overt definiteness markers, both syntactical and morphological, adjectives seem to play a significant role. In the literature analysing the Swedish NP much attention has been paid to the interplay between the free and the bound definiteness marker. However, I think that the absence of these markers (the previously mentioned definiteness mismatches) provide even more important insight into the interplay between the adjectival marking of definiteness and the definite status of an NP.

3.2.2. Lithuanian

The detailed description of the definiteness marking in modern Lithuanian is presented in Section 2.5. The repeated table below contains the summary of it:

[35] There are very rare instances where the short form is used. See examples (41b) and (41c) for more information.

[36] A detailed overview of the variation in definiteness marking across the Continental Scandinavian standard languages is available in an article by Dahl on definite articles in Scandinavian (Dahl 2004, 154).

Table 1. Definiteness marking in Lithuanian (repeated as on page 32)

Prenominal attributes	**NP$_{DEF}$**	
Adj	**Adj$_{INDEF}$ + N**	**Adj$_{DEF}$ +N**
	sena knyga 'the old book'	*senoji knyga* 'the old book'
Other definite attributes[37]	**Det (+Adj$_{INDEF}$)+ N**	**Det (+Adj$_{DEF}$) +N**
	ta sena knyga 'the/that old book'	*ta senoji knyga* 'the/that old book'

The Lithuanian category of definiteness has the following expressions:

1. In an NP with an adjectival modifier, it is marked by the presence of the suffixal ending on the Adj;
2. Otherwise, it is conveyed through usage of definite attributes, incl. demonstratives, possessive genitives and other, as will be outlined in the following sections.
3. Sometimes, the so-called definiteness effects (Lyons 2003, 227-251) come into play, e.g., mass nouns and plurals as objects of perfective verbs are interpreted as definite (Holvoet, Tamulionienė 2006, 30-32); certain word order models in which the thematic (topicalised) NP gets a definite reading as seen from the functional sentence perspective. Also, according to Lyons, existential sentences, property predication and superlatives, as well as a number of other sentence syntactical constructions are to be treated as definiteness effects.

In addition to what has been said, a couple of important observations based on recent studies in functional sentence perspective and translation studies will be discussed here to complete the picture with linguistic strategies used to compensate for the absence of prototypical definiteness markers, viz. articles.

From recent studies in functional sentence perspective (Mickūnaitė-Griškevičienė 2004), definiteness marking in Lithuanian (Sprauniené 2011) and translation studies (Vaitkutė 2017), it has become apparent that a number of linguistic strategies are used to convey the definite status of NPs. Certain compensation mechanisms are employed to make up for the absence of (in)definite articles in Lithuanian when translating from e.g., Scandinavian languages. They mostly include the process of addition whereby certain elements are added to an NP to clarify or underline its definite status, e.g.,

[37] The types of definite attributes will be described in detail in section 3.5.

demonstratives, possessives, quantifiers, participles, adjectives and other, as shown in the examples below:

Table 3. Compensation for the absence of the definite article in Lithuanian

No	Swedish	Lithuanian translation
(13)[38]	Förr brukade hon säga att de skulle bo där i *hus**et*** en dag [...]	Anksčiau ji sakydavo, kad juodu vieną dieną apsigyvensią ***tuose namuose***.
	house.SG.[NEUTR].DEF	that.LOC.PL.M house.LOC.PL.[M]
	'Before she used to say that they would live there in the house one day'	
(14)	*Familj**en*** bor i en stor lägenhet vid Vanadisplan.	***Jų šeima*** gyvena didžiuliame bute Vanadžio aikštėje.
	family.SG.[COMM].DEF	they.GEN.PL.M family.NOM.SG.[F]
	'The family lives in a big apartment on the Vanadis Square.'	
(15)	Rakel ligger i sängen den första natten i ***det främmande nya hemmet*** och lyssnar till husets alla ljud.	Rakelė pirmąją naktį guli lovoje, klausydamasi namo garsų ***savo naujame svetimame bute***.
	the.ART.NEUTR.DEF alien home.SG.[NEUTR].DEF	her_own.POSS new.LOC.SG.NONDEF alien.LOC.SG.NONDEF flat.LOC.SG[M]
	'During the first night, Rachel lies in bed listening to the sounds in her new alien home.'	
(16)	Han mindes den tysta vintern, då ***ljuden*** sögs in i snön och de vuxna höll sig inomhus.	Prisimena tylias žiemas, kai sniegas sugerdavo ***visus garsus***, kai suaugusieji daugiausia būdavo namuose.
	sound.PL.[NEUTR].DEF	all.ACC.PL.M sound.ACC.PL.[M]
	'He remembered the silent winter when the snow used to suck in the sounds and the grown-ups stayed inside.'	

As shown in the examples presented in the table above, additional constituents are added to the Lithuanian NPs to emphasise their definite status, increasing the number of prenominal modifiers.

[38] Examples (13) – (16) are taken from Vaitkutė (2017).

3.3. Prenominal attributes

3.3.1. What they are and how they are organised

The grammatical elements that can precede a noun in Lithuanian and Swedish are articles (only in Swed.), quantifiers, demonstratives, pronouns, genitives, and adjectives[39], e.g.:

(17) *alla* *dessa* *hans* *många* *andra* *norska* *vänner*
all.PL this.PL he.GEN.SG many other.PL Norwegian.PL friend.PL.[COMM]
'all these many others of his Norwegian friends'[40]

(18) *visi* *tie* *jos* *kiti*
all.NOM.PL.M that.NOM.PL.M she.GEN.SG.F other.NOM.PL.M
lengvi *pinigai*
easy.NOM.PL.M money.NOM.PL.[M]
'all that other easy money of hers'[41]

As Lyons rightly notes, "no investigation into the nature of definiteness can proceed far without consideration of the place of articles and other determiners within noun phrase structure" (Lyons 2003, 41). It is evident that elements comprising a definite NP in any given language come in a particular order (in some languages more rigid, e.g., Scandinavian languages, and in some more liberal, e.g., Baltic languages). A noun can be preceded by certain elements; it can likewise be followed by these. However, in the case of unmarked word order, regular patterns can be found in the linear structure of these elements, e.g., adjectives in the Lithuanian NP are most likely to be found in prenominal position, likewise in Slavonic and Scandinavian languages; in Romance languages, they will be placed postnominally.

Variation in the placement of prenominal attributes is observed as well. However, some of these elements have a tendency to move and take up different slots in an NP more freely than others, as will be shown. It is obvious that the particular placement of such elements might have implications for the reading of the NP, as well as for the NP structure, e.g., a Swedish NP with the

[39] And other noun or adjective-like elements, e.g., participles, cardinals, ordinals, and certain pronouns.

[40] This example is from SAG (SAG, Vol 2, 249).

[41] Adopted from the corpus example *visi tie jos pinigai* 'all that money of hers' and expanded by adding modifiers *kiti* 'other' and *lengvi* 'easy'.

leftmost element being a possessive pronoun allows only one placement for this pronoun (or a genitive, e.g., *Peters → Peters många nya hästar*):

(19) ***hans*** *många* *nya* *hästar* vs
he.GEN.SG.M many new.PL horse.PL.[COMM]
'his many new horses'[42]
**nya hans många hästar*
**många hans nya hästar*
** nya många hans hästar*
**många nya hans hästar*

In Lithuanian, more variation is possible, e.g., an NP in the leftmost position containing a collective cardinal number allows variation both in terms of placement of the genitive pronoun *his/her* and in terms of semantic interpretations:

(20) a. *trejetas* ***jo*** *draugų*[43]
threesome.NOM.SG.[M] he.GEN.SG.M friends.GEN.PL.[M]
'three of his friends'[44]

b. ***jo*** *trejetas* *draugų*
he.GEN.SG.M threesome.NOM.SG.[M] friends.GEN.PL.[M]
'three of his friends'

c. ***jo*** *draugų* *trejetas*
he.GEN.SG.M friends.GEN.PL.[M] threesome.NOM.SG.[M]
'three of his friends'

[42] This example is taken from SAG (SAG, Vol 3, 5).

[43] This example already poses a question – what should be considered the head of the NP? Is it *a threesome* or *friends*? I am inclined to think that *friends* should be considered a head because of its rightmost position. Whether this would then imply that *trejetas jo draugų ≠ jo draugų trejetas* is a separate broad discussion subject; and it will not be taken up in this thesis.

[44] This example is based on the one containing an NP *trejetas mano draugų* 'three of my friends', which I found online: *Jau* dabar *trejetas mano draugų susidomėjo WM Nokiom (ir 925*)...
http://www.mobili.lt/lt/forumas/mobilieji_telefonai/nokia/nokia_lumia_925.html?&psl=5, 2018-10-14. Likewise, I found online structures similar to (20b) and (20c) with different lexemes, however, I chose to use the same lexemes in these examples for the purpose of illustration.

Reading (20a) is the most neutral one, while some speakers would consider (20b) and (20c) possible only in a marked structure as demonstrated below:

(21) ***Jos*** *draugių* *trejetas* *jau*
she.GEN.SG.F friend.GEN.PL.[F] threesome.NOM.SG.[M] already
užsiregistravo *į* *komandines*
register.3PST to team.ADJ.ACC.PL.F
varžybas. *O* ***jo***[45] *trejetas*
competition.ACC.PL.[F] and he.GEN.SG.M threesome.NOM.SG.[M]
draugų?
friend.GEN.PL.[M]

'Three of her girlfriends have already registered for the team competition. What about a threesome of his friends?'[46]

In terms of the possible readings, in (20a) the numeral is the leftmost element quantifying over an NP *his friends*; in (20b) the leftmost element is a possessive specifying an NP *a threesome of friends*; (20c) is the most complicated to interpret, as the rightmost element *threesome* syntactically could be considered the head of the NP with two genitival modifiers.

One could thus say that Swedish has a much more rigid word order when it comes to combinatory variations amongst the prenominal modifiers; but even Lithuanian, with its much freer word order, has its limitations and impossibilities.

It is known that linear relations between components of a syntactic structure are of importance (Haspelmath 1999; Croft 2004; Langacker 2008). Hence, it is the order of linear precedence of noun attributes and the interplay between them that will be closely looked at first and foremost in relation to the Lithuanian NP and its Swedish counterpart, sometimes also to English.

3.3.2. Definite attributes (determiners) and modifiers

Aiming to describe the topology of the definite NP, be it a Scandinavian or a Lithuanian one, we must identify the potential functional positions (slots, fields, types of attributes) that are found in the NP. Examining the Scandinavian patterns, one can notice that there is clearly a difference between the prenominal attributes, in that even though they all contribute to the

[45] The genitive in this NP bears the stress.

[46] I am the author of this example, however, as in (20), usage of similar structures is attested.

definiteness reading of the noun, not all of them make an NP definite. These attributes can be broadly divided into two groups, based on their function: a) functional modifiers; b) descriptive modifiers (SAG, Vol 3, 4). The so-called functional modifiers are the ones that perform the function of determiners. In Perridon's terminology, they constitute the field of determination (Perridon 1989, 195). In SAG, they are called definite attributes. In CGEL, they are called determiners, dependents that perform the function of a determiner, as opposed to a lexical class of determinatives (CGEL, 330, 355-356), namely those "whose distinctive syntactic properties concerns their association with the determiner function" (CGEL, 355), e.g., the English singular noun *book* only becomes an NP if articles *a* or *the* are added to it. The same could be said about its Swedish counterpart *bok*, which in order to become an NP needs to be either *en bok* 'a book' or *boken* 'the book'. Lithuanian, as opposed to English and Swedish, lacks articles. Yet it is not always the case that determinatives perform a determiner function[47]. The CGEL provides a finite list of words which it calls *basic determiners* (CGEL, 356), which includes articles and a number of various types of determinatives (e.g., *this*, *that*, *each*, *all*, cardinal numbers, etc.). In the Swedish tradition these are called *definita attribut* 'definite attributes' (SAG, Vol 3, 13); in Danish, they are called *bestemmere* 'determiners' (GDS, Vol 1, 181). In Norwegian grammars, this lexical class is recognised and called *determinativer* 'determinatives'. They include 5 types of determinatives, namely articles, possessives, demonstratives, quantifiers and intensifiers like *selv* 'self' and *egen* 'own' (Norwegian *forsterkende ord*) (NRG, 202; NSF, 29). In grammars of Swedish and Lithuanian, there is not such a tradition of singling them out as a lexical class. Therefore, I will refer to them respectively as definite attributes (definite articles, demonstratives, quantifiers, possessive-genitives). They all are known to perform determiner function in an NP.

Like SAG, Perridon singles out a special field for quantification, known in SAG as quantitative attributes that follow definite attributes, but precede descriptive attributes. In grammars that acknowledge determinatives as a lexical class, quantifiers of all types are considered to be determinatives. Yet, as will be shown later, only some quantifiers are determiners, while others

[47] To clarify the difference between the two, it must be said that "while determinatives function most distinctively as determiners in NP structure, most of them are not restricted to that function" (CGEL, 330), e.g., *all* in *all children* is a determiner, but in *all the children* it is a modifier, while *the* functions as a determiner. Likewise, *this* in *this girl* functions as a determiner, but in *She is about this tall* is clearly a modifier (CGEL, 25).

behave like adjectival modifiers, assigning cardinal quantitative properties to the head.

Besides definite attributes, there is another group of prenominal attributes that is often described as performing modifier function, describing and specifying the content of the head. In Perridon's terms, they constitute the field of description (Perridon 1989, 195). In SAG, they are called adjectival attributes and include adjectives, participles, and some types of pronouns (SAG, Vol 3, 13). This group is of significance as it has an important implication for the structure of the Lithuanian NP. This is because definite adjectival forms, traditionally considered to represent canonical modification, perform the function of determiners, in addition to their descriptive function. Also, genitival constructions, widespread in Lithuanian, would fall into this category; more specifically, non-determiner genitives of the type illustrated below would belong to this group, e.g.:

(22) *linksmų* *plaučių* *jaunuolis*
jolly.GEN.PL.M lung.GEN.PL.[M] lad.NOM.SG.[M]
'a jolly lad, a lad with a good sense of humour'

(23) *mano* *vaikystės* *draugė*
I.GEN.SG childhood.GEN.SG.[F] friend.NOM.SG.[F]
'my childhood friend'

The genitives in (22) are non-determiner genitives, in a metaphorical way describing a quality of a young man, viz. having good lungs, i.e., prone to laughter, having a good sense of humour. The genitive in (23) refers to the time frame, i.e., *my childhood friend* is a friend I made in my childhood.

To sum up, in Swedish, the following attributes serve as definite attributes (performing the determiner function): definite articles, both preposed and postposed; demonstratives; possessives and genitives; quantifiers and certain pronouns (e.g., *samma* 'same', *nästa* 'next', etc.) (SAG, Vol 3, 15-43). The modifier function is served by modifier attributes, which include adjectives, participles and some types of pronouns. In Lithuanian, it is demonstratives (incl. the arthroid *tas*, on which see below), possessives and determiner-genitives, quantifiers and definite adjectival modifiers that function as determiners. As previously mentioned, even indefinite adjectival forms can sometimes serve as determiners (see example (26)). This will be discussed in section 3.6.2.

3.4. Determinatives, determiners and definiteness carriers

It is evident that elements in grammars considered to be determinatives include both markers of definiteness and indefiniteness. As the main interest of this section is the expression of definiteness, I will only focus on the determinatives that in their determiner function assign definiteness, viz. determinatives as indefinite articles, also the ones like *some, few, several, any*, etc. will not be looked at here.

What is the relationship between determinatives, determiners and modifiers when we speak about the category of definiteness?

Grammars of the Baltic, Scandinavian and English languages all make a clear distinction between common nouns and proper nouns. This is important as it adds to the understanding of the category of definiteness. The difference between common nouns and proper nouns lies in that the latter have little or no descriptive content (SAG, Vol 2, 9); they denote a category consisting of one individual, as names are prototypically assigned to unique individuals.[48] Common names, in contrast, refer to categories (or types) that include more than one individual member (GDS, Vol 2, 450; Holvoet, Tamulionienė 2006, 12). It is through various processes of individualisation, through modification and quantification that definiteness arises as a multi-layered phenomenon allowing a noun to become a part of a nominal unit, namely an NP that can be used in discourse (GDS, Vol 2, 464; Holvoet 2009, 19):

[48] CGEL makes a distinction between proper nouns and proper names. The main use of proper names is to refer to the particular entities that they name; in this use, they constitute NPs. Proper nouns, on the other hand, are word-level units that belong to the category of noun, e.g., *Zealand* is a proper noun, whereas *New Zealand* is a proper name (CGEL 2002, 516).

Table 4. Stages of NP modification

	Noun[49]	**NP**	**Quantification**	**Quantification + modification**	**Determination + quantification + modification**
Lithuanian	*berniukas* → boy	*berniukas* → boy.NOM.SG.[M]	*trys berniukai* → three boy.NOM.PL.[M]	*trys pavargę, bet laimingi berniukai* three tired.NOM.PL.M.INDEF but happy.NOM.PL.M. INDEF boy.NOM.PL.[M]	*tie trys pavargę, bet laimingi berniukai* that.DEM.NOM.PL.M three tired.NOM.PL.M.NONDEF but happy.NOM.PL.M. NONDEF boy.NOM.PL.[M]
Swedish	*pojke* → boy	*en pojke* → a boy.SG.[COMM].INDEF	*tre pojkar* → two boy.PL.[COMM].INDEF ‘three boys’	*tre trötta men lyckliga pojkar* two tired.PL but happy.PL boy.PL.[COMM].INDEF	*de tre trötta men lyckliga pojkarna* those.DEM.PL three tired.PL but happy.PL boy.PL.[COMM].DEF
Transl	‘boy’	‘a boy’	‘three boys’	‘three tired but happy boys’	‘those three tired but happy boys’

[49] CGEL recognises a category of *nominals*, an intermediate category between nouns and NPs, e.g., in the phrase *the **old man*** the highlighted part is considered to be a nominal. These are usually not single words, but expressions, unable to function as arguments on their own.

Also, in languages where the marking of (in)definiteness is obligatory, it is determinatives that turn nominal units into NPs. Determinatives are important in reference assignment, in signalling to the speaker that he/she knows/is able to identify the referent of an NP. However, it is also known that many definite NPs are *non-referential*, e.g., the phrase *the many thousands of people who live and work in a large and congested area*; moreover, even if some processes, like quantification using cardinal numbers, do help out in narrowing down the individual members of a certain category, e.g. *three boys* in the example above, as opposed to having a choice of all the hypothetical boys of the real (or imaginary) world, they do not provide a satisfactory solution to singling out a nominal referent. However, adding, e.g., *all* or *those* in the example above, makes the NP definite.

In the tradition of Cognitive Grammar (CG), this function is described as *grounding*, which "is not a grammatical category (like noun, verb, or preposition). It is rather a semantic function, an aspect of conceptual organization by which an expression qualifies as a nominal[50] […]" (Langacker 2008, 272). Besides making a noun phrase a nominal (or an argument, e.g., a subject or an object), the grounding elements are crucial in singling out a nominal referent; they can act alone as full nominals, e.g.: ***These** are not suitable*; they do not underlie predicate constructions of the type:

(24) **The politicians who can be bought are **all/most/every/each**.* [51]

as opposed to other word-categories, e.g., adjectives:

(25) *The politicians are corrupt.*

The above examples are important as they reveal the difference between grounding and nongrounding determinatives like quantifiers. Evidently, there is a difference between the universal quantifiers like *all* and *every* and cardinal numbers like *three* in that the constructions of the type *The politicians were three* is grammatical, as opposed to (24).

That is why I propose rather to use the term definiteness carriers to refer to elements that trigger the definite reading of the NP; also, that is why further in my analysis I suggest distinguishing between two fields of quantification, namely, the Q_1 to be assigned to universal quantifiers (definiteness carriers) and Q_2 to cardinal expressions (modifiers, specifying the cardinal quantification of the set).

[50] A nominal in the CG tradition refers to what I call an NP in this section.
[51] This example is taken from Langacker (2008, 274).

Apart from articles (as described in 3.5.1), definiteness carriers can combine with one another, and in languages that share adjectival marking of definiteness, they most often precede the noun and the adjectival modifiers (e.g., Swed. *hela denna osedvanligt sorgliga historia*, Lith. *visa ši neįprastai liūdna istorija* 'all this unusually sad story'), with one of the two, either a definite attribute or a universal quantifier, usually taking the first position in an NP.

Definite attributes by default trigger definiteness marking on preposed adjectival modifiers in Scandinavian languages, and quite often in Baltic languages (more consistently in Latvian, where definiteness marking is more grammaticalised compared to Lithuanian).

Also, it has to be noted that elements classified as determinatives have a very high usage frequency in languages that have them in their inventory. Articles (both definite and indefinite) are amongst the highest-ranking words in English and Swedish in terms of relative frequency. Also in Lithuanian, based on the data available in the Frequency Dictionary of the Written Lithuanian Language (Utka, 2009), it is evident that demonstratives and other elements qualifying as determinatives rank very high on the frequency list. The list appears to be quite extensive with 12 determiners within the top 100 positions; and 25 determiners within the top 500 positions. It can be observed that demonstrative (deictic) *tas* 'that/the' leads this list, being in the fourth position. The dictionary provides no data on the stress pattern. However, Rosinas (Rosinas 1996, 2009) and Tumėnas (Tumėnas 1988) state that the unstressed bleached deictic *tas*, also called the arthroid, known for its high frequency (which is re-confirmed here as well), behaves in an article-like function (see 3.5.3). One could assume that this high position on the list is due to arthroid usage. Yet, it should be acknowledged that some of these uses might be accounted for as recognitional uses of demonstratives (Diessel 1999, 105-109); these are prenominally used unstressed demonstratives that introduce information that is new in the discourse, yet "old" to the hearer and private, viz., shared between the speaker and the hearer due to common experience, e.g.:

(26) *I could not sleep last night.* ***That*** *dog (next door) kept me awake.* [52]

[52] This example is from Diessel (1999, 106).

It would seem that this use would be more common in the spoken language, and hence its statistical significance in the corpus of written language would be negligible[53]. This, however, remains to be examined.

Summarising, it could be said that even though determinatives often serve as determiners and have interesting features, like high frequency usage, they do not neatly map onto the category of definiteness; not all definite determinatives are definiteness carriers.

3.5. Definite attributes

3.5.1. Definite articles

Lyons introduced the terms of *simple definites* for NPs where definiteness arises due to the presence of definite articles vs *complex definites* for NPs whose definiteness "is due to something other than presence or absence of an article" (Lyons 2003, 107). In this regard, Swedish and Lithuanian differ significantly, as Swedish has two definite articles, namely SDEF and PDEF, whereas Lithuanian has none. It is important to highlight that even though Lyons calls the Lithuanian pronominal adjectival morpheme a phrasal clitic article, apart from the fact that this is only historically accurate, it is easy to demonstrate that the function of the definite ending differs from that of the typical article. First of all, it appears on an adjective and therefore contributes to the definiteness reading of an NP that contains an adjectival modifier; whereas the Scandinavian postposed article and the English *the* modify NPs with or, most importantly, without prenominal modifiers. The Swedish free article is only used when an adjectival attribute is present in an NP. Secondly, one of the special features of articles is that they do not combine with other determiners[54] in many languages, while the Lithuanian pronominal adjectival morpheme does not prevent other determinatives from appearing alongside the definite adjectival modifier in an NP. It is the definite article that transforms a nominal in the CGEL tradition[55] (a noun with a prenominal attribute, e.g., *old book*) into a complete NP *the old book*. In Lithuanian, both

[53] It should be noticed that the recognitional use of *tas* in Lithuanian may have an implication for the structure of the NP in that that it could insert itself into a slot between an adjectival attribute and a noun like in the following example: *Kurį laiką blokuodavau vis naujas tas anketas […]* lit. 'For quite some time I kept on blocking ever new those questionnaires'; found online: https://www.vinted.lt/forumai/sirdies-reikalai/1646572-mane-seka-ig., 2018-11-23. This is an unusual position for a demonstrative to occupy.

[54] Universal quantifiers behave differently, see 3.5.2.

[55] For a definition of a nominal in CGEL, see footnote 50.

sena knyga ‘old book’ and *senoji knyga* ‘the old.DEF book’ are fully realised NPs that could both have a definite reading in particular contexts:

(27)	***Sena***	*knyga*	*pirmiausia*	*jam*
	old.NOM.SG.F	book.NOM.SG.F	first.ADV	he.DAT.SG.M
	krito	*į*	*akis.*	
	fall.3PST	into.PREP	eye.ACC.PL.F	
	‘The old book caught his eye first.’[56]			

(28)	***Senoji***	*knyga*	*pirmiausia*	*jam*
	old.NOM.SG.F.DEF	book.NOM.SG.F	first.ADV	he.DAT.SG.M
	krito	*į*	*akis.*	
	fall.3PST	into.PREP	eye.ACC.PL.F	
	‘The old book caught his eye first.’			

In (27), the context allowing the definite reading would be the following: the viewer is in a room full of books, predominantly new. Hence, his eyes focus on the old book, which is an unusual object. The feature *old* singles out the referent. In (28), the usage of the definite form could be explained by an anaphoric function, referring to a particular book that was mentioned in the previous context, e.g., there were two editions of the same book, an old and a new one; and now in a room, the viewer identifies the old edition. This reflects the key notions associated with definiteness, namely *uniqueness* and *familiarity*, described by Lyons (Lyons 2003, 2–12). In (27), the book is identifiable because of its unique feature *old*; whereas in (28) it is implicated that the reader is already familiar with the object due to some previous encounter or knowledge about it. It has to be noted that the short adjectival form is the choice to indicate uniqueness, and the long one to indicate familiarity.

The synchronic data shows that there is a need to re-interpret the historical phrasal clitic, since in modern Lithuanian it appears alongside other determinatives, e.g., *ta mano sen**oji** knyga* ‘that/this old.DEF book of mine’. The reading of this NP implies several loci of definiteness in the structure of the NP.

Another important remark is on the difference between the two kinds of Scandinavian articles, in particular in Swedish. A Swedish NP that contains prenominal modifiers will have two articles – a PDEF and a SDEF attached to the noun itself and a bound definite article, e.g., ***den*** *gamla bok**en*** ‘the old book’. In this regard Danish is different, as it only has one definite article in an NP with prenominal modifiers, namely the free preposed definite article,

[56] I am the author of the examples (27) and (28).

e.g., ***den*** *gamle bog* 'the old book'. Yet, the usage of PDEF and SDEF, as will be shown in section 3.6.1, also seems to be linked to the concepts of uniqueness and familiarity.

3.5.2. Quantifiers (universal and other)

Another clearly identifiable group functioning as determiners and exhibiting specific semantic properties is the group of universal quantifiers (Lith. *bendrumo įvardžiai*[57]) as they quantify over the totality of objects (in a set); with the exception of *half*, which could be described as quantifying over the totality of one of the two halves. The language-specific inventories of universal quantifiers differ; however, some prototypical ones like *all* and *whole* can be identified in English, Lithuanian and Swedish.

Universal quantifiers are peculiar in that they are the only group of determinatives that combine with definite articles, both preceding (the case in Scandinavian languages) and following them, e.g., English *the whole village*, but *all the villagers*; in Scandinavian languages, they can only precede the determinative, e.g., Danish *hele det danske samfund* lit. 'whole the Danish society'; Swedish both *hela den tiden* lit. 'whole the time.DEF' and *hela tiden* lit. 'whole time.DEF' are possible.

The Swedish universal quantifiers are *all* 'all/whole/every', *samtliga* 'all/all together', *hela* 'whole', *halva* 'half', *båda* 'both'. Also *själva* 'self' and sometimes *varje* 'every' are included here (SAG, Vol 3, 24–25). It has to be noted that in Modern Swedish only universal quantifiers can appear in the leftmost position in a definite NP. In older texts, however, all kinds of quantifiers could occupy this position, cf. Old Norse *drap eg þá marga vargana* 'killed I then many wolves.DEF' (Perridon 1989, 197).

Lithuanian *visi* 'all', *visas* 'whole', *kiekvienas* 'every' seem to behave similarly with regard to syntax, e.g., *visus tuos metus* 'all the/those years'. However, if in English and Swedish universal quantifiers show rigidity with regards to their placement – either strictly preceding or strictly following demonstratives, their Lithuanian analogues show more flexibility in combinatorial possibilities, e.g., *visi tie virusai* 'all the/those viruses' and *tie visi virusai* 'the/those all viruses' are equally possible. The frequency of their

[57] Rosinas calls them *bendrumo įvardžiai* (1996, 121). However, he later specifies and calls them *egzistavimo ir bendrumo įvardžiai* (1996, 131). In the Lithuanian Grammar (Ambrazas et al. 2006, 270), they are called *teigiamieji apibendrinamieji įvardžiai*. The suggested translation is *generalising pronouns*.

usage will differ significantly, with the universal quantifiers preceding the arthroid/demonstrative being a clearly predominant type.[58]

Since *half* seems to belong this group, what about *one third*? This type of prenominal modifiers is called fractions according to CGEL (CGEL, 434). *Half* seems to be a special case amongst fractions in that it can form phrases like *half the village*, whereas **third the village* is impossible. Instead, we have a prepositional phrase as in *one third of the village*, where the noun will be marked with a definite article: *the village*. The situation is similar in Swedish. Meanwhile in Lithuanian, *pusė kaimo* (half the village) and *trečdalis kaimo* (the third of the village) do not differ syntactically; both contain the genitive-marked noun *kaimo* 'of the village'. It has to be noted that *pusė*, likewise *trečdalis,* morphologically are nouns, but consistently appear prenominally. To sum this up, we could say that fractions are not typical prenominal modifiers but may marginally be drawn into their orbit.

A different type of quantifiers[59], speaking in semantic terms, is cardinal numbers, the primary function of which is to provide an exact cardinality of objects (the set). In this, they are instrumental in enabling reference identification, yet insufficient to make a referent definite, as described in Table 4. It is often said that appearing as prenominal modifiers, they tend to behave like adjectives. This is, however, not the case since syntactically descriptive adjectives can combine with one another; whereas the cardinal quantification can only be expressed once, and hence, occupies only one position in the prenominal structure of modifiers[60]. Identifying its position on an axis running from left to right, it will be found following determiners such as articles and universal quantifiers, but preceding other adjectival attributes, e.g., *the three big boys*, *all the four children*, *three big boys*.

An interesting case where both types of quantification, universal and cardinal, are fused, is reflected in the usage of the dual demonstratives in Lithuanian. The duality concept is also reflected in the pronoun *abu* 'both', which is a universal quantifier, and behaves like *visi* 'all'. In Lithuanian, there

[58] A simple search in the Corpus of Contemporary Lithuanian Language (tekstynas.vdu.lt) will show the following results: *visi tie* 'all the/those' 1015 hits vs. *tie visi* 'the/those all' 199; *visi šie* 'all the/these' 800 vs. *šie visi* 'the/these all' 28. They are undoubtedly statistically significant.

[59] Yet another semantic type of quantification, namely existential quantification, should be mentioned here. However, as it is closely related with the grammatical marking of indefiniteness, it will not be discussed here. More about existential quantification can be found in CGEL (CGEL, 358-359).

[60] This is a very simplified account of the difference between cardinals and adjectives. It seems that cardinals could best be described as being on the borderline between quantifiers and adjectives. This is in need of further exploration.

exists a category of dual demonstratives such as *šiedu* 'these two', *tiedu* 'those two', *aniedu* 'those two', reflecting the trinomial system of Lithuanian with the demonstratives *šie* and *anie* indicating distance (close vs. far) and *tie* indifferent to the concept of distance[61], e.g.:

(29)	*tiedu*	*solidūs*	*pirkėjai*
	those.DEM.NOM.DUAL.M	solid.NOM.PL.M	buyer.NOM.PL.[M]
	'those two solid buyers'		

Even though there is a strong link between quantification and definiteness, since quantification *per se* is a reference-assigning mechanism, as "it derives from the ability to perceive something as a token, an instance of a class of referents, and the ability to differentiate between one and more than one (i.e. the 'plurality' of) instances of the referent" (Kibort, Corbett, 2008)[62], this does not mean that the category of quantification directly maps onto the category of definiteness. It only does in the case of universal quantifiers.

I will conclude that with regard to quantifiers, two separate positions need to be established in the linear structure of prenominal modifiers, namely that of universal quantifiers (a determiner category) and that of cardinal quantifiers (a modifier-like, reference-narrowing category), with the latter occupying the position to the right of the definite attributes on the left-to-right axis.

3.5.3. Demonstratives

Yet another type of prenominal attributes that perform the function of determiners in both Lithuanian and Swedish is demonstratives. In SAG, demonstratives are classified as a special type of definite pronouns that obtain their definite function through deictic or anaphoric use (SAG, Vol 2, 255). They help identify a referent that is relatively pronounced in discourse. Often it is through the process of contrasting the referent with other possible referents that their meaning is construed, e.g.:

(30)	*Den*	*förklaringen*	*gäller*	*inte*	***denna***
	the.ART.DEF	explanation.SG.[COMM].DEF	apply.3PRS	not	this.DEM.COMM
	gång.				
	time.SG.[COMM]				
	'the/that explanation is not valid this time'				

[61] According to Rosinas, in Lithuanian, there are only three demonstratives that form a trinomial system, namely *šis/šitas, anas* and *tas* (Rosinas 1996, 58-59).

[62] Cf. the distinction between common and proper nouns (see section 3.4).

Here *this time* clearly refers to a particular time as opposed to many other times when the same explanation was used.

Speaking of their usage in constructions denoting different types of definiteness, it has to be mentioned that demonstratives, as opposed to other definite attributes, cannot be used in associative anaphora, nor in larger-situation uses (a term introduced by Hawkins in 1978 to denote instances where a referent is identified on the basis of a large common context/shared knowledge), e.g.:

(31) *I bought a house.* ***The*** *roof was completely new.* vs **I bought a house.* ***This*** *roof was completely new.*

(32) ***the*** *Houses of Parliament* vs *****these*** *Houses of Parliament*

However, (32) is possible with deictic reference if particular buildings are being singled out in opposition to the previous buildings erected on the site. It is not possible to use the demonstrative to indicate the buildings as an object of general knowledge. This is important as it relates to the two key concepts associated with definiteness, viz. uniqueness and familiarity (see the analysis of examples (27) and (28)). Demonstratives by their deictic nature implicate familiarity (or something that a reader/hearer can familiarise himself/herself with) and not uniqueness.

In Swedish, there are 4 demonstratives: *denna, den här* 'this', *den* and *den där* 'that'. In Lithuanian, there are 3 core demonstratives used in definite NPs, namely *šis/šitas* 'this', *anas* 'that' and *tas* 'this/that', e.g., *car* (Swed. *bil*, Lith. *mašina*):

Table 5. Demonstratives in Swedish and Lithuanian

	'this car'		'that/the car'	'that car'
Swed.	*denna bil-ø*	*den här bil-en*	*den bil-en*	*den där bil-en*
Lith.	*ši/šita mašina*		*ta mašina*	*ana mašina*

In both languages, the demonstratives *this* (Swed. *den här*, Lith. *šis/šitas*) and *that* (Swed. *den där*, Lith. *anas*) make a distinction between proximal and distal, referring to nearby and remote objects respectively, whereas *den* and *tas* are indifferent to distance. Both Swedish and Lithuanian demonstratives are adjectival.

It has to be noted that *denna* differs from other Swedish demonstratives in that it does not require the suffixed definite article on a noun (SDEF), whereas all three other exhibit Dem+SDEF behaviour, namely the

demonstrative will be used alongside the suffixed definite article (see section 3.2.1 for a detailed account).

If an NP contains a preposed adjectival modifier, both in Swedish and Lithuanian, it will occupy the same slot in an NP structure, namely following the demonstrative. Adjectives cannot precede demonstratives in either language, e.g.:

(33) a. *denna* *nya* *bil*
this.DEM.COMM new.SG.DEF car.SG. [COMM]
'this new car'

b. *den* *här* *nya* *bilen*
this.DEM.COMM (here).DEM new.SG.DEF car.SG.[COMM].DEF
'this new car'

c. **nya* *denna* *bil*
new.SG.DEF this.DEM.COMM car.SG.[COMM]

d. **nya* *den* *här* *bilen*
new.SG.DEF this.DEM.COMM (here).DEM car.SG.[COMM].DEF

(34) a. *ši/šita* *naujoji* *mašina*
this.DEM.NOM.SG.F new.NOM.SG.F.DEF car.NOM.SG.[F]
'this new car'

b. **naujoji* *ši/šita* *mašina*
new.NOM.SG.F.DEF this.DEM.NOM.SG.F car.NOM.SG. [F]

Finally, a few words need to be said about the Lithuanian *tas*, which some scholars describe as an arthroid, i.e., a unit with function coming close to that of a definite article (Rosinas 1996; Tumėnas 1988). The preposed, unstressed, bleached deictic *tas* differs from a true demonstrative in that in a prenominal position it has lost its distance-related opposition, it cannot be replaced by a demonstrative *šis* 'this' or *anas* 'that'; the meaning of an NP would change (Rosinas 1996, 67). Also, its optionality is another criterion separating it from other demonstratives, namely, if *tas* is omitted and the referentiality of the NP does not change, it clearly indicates that it is used as an arthroid (Rosinas 1996, 68). Moreover, unlike demonstratives, the arthroid can be used to express inferential and context-based (general knowledge-induced) definiteness. Yet another argument for separating the arthroid from true demonstratives is the fact that in recent translation studies (Vaitkutė 2017) it has been proven that while translating from languages

with fully grammaticalised definiteness marking, e.g., Swedish, NPs containing definite articles are often translated into Lithuanian as NPs with preposed demonstratives, especially often with the arthroid *tas*. And finally, its exceptionally high frequency (it occupies the fourth place in the Lithuanian word frequency list (Utka 2009, 1)) strengthens this insight.

3.6. Modifier attributes

3.6.1. Adjectival attributes in Swedish

As mentioned, Lithuanian and Swedish share a typologically rare feature, namely ADM, whereby the definiteness marker occurs on the adjectival modifier in an NP.

The question of whether this type of definiteness marker should be treated as independently performing a determiner function is often raised, most often in the case of the well-studied Scandinavian NPs (Perridon 1989; Delsing 1993; Börjars 1994, Julien 2005, Lohrmann 2011). Even today, conflicting analyses of the role of multiple definiteness exponents in an NP, the adjectival marker in particular, are proposed. As mentioned, much of the research is focused on the interplay between the proposed free and the postposed affixed definiteness articles; and much less on the role of the weak adjectival forms in an NP.

Many of the double-definiteness accounts do not consider the use of the weak (definite) adjectival forms as locus of definiteness but rather as a case of agreement (Börjars 1994; Lyons 1999). SAG claims that definite adjectival forms in most of the cases do not mark definiteness on their own (SAG, Vol 2, 220), but rather agree with the definiteness that is marked in another way elsewhere in an NP; thus allowing for cases where they could carry the [+DEF] feature autonomously (SAG, Vol 3, 15). Börjars (1994) argues that a distinction should be made between two terms, namely 'double definiteness' and 'double determination', whereby the term 'double determination' is used when both elements operate independently as semantic determiners; the term 'double definiteness' is used to denote a form of agreement. Weak adjectival forms as definiteness markers can only contribute towards double (or multiple) definiteness, but cannot function as determiners, only as agreement markers. Börjars maintains that a definite adjectival modifier does give rise to multiple definiteness, but on their own they are "unable to determine a nominal in the sense that its presence is not sufficient to allow a nominal to function as a full noun phrase" (Börjars 1994, 222).

However, in some cases the obligatory preposed article co-occurring with an adjectival modifier is omitted. Similarly to Delsing (1993, 118-119), Julien says that these occurrences are mostly restricted (Julien 2005, 30-34) to the domains of vocatives, proper names, non-referential NPs and a couple of cases with referential NPs as illustrated below:

(35) *Ta* ***stora*** ***kniven!***
take.IMP big.SG.DEF knife.SG.[COMM].DEF
'Take the big knife!'[63]

(36) *Ibland* *måste* *jag* *ha* *ringen* *på*
sometimes must I have.INF ring.SG.[COMM].DEF on
högra ***handen***.
right.SG.DEF hand.SG.[COMM].DEF
'Sometimes I must wear the ring on the right hand.' [64]

In example (35), the omission of the PDEF implies almost a deictic use, a strong familiarity with a referent. If a PDEF was inserted here, the interpretation would evoke the concept of uniqueness. The example (36) is different in that it exhibits the so-called associative anaphora; a person has one unique right hand, which, once the referent of "I" is established, is implicit.

In the cases where the definite articles PDEF and SDEF are omitted, the question may be posed what licences the definite reading in the two NPs below:

(37) *Slutar* *auktionen* *under* *detta* *ska*
end.PRS auction.SG.[COMM].DEF below this.SG.NEUTR shall
säljaren *godkänna* ***högsta***
seller.SG.[COMM].DEF approve.INF highest.SG.DEF
bud.
offer.SG.[NEUTR]
'If the auction finishes below this [amount], the seller shall accept the highest bid.'

[63] This example is from Delsing (1993, 118).

[64] Examples (36), (37) and (38) are from https://spraakbanken.gu.se/, corpora on social media texts.

(38) *människor* *som* *går* *omkring* *med*
people.PL.[COMM].INDEF who go.PRS around with
en *trumma* *runt* *halsen,*
a.ART.COMM.INDEF drum.SG.[COMM].INDEF round neck.SG.[COMM].DEF,
en *flöjt* *i* ***vänstra*** ***hand***[65]
a.ART.INDEF flute.SG.[COMM].INDEF in left.SG.DEF hand.SG.[COMM]
och *en* *bibel* *i* ***högra***
and a.ART.INDEF bible.SG.[COMM].INDEF in right.SG.DEF
hand
hand.SG.[COMM]

'people that wander round with a drum around their neck, a flute in their left hand and a bible in their right hand'

It is evident that apart from the definite adjectival form in (37) and (38), no other morphological or syntactic determiner is present; yet definite readings are obtained. Nevertheless, these two examples are different from (35) and (36) in that they resemble the cases of generic definiteness referencing the kind rather than individual objects. It must be noted that example (37) features a superlative. Example (38) features the so-called perspectival pronouns *left* and *right*. All of these fall into the category of so-called *selectors* (Dahl 2004, 153), which share common semantics, namely they are inherently definite. However, in these and similar constructions (e.g., *nedersta deck* 'bottom deck', *första pris* 'first prize', etc.) it is the AdjDEF that is used with a bare noun rather than an indefinite NP (e.g., *a highest offer*, *a left hand*, etc.) to achieve definite readings due to the concept of uniqueness (the presence of one unique referent) rather than familiarity. However, the formal marking is present as a definite adjectival form.

Börjars notes that a small set of adjectives, some of the above-mentioned selectors, seem to function as determiners themselves because a) they behave like syntactic determiners in that they can license a definite adjective, e.g., *sista misslyckade försöket* 'the last failed attempt', but **misslyckade försöket*; b) they select the same morphological marking of the head noun as syntactic determiners. Therefore they are no longer functioning as adjectives, but rather as adjectival determiners (Börjars 1994, 224-225).

[65] The use of indefinite *vänster hand* and *höger hand* would be more neutral here, yet many other examples of the kind *högra/vänstra hand*, *sida* 'side', *fot* 'foot' are to be found mostly in the spoken language domains: weblogs, online forums and social media.

Adjectival definiteness carriers differ from the determinatives described in section 4 in that that they clearly have descriptive content that modifies the noun (or a noun-like component) in an NP. Also, they can, but not necessarily do perform a determiner function, and, thus, differ from determinatives in one more respect. As already demonstrated in numerous examples above, adjectival marking of definiteness is fully compatible with other determinatives (obligatory in Scandinavian languages, and optional in Lithuanian). More details about the language-specific usage of definite adjectival forms can be found in sections 3.2.1 and 3.2.2. Regardless of whether they do or do not generate the definite reading, they do contribute to narrowing down the referential mass.

It has been shown that omissions of either PDEF or SDEF seem to be linked to definiteness due to familiarity or due to uniqueness. Also, it has been shown that a restricted set of adjectives could be considered to be determiners. Yet, the question could be posed what the function of all other adjectives marked for definiteness in Swedish NPs is. Is it merely agreement?

In an attempt to provide a unified structure analysis for Scandinavian NPs, Lohrmann argues that "the notion of definiteness in Scandinavian DPs is made up of three particular components, which are expressed by three distinct morphemes: discourse reference, identity, and specific reference" (Lohrmann 2011, 124). The suffixed definite article (SDEF) brings about specific reference. Specificity as the content of the SDEF was suggested by Julien (Julien 2005, as cited in Lohrmann 2011), which was then expanded by Lohrmann to mean *referential*, as in "denotation of N + DEF yields a referential reading and that the denotation is identifiable and locatable by the hearer" (Lohrmann 2011, 116), because, as she rightly notes, indefinite NPs can also be *specific*. If a non-specific non-referential reading is intended, the SDEF can be omitted, e.g.:

(39)	*Du*	*uppför*	*dig*	*som*	*den*	*värsta*
	you	behave.PRS	yourself	like	the.ART.COMM.DEF	worst.DEF
	buse*!*					
	toughie.SG.[COMM]					
	'You behave like the worst kind of tough guy!'					

(40)	*De*	*vackra*	*färgerna*	*lyser*
	the.ART.DEF.PL	beautiful.PL	colour.PL.[COMM].DEF	brighten.PRS
	upp	*den*	*gråaste*	***dag***.
	upp	the.ART.DEF.SG.COMM	greyest.DEF	day.SG.[COMM]

'The beautiful colours brighten up the greyest day.'

The preposed article PDEF introduces a new, modified[66] discourse referent and the adjectival inflection contributes to identifying members of the AdjDEF+ N denotation. Lohrmann claims that multiple definiteness marking in Scandinavian NPs is not a mere agreement phenomenon, but has an interpretive value. This is also valid for weak adjectival inflections. To support her claim, Lohrmann provides several examples. The one below is Swedish (Lohrmann 2011, 118):

(41)[67] a.

den	*egn**a***	*torvan*
the.DEF.ART.COMM	own.SG.DEF	garden.SG.[COMM].DEF

'one's own plot of land'

b.

hans	*egen-**ø***	*hemlighet*
his	own.SG.COMM	secret.SG.[COMM]

'his own secret'

c.

deras	*eget-**ø***	*fin**a***	*hus*
their	own.SG.NEUTR	fine.SG.DEF	house.SG.[NEUTR]

'their own fine house'

d.

hans	*egn**a***	*uppträdande*
his	peculiar.SG.DEF	behaviour.SG.[NEUTR]

'his peculiar behaviour'

Example (41a) exhibits the canonical case of a definite NP with 3 definiteness exponents. In example (41b) and (41c) *egen* follows possessives and preserves its strong form, while in (41d) it carries a long form also following a possessive. The explanation is that in (41b) and (41c) it refers directly to the possessor and hence does not need to carry additional information to identify the referent, while in (41d) it refers to the noun it

[66] See 3.2.1, this article is also called the adjectival definite article, as it is only used when an adjectival modifier is inserted in front of an N.

[67] I have encountered several examples containing both weak and strong forms of indefinite demonstrative *sådan* 'such', e.g. *det första sådana mötet* 'the first.DEF such.DEF meeting.DEF' and *mitt andra sådant lur-mejl* 'my second.DEF such.INDEF fake mail'. This requires a further analysis.

describes and helps identify the member of Adj + N denotation, contributing to the reading, namely the possessive *his* scopes over the *peculiar behaviour* (also note a slight difference in the meaning of the adjective[68]).

As has been demonstrated, the function of adjectival attributes in a definite NP, as well as their relation to the category of definiteness, is a complex one. The semantics of the category of definiteness might help to explain some of the irregular patterns (omissions or mismatches between the form and the content) in the marking of the definite Swedish NP; so would the mapping of their intrusion into the domains of the generic use and indefiniteness. Looking at the Lithuanian equivalents of phrases like *left/right hand*, *last attempt*, *first prize*, *the above-mentioned person*, etc., might yield interesting results. This is, however, not within the scope of this section.

3.6.2. Adjectival attributes in Lithuanian

In Lithuanian, the adjectival suffixed morpheme is the only dedicated marker of the definite status of an NP (more about this in section 2.2); hence the NPs containing AdjDEF modifiers are always definite, while the ones with short adjectival forms may be definite or indefinite depending on discourse, as shown in example (22). According to Sprauniené, short adjectival forms are neutral with regard to the definiteness marking of an NP (2008b, 119): AdjNONDEF = [±DEF NP].

She even provides an example where short adjectival forms serve as heads of definite elliptic NPs:

(42) *Turguje* *pirkau* ***raudoną*** *ir*
market.LOC.SG.[M] buy.1PST red.ACC.SG.F.NONDEF and
baltą *rožę.* ***Raudoną***
white.ACC.SG.F.NONDEF rose.ACC.SG.[F] red.ACC.SG.F
pasiliksiu *sau,* *o* ***baltą***
leave.1PST.REFL myself.DAT.SG. while white.ACC.SG.F
nuvešiu *mamai.*
take.1FUT mum.DAT.SG.[F]
'I bought a red and a white rose in the marketplace. The red (one) I will keep for myself, while the white (one) I will bring to my mum.'[69]

[68] It could be argued that this is a lexicalisation of some sort. Also, if in (41c) a modifier *alldeles* 'entirely' is inserted, the form *eget* would change to *egna*.SG.DEF, viz. *deras alldeles egna fina hus* 'their entirely own fine house'. This requires further analysis.

[69] This example is a modified example by Sprauniené (2008b, 118). The original example is: *Turguje pirkau raudoną ir baltą rožę. Raudoną pasisodinsiu savo*

These elliptic NPs in (42) could also be considered nominalisations, and as such they are more likely, but not necessarily, to contain definite adjectival forms. Frequently, in nominalisations long adjectival forms are often encountered in generic uses[70], referring to kinds rather than individuals, e.g., *balt**oji** meška* lit. 'white.DEF bear' meaning 'polar bear'; *raudon**oji** arbata* 'red.DEF tea' meaning 'red tea or rooibos'. The example (42) above, however, is a clear case of anaphoric use and the long adjectival forms could as well be expected and employed:

(43)	*Turguje*	*pirkau*	***raudoną***	*ir*
	market.LOC.SG.[M]	buy.1PST	red.ACC.SG.F.NONDEF	and
	baltą	*rožę.*	*Raudon**ąją***	
	white.ACC.SG.F.NONDEF	rose.ACC.SG.[F]	red.ACC.SG.F.DEF	
	pasiliksiu	*sau,*	*o*	
	leave.1PST.REFL	myself.DAT.SG.	while	
	*balt**ąją***	*nuvešiu*	*mamai.*	
	white.ACC.SG.F.DEF	take.1FUT	mum.DAT.SG.[F]	

Yet, short adjectival forms sometimes may serve in cases of anaphoric definiteness, as illustrated in (42). Such cases, nevertheless, are seldom encountered

It should be emphasized that the main function of adjectival attributes is that of modification, of assigning particular properties to the head. Yet, in Lithuanian, due to the morphological reasons, the adjectival attributes, if marked for definiteness, serve as determiners as well, thus combining the two main functions of attributes.

Lithuanian differs from Swedish in one more aspect, namely in how multiple adjectival attributes can be marked for definiteness in an NP. Apart from very rare cases (see footnote 68, for example), multiple adjectives in Swedish exhibit uniform marking for definiteness, viz. they all are consistently definite. This is not necessarily the case in Lithuanian. Examples containing multiple adjectives are few and difficult to find. Amongst the ones found in the corpus, the pattern below, where the first adjectival attribute is definite, is a predominant one:

---

*daržely**je**, o baltą nuvešiu mamai lauktuvių* 'I bought a red and a white rose in the marketplace. The red (one) I will plant in my garden, while the white (one) I will take to my mum as a gift'. It must be noted that the example is generated by the author and not found in the corpus.

[70] Examples of short adjectival forms used in generic NPs are to be found, e.g., *juoda duona* 'black bread' (bread made with flour from rye grain), *geltonas sūris* 'yellow cheese' (a type of fermented cheese, yellow in colour). The topic of nominalisations and generic definiteness is a broad one and deserves a separate study.

(44) *ši* *nauj**oji*** *svarbi* *užduotis*
this.NOM.SG.F new.NOM.SG.F.DEF important.NOM.SG.F task.NOM.SG.[F]
'this new important task'

(45) *nauj**asis*** *sukurtas* *aromatas*
new.NOM.SG.M.DEF created.NOM.SG.M fragrance.NOM.SG.[M]
'(the) new created fragrance'[71]

Other similar examples, e.g., *ši nauj**oji** dviguba žmogžudystė* 'this new.DEF double homicide', *aklojo tikr**asis** nemeluotas šauksmas* 'lit. blind's true.DEF veracious outcry', *tikr**asis** išlavintas proletariatas* 'the true.DEF educated proletariat' are to be found in the corpus. With or without a definite attribute, as is shown, the cited examples follow the pattern where the first adjectival attribute also seems to be assigning or strengthening the definiteness of the NP, while the second attribute purely modifies the head. If example (45) was to be re-written *naujas sukurt**asis** aromatas* 'new created.DEF fragrance', the NP would become indefinite with an interpretation of 'a newly created fragrance'.

Thus, in case of multiple adjectival attributes, the ones serving determiner function seem to be found on the left periphery from the head, whereas the modifiers are to be found closer to the head noun. Evidently, examples where multiple adjectival attributes present themselves in uniform usage (e.g., example (54)) are to be found, but they represent little interest for this section.

3.6.3. Genitives and possessives

I will use two terms for the possessive constructions, viz. possessives to speak about the pronoun-derived prenominal attributes like *my*, *mine*, *his*, *their*, etc. that most often are considered to serve as determiners; and genitives derived from the genitive case of nouns (or words with noun-like function), e.g., *Peter's*, *man's*, *one's*, etc. that in the case of Lithuanian (and certain Swedish genitive constructions) often serve as modifiers rather than determiners.

This is a group of prenominal attributes that can, but do not necessarily, act as determiners. Swedish and Lithuanian belong to two different groups, namely, Swedish is a DG[72] language, whereby possessives and genitives

[71] This example was found online: https://www.alio.lt/skelbimai/lancome-la-vie-est-belle-l%E2%80%98eclat-edp-75-ml---kvepalai-moterims/ID58143106.html, 2018-10-15.

[72] DG = determiner genitive; another type of language is known as 'AG' languages, where AG = adjectival genitive, e.g., Italian: *il mio sole* 'the my sun' (Lyons 2003, 24, 130-134).

occupy the position of a determinative and do not combine with the definite articles; Lithuanian is clearly not a DG language as genitives can freely combine with indefiniteness markers without changing the syntactic structure, e.g., *mano pažįstamas* 'my acquaintance' vs *vienas toks mano pažįstamas* lit. 'one such acquaintance of mine' meaning 'an acquaintance of mine', which would be the case in English and Swedish, e.g. *my friend*, but *a friend of mine*.

However, as Lyons noticed: "The traditional assumption that possessives are definite determiners, stated without further comment in many descriptive grammars and in much recent theoretical work – presumably because possession is assumed to entail definiteness – is misguided." (Lyons 2003, 24). In Swedish, the so-called *inserted genitives*, *swear genitives*, and *measure genitives* are evidently non-determiner genitives (Koptjevskaja-Tamm 2003, 515-558), e.g.:

(46)	*detta*	*tre*	***timmars***	*skriftliga*	*prov*
	this.DEM.SG.NEUTR	three	hour.GEN.PL	written.DEF	test.SG.[NEUTR]

'this three-hour long written test'[73]

Moreover, there is a difference between possessive genitives, which usually do function as determiners, and non-determiner genitives that do not necessarily do so[74].

In Swedish, as illustrated by the example below, certain types of possessive constructions have the same distribution and function in the same way as the suffixed definite article, namely they are considered to be definite attributes or true determinatives:

(47)	a.	*Jag*	*ställde*	***bilen***	*på*
		I	park.PST	car.SG.[COMM].DEF	in
		gatan.			
		street.SG.[COMM].DEF			

'I parked the car in the street.' [75]

[73] This example is from Koptjevskaja-Tamm (2003, 524).

[74] Lyons speaks of examples like *a woman's drink*, which is structurally ambiguous. Yet, in English, like in Swedish, "a possessive NP, whether itself definite or indefinite, renders its matrix noun phrase definite" (Lyons 2003, 23). Interestingly, in the Swedish corpus based on online blogs (Bloggmix 2008), the examples containing *a + possessive his* are plenty – 590, e.g., *en hans bästa kompis* lit. 'a his best buddy' in *Ringde till en hans bästa kompis* lit. 'I called a his best buddy.'

[75] This example is taken from Ekerot (2011, 7.3.4).

b.	*Jag*	*ställde*	***min***	*bil*	*på*
	I	park.PST	my	car.SG.[COMM]	in
	gatan.				
	street.SG.[COMM].DEF				
	'I parked my car in the street.'				

If no previous context is provided, example (47a) above is clearly a case of inferential definiteness, based on the assumption that the person that has parked the car is the owner of the car, as usually people own and drive cars belonging to them. The SDEF helps identify the referent – the driver and, hence, most probably the owner of the car, whereas in (47b) a possessive construction is used, directly pointing out the possessor of the car. So, in (47a), a morphological determinative is present, whereas in (47b) a lexical determinative is employed.

The same type of constructions is available in Lithuanian:

(48)	a.	*Aš*	*pa**si**stačiau*	*automobilį*	
		I	park.1PST.REFL	car.ACC.SG. [M]	
		gatvėje.			
		street.LOC.SG.F			
		'I parked the car in the street.'			
	b.	*Aš*	*pa**si**stačiau*	***savo***	*automobilį*
		I	park.1PST.REFL	my	car.ACC.SG.[M]
		gatvėje.			
		street.LOC.SG.[F]			
		'I parked my car in the street.'			

In the case of Lithuanian, the referencing mechanism is further strengthened by the presence of the reflexive verb *pasistatyti* 'to park one's car', which points to the subject, which is also cross-referenced as the owner of the car (object) by the presence of the possessive pronoun.

Possession has a direct link with the animacy hierarchy[76], as a prototypical possessor is always animate. It is evident that *Petro žiedas* 'Peter's ring' differs from *aukso žiedas* lit. 'gold's ring'. In English the latter would be replaced by an adjective *golden*. In Swedish, yet another strategy, viz. compounding, would be used to disable the possessive reading: *en guldring*.

A typologically interesting case in both Swedish and Lithuanian is the insertion of an adjectival modifier into an NP with a possessive:

[76] According to Croft, the *Extended Animacy Hierarchy* looks like this: 1st/2nd person pronouns < 3rd person pronoun < proper names < human common noun < nonhuman animate common noun < inanimate common noun. (Croft, 2004, 130).

(49) *Adams* *första* *bok*
Adam.GEN.SG first.DEF book.SG.[COMM.]
'Adam's first book'

(50) a. *Adomo* *pirm**oji*** *knyga*
Adam.GEN.SG.[M] first.NOM.SG.F.DEF boook.NOM.SG.[F]
'Adam's first book'

b. *pirm**oji*** *Adomo* *knyga*
first.NOM.SG.F.DEF Adam.GEN.SG.[M] boook.NOM.SG.[F]
'Adam's first book'

Even though on the surface the examples like (50a) *Adomo pirm**oji** knyga* 'Adam's first.DEF book' and (50b) *pirm**oji** Adomo knyga* 'first.DEF Adam's book' appear to exhibit little difference in terms of semantics, the organisation of the prenominal modifiers, implying different readings of the definiteness of the NP, plays a role. (50b), *pirmoji [Adomo knyga]*, seems to imply that there exists a category of books written by/belonging to Adomas, and of them is first (e.g., was written first, was acquired first, etc.). (50a), *Adomo [pirmoji knyga]*, seems to imply that there exists a category of *first books*, e.g., a writing competition where beginning writers are asked to submit their first books, and one of them is authored by Adomas. That is, the example (50a) seems to introduce the category of *first books*, whereas (50b) introduces the category of *Adam's books*. The reading of (50a) almost implies a previous mentioning of *first book*, making this phrase anaphoric-like in terms of the usage of the definite adjectival form, while (50b) implies that it is probable that there were other *books belonging to Adam*. Vaičiulytė-Semėnienė (2006, 165, footnote 20), provides a similar analysis of the phrases *mano nauj**oji** suknelė* (my new.DEF dress) versus *nauj**oji** mano suknelė* (new.DEF my dress). This only strengthens the assumption that multiple definiteness carriers mirror the multi-layered—in terms of definiteness—structure of the NP.

It has to be noted that the following example shows yet another possibility of expression in Swedish:

(51) *det* *första* *försöket* *mitt*
the.ART.NEUTR.DEF first.SG.DEF attempt.SG.[NEUTR].DEF my.GEN.SG.NEUTR
'the first attempt (of) mine'

Here the possessive genitive is placed postnominally; hence the slot in the determiner position is occupied by the PDEF, followed by the AdjDEF and SDEF on the noun.

Taking into account the analysis of the Swedish adjectival modifiers in 3.6.1, as well as Vaičiulytė-Semėnienė's notion of a multi-layered reference-assignment (Vaičiulytė-Semėnienė 2006, 162–163), with the leftmost

exponent of an NP providing discourse definiteness and those on the right end of the axis closest to the head helping to identify the member(s) of the subset denoted by the modifier + N, we can conclude that the linear positioning of definiteness markers follows some sort of a definiteness hierarchy where the exponents on the left reflect the discourse-bound definiteness, usually based on familiarity, and those closest to the head have identifiability-based definiteness, associated with the concept of uniqueness.

In conclusion, we may note that there seem to be four major differences in the NP structure between Swedish and Lithuanian, once a genitive construction is introduced into an NP:

1. There are only two ways to express canonical possession, that is, through the use of possessive pronouns and proper nouns or through animate common nouns, since prototypically the semantics of possession require an animate possessor. The constructions with genitives of inanimate common nouns do not express semantic possession, but rather different types of relations, e.g., proprietorship, e.g., *valstybės miškas* lit. 'state's forest' meaning 'state-owned forest', origin/material, e.g., *aukso žiedas* lit. 'gold's ring' meaning 'golden ring', and other non-anchoring relations.
2. If a possessive pronoun is used, both languages allow two alternatives of expression: 1) a possessive occupying the first slot in the NP and functioning as a determiner (in this case in Swedish, SDEF on the noun is omitted, see (49)); 2) if a possessive genitive is occupying a slot that is not the first in the NP, then the determiner slot retains [+Def] marking either on the adjective (in Lithuanian (50a)) or by introducing both PDEF and SDEF in Swedish (AdjDEF is always maintained, see (51)).
3. If a proper noun or an animate common noun is used, both Swedish and Lithuanian have rigid structures in which the possessive occupies the first available position and cannot be moved in Swedish (the alternative in (48b) is impossible), whereas in Lithuanian the genitive may remain attached to the noun and the first slot is occupied by AdjDEF.
4. In Lithuanian non-determiner genitives are common; they do not express possession and hence function as modifiers. Even though genitive constructions are used, entirely different structures are employed in Swedish to reflect this, namely compounding and the use of a periphrastic prepositional construction. However, here again in Lithuanian, the genitive remains next to the noun and the first slot is occupied by AdjDEF. One could argue that two variants are available here as well, e.g.: *sen**osios** buto durys* vs *buto sen**osios** durys* 'the old.DEF doors to the apartment' where the genitive *buto* 'apartment' could be placed in front of the adjective, however, this structure should be

phonologically marked (stressed) with *buto* stressed and *old* carrying the contrastive stress and implying that there are/were more than one door to the apartment, as opposed to the one with genitive remaining close to the N, which could be viewed as neutral, e.g.:

(52) O *kur* *buto* *sen**osios*** *durys?*
So where apartment.GEN.SG.[M] old.NOM.PL.F.DEF doors.NOM.PL.[F]
'So where is the old door to the apartment?'

3.7. Topology of the definite Lithuanian NP

3.7.1. Possible formulas for the linear NP structure

As mentioned, the aim of this section is to examine and describe the linear positional structure of the Lithuanian NPs, a sequence of elements lining up on the left to right axis with the head (in Lithuanian terminology also known as *nucleus*) being the rightmost element.

In a well-known function-oriented Lithuanian grammar (Valeckienė 1998, 118–130), the following structure for the NP is suggested (parentheses signal that these elements are optional):

NP= (modifiers) + nucleus +(qualifiers)

Speaking of the modifiers that predominantly appear in the prenominal position, Valeckienė notices that these include lexical classes of words, mostly adjectives and participles, and, to some extent, numerals and pronouns. Modifiers show agreement with the head, whereas qualifiers are governed by the head (Valeckienė 1998, 121). This analysis does not provide a comprehensive description of the line-up of the elements preceding the head.

In the CG, the formula below was suggested to describe the organisation of English nominals (Langacker 2008, 312):

NP = [Grounding [(Modifiers) [Head Noun] (Modifiers)]]

The term *nominal* in CG refers to the term *NP* as used in this thesis. It is not synonymous with CGEL's *nominal*, an interim category between a noun and an NP (CGEL, 329), as illustrated in Table 4. I find it useful in that that it makes a distinction between the grounding elements and modifiers. It also explains why some elements that in grammars are often classified as determinatives, e.g., ordinal numbers, are not functioning as definiteness carriers. CG refers to these as nongrounding quantifiers. Speaking about the linear placement of the elements, Langacker says that in many languages, a

grounding element is generally the one occupying the leftmost position in the structure of a nominal. As mentioned in the section on quantifiers 3.5.2, two separate positions for quantifiers (universal or grounding) and others should be recognised. Moreover, the grounding elements provide the least information about the referent per se; they indicate the discourse status of a referent (Langacker 2008, 275).

From SAG (SAG, Vol 3, 5), the subsequent formula can be retrieved:

NP = ((Attr)N / N(Attr)) / ((Attr) +N + (Attr))

All of these formulas show great similarity. However, as the focus of this section is the prenominal attributes, only the *NP = (Attr)N* will be described here.

In his thorough study of the Swedish NP in relation to reference and definiteness, Perridon suggests the field model below (based on the works of Diderichsen) to describe the linear structure of an NP; it contains only three prenominal slots in an NP in modern Swedish (1989, 201)[77]:

(Attr)N = Field of Determination (Det) + Field of Quantification (Qu) + Field of Description (Descr)[78]+Nucleus

As previously mentioned, the number of the fields identified by Perridon (and SAG, Vol 3, 13) needs to be further expanded to include the two distinctive fields of quantification, namely that for the universal quantifiers, Q_1, and that for the cardinal quantifiers, Q_2. Also, the field of determination needs to be further specified to reflect the difference between various types of determiner modifiers, to include positions D_1 for definite articles or demonstratives, D_2 for adjectival modifiers, D_3 for genitives, etc. A field M for modifier attributes (should an NP contain multiple modifiers, the number of M positions could be increased, viz. M_1, M_2, etc.) should be next to the head. Also, an additional field, called *peripheral modifier*, PM, needs to be established, to include "external modifiers occurring at the periphery of the NP, mainly in initial position […]" (CGEL, 436), e.g., to accommodate prenominal elements like *even*, *only*, *too*, *such*, etc.

[77] GDS offers yet another, but similar outline, namely 3 slots: 1) the framing slot (DA *rammeplads*), 2) the determiner slot (DA *bestemmer*); 3) the description slot (DA *beskriver*) (GDS, Vol 3, 478).

[78] "Each of these fields may contain syntagms with an internal structure of their own." (Perridon 1989, 201).

The two examples below prove the potential complexity of the definite NPs with its various prenominal modifiers, modifying the Swedish head *books* and the Lithuanian one *powers*:

(53) *alla* *dessa* *hans* *många* *andra* *sådana*
all.PL these.PL he.GEN many other.PL such.PL
danska *böcker*
Danish.PL book.PL.[COMM]
'all those his many other such Danish books'[79]

(54) *tos* *abi* *slėpingosios* *ir*
those.NOM.PL.F both.NOM.PL.F mysterious.NOM.PL.F.DEF and
galingosios *žmogaus* *egzistencijos*
prevailing.NOM.PL.F.DEF human.GEN.SG.[M] existence.GEN.SG.[F]
jėgos
powers.NOM.PL.[F]
'those both mysterious and prevailing powers of human existence'

3.7.2. Topology of the definite Lithuanian NP

Based on the previous sections, the following positions in the linear structure of the Lithuanian NP have been identified:

1) a peripheral modifier – PM;
2) a universal (grounding) quantifier – Q_1;
3) a first determiner – D_1, containing either a demonstrative (in languages with articles, like Swedish and English, this would be occupied by the definite article in absence of a Dem);
4) a cardinal (nongrounding) quantifier – Q_2;
5) a second determiner – D_2, containing an adjectival modifier with a determiner function;
6) a third determiner – D_3, containing a possessive or a determiner-genitive;
7) a modifier – M, containing modifiers not marked for definiteness (this position is needed to reflect the cases where several adjectival attributes are used in the same NP with different definiteness values).

Reflecting the analysis presented in this section, the positional structure of the definite Lithuanian NP could best be described in the example below:

[79] This example is borrowed from Perridon (1989, 183), who in his turn has borrowed it from Loman (1956).

Table 6. The structure of the definite Lithuanian NP

Abbr.	**PM**	**Q_1**	**D_1**	**Q_2**	**D_2**	**D_3**	**M**	**Head**
Description	*peripheral modifier*	*universal quantifier*	*demonstrative/ definite article*	*cardinal quantifier/ multal*	*AdjDEF*	*possessive/ determiner-genitive*	*modifier*	*noun/ noun-like word*
Example	net	visos	tos	trys	saldžiosios	tėvo	žieminės	kriaušės
Gloss	even	all.NOM.PL.F	that.NOM.PL.F	three	sweet.NOM.PL.F.DEF	father.GEN.SG.[M]	winter.NOM.PL.[F]	pear.NOM.PL.[F]
Translation	lit. even all those three sweet father's winter pears [did not bring her any pleasure] i.e., even all those three sweet winter pears of her father's [did not bring her any pleasure]							

The following rules can describe variation in slot occupancy that is attested:

1. The very first slot in the definite NP is occupied by a peripheral modifier. In its absence, it is the universal quantifiers that take up the first slot.
2. Universal quantifiers are the ones that can occupy the Q_1 slot. They can be, but most often are not preceded by other determiners, unless they move to Q_2, as they can freely move between the slots Q_1 and Q_2. Yet, Q_2 is the customary host for other quantifiers, viz. cardinal numbers and multal quantifiers, e.g., *many*, *three*, etc. By contrast, movement in the opposite direction from Q_2 to Q_1, if Q_1 is occupied by a universal quantifier, is impossible, e.g., *visi tie trys berniukai* 'all the three boys' versus **trys tie visi berniukai* 'three the all boys'. However, if a universal quantifier is absent, the traditional occupant of slot Q_2 can move up to Q_1, e.g., *trys šie berniukai* lit. 'three these boys' cf. *šie trys berniukai* 'these three boys' in both Lithuanian and Swedish. If an NP contains an adjective, the adjective will follow both Q_1 and Q_2, e.g., *visi šie trys gerieji vaikai* 'all these three kind children', *trys šie gerieji vaikai* lit. 'three these kind children', *šie trys gerieji vaikai* 'these three kind children', etc. in both Lithuanian and Swedish. A peculiar case in Lithuanian is that of constructions like *tie visi trys komponentai* lit. 'those all three components', where we could claim that the universal quantifier *all* has moved into the Q_2 slot to fuse with the cardinal *three*, in a manner similar to the above-mentioned dual demonstratives in 3.5.2. The demonstrative *those* remains in the D_1 slot, not preceded by any other elements (see conclusion No 4 below).
3. NPs containing possessives (or determiner-genitives) exhibit the most variation between the languages under comparison. Also, their analysis offers the most complexities due to the interplay between animacy and possession. With possessives high in animacy, Lithuanian allows variation in the placement of adjectival modifiers, as described in section 3.6.3., viz. they may precede or follow the possessive as opposed to Swedish, which offers only one possibility: the adjectival modifier must directly follow the genitive.

4. If an NP starts with slot D_1 and it is occupied by a demonstrative, adjectives can only follow the occupants of D_1. The reverse order is impossible in both Lithuanian and Swedish.[80]
5. Once both Q_1 and Q_2 are occupied, all the other determinatives and determiner modifiers can only stand between these and the head of an NP, regardless of whether the D_1 slot between Q_1 and Q_2 is occupied or not. Starting from D_2 and moving in the direction of the head, the number of slots D may increase but is subject to limitations if the process of referent identification is to be successful.
6. The slot closest to the head is M, containing modifiers – adjectival and non-determiner genitive attributes. The number of M may also be more than one. If a Lithuanian NP contains two or more adjectival attributes with different definiteness-marking status, viz. both long and short forms, the short ones will be found closer to the head than the long ones.
7. Yet, it appears that this slot M may also be occupied by an adjectival modifier of the classifying kind (including the non-determiner genitive attributes, mentioned in 6.), both on the level of generic reference, e.g., Table 6. *žieminės kriaušės* 'winter pears', and on that of context-based *ad hoc* categorisation, as in (50a) *pirmoji knyga* 'first book'. This implies that this slot as well may host a determiner acknowledging the fact that adjectives having the ability to establish *ad hoc* or more permanent categories (generic nominals) are definiteness carriers (see 4.2.1 and 4.2.2 for detailed discussion).

The table below reflects the variation in slot occupancy in a definite Lithuanian NP:

[80] A peculiar case of a demonstrative (recognitional use) that is placed between an adjective and a noun in Lithuanian is illustrated in footnote 53.

Table 7. Variation in the slot occupancy of the definite Lithuanian NP

PM	(Q_1)	D_1	(Q_2)	D_2	$D_{3(D2+1)}$	$D_{3((D2+1)+1)}$	$D_{4(((D2+1)+1)+1)))}$	$D_{x((((D2+1)+1)+1)+1+...))))}$	M	Head
										vaikai ‘children’
									išdykę ‘mischievous’	*vaikai* ‘children’
								mažieji ‘little.DEF’	*išdykę* ‘mischievous’	*vaikai* ‘children’
							mažieji ‘little.DEF’	*draugės* ‘friend’s’	*išdykę* ‘mischievous’	*vaikai* ‘children’
						tie ‘that/the’	*mažieji* ‘little.DEF’	*draugės* ‘friend’s’	*išdykę* ‘mischievous’	*vaikai* ‘children’
					tie ‘that/the’	*mažieji* ‘little.DEF’	*mano* ‘my’	*draugės* ‘friend’s’	*išdykę* ‘mischievous’	*vaikai* ‘children’
				mažieji ‘little.DEF’	*mano* ‘my’	*ištikimosios* ‘loyal.DEF’	*vaikystės* ‘childhood’	*draugės* ‘friend’s’	*išdykę* ‘mischievous’	*vaikai* ‘children’
			abu ‘both’	*mažieji* ‘little.DEF’	*mano* ‘my’	*ištikimosios* ‘loyal.DEF’	*vaikystės* ‘childhood’	*draugės* ‘friend’s’	*išdykę* ‘mischievous’	*vaikai* ‘children’
		tie ‘those/the’	*abu* ‘both’	*mažieji* ‘little.DEF’	*mano* ‘my’	*ištikimosios* ‘loyal.DEF’	*vaikystės* ‘childhood’	*draugės* ‘friend’s’	*išdykę* ‘mischievous’	*vaikai* ‘children’
	visi ‘all’	*šie* ‘these’	*trys* ‘three’	*mažieji* ‘little.DEF’	*mano* ‘my’	*ištikimosios* ‘loyal.DEF’	*vaikystės* ‘childhood’	*draugės* ‘friend’s’	*išdykę* ‘mischievous’	*vaikai* ‘children’
tik ‘only’	*visi* ‘all’	*šie* ‘these’	*trys* ‘three’	*mažieji* ‘little.DEF’	*mano* ‘my’	*ištikimosios* ‘loyal.DEF’	*vaikystės* ‘childhood’	*draugės* ‘friend’s’	*išdykę* ‘mischievous’	*vaikai* ‘children’

3.8. Conclusions

1. The structural pattern described in section 3.7.2 is obviously a simplified attempt at providing a formula that would capture the basic topology of the definite Lithuanian NP. A more detailed and fine-grained analysis needs to be carried out.
2. The cross-linguistic approach has been fruitful in that it has enabled me to identify and outline the potential prenominal attribute positions on a linear axis moving from left to right, from the periphery to the head of an NP.
3. It is the category of definiteness manifesting itself through multiple exponents that dictates the outline of the prenominal attributes and therefore the reading of the NP.
4. Even if Lithuanian NP structure exhibits more freedom of variation amongst the various slots in an NP, there are certain limitations and certain preferences clearly predominate.
5. Definite adjectival modifiers can function as determiners in both Lithuanian and Swedish.
6. Other definite attributes and quantifiers can also assume the function of determiners in an NP. Their *modus operandi* in the reference-assigning process is very similar; and, also, syntactically they behave similarly.
7. It is through establishing the positions of the most common determiners, quantifiers, and modifiers in an NP that the structure of an NP can be described in its entirety.
8. The structure of a definite Lithuanian NP is a multi-layered structure where all the prenominal determiners contribute to the definiteness reading of the NP, with those on the left carrying the most powerful load and having the broadest scope of action in the process of discourse building, and those closest to the noun contributing to the ultimate identification of the referent. In other words, the feature [+Def] may and often is encoded in several loci with different degrees of impact where the elements in the left periphery encode discourse-bound definiteness while elements in the right (closer to the head) denote identifiability-based definiteness.
9. Moreover, it has to be noted that the structure of the definite Lithuanian NP perfectly reflects the Prepositional Noun Modifier Hierarchy (Croft, 2004, 122), with the exception of NRel:
NNum > NDem > NA > NG > [NRel]
10. The functional differences between the two groups of prenominal attributes (referred to in this section as definite attributes (determiners) and modifiers), show two different types of definiteness marking, strong and weak, associated either with familiarity or uniqueness, discourse, and specificity (referentiality). These differences stand in need of further investigation.

4. ADJECTIVAL DEFINITENESS MARKING IN LITHUANIAN – ONE MORE PUZZLE PIECE: QUALITATIVE ADJECTIVES THAT COULD BUT DO NOT TAKE DEFINITE FORMS[**]

4.1. Introduction

4.1.1. The data puzzle

Both section 4 and section 5 of this dissertation examine the use and distribution of long and short forms in contemporary Lithuanian in an attempt to provide to provide tangible quantified data shedding some light on the actual use and distribution of long and short adjectival forms in Lithuanian. The following Section 4 will address the issue of the evident absence of long adjectival forms within a large group of qualifying adjectives that could but do not take definite forms.

It has been argued that long adjectival forms (LFs) always encode definiteness (Sprauniene 2011, 74–76), both on the level of individual reference, e.g., *balt**asis** katinas* 'the white.DEF cat' and on that of generic use, referencing a kind rather than individual objects, e.g., *balt**asis** lokys* literally 'the white.DEF bear = polar bear'. The use of both *baltas katinas* 'a/the white.NONDEF cat' and *balt**asis** katinas* 'the white.DEF cat' is attested and frequent in both written and spoken Lithuanian. While *balt**asis** katinas* will always get a definite reading, *baltas katinas* may or may not get a definite reading, depending on the context. Nevertheless, there appears to be a group of adjectives that, even though they may in principle assume definite forms, never or seldom do so in the contemporary Lithuanian language (see Table 8 for zero counts, as well as very low counts of long adjectival forms), e.g., *įvairus* 'various, varied, diverse', *panašus* 'similar, alike, analogous, resemblant', *skirtingas* 'different, separate, distinct', *nemažas* 'considerable, not small', *menkas* 'insignificant, meagre, poor', *reikalingas* 'necessary, needed, required', *optimalus* 'optimal, optimum, superb', *gausus* 'abundant, ample, bountiful', *aiškus* 'apparent, evident, transparent', *švarus* 'clean, pure' and others, e.g.:

(55)	a.	*įvairus*	*maistas*	vs
		diverse.NOM.SG.M.NONDEF	food.NOM.SG.[M]	
		'a varied diet'		

[**]This section, with minor updates and modifications, is based on this article: Trakymaitė R. (2021). "Adjectival definiteness marking in Lithuanian – one more puzzle piece: Qualitative adjectives that could but do not take definite forms", Baltistica, Vol. 56. No. 1 (2021), 19–79.

	b.	*įvair***usis**[81]	*maistas*	vs
		diverse.NOM.SG.M.DEF	food.NOM.SG.[M]	
		'the diverse diet'		
(56)	a.	*menka*	*nauda*	vs
		meagre.NOM.SG.F.NONDEF	benefit.NOM.SG.[F]	
		'a meagre benefit'		
	b.	**menk**oji***	*nauda*	
		meagre.NOM.SG.F.DEF	benefit.NOM.SG.[F]	
		'the meagre benefit'		

Why do they exhibit this particular behaviour? Do they share other characteristics that allow them to be assigned to a particular group/class of adjectives? Is it a homogeneous group? Does this behaviour signal their peculiar relationship with (in)definiteness? In this section I will suggest that the absence of long adjectival forms in NPs is due to several reasons. For some adjectives, it is their semantic-pragmatic properties that account for the absence of long forms, while for some others, it is their properties, similar to those of determiners and quantifiers, that disable the use of long forms.

This section is structured as follows. In section 4.1, the background, data, and method are introduced. Section 4.2 contains a discussion on two types of Lithuanian adjectives, relational and qualifying, and the implications of these for the paradigms of LF. Section 4.3 is devoted to the Lithuanian definiteness-marking system and the role which the short adjectival modifiers (SF) play in the structure and the interpretation of a noun phrase (NP). Section 4.4 contains the analysis of the data. Both common properties shared by all adjectives not used in LFs and distinctive properties of individual sub-groups of the selected adjectives are examined, and a classification is proposed based on their semantic-pragmatic and functional properties, following the results of the qualitative analysis. Where relevant, for illustrative purposes, examples of other languages having paradigms of strong and weak adjectival forms, e.g., Swedish, will be given to show the cross-linguistic nature of the phenomenon. Finally, section 4.5 contains some concluding remarks and prospects for future research.

4.1.2. Background

In trying to establish the extent of the use of LFs in contemporary Lithuanian, I studied the data of the Frequency Dictionary of the Written Lithuanian

[81] Both in (55b) and (56b), the definite forms are attested, yet the NPs in the given examples are not possible.

Language[82] (Utka, 2009) and compiled a list of the most frequently used adjectives that can take LFs. Under every individual entry published in the Dictionary (also referred to as FrD in this thesis), an inventory of paradigm forms with usage frequencies is displayed in descending order. This enabled me to compile a frequency list of the Lithuanian adjectives that can have a paradigm of definite forms. Since the actual usage counts of each paradigmatic form (both long and short) are displayed under each individual entry of an adjective, I was able to calculate and compare the percentage of LF versus SF used for each adjectival entry. Among the 111[83] most frequently used adjectives that can have definite forms, 43 had between 0 and 2% of long forms (see Appendix A for the original counts of long forms in FrD, as well as comments on individual entries). In view of the fact that the Dictionary was compiled on the basis of just 1 million morphologically annotated words, I then proceeded to run checks for the singled-out adjectives in the Corpus of the Contemporary Lithuanian Language (CCLL)[84]. In the process it became evident that amongst the 111 most frequently used adjectives, approximately one-third do not appear in their LFs at all (0 instances) or have very few instances of LFs (less than 1%). This was a significant finding. Much has been written on the use of the long adjectival forms as definiteness markers, yet very little data is available[85] on the absence of long adjectival forms where they are to be expected. The absence of a grammatical phenomenon is as important as its presence.

The aim of this data-driven section is to closely examine the list of Lithuanian adjectives that do not show long forms, though these could in principle be derived, and to seek an explanation for this phenomenon.

4.1.3. Data and method

In the Frequency Dictionary, having identified adjectives that take less than 2% of LFs (all these cases included actual counts between 0 and 2 of long

[82] It is available online at http://donelaitis.vdu.lt/publikacijos/Dazninis_zodynas.pdf.

[83] I originally intended to compile a list of the 100 most frequently used adjectives that take long forms, but I realised that adjectives following the 100th example, *sausas* 'dry', display very similar statistics (similar use counts, similar numbers of long/short forms, close to each other in sequence on the list, etc.). Therefore, I decided to include 11 more.

[84] Accessible online: http://tekstynas.vdu.lt/tekstynas/.

[85] An article by Šereikaitė on strong and weak definites in Lithuanian slightly touches upon this question (Šereikaitė 2019).

forms), I ran checks on these adjectives in the CCLL. I searched for the paradigms of both short and long forms and then counted the percentage of long forms. A few adjectives, e.g., *sunkus* 'heavy, difficult, hard', *puikus* 'great, excellent, fine', *ramus* 'calm, peaceful, tranquil', etc., were removed from the final shortlist because they had a higher count of LFs in the corpus than in the Frequency Dictionary (above 1%, see Appendix A for statistics and comments on individual adjectives).

While working with the data from the CCLL, the following three major issues had to be dealt with:

1) All the searches had to be done manually, for each form of both short and long paradigms identified; the data extracted from the CCLL was not annotated, hence it contained a high number of homonyms in the paradigms of SFs, e.g., the results for the short form paradigm of the adjective *vertas* 'worth/worthy/deserving/valuable' (total count 36163), contained the following homonyms (with different stress patterns when pronounced):

- *vertai* (ADVB) ≠ *vertai* (ADJ, DAT.SG.F) – 77 instances;
- *verta* (ADJ, NEUTER) ≠ *verta* (ADJ, NOM.SG.F) – 9788 instances;
- *vertus* (VERB, GER.PST) ≠ *vertus* (ADJ, ACC.PL.M) – 19441 instances.

Not all the adjectives had so many homonyms as the examples above. However, most of them did have an adverbial form homonymous with the dative singular feminine.

2) Since it was physically impossible to fine-tune data because of the high numbers and lack of annotation, I chose to remove the counts of homonymous forms from the short form paradigms, e.g., the adjusted overall number of instances of the adjective *vertas* (above) was 6857. As the goal was to identify the percentage of LFs used, the logic behind removing the counts of homonyms that could not be dealt with manually due to high counts was that it would potentially increase the percentage of LFs (as the total count of short ones would decrease, the total count of the long ones would automatically increase, increasing the chance of them being removed from the list of the atypically behaving adjectives, viz., not assuming LFs while being capable). Therefore, those adjectives that still showed a very low percentage of LFs would be of significance.

3) With the aim of assessing the paradigms of LFs, lists of collocations of each lexeme had to be produced in order to eliminate the counts of nominalisations and terminology-based uses. It is well-known from the literature that the long forms of qualitative adjectives act as noun-replacements in nominalisations (Mikulskas 2006, 59-60) and are also used in terminology, viz., in NPs containing modifying adjectives, e.g., in linguistics, *konkret**ieji** daiktavardžiai*

‘concrete.DEF nouns’, in maths, *normal**usis** skirstinys* ‘normal.DEF distribution’ or in geometry, *panaš**ieji** trikampiai* ‘similar.DEF triangles’ (more about these phenomena in 4.2.3, see Appendix B for a sample of collocation lists).

To sum up, two types of data manipulation were carried out, viz., removal of homonymous forms from the paradigms of SFs; and removal of the counts of nominalisations and terminology-based uses from the counts of the paradigms of LFs.

Finally, two additional adjectives, even though they had more than 2% of long forms in the original FrD list, were included, viz., *individualus* ‘individual/separate/distinctive/special’ (total count in FrD – 5 or 3.16% of long forms) and *paprastas* ‘simple/ordinary/normal/average’ (the total FrD – 17 or 6.42% of long forms). This was done because of: a) their relatively high ranks in FrD; b) their semantic similarity to other adjectives in the list, e.g., *individualus* ‘individual/distinctive’ ≈ *atskiras* ‘separate/distinct/individual’, *paprastas* ‘simple/ordinary’ ≈ *normalus* ‘normal’/*vidutinis* ‘average’; c) an intuition that the higher count of LFs in FrD would be the result of their frequent use in terminology. Also, similar to a few other adjectives on the shortlist that included antonym pairs, e.g., *menkas* ‘meagre’ vs *gausus* ‘abundant’, *panašus* ‘similar’ vs *skirtingas* ‘different’, *paprastas* ‘simple’ would pair up with *sudėtingas* ‘complex’.

The final empirically observed patterns of the FrD and CCLL combined are presented in this table, arranged alphabetically:

Table 8. The alphabetical shortlist of adjectives with less than 1% of use of long forms

No	Adjective	Translation	FrD – count of long	CCLL – count of short (adjusted)	CCLL – count of long (raw)	CCLL – count of long (adjusted)	CCLL - % of long
1	aiškus	clear/understandable/explicit/evident	0	14828	27	5	0.03
2	atskiras	sundry/separate/individual/special/ distinct/detached	2	31971	695	20	0.06
3	būdingas	typical/characteristic/specific	2	23616	74	23	0.10
4	dažnas	frequent/habitual/periodic/repeated	0	6539	7	3	0.05
5	galutinis	final/ultimate/terminal	0	10964	4	4	0.04
6	gausus	abundant/numerous/plentiful/ample/ bountiful	0	8190	8	8	0.10
7	įdomus	interesting/exciting/entertaining	0	15919	70	23	0.14
8	individualus	individual/separate/distinctive/special	5	15169	647	69	0.44

No	Adjective	Translation	FrD – count of long	CCLL – count of short (adjusted)	CCLL – count of long (raw)	CCLL – count of long (adjusted)	CCLL - % of long
9	įvairus	various/varied/miscellaneous	0	83278	17	15	0.02
10	konkretus	concrete/particular/specific	0	26824	127	36	0.13
11	menkas	meagre/insignificant/poor	0	7135	10	8	0.11
12	neaiškus[86]	unclear/uncertain/obscure/vague/ indistinct	0	5476	9	2	0.04
13	nemažas	considerable/not small	0	12852	0	0	0.00
14	normalus	normal/regular/ordinary/average	0	11174	120	16	0.14
15	optimalus	optimal/optimum/superb/top-notch	0	2585	4	3	0.12

[86] A reasonable question could be posed whether adjectives with a negation prefix *ne-* as in ***ne****aiškus* could form a distinct group (see also the following example of ***ne****mažas*) of adjectives that do not take long forms due to their specific derivational and morphological characteristics. Yet, a quick search in ltTenTen14 yields a number of examples like *neskan**ieji** sausainiai* 'tasteless.DEF cookies', *nemalon**usis** įspūdis* 'unpleasant.DEF impression', *nemiel**asis** Artūras* 'unattractive.DEF Arthur', etc.

No	Adjective	Translation	FrD – count of long	CCLL – count of short (adjusted)	CCLL – count of long (raw)	CCLL – count of long (adjusted)	CCLL - % of long
16	panašus	similar/like/alike/analogous/resemblant	0	48269	7	1	0.002
17	paprastas	simple/ordinary/normal/average	17	13730	1490	100	0.72
18	patogus	convenient/comfortable/handy	0	3977	7	2	0.05
19	privatus	private/personal/own/proprietary/individual	1	21997	126	25	0.12
20	reikalingas	needed/required/necessary/requisite	0	25876	17	7	0.03
21	reikšmingas	significant/meaningful/important/weighty	0	8238	40	40	0.48
22	ryškus	bright/stark	1	6778	44	38	0.56
23	skirtingas	different/unlike/separate/distinct/diverse	0	30471	4	1	0.003
24	sudėtingas	complex/complicated/multiplex/elaborate	0	11123	39	20	0.18

No	Adjective	Translation	FrD – count of long	CCLL – count of short (adjusted)	CCLL – count of long (raw)	CCLL – count of long (adjusted)	CCLL - % of long
25	švarus	clean/pure/clear/fresh/immaculate	0	6358	59	41	0.64
26	tolesnis	further/subsequent/successive	1	10592	14	14	0.13
27	vertas	worth/worthy/deserving/valuable	0	6857	27	2	0.03
28	vidutinis	average/medium/middle/moderate/normal	2	22517	175	21	0.09
29	vienodas	uniform/equal/same/homogeneous/like	0	9551	0	0	0.00
30	visiškas	complete/total/full/absolute/superior/superb	0	9677	82	3	0.03

As can be seen, all the adjectives in the table show less than 1% use of LFs. Their values differ between 0% (the lowest) and 0.73% (the highest value). This is statistically significant.

4.2. Lithuanian adjectives and their paradigms of long forms

As mentioned in 1.1, only qualitative adjectives have paradigms of definite forms. This fact deserves a few remarks, as I believe this difference is essential in understanding the use of long adjectival forms in Lithuanian.

4.2.1. Relation versus property

In descriptive grammars of Lithuanian, adjectives are defined as a separate part of speech consisting of two different types, viz., *qualitative* adjectives and *relational* adjectives. While both groups denote properties, the distinction between the two is based on semantic and morphological differences. The qualitative adjectives denote properties "directly by their lexical meaning" (Ambrazas et al. 2006, 134), while the relational adjectives denote properties arising "through their relation to a basic word" (ibid.). In other words, relational adjectives express relation to corresponding nouns. Morphologically, this makes relational adjectives mostly derivatives, while qualitative adjectives are primary words (root-based):

(57)	a.	*balt-as*	*balt-a*		
		white.NOM.SG.M	white.NOM.SG.F		
		'white'			
	b.	*auks-inis*	*auks-inė*	cf.	*auks-as*
		golden.NOM.SG.M	golden.NOM.SG.F		gold.NOM.SG.[M]
		'golden'			'gold'

Relational adjectives denote a property arising in relation to another object or occurrence, as illustrated above in *auksinis* → 'of gold' (Paulauskienė 1994, 175). They often indicate the material from which the modified object is made, or the purpose of the object, the suitability of the object, etc. They denote objective reality-based qualities that do not change, e.g., *beržinės malkos* 'birch firewood', *keramikinės plytelės* 'ceramic tiles', *pernykščiai obuoliai* 'last year's apples'. Most relational adjectives are formed with the aid of the suffix *-inis*, which is a very productive pattern to replace

the use of the non-determiner genitives[87] (Kniūkšta 1976: 3) with adjectives:

(58)	*medžio*	*stalas*	cf.
	wood.GEN.SG.[M]	table.NOM.SG.[M]	
	med-inis	*stalas*	
	wooden.NOM.SG.M	table.NOM.SG.[M]	
	‘table (made) of wood’ → ‘wooden table’		
	[N → Ajd]		
(59)	*darbo*	*drabužiai*	cf.
	work.GEN.SG.[M]	clothing.NOM.PL.[M]	
	darbiniai	*drabužiai*	
	work.NOM.PL.M	clothing.NOM.PL.[M]	
	‘work (N) clothing’ → ‘work (Adj) clothing’		
	[N → Adj]		

To sum up, as the name implies, relational adjectives characterise a relation; and through it they imply an association with classes of objects, e.g., *medinis* ‘wooden’ → belonging to a class of objects made of wood; *mokyklinis* ‘school’ (e.g., *mokyklinis autobusas* ‘school bus’) → belonging to a class of objects related to school, etc. This explains the lack of gradation and scalarity. Also, once attributed to a certain class denoted by a relational adjective, an object acquires a permanent property, e.g., work clothing (*darbiniai drabužiai*) always refers to a specific class/type of clothing worn for work as opposed to, e.g., party wear (*šventiniai drabužiai*), while white clothing (qualitative adjective) can go grey/yellow/dirty over time; it is a matter of perception. This, in the case of relational adjectives, renders category marking by means of definite forms redundant (**darbin**ieji** *drabužiai*), while for qualitative adjectives morphological marking remains the preferred mode of marking a taxonomic or *ad hoc* category, e.g., the white clothing (*balt**ieji** drabužiai*) as opposed to the coloured clothing (*spalvot**ieji** drabužiai*). It could also be said that qualitative adjectives reflect human perception, while relational adjectives are knowledge-based.

[87] More about genitives and possessives in the Lithuanian NP in Trakymaitė (2018, 117-122).

4.2.2. Establishing a category (classifying adjectives)

As mentioned above, the ability to establish a category (based on a well-established taxonomy or *ad hoc*) is an inherent property of long adjectival forms in Lithuanian. This reflects the fact that "the definite adjectival form can only be used in Lithuanian if the modified NP can yield a definite interpretation either on the level of individual or categorical (taxonomic reference)" (Spraunienė 2011, 12).

Rutkowski and Progovac state that classifying adjectives differ from attributive (or qualifying) ones in that "they do not merely describe a property of the entity denoted by the noun, but categorise that entity as belonging to a certain class/type" (Rutkowski, Progovac 2006, 265), hence *de facto* imply a definite interpretation on the level of categorical or taxonomic definiteness. It is important to highlight that a classifying adjective serves as a restrictive modifier since it limits the denotation of the noun (ibid.). Based on this, I would draw a parallel with the relational adjectives and say that the latter denote objects as attributed to a certain class/type, i.e., a wooden spoon cannot be plastic but can be old, while an old spoon can be either wooden or plastic[88]. Therefore, morphological marking of these adjectives for definiteness is redundant and hence non-existent. In the case of qualifying adjectives, on the other hand, we need tools for disambiguating the interpretation of contextual uses of NPs containing these adjectives to see whether they are restrictive attributes. Subsequently, using a LF would imply this either on the level of the individual (context-based, *ad hoc* taxonomy) or the categorical (generic/taxonomic level).

While taxonomy or generic definiteness is a familiar term (*žal**ioji** arbata* 'green.DEF tea', a kind reference where the adjective is classifying vs *žal**ioji** suknelė* 'green.DEF dress', a context-based reference where the adjective is qualitative) (Rutkowski, Progovac 2006), the term '*ad hoc* category' requires an explanation. An *ad hoc* category is a pragmatic category with an overt linguistic encoding constructed instantaneously to achieve communication goals. It is not available as a structure in long-term memory; it is highly context-dependent and based on an exemplar. Yet, the category itself is more relevant in discourse than the mentioned exemplar. The *ad hoc* categories do not appear with ready-made linguistic labels, but rather by means of complex expressions, e.g., *things to do on a rainy Sunday afternoon* (Mauri 2014).

[88] *medinis šaukštas* 'wooden spoon' → [classifying]; *medinis šaukštas* 'wooden spoon' → ? [qualifying].

Speaking of categorisation, as in establishing *ad hoc* categories, we evoke the concept of inclusiveness, viz., we assign a subject/object to a category based on particular properties and ascribe those properties to all those belonging to the same category inclusively, e.g.:

(60) *Tikrai nesu iš [tų][89] rūpesting**ųjų** tėvų, kurie niekada nepamiršta vaikams sudėti priešpiečių dėžutės.*
'I am definitely not one of [those] caring.DEF parents who never forget to pack a lunch box for their kids.'

In (60), the *ad hoc* category of *caring parents who never forget to pack a lunch box for their kids* is established by ascribing 2 properties to the referentially heterogeneous group: 1) being caring; 2) always performing the duty of packing lunches for their kids.

Inclusiveness is also known to be a semantic feature attributed to the category of definiteness, especially when dealing with plural and mass NPs. The inclusiveness condition entails "the reference to the totality of the objects or mass in the context which satisfy the description" (Lyons 2003, 11). Moreover, in case of a singular NP, "uniqueness can be assimilated to inclusiveness" because there is only one object that satisfies the description used (ibid.), e.g.:

(61) *The winner of the 17th series of The Voice is to be announced tonight.*

It is obvious that there can be only one winner of the TV singing completion. Even though the NP is non-referential, it is nevertheless definite because the condition of inclusiveness is met. This is an important concept for Lithuanian as the difference between the use of long and short forms in case of nominalisations can be explained by the notion of maximal inclusivity (see 2.3 for detailed discussion), e.g.:

62) a. *alkani žmonės* 'hungry.NONDEF people' → some hungry people/the hungry people (if used anaphorically) vs
b. *alkan**ieji*** 'hungry.DEF [ones]' → ALL those who are hungry

Recently, the terms 'weak definites' vs 'strong definites' have been applied to the use of short and long adjectival forms in Lithuanian NPs (Šereikaitė 2019). The notions of *weak vs strong definites* were proposed by Florian Schwarz in 2009 based on his analysis of definite articles in German and Germanic dialects (Schwarz 2009). Weak definites are referential expressions "that presuppose that there is a unique entity meeting the description of the noun phrase", whereas strong definites "involve an

[89] This demonstrative here is optional, and can be omitted.

additional anaphoric component, captured by a (pronoun-like) index introduced as a syntactic argument of the strong article" (Schwarz 2019:11). The concept of uniqueness encoded by weak articles (or short adjectival forms for Lithuanian) is contrasted with the concept of anaphoricity (familiarity) encoded by strong articles (long adjectival forms for Lithuanian). These notions are explored in the article. Ultimately, both uniqueness and familiarity contribute to achieving the inclusiveness and identifiability needed for grasping the notion of definiteness, which may differ in their linguistic expressions.

4.2.3. Other instances of long form uses: nominalisations, terminology, NPs with the emphatic pronoun *pats* 'self'[90]

Besides dominating in the nominalised uses, the long forms of adjectives are compulsory in terminology, where a term comprises an NP containing a modifying adjective. Their uses in such instances are taxonomic[91]. They are used in terms in all fields of science, e.g., botany, biology, chemistry, medicine, technology, etc., as modifiers to nouns, establishing kind reference, e.g.:

(63) *paprast**ieji** spuogai* (med.) – 'ordinary.DEF acne' – Lat. *Acne vulgaris* (in medicine)

(64) *atskir**oji** nuomonė* (jur.) – 'dissenting.DEF opinion' (in law)

(65) *individual**usis** akcizas* – 'personal.DEF excise duty' (in finance)

This is based on the notion that long forms in an NP establish a category, a reference to the kind, viz., an NP with generic reference, as demonstrated in (63) – (65). Yet, as Holvoet and Sprauniené rightfully notice, "if a combination of adjective and noun does not form a unitary concept referring to a more or less established kind or type of individual, the possibility of using definite adjectives in generic and indefinite contexts is lost in Lithuanian" (Holvoet, Sprauniené 2012, 51). They use the concept of *šaltas maistas* 'cold.NONDEF food/meal' to illustrate this. The concept of *hot* vs *cold meal* is rather well-established. Yet, in Lithuanian, the taxonomic NP *šalt**asis** maistas* 'cold.DEF food/meal' is impossible due to the lack of conceptual prominence of this *ad hoc* category (ibid., 51-52). On the other hand, *šalt**oji** kava*

[90] In Lithuanian, this pronoun is multifunctional, exhibits a peculiar morphosyntactic behaviour and consequently deserves special attention and analysis.

[91] Cases where an adjective and a noun form a unitary concept referring to a kind, a class, or a type (Sprauniené 2011).

'cold.DEF coffee', referring to the Italian-style *frappé coffee* seems to be functional and unitary enough to be used as a definite generic.

While working through the lists of collocations of the selected adjectives, I came across numerous dubious examples of terminology-like use of long forms in NPs as they included several instances of uses, e.g.:

(66) *būding**ieji** bruožai* – 'characteristic.DEF features'
(67) *įdom**iosios** užduotys* – 'interesting.DEF tasks'
(68) *vidutin**ioji** karta* – 'middle.DEF generation'
(69) *normal**ieji** mokiniai* – 'ordinary.DEF pupils' (cf. *normal**iosios** mokyklos* 'ordinary.DEF schools', as opposed to *special**iosios** mokyklos* 'special.DEF schools', as in schools for children with special needs)

They were all represented by numerous counts proving their relatively frequent use and formed seemingly unitary concepts, hence I consider them to be cases of generic definiteness[92]. Similarly, there were cases of adjectives where nominalised uses were prevalent, e.g., *skirting**asis*** 'different/distinct.DEF', where out of 4 instances, 3 were nominalisations; *reikaling**asis*** 'needful.DEF' as in *pagalbos reikaling**ieji*** 'those needful.DEF of help', where out of 17 instances, 10 were nominalisations; and *vert**ieji*** 'worthy.DEF; as in *vert**ieji** valdyti* 'those worthy of rule', where out of 27 instances, 25 were nominalisations. Most nominalisations occur predominantly in the plural. As Mikulskas notes, the bare use of long adjectival forms in the plural could be considered definite NPs *per se*, as the referents they denote are a well-defined group of people due to the inclusiveness condition being satisfied (Mikulskas 2006, 60). The condition of inclusiveness foresees "the reference to the totality of the objects or mass in the context which satisfy the description" (Lyons 2003, 11). Mikulskas further infers that this inclusive definiteness in cases of long-form nominalisations is *inherited* from the corresponding complex NPs (ibid), e.g., *pagalbos reikalingi žmonės* 'people needful.NONDEF of help' = 'those who are needful of help' vs. *pagalbos reikaling**ieji*** 'those needful.DEF of help' = 'ALL those who are needful of help'. He considers such nominalisations, the characterising feature of which is the notion of the maximal inclusivity (note the inserted ALL), to be elliptical structures. Yet, both nominalisations and term-like nominals are triggered by the same semantic mechanism, viz., the

[92] In this dissertation, I considered them to be cases of generic definiteness for purely pragmatic reasons, in order to deal with high numbers that needed to be assessed manually. The general rule was that if a collocation was repeated twice or more, it was considered to be a case of generic definiteness. I am aware of the limitations of this approach.

use of long adjectival forms as modifiers in nominals licenses the establishment of categories (both singular and plural), that can and often do become term-like phrases, independent lexical units with established meanings.

However interesting and worth analysing, these three types of uses of long adjectival forms belong to the periphery of the definiteness phenomenon. These uses have been disregarded and eliminated from the data used in this article and will not be further discussed, apart from the discussion concerning the semantics of definiteness.

A third type of construction containing long adjectival modifiers needs to be discussed here as it also presents a special type of use of long forms, namely, in adjectivally modified NPs containing a pronominal intensifier *pats*. This type of use corresponds to the superlative constructions in gradation by specifying that this uniqueness/inclusiveness is based on the said quality to a higher degree than all the others, e.g.: *būdingie**ji*** 'typical.DEF' vs *būdingiausi* 'typical.SUPL' or *įvair**iosios*** 'various.DEF' vs *įvairiausios* 'various.SUPL' (Paulauskienė 1994, 232):

(70)	*patys*	*būdingie**ji***	*raštai*
	self.NOM.PL.M	typical.NOM.PL.M.DEF	pattern.NOM.PL.[M]
	'the most typical patterns'		
(71)	*pačios*	*įvair**iosios***	*priešpriešos*
	self.NOM.PL.F	various.NOM.PL.DEF	contradiction.NOM.PL.[F]
	'the most different contradictions'		

Very few examples of this type were encountered in the data. One interesting case is the example below, where the long adjectival form *įvair**ioji*** seems to denote the property of the superlative degree without the intensifier pronoun, as in examples (70) – (71), e.g.:

(72)	*Sostinės*	*gatvės*	*ir*	*skverai*
	capital.GEN.SF.[F]	street.NOM.PL.[F]	and	square.NOM.PL.[M]
	mirga	*įvair**iosiomis***		*reklamomis.*
	flash.3.PRS	various.INSTR.PL.F.DEF		advertisement.INSTR.PL.[F]

'The streets and the squares of the capital flash with the widest variety of ads.'

These examples, though very few, were included in the statistics of the data presented in this article because they resemble the construction illustrated by (70) and (71) except for the absence of *pats*. I think this is the same construction.

4.3. Uses of short adjectival forms

4.3.1. Short adjectival forms in the definiteness marking system

A few words need to be said about the short adjectival forms and their place in the definiteness marking system of a Lithuanian NP. A full overview of the definiteness marking in Lithuanian is presented in Section 2.5.

As mentioned, SFs may, but do not necessarily, get a definite reading. Spraunienė (2011, 4) and Šereikaitė (2019, 97) both note that it is only short adjectival forms that can introduce a new discourse referent, which is a function typically attributed to indefinite markers; the long forms seem to be impossible in this context, e.g.:

(73)	*Ant*	*palangės*	*tupėjo*	*[*balto**ji**]*
	on	windowsill.GEN.SG.F	sit.3PST	[white.NOM.SG.F.DEF]
	balta		*katė.*	
	[white.NOM.SG.F.INDEF]		cat.NOM.SG.[F]	

'On the windowsill there sat a white cat.'[93]

It has been argued that anaphoric definiteness, known as the only type of linguistic definiteness where the referent is to be found in the linguistic, rather than the extralinguistic, context (Lyons 2003, 158), is considered to represent strong definiteness, as opposed to weak definiteness associated with the notion of uniqueness (Schwarz 2009, 2019) (see 2.2). The subsequent examples will provide a context for the anaphoric use, in which both long and short forms are eligible:

(74) *Ant palangės tupėjo dvi katės, juoda [*juodo**ji**] ir balta [*balto**ji**].*
*Pamačiusi mane, juoda/ juodo**ji** nušoko žemėn, o balta/ balto**ji** liko tupėti.*
'On the windowsill there sat two cats, a black one and a white one. Upon seeing me, the black one jumped down, whereas the white one remained [on the windowsill].'

Further, Šereikaitė says that "nominals with short form adjectives occur in indefinite environments. In contrast, the presence of a long adjective in nominal expressions is incompatible with an indefinite context […]" (Šereikaitė 2019, 98). Similar types of anaphoric constructions containing long adjectival forms and associated with strong definites seem to be incompatible with *įvairus* 'various, varied, diverse' or *gausus* 'abundant, ample, bountiful', or *nemažas* 'considerable, not small', e.g.:

[93] This example is cited from Spraunienė (2011, 74).

(75) *Jie valgo įvairų maistą. [* Įvair**usis**] Įvairus maistas jiems teikia malonumą.*
'They eat a varied diet. The varied diet gives them pleasure.'

(76) *Staiga jam prasidėjo gausus kraujavimas. [*Gaus**iojo**] Gausaus kraujavimo niekaip nepavyko sustabdyti.*
'Suddenly, he started bleeding profusely. There was no way to stop the profuse bleeding.'

(77) *Prie bibliotekos buvo susirinkęs nemažas būrys vaikų. Po ilgų diskusijų [*nemaž**asis**/nemažas] būrys patraukė link stadiono.*
'A considerable crowd of children gathered by the library. After long discussions, the [*considerable] crowd moved towards the stadium.'

(78) *Jie gyvena nuosavame name. [*Nuosav**asis**] Nuosavas namas jiems nepigiai atsiėjo.*
'They live in a private house. The privately-owned house did not come cheap to them.'

As we see, in these clearly definite anaphoric contexts, the use of long forms is unacceptable. Moreover, as shown in (77), the repeated use of the adjectival modifier is ungrammatical as well. What are the implications of this? Šereikaitė notes that "short adjectives pattern in a similar way to the weak definite that is associated with uniqueness" (Šereikaitė 2019, 85), implying a relationship between the adjectival form and the semantics of definiteness. Yet, when used with proper names, which is evidently a case of uniqueness-motivated definiteness, long adjectival forms are compulsory, e.g., *drąs**usis** Nelsonas Mandela/*drąsus Nelsonas Mandela* 'the courageous.DEF Nelson Mandela' or *jaun**ieji** Petrauskai/*jauni Petrauskai* 'the young.DEF Petrauskas family' (as opposed to *sen**ieji** Petrauskai* 'the old.DEF Petrauskas family' referring to the parents of the young Petrauskas). Short adjectival forms would not be possible in these instances.

Following our discussion in 4.2.2, we could say qualitative adjectives that do not take LFs (as in (75) – (78)) cannot be restrictive attributes limiting the denotation of the noun. Our hypothesis is that they are not typical qualitative adjectives because the properties they imply do not allow the establishment of categories, viz., they function as classifying adjectives, cf. with (79) where a rather common negating construction is used to establish an *ad hoc* category:

(79) *Jūsų klausimas – ne iš lengv**ųjų**.*
'Your question is not an easy one.' → literally 'is not from the easy.DEF [ones]'

(80) **Jūsų dieta – ne iš įvair**iųjų**.*
'Your diet is not a varied one.' → literally 'is not from the varied.DEF [ones]'

4.3.2. Attribution versus predication

It is an established fact that Lithuanian adjectives are assigned three different functions: 1) attributive as a primary function, 2) a predicative function (as a complement) and 3) substantivised (nominalisations) as a secondary function (Kamadulytė-Merfeldienė, Balčiūnienė 2016, 128). Since nominalisations were discussed in detail in 2.3, a few words need to be said about the first two. The difference between attributive and predicative adjectives is of importance here because only short forms in Lithuanian can occur in the predicative function. We know that Slavic languages, i.e., Russian, have had a similar long-short form opposition. Today, these two groups could be considered as belonging to different word classes, namely, adjectives and predicatives.[94] Short forms might differ slightly in meaning (compared to LFs); they are restricted to predicate position only; and only they can govern direct objects (Hansen 2004, 62-63). In this, they are more verb-like than adjective-like.

If adjectives are used predicatively in Lithuanian, they too may appear only in their SFs both in cases of primary and secondary predication. All the adjectives listed in Tables 8 and 9 can be used both attributively and predicatively (primary and secondary predication in (82a) and (82b)), e.g.:

(81) *Šįvakar prognuozuojami **gausūs** krituliai.*
'Heavy precipitation is forecast tonight.'

(82) a. *Krituliai buvo **gausūs**.* vs **Krituliai buvo **gausieji**.*
'Precipitation was heavy.'

b. *Krituliai iškrito **gausūs**.* vs **Krituliai iškrito **gausieji.***
'Precipitation was heavy.' → literally 'Precipitation fell heavy.'

There are several other properties that distinguish attributive adjectives from predicatives, according to Belk, including ordering restriction and scopal implications governing attributive uses (Belk 2017, 17-30) that pose some very interesting challenges and implications, but these will not be analysed in this dissertation.

4.3.3. A few final comments on the data

The table below presents a value-sorted list (from smallest to largest) of the 30 selected adjectives displaying less than 1% use of long forms as opposed to short forms.

[94] For a discussion on this, see Belk (Belk 2017, 17-22).

Table 9. The shortlist of adjectives predominantly used in the short forms, sorted by values: the percentage of long forms in CCLL (smallest to largest)

No	Adj	Translation	FrD – count of long	CCLL – count of long (raw)	CCLL – count of long (adjusted)	CCLL - % of long
1	nemažas	considerable/not small	0	0	0	0
2	vienodas	uniform/equal/same/homogeneous/like	0	0	0	0
3	panašus	similar/like/alike/analogous/resemblant	0	7	1	0.002
4	skirtingas	different/unlike/separate/distinct/diverse	0	4	1	0.003
5	įvairus	various/varied/miscellaneous	0	17	15	0.02
6	aiškus	clear/understandable/explicit/evident	0	27	5	0.03
7	reikalingas	needed/required/necessary/requisite	0	17	7	0.03
8	vertas	worth/worthy/deserving/valuable	0	27	2	0.03
9	visiškas	complete/total/full/absolute	0	82	3	0.03
10	galutinis	final/ultimate/terminal	0	4	4	0.04
11	neaiškus	unclear/uncertain/obscure/vague/indistinct	0	9	2	0.04

No	Adj	Translation	FrD – count of long	CCLL – count of long (raw)	CCLL – count of long (adjusted)	CCLL - % of long
12	dažnas	frequent/habitual/periodic/repeated	0	7	3	0.05
13	patogus	convenient/comfortable/handy	0	7	2	0.05
14	atskiras	sundry/separate/individual/special/	2	695	20	0.06
15	vidutinis	average/medium/middle/moderate/normal	2	175	21	0.09
16	būdingas	typical/characteristic/specific	2	74	23	0.1
17	gausus	abundant/numerous/plentiful/ample/	0	8	8	0.1
18	menkas	meagre/insignificant/poor	0	10	8	0.11
19	privatus	private/personal/own/proprietary/individual	1	126	25	0.12
20	optimalus	optimal/optimum/superb/top-notch	0	4	3	0.12
21	konkretus	concrete/particular/specific	0	127	36	0.13
22	tolesnis	further/subsequent/successive	1	14	14	0.13
23	įdomus	interesting/exciting/entertaining	0	70	23	0.14

No	Adj	Translation	FrD – count of long	CCLL – count of long (raw)	CCLL – count of long (adjusted)	CCLL - % of long
24	normalus	normal/regular/ordinary/average	0	120	16	0.14
25	sudėtingas	complex/complicated/multiplex/elaborate	0	39	20	0.18
26	individualus	individual/separate/distinctive/special	5	647	69	0.44
27	reikšmingas	significant/meaningful/important/weighty	0	40	40	0.48
28	ryškus	bright/stark	1	44	38	0.56
29	švarus	clean/pure/clear/fresh/immaculate	0	59	41	0.64
30	paprastas	simple/ordinary/normal/average/usual	17	1490	100	0.72

Evidently, some values in the column "CCLL – count of long (adjusted)" differ significantly from the column to its left, presenting the actual ("raw") count of the long forms. In most of the cases, e.g., *atskiras*, *privatus*, *konkretus*, *normalus*, *individualus* and *paprastas*, as explained in section 1.2, this is due to the extraordinarily high number of uses in either terminology or nominalisations. The selected sample of the collocation analyses of these adjectives is included in Appendix B. Another fact worth mentioning here is that there seems to be a gap in values between 0.18 and 0.44. Whether or not this is significant is worth analysing; however, it will not be done in this thesis.

Finally, I would like to mention the fact that certain adjectives are known to function as determiners signalling the definiteness of an NP, e.g., in Swedish and Danish, morphologically marked long adjectival forms seem to license the definite reading and the necessary morphological marking for definiteness elsewhere in an NP (e.g., on the stacked adjectives) without having a preposed definite article or another acknowledged determiner (Börjars 1994, Van de Velde 2011):

(83)	Swed.:	*sist**a***	*misslyckad**e***	*försök**et***
		last.DEF	failed.DEF	attempt.DEF
		'the last failed attempt'[95]		
(84)	Dan.:	*nederst**e***	*højr**e***	*skrivebordsskuffe*
		lowest.DEF	right.DEF	desktop drawer
		'the bottom-right desktop drawer'[96]		

Börjars calls them *adjectival determiners* (1994, 225). Taking our clue from such parallels, we would like to suggest that, at least in some cases, the lack of an opposition between short vs long form is indicative of a shift towards a determiner-like function.

4.4. Data analysis

4.4.1. "Adjective-like" adjectives

At first glance at Table 9, one major group stands out, viz., qualitative adjectives that, besides the absence of uses with long forms, display all the typical features characterising this type of adjective in Lithuanian: they act as attributive modifiers in NPs; they are gradable (or scalar); they exhibit complex word-formation patterns; they are used in nominalisations; and they

[95] This example is cited from Börjars (Börjars 1994, 224).

[96] This example is from ordnet.dk: https://ordnet.dk/korpusdk/teksteksempler/kontekst, accessed 2021-05-24.

are used in terminology. These adjectives include the following items from the table: *įdomus* 'interesting, exciting, entertaining', *patogus* 'convenient, comfortable, handy', and *švarus* 'clean, pure, clear, fresh, immaculate'.

These adjectives seem to denote properties of unstable, temporary character. As we know, LFs establish *ad hoc* categories. The unstable nature of the properties denoted by these adjectives seems to make them unfit to serve as classifying properties. While *šilt**ieji** drabužiai* 'warm.DEF clothes' denotes a functional property of clothes that is of a permanent nature, **švar**ieji** *drabužiai* 'clean.DEF clothes' cannot establish a category due to the regular shifts between the categories *clean* and *dirty*, as all clothes become dirty over time and can be made clean again. Most certainly, labels such as *švar**iųjų** *drabužių pintinė* 'basket for clean.DEF clothes' as opposed to *nešvar**iųjų** *drabužių pintinė* 'basket for dirty.DEF clothes' can be created for containers in a laundry sorting room. However, the use of long forms is not attested in the CCLL. Trost (1966, cited from Riessler 2016, 48) notes that there is a correlation between the use of long versus short adjectival forms and the permanent versus non-permanent properties denoted by adjectives.

To further explore this hypothesis, I expanded the list of adjectives with similar temporary properties to include the following: *aktualus* 'actual, relevant', *alkanas* 'hungry', *naudingas* 'useful, beneficial, valuable', and *tuščias* 'empty, blank, dummy'. They exhibit the same behaviour, viz., uses with long forms are very few, mostly as nominalisations or as modifiers in terms. It seems that an inherent semantic property of [+IMPERMANENCE] disables morphological marking of the NP as a representative of an *ad hoc* category. It is possible to have an established category of *nauding**osios** *iškasenos* 'natural resources' (in Lith. literally 'useful.DEF resources'), but impossible to have one of **naudingieji** žmonės* 'useful.DEF people'. Likewise, *tuščias* 'empty' allows long forms in terms like *tušč**ioji** žarna* 'jejunum' (in Lat. *Jejunum*, in Lith. literally 'empty.DEF intestine') and *tušč**iosios** avižos* 'a species of grass in the oat genus' (in Lat. *Avena fatua*, in Lith. literally 'empty.DEF oats'), but disallows nominals like *tušč**ioji** lėkštė* 'empty.DEF plate'.

Besides the above-mentioned adjectives, two other sub-groups seem to belong to this section of the "adjective-like" adjectives: 1) a group that is called the 'absolute' adjectives in CGEL[97], e.g., *optimalus* 'optimal, superb, top-notch', and 2) the group that I choose to call the 'mid-class' adjectives, e.g., *normalus* 'normal, regular, ordinary, average', *paprastas* 'simple,

[97] *The Cambridge Grammar of the English Language* 2002.

ordinary, normal, average, usual', *vidutinis* 'average, medium, middle, moderate, normal', and *būdingas* 'typical, characteristic, specific', since they represent values located in the middle of the scale, as opposed to absolute adjectives, which usually can be found at either end of the scale.

The so-called 'absolute' adjectives, like *optimalus* 'optimal' but also *unikalus* 'unique/alone' and *totalus* 'total' (here again, I have chosen to expand the category and include the latter two to test the hypothesis), are traditionally viewed as non-gradable because they denote the endpoints of the scale and hence are non-scalar *per se*. The properties they denote are perceived as of the absolute (superlative) degree. Even though it is technically possible to form (and use) gradation forms, e.g., *optimalus* (POSITIVE) – *optimalesnis* (COMPARATIVE) – *optimaliausias* (SUPERLATIVE), they are seldom used. One of the very few examples of the long form use is the type of construction with the emphatic pronoun *pats* 'self', as described in 2.3., where the combination [*pats* + long form] seem to license the superlative meaning:

(85) *Šešeri metai - pats optimal**usis** amžius pradėti formuoti balsą.*
'Six years is the optimal.DEF age to start working on voice formation.'

The mid-class adjectives are slightly different from the others listed above, in that they do not establish categories based on particular properties – these remain undisclosed. What they denote is a proximity to the norm/average/medium. In terms of semantics, this group resembles similarity expressions, but rather than expressing proximity in similarity between objects, as similarity expressions do, these adjectives express proximity between an object and the norm/medium, instead of another object. This is the group where the counts of the long forms in the CCLL had to be significantly adjusted, as they contained many terminology-like uses, precisely because of the semantics of the norm/type. The adjectives in Table 9 that belong here are *normalus* 'normal, regular, ordinary, average', *paprastas* 'simple, ordinary, normal, average, usual', *vidutinis* 'average, medium, middle, moderate, normal', and *būdingas* 'typical, characteristic, true to type'. Here again, I have chosen to include an additional example, *tipiškas* 'typical, characteristic, true to type', which, unsurprisingly, also exhibits the absence of long forms. Otherwise, apart from the adjective *vidutinis* 'average, medium, middle', all of these are gradable (or scalar[98]); they exhibit complex word formation patterns; they are used in nominalisations; and they are used in terminology. The exception of *vidutinis* could be explained morphologically, i.e., it is a

[98] With certain restrictions, e.g., *paprastesnis*, a comparative degree of *paprastas* 'simple' means 'simpler in structure' rather than 'more usual'.

derived adjective constructed with the suffix *–inis*, which seems to be incompatible with qualitative adjectives[99]. Hence, features typical of qualitative adjectives, such as gradation, are disabled. Due to their semantics, these adjectives seem to be able to have multiple antonyms depending on the context; one could say that they establish *ad hoc* paradigms of antonyms, e.g.: *paprastas* 'simple, ordinary, usual':

(86) *paprastas butas – tarnybinis butas* 'normal accommodation – tied accommodation'
paprasta mokykla – speciali mokykla 'ordinary school – special school'
paprasti agentai – įtakos agentai 'ordinary agents – agents of influence'
paprasta sąskaita – taupomoji sąskaita 'an ordinary bank account – a savings account'
paprastas skrydis – skrydis su persėdimu 'a direct flight – a transfer flight'

Summing up, I would like to say that despite some peculiarities, all of these adjectives behave like true adjectives in that they: 1) function as descriptive modifiers, assigning properties to heads of NPs; 2) exhibit the full set of features characteristic of qualitative adjectives, with the exception of assuming LFs, which is due to the semantics of the three sub-groups, viz., denoting properties that are either non-stable (impermanent), undisclosed or 'absolute' and therefore not instrumental in establishing classifications or categories.

4.4.2. Displaced modifiers (or adverbiatives)

Analysing the 30 adjectives in the table, yet another group of adjectives stands out, i.e., *dažnas* 'frequent, habitual, periodic, repeated', *aiškus* 'clear, understandable, explicit, evident' and *ryškus* 'bright, significant'. As in the sub-groups above, I expanded the category by including two additional adjectives here to test the hypothesis, viz. *akivaizdus* 'obvious, evident, apparent' and *retas* 'rare, scarce, infrequent', which are also characterised by the absence of long forms. I call these 'displaced modifiers', a term partly borrowed from Mel'čuk's concept of displaced categories (Rus. *смещенная категория*) (Mel'čuk 1998, 2:29-31) implying that information encoded in them is displayed "in the wrong place", viz., sentence-level modification is downgraded to a nominal level. As shown below, the manner adverb *angrily*

[99] With the exception of very few like *galutinis* 'final, ultimate, terminal, end', *paskutinis* 'last, final, ultimate', *žemutinis* 'lower, low, ground', *aukštutinis* 'upper, high', *vidurinis* 'middle, mid, secondary' and similar.

is in fact a subject-oriented adverb that is, it attributes a property *angry* to a female person leaving the room.

(87) *She angrily left the room.*[100]

In a similar way, in examples (88) – (91) the sentence-level modification typically expressed by adverbials is relocated to the phrasal level and encoded in the NP through the modifying adjective:

(88) *Jis yra dažnas svečias šiuose namuose.*
'He is a frequent guest in this house.'

(89) *Lietuvos nacionaliniame muziejuje yra tokių vietų, kuriose tik retas lankytojas tėra pabuvojęs.*
(literally) 'There are places in the Lithuanian national museum which only a rare visitor has stumbled upon. → There are places in the Lithuanian national museum rarely stumbled upon by visitors.'

(90) *Auditas atskleidė ryškų piktnaudžiavimą sistema šalyje.*
'The audit revealed a significant abuse of the system throughout the country.'

(91) *Norėčiau pranešti apie akivaizdų (aiškų) pažeidimą.*
(literally) 'I would like to report an obvious infraction. → I would like to report what is obviously an infraction.'

Dažnas svečias in (88) is not a property of a guest, but rather a modifying predication informing the frequency of a male person's visits to a particular home. In (89) *retas lankytojas* is not a property of a visitor, but a modifying predicate informing the reader that people rarely visit certain places/locations. Likewise, in (90) the modifier 'significant' does not specify a kind of *abuse* but rather the level/degree of its prevalence in the country. In (91) the evident or explicit violation does not entail a property of *violation*, but rather the level of it. The latter case could be called an evidential adverbiative. In CGEL, these uses of adjectival modifiers are called *modal attributives* (2002, 557). Similar examples would be *a potential winner of the Nobel Prize*, *a plausible explanation*. Due to the fact that it is a displaced sentence-level modification, the grammatical features of adjectival modification on a nominal level are disabled, eliminating the possibility of uses of long forms.

In a way, these constructions could be treated as nominalisations, e.g., *dažnai lankosi* 'frequently visits' → *dažnas lankytojas* 'frequent visitor'; *aišku, kad tai – pažeidimas* 'it is obvious that this is an infraction' → *aiškus pažeidimas* 'obvious infraction'. The degree of membership in the category denoted by the noun depends on the validity of what is expressed by the

[100] Example provided by Kees Hengeveld at *Academia Grammaticorum Salensis Septima Decima*, Lithuania, July 29, 2020.

modifier: in the latter example, the less obvious an infraction, the less assuredly we can classify the event involved as an infraction. They qualify the belonging of the noun to the category denoted, but do not establish the category. These adjectives behave differently from the typical adjectives like *juodas* 'black' (e.g., *juodas švarkas* 'black jacket' is a sub-category of all jackets). The modifier *juodas* 'black' can establish a sub-category of black jackets, but it is not essential in identifying a jacket as member of the category of *jackets*; while *dažnas* 'frequent' speaking of frequent visitors does define the membership degree to which a visitor can be considered to belong to the category of visitors (the more often/frequently one comes, the more likely he/she will be considered a visitor; similarly, the more obvious the infraction, the more likely it can be classified as one).

Even though the observation below does not fall under the label of displaced modification, but rather under the label of some type of quantification (this will be discussed in Section 4.4.5), I would briefly like to comment on two adjectives discussed above, viz., *dažnas* 'frequent' and *retas* 'infrequent'. They seem to belong to more than one group of adjectives used predominantly with SFs, viz., they may also function as quantifiers with *dažnas* meaning 'more than one, a few, many' and *retas* meaning 'few' as in an unspecified quantity, as demonstrated in these examples:

(92) *Dažnas žmogus, išgirdęs žodį* Belgija*, pagalvos apie šokoladą.*
'Upon hearing the word *Belgium*, many people will think of chocolate.'

(93) *Reta moteris praranda savitvardą konflikto metu.*
'Few women lose control in a conflict situation.'

They show behaviour similar to that of multal (expressing quantification by items such as *many*, *much*, *a lot*, etc.) and paucal (expressing quantification by items such as *a few*, *several*, *a little*) quantifiers. Other languages, e.g., Swedish, also have these types of quantifiers expressed through adjectives, e.g., *åtskilliga* 'several' and *enstaka* 'single, isolated'.

4.4.3 Adjectives with complex verb-like argument structures

Yet another feature shared by quite a few adjectives in Table 2 is worth special attention. While qualitative adjectives do not traditionally take complements, some of the examined ones do. In this regard they exhibit verb-like behaviour and often take more than one argument, e.g.:

(94) *globos* *reikalinga* *būklė*
care.GEN.SG.[F] in need.NOM.SG.F.NONDEF state.NOM.SG.[F]
'a state requiring care'

(95) *vaikų* *dvasios* *sveikatai* *didžiai*
children.GEN.PL.[M] spirit.GEN.SG.[F] health.DAT.SG.[F] greatly.ADV
reikalingas *mokslas*
necessary.NOM.SG.M.NONDEF education.NOM.SG.[M]
'education, much needed for the mental health of children'

(96) *verti* *įrašo* *knygoje*
worthy.NOM.PL.M.NONDEF inscription.GEN.SG.[M] book.LOC.SG.[F]
'worthy of an inscription in the book'

(97) *psichikos* *sutrikimams* *būdingi*
psyche.GEN.SG.[F] disorder.DAT.PL.[M] characteristic.NOM.PL.M.NONDEF
pokyčiai
change.NOM.PL.[M]
'changes inherent in mental disorders'

As shown in (94) and (95) the adjective *reikalingas* 'needed, required, necessary' takes two arguments: 1) in need of something – an argument in genitive (*in need of care*); 2) necessary for something – an argument in the dative case (*necessary for health*). In (96) the adjective *verti* 'worth, worthy, deserving, valuable' takes an argument expressed in the genitive – worthy of something (*worthy of inscription*). In (97) the adjective *būdingas* 'typical, characteristic, inherent' takes an argument in the dative (*inherent in* or *characteristic of*).

Besides the three adjectives mentioned in the examples above, this group includes other items listed in Table 9:

(98) *atskiras* 'separate, individual, distinct' – *atskiras nuo ligoninės administracijos žmogus* 'a person independent of the hospital administration' (literally 'a person detached from the administration') - argument with preposition *nuo* 'from' + genitive

(99) *panašus* 'similar, like, alike, analogous, resemblant' – *panašūs į riedulius koralų gabalai* 'boulder-like pieces of coral' (literally 'pieces of coral similar to boulders' – argument with preposition *į* 'to' + accusative

(100) *vienodas* 'uniform, equal, same' – *vienodos su kitais piliečiais galimybės* 'opportunities equal with other citizens' – argument with preposition *su* 'with' + instrumental

(101) *skirtingas* 'different, unlike, separate, distinct' – *nuo žydų skirtingi krikščionys* 'Christians different from Jews' – argument with preposition *nuo* 'from' + genitive

All these adjectives can be used both attributively and predicatively. It is significant that even predicates can keep with their argument structure, e.g., *Jis buvo reikalingas gydymo,* 'He was in need of treatment', where the adjective *reikalingas* 'in need' takes a genitival complement *gydymo* 'treatment'.

As in the Slavic languages (see 4.3.2), it is only SFs that can function as clausal predicates in Lithuanian, a function usually performed by verbs and verb-like elements, where the functions and properties of attribution are no longer important. It is also known from studies of Russian adjectives that adjectives taking (or implying) arguments strongly favour short forms; they behave more like verbs while long forms behave like nouns (Corbett 2004, 207). Therefore, it is not surprising that these verb-like adjectives do not engage their attributive properties and are used mostly in their SF. Unquestionably, the verb-like argument structure associated with these adjectives alone cannot explain why they are predominantly used in their SFs, yet I felt that it was an evident shared feature allowing me to group them in an attempt to organise the data evidence.

4.4.4. Adjectives between determinatives and pronouns: Quasi-determiners

While definiteness marking has received some attention, we still have no comprehensive description of Lithuanian indefiniteness marking strategy. We know from linguistic studies that languages that have prototypical indefinite markers (i.e., indefinite articles) are uncommon, yet many have other means to signal the indefinite status of an NP. As Lyons notes: "Real indefinite articles – encoding [- DEF], and in part identifiable by not being the same as or readily derivable from a cardinality word – are rare, if they genuinely exist at all." (Lyons 2003, 89). Instead, languages use other markers, which, often, are optional in NPs to indicate specific indefinite reference, e.g., *any*, *some*, *certain*, etc. It is often indefinite pronouns (and certain adjectives) that take on the role of signalling the indefiniteness of an NP.

Several attempts have been made to create a comprehensive classification of Lithuanian pronouns. In 1984, Rosinas published a monograph on their semantic structure. In 1996, a new revised edition of the book was published, providing a novel and, to date, the most comprehensive overview of Lithuanian pronouns. In 1997, a major study based on a sample of 40 languages, including Lithuanian, of indefinite pronouns and their formal and semantic properties was written by Haspelmath. It modernised and completed the analysis of Lithuanian pronouns, which in turn was further fine-tuned by Kozhanov in 2010, focusing on certain series of indefinite pronouns. Yet, the major issue of differentiation between certain pronouns and adjectives, in terms of word class assignment, remains open, as is shown in this section.

Reviewing Rosinas' monograph, Tekorienė (1987, 88-89) justly notices and questions the relationship between certain pronouns and adjectives, saying that their referential functions are fairly similar, e.g.: *visas*

'whole' and *pilnas* 'full, complete'; *tam tikras* 'certain' and *nustatytas* 'given, established', *ypatingas* 'particular' and *specialus* 'special'; *visoks* 'any, all sorts' and *įvairus* 'various, varied, diverse'; *toks pat* 'same' and *vienodas* 'same, uniform, one'; *toks* 'such' and *panašus* 'like, alike, similar'. Paulauskienė makes some new additions to this list, viz.: *dažnas* 'frequent/manifold/numerous', *tūlas* 'frequent/manifold/various', *ištisas* 'whole/entire', *kiauras* 'whole/entire' (Paulauskienė 1994, 44).

Many of the above fall within the four categories explicitly listed by Haspelmath as commonly associated with, but not belonging to, the class of indefinite pronouns (1997, 11-12), viz., mid-scalar quantifiers like *few*, *several*, *many*; generic pronouns like French *on*, German *man*, English *one*; universal quantifiers *all* and *every*; and identity pronouns like *other* and *same*. According to Haspelmath, the mid-scalar quantifiers "express quantity and have nothing to do with indefinites" (ibid.). Universal quantifiers are semantically definite, even though they sometimes lack formal definiteness markers. And finally, the identity pronouns, which, according to Haspelmath "express identity and non-identity and do not show any affinity to indefiniteness at all" (1997,12). The generic pronouns, like those in French or German mentioned above, are lacking in Lithuanian.

Yet, adjectives with meanings and functions very similar to those of the three above-listed groups behave in a peculiar way with regard to the definiteness marker (or rather, the absence of it) in Lithuanian. Whether or not this can be linked to indefiniteness will be discussed later while analysing the material. I assume that they act as quasi-determiners because, besides their modifying function, they perform additional functions in the NP similar to those of determiners.

4.4.4.1. Similatives, dissimilatives and variatives (and multipart modifiers)

It has been argued that there exists a cross-linguistic category containing nominal expressions of similarity, which creates *ad hoc* categories in discourse (van der Auwera, Sahoo 2019). The prototypical representative of this category is the word *such* (Lith. *toks*, Swed. *sådan*, Dutch *zulk*, Germ. *solcher*). This category is known to contain words attributed in grammars to various word classes, e.g., pronouns, adjectives, determinatives. Their meanings entail a combination of semantic categories of similarity and demonstration. That is why this category is also known as

'similative demonstratives'[101]. Examining the list of adjectives in Table 9, it became apparent that Lithuanian adjectives *panašus* 'similar, like, alike, analogous, resemblant' and *vienodas* 'uniform, equal, same, homogeneous, like' could be considered as candidates for this category, especially when used without complements, as their main function is to express similarity or comparison and point to the object of comparison.

Even though inherently indefinite, in many languages pronominal similatives used as modifiers, e.g., English *such*, Swedish *sådan*, appear in NPs with indefiniteness markers. This seems to be the case in Lithuanian, too (Vaitkutė 2019). Interestingly, out of 7 examples of *panašusis* 'similar.DEF' in the CCLL, 5 were terminology uses, viz., in geometry *panašieji trikampiai* 'similar.DEF triangles' and *panašiosios figūros* 'similar.DEF figures', and 2 nominalisations. The adjective *vienodas* 'uniform, same, equal' has 0 instances of long forms in the CCLL.

Following this line of thought, I added a couple of additional adjectives to this group, naming them 'dissimilatives' and 'variatives'. While *toks* 'such' creates an *ad hoc* category in the utterance, *įvairus* 'various' also creates a category for the purpose of what is being stated in the utterance while at the same time emphasising that the categorised objects are dissimilar in other essential aspects. Adjectives like *skirtingas* 'different, unlike, separate, distinct, diverse' and *atskiras* 'distinct, non-identical, unlike' naturally belong here. Also, words expressing variety or complexity (being multipart), the prototypical member of which is *all kinds/all sorts* (Lith. *visoks*, Swed. *alla slags*, Dutch *allerlei*), could fit in this group.

The dissimilative *skirtingas* 'different' has 4 instances of long forms in CCLL, 3 of which are nominalisations. I think that these adjectives do not capture the notion of identity (apart from some instances of *vienodas* that has a meaning of 'same' in its inventory). Rather than denoting a criterion from which a category is established, variatives characterise the heterogeneous structure of a set of objects. Therefore, since the long adjectival forms denote a category (either taxonomic or *ad hoc*) based on a qualifying property, they are incompatible with variatives. Of the adjectives found in Table 9, the following two belong here: *įvairus* 'various, varied, miscellaneous' and *sudėtingas* 'complex, complicated, multiplex, elaborate, multipart'. I have expanded this category by adding here *tūlas* 'various, of all sorts', which also shows a complete absence of long forms in CCLL, strengthening the

[101] The term 'similative' referring to a linguistic category was coined by van der Auwera. In 2018, this term was modified to 'demonstrative similatives' or 'similative demonstratives' (van der Auwera, Sahoo 2018).

hypothesis that these items form a particular group, and are more than just mere attributive adjectives.

Since long forms of adjectival modifiers establish categories based on particular properties, variatives like *įvairus* 'various' and multipart modifiers like *sudėtingas* 'complex' (in Lith. literally 'comprising different bits') cannot establish a category based on a particular property – this property is undisclosed. It clearly establishes a category, the individual members of which are different. Hence, as opposed to *balt**ieji*** '[the] white.DEF [ones]' = 'ALL those who are white', *įvair**ieji*** '[the] various.DEF [ones]' = 'ALL those who are ?':

(102) *įvairių tautybių žmonės*
'people of various.NONDEF nationalities'

(103) *Jis po tūlas parduotuves vaikščiojo, bet ko reikėjo, taip ir negavo.*
'He walked around various.NONDEF shops, but did not get what he needed.'

(104) *Esi laimingas, patyręs labai sudėtingas emocijas.*
'You are happy having experienced very complex.NONDEF emotions.'

(105) *Jis susirgo komplikuota alergijos karštam klimatui forma.*
'He contracted a complicated.NONDEF form of allergy to hot climates.'

These adjectives are non-singular due to their semantics, i.e., one cannot be various; multipart implies composition of more than one part. Another characteristic feature of variatives and multipart modifiers is that they are not bipolar. *Various* and *multipart* or *complex* do not have clearly defined antonyms. To sum up, one could say that their primary function is other than just to modify. They express variety and complexity, hence have a correlation with quantification: an expression of multal quantification.

4.4.4.2. Particularising attributives

In parallel to the above discussion, one could say that words expressing specificity or particularity form a minor category across languages, the prototypical member of which is *certain* (Swed. *somlig* or *viss*, Dutch *een zekere*). The term used in CGEL is 'particularising attributives' (2002, 558). In Lithuanian, its counterpart is a pronoun *tam tikras*[102]. Adjectives belonging to this group "serve to pick out a specific member or group of members of the set denoted by the head" (ibid.). The shared property of these adjectives is that, essentially, they do not denote any property, but rather they specify, point

[102] In the Lithuanian Grammar (Ambrazas et al. 2006, 188), it is classified as an indefinite differentiating pronoun. Rosinas does not consider this to be a pronoun, but rather an adjective (1996, 11).

out or particularise a member or a group of members belonging to the category denoted by the NP. Adjectives belonging to this group are *konkretus* 'concrete, particular, specific' (only this adjective is included in Table 9), *specialus* 'special, particular, individual', and *ypatingas* 'special, particular, peculiar, extraordinary' (these two have been added by me to test the hypothesis, based on the comments of Tekorienė, see 4.4.4). The adjective *konkretus* has a very low count of long forms, while both *specialus* and *ypatingas* show a different pattern with a much higher percentage[103] of long form uses, e.g.:

(106) *Surinktos lėšos numatytos konkretiems tikslams.*
'The funds raised are earmarked for specific.NONDEF purposes.'

(107) *Kunigas daug dėmesio skyrė ypating**iesiems** sielovados darbams.*
'The priest paid much attention to the special.DEF pastoral tasks.'

While in (106) a short form of *konkretus* is found, in (107) we see a long form of *ypatingas*. In both cases, the translation of the adjective is 'specific', which in these cases could be interpreted as a direct synonym of *certain*. Yet, LF in (107) could have been chosen deliberately to avoid the determiner-like reading and to demonstrate that special pastoral tasks do not just mean 'certain tasks', but rather 'tasks specific to the pastoral vocation'. I cannot offer an explanation why *specialus* and *ypatingas* behave differently than *konkretus*. What we do know from the grammars of other languages, e.g., Swedish, is that this type of adjective is predominantly used in indefinite NPs. A simple search in the BNC[104] for the string 'a certain' *versus* 'the certain' yields a result of 5100 instances versus 56, which is significant.

Summing up, one could say that this group of adjectives establishes a category not based on a particular property denoted by them (the property is not disclosed); rather, they seem to describe the structure of the category. They could be considered to be quasi-determiners that appear with indefinite NPs. In discourse, these adjectives seem to function as anonymity guardians, allowing the speaker to indicate that a set is not arbitrary without disclosing the feature that constitutes it. In a nutshell, like similatives and dissimilatives, these adjectives, besides their main function to serve as modifying attributes, perform other functions; in this case, that of particularising while leaving the referents unidentified.

[103] *Specialus* has appr. 50% of LFs, while *ypatingas* has appr. 6%. However, collocation analysis needs to be carried out to eliminate cases of terminology-like uses, e.g., *special**iosios** pajėgos* 'special forces', *special**ioji** mokykla* 'special school → school for children with special needs', etc.

[104] The BNC stands for the British National Corpus, accessed on 2020-08-28: english-corpora.org/bnc/.

4.4.4.3. Possessives

Possessives[105] are inherently definite and, in languages with determinatives, are incompatible with other determinatives, like articles. Even though mostly expressed by pronouns or genitives, sometimes they can be expressed by adjectives. The prototypical member of this category is *own* (Swed. *egen*, Dutch *eigen*). In English, it often appears following a possessive pronoun, i.e., *He cooked his own dinner*. In Table 9, we find an adjective *privatus* 'private, own, personal, individual'. This group could be expanded by adding yet another adjective *nuosavas* 'own, private, one's very own', which is a prototypical member of this category. It also has 0 long forms in the FrD and the CCLL but has not been included in the data here since its ranking number is 5280. Moreover, in languages that have an adjectival marking of definiteness alongside determinatives, a prenominal modifying adjective in an NP is always marked for definiteness, viz., used in its long form, e.g.:

(108) Swed. *hans* *lill**a*** *hus*
'his little.DEF house'

However, interestingly, if the possession is expressed with the help of an adjective, like the ones mentioned above, the marking on the adjective disappears, e.g.:

(109) Swed. *hans* *eget* *hus*
'his own.NONDEF house'[106]

In both (108) and (109) the noun *hus* is not marked with a postposed definite article *hus**et***. This seems to correspond to the Lithuanian use of these adjectival modifiers, viz., predominantly in their short forms.

(110) *jo nuosavas namas* ≈ *savas namas*
'his own house'

In (110) the use of a long form **nuosav**asis*** 'own.DEF' is impossible. Rather than expressing a property, these adjectives express possession and ownership; and while they are inherently definite, they are seldom marked for it.

---

[105] Possessives here are to be understood as pronoun-derived possessives, like *my, your,* and determiner-genitives like *Peter's, mother's,* etc., as opposed to non-determiner genitives, like *aukso žiedas* 'gold.GEN.SG.[M] ring.NOM.SG.[M]' (Trakymaitė 2018, 117-122).

[106] If in (56) a modifier *alldeles* 'entirely' is inserted, the form *eget* would change to *egna.DEF*, viz. *hans alldeles egn**a** hus* 'his entirely own house'. This requires further analysis.

4.4.5. Quantifiers

A rather large group of adjectives in Table 9 seems to have something to do with the notion of quantification. Not all languages acknowledge quantifiers as a word class. In some, words expressing quantification are considered members of the pronoun class; in some others, members of the classes of determinatives or adjectives. In Lithuanian linguistics, little attention has been paid to quantification and its expression (some insights into Lithuanian universal quantifiers are to be found in Rosinas (1996, 121-131). There are two main types of quantification, viz., *existential quantification*, which "indicates a number greater than zero, and has *some* as its most straightforward expression" (CGEL 2002, 358); and *universal quantification*, which is expressed by numerous quantifiers of which *all* is the most prototypical one (CGEL 2002, 359). On the basis of the empirical findings presented in Table 9, I will split the adjectives that in some way express quantification into 4 separate groups, based on the semantics of their quantification, which partly match the two main known types of quantification: 1) approximatives, which overlap with existential quantification; 2) 'dispersed' quantification[107]; 3) ordinatives[108], which borrowed their name from the Swedish tradition to refer to words like *nästa* 'next, further, subsequent', *första* 'first, initial, prime', *sista* 'last, final, ultimate' and *förra* 'previous, preceding' as 'ordinative pronouns' due to their partial resemblance to ordinal numbers; and 4) and universal quantifiers.

Speaking of the semantics of quantification, there is a close correlation between quantifiers and the notions of uniqueness and inclusiveness, attributed to the category of definiteness (Lyons 2003). Uniqueness implies that the number (both at individual and at generic reference level) is one. Speaking about uniqueness, Lyons notes that "the definite article signals that there is just one entity satisfying the description used. This uniqueness is generally not absolute but is to be understood relative to a particular context." (2003, 8). Yet, in instances of the use of count nouns in plural or mass nouns, or collective nouns in the singular, but referring to non-singular concepts, we evoke the concept of inclusiveness rather than uniqueness: "the reference is to the totality of the objects or mass in the context which satisfy the description" (Lyons 2003, 11). Summarising, it can be said that "the uniqueness clause

[107] I am grateful to Axel Holvoet for the suggested term.

[108] It could be argued that ordinatives deal not with quantification, but rather with location modification, as they specify the placement in a specified order or rank in a series (or taxonomy). As will be shown, ordinatives differ from the other sub-groups in their use of long forms.

can be reformulated as inclusiveness or totality" (Lyons 2003, 265). Moreover, quantification *per se* is a reference-assigning mechanism, as "it derives from the ability to perceive something as a token, an instance of a class of referents, and the ability to differentiate between one and more than one (i.e., the 'plurality' of) instances of the referent" (Kibort, Corbett 2008). Turning these concepts around, one could say that indefiniteness is associated with non-totality of objects or mass, approximative values and cardinality, which singles out a certain known number of referents of a class, or possibly one, but does not necessarily make them definite.

4.4.5.1. Approximatives (resembling multal and paucal quantifiers)

Six adjectives in Table 2 seem to denote quantities or approximative values. These six consistently take only paradigms of short forms. They resemble the mid-scalar quantifiers mentioned by Haspelmath (see 4.4) in that they denote properties of unidentified degree that are scalar both in larger and smaller quantities. CGEL calls these two types of quantification multal and paucal quantification, respectively (CGEL 2002, 365–366). This group includes *gausus* 'abundant, numerous, plentiful, ample, bountiful', *nemažas* 'considerable, not small', *reikšmingas* 'significant, meaningful, important, weighty' and *menkas* 'meagre, insignificant, poor'. To expand the group, I have included an additional adjective, similar in meaning and function, viz., *pakankamas* 'sufficient, adequate, enough', which does not show long forms, and also but belongs to this group by virtue of its meaning.

In NPs modified by approximatives, reference is made to quantity and not to an attributive property of the N, e.g., *nemažos pajamos* 'significant.PL income.PL'[109] . Their values are truly mid-scalar, lining up between little and much, e.g.:

LITTLE > menkas > pakankamas > nemažas > gausus > reikšmingas >MUCH

(111) *Reikšmingas susirinkusiųjų skaičius siekė kelis šimtus.*
'The significant number of attendees reached several hundred.'

In (111), 'significant number' means, simply, a rather high number, approaching the scalar endpoint 'much'.

Even though, formally, some members of this category, e.g., *gausus* 'abundant', *menkas* 'meagre', seem to be able to form grade-like expressions

[109] In Lithuanian *pajamos* 'income' is a *plurale tantum* noun.

(see the comparative degree of *gausus* below in (112)), their values still seem to remain approximative and mid-scalar, with no defined values, as there is no fixed reference point:

(112) *Gaus**esn**is derlius kainų augimą pristabdys.*
'Higher yields will slow down price growth.'

On the above-mentioned scale, they would just take a place close to one of the relative end-points:

LITTLE > **menkesnis** > menkas > nemažas > gausus > **gausesnis** > ryškus > reikšmingas >MUCH

Observing the adjectival quantifiers in the sub-group of approximatives, one gets the impression that their predominant use with short forms has to do with their ability to establish categories, the semantics of which are [NUMBER] + [APPROXIMATE VALUE]. As mentioned above, they partially overlap with existential quantification in that that they, too, always indicate a number higher than zero; in most of the cases, higher than one, as they often appear in collocations with plural nouns or mass nouns. These adjectives are incompatible with the sole Lithuanian definiteness marker in that it always carries the [+DEF] value and signals a definite referent because the referent of an NP modified by an approximative can never be definite. They cannot establish categories like *gaus**ieji*** 'abundant.DEF' → 'ALL those that are abundant' because the property denoted by them is too vague to serve as a basis for classification. It could be said that this category is the embodiment of the semantics of indefiniteness. An adjective like *ryškus* 'bright/significant' belongs to this category because of its intensifying properties, as in *ryški pergalė* 'a significant victory' and not because of its brightness, as in *ryšk**ioji** žvaigždė* 'the bright.DEF star'. The two displaced modifiers *retas* 'rare, infrequent' and *dažnas* 'frequent, repeated, periodic' also belong here in their quantifying meanings (NB in the examples the singular NPs *retas vaikas*, literally 'a rare chid' and *dažnas darbdadvys*, literally 'a frequent employer' refer to multiple referents, as demonstrated in translation), e.g.:

(113) *Retas [kuris] vaikas suprato egzamino užduotį.*
'Few children understood the exam task.'

(114) *Dažnas darbdadavys nuolat skubina darbuotojus.*
'Many employers are always rushing their employees.'

4.4.5.2. ‘Dispersed’ quantification

‘Dispersed’ quantification is a term that I will use to describe a category of quantifiers that typically indicate a small number and a dispersed occurrence. An adjective like *pavienis* ‘isolated, single, solitary’ (Swed. *enstaka*), which I added to this group, could act as a prototype for these quantifiers. The two similar adjectives found in Table 9 are *atskiras* ‘separate, individual, special, distinct, detached’ and *individualus* ‘individual, separate, distinctive, special’:

(115) *Tyrinėtojai dažniausiai rėmėsi pavieniais pavyzdžiais.*
‘Researchers mostly relied on individual/isolated examples.’

(116) *Jie nagrinėjo atskiras bylas.*
‘They dealt with individual cases.’

(117) *Komanda bandė varžovus įveikti individualiais veiksmais.*
‘The team tried to beat the opponents with individual actions.’

(118) Swed. *Enstaka diabetesläkemedel har även visats skydda mot kardiovaskulära och renala komplikationer.*
‘Occasional diabetes drugs have also been shown to protect against cardiovascular and renal complications.’

Rather than denoting a criterion from which a category is established, these quantifiers (like similatives and variatives) epitomise the structure of the category. Therefore, since the long adjectival forms denote a category (either taxonomic or *ad hoc*) based on a qualifying property, they are incompatible with dispersed quantifiers and other groups of quasi-determiners mentioned herewith.

4.4.5.3. Ordinatives

This group of adjectives is discussed among quantifiers due to their parallels with ordinal numbers. Ordinatives are one of the categories of quantifiers that differ in terms of the use of long forms. This category exhibits mixed behaviour, with some of its members, viz., *galutinis* ‘last, final, end’ and *tolesnis* ‘further, subsequent’, predominantly used in their short forms (the count for long forms is less than 1%). These are the two adjectives to be found in Table 9. As is customary, I have added a few more adjectives with similar meanings to expand the group, i.e., *ankstesnis* ‘previous, former, preceding’, *paskutinis* ‘final, last, ultimate, end’ and *vėlesnis* ‘later, subsequent, posterior’, which have a significantly higher count of long forms. Morphologically, they are different from others, too, because they are formed with either the comparative degree suffix *–esn-* (*tol-esn-is, anskt-esn-is, vėl-esn-is*), which allows for long-form paradigms, or with the suffix typically used to form

relational adjectives (which do not have long form paradigms), *–in-* (*paskut-in-is*, *galut-in-is*)[110]. Yet, in this case it does not seem to prevent ordinative adjectives from developing paradigms of long forms, as in *paskutin-ysis*, *galutin-ysis*, *aukštutin-ysis* 'upper', *žemutin-ysis* 'lower'. This could be explained by the inherent semantic definiteness of ordinatives with or without added definiteness markers. The mixed morphological pattern could partially explain the mixed use of short vs long forms in NPs containing these adjectival modifiers[111]. In other languages with elaborate (in)definiteness marking and adjectival definiteness marking, e.g., continental Scandinavian and Dutch, ordinatives serve as a potential source for acquiring new determinatives (see 4.3.3), e.g.:

(119)	Swed.	*Jag*	*längtar*	*efter*	*först**a***	*skön**a***	*sol**en***.
		I	long.PRS	for	first.DEF	beautiful.DEF	sun.DEF
		'I long for [the] first beautiful sun.'					
(120)		*näst**a***	*lång**a***	*etapp*			
		next.DEF	long.DEF	stage			
		'[the] next long stage'					

In (119) and (120) ordinatives are used instead of definite articles and all the following attributes are compulsorily marked for definiteness (used in their weak forms) as well.

4.4.5.4. Universal quantifiers

Universal quantifiers as a sub-group also denote semantically definite referents due to their semantics of totality (inclusiveness). This group includes the adjective *visiškas* 'complete, total, full, absolute', to be found in Table 9. I have added two other adjectives, viz., *ištisas* 'whole, entire, all' and *kiauras* 'whole, entire, all'[112]. These, however, are consistently used in the short form, which is not a typologically rare feature according to Haspelmath (1997, 11-12). Šereikaitė (2019, 85) explains this with a parallel between the use of short adjectival forms in Lithuanian and expressions of weak definiteness,

[110] As was rightly noted by a reviewer, *galutinis* differs from *auksinis* in its morphological composition, viz., it contains a morphologically complex suffix with a different prehistory; also, *auksas* + *-in*= *auksinis*, but *galas* + *-in*= *galinis*, not *galutinis*.

[111] Evidently, ordinal numbers, like *first*, etc., which, in Lithuanian, allow for paradigms of long forms, belong to this group.

[112] The adjective *kiauras* has another direct meaning, viz., 'holey, full of holes', which is irrelevant in this context.

which is typically linked to the notion of uniqueness; this seems quite plausible because, as demonstrated, totality is a mirror of uniqueness.

(121) *Lijo kiaurą naktį.*
'It rained all night.'

A remark needs to be added here concerning the relation existing between totality and distributivity, e.g.:

(122) *Ištisas kaimas žinojo jo paslaptį.* ≈ *Kiekvienas to kaimo gyventojas žinojo jo paslaptį.*
'The entire village knew his secret.' 'Every villager [of that village] knew his secret.'

While *the entire village* gets a definite reading, *every villager* gets a distributive reading that need not necessarily be interpreted as definite. However, both universal quantifiers and distributives like *every* exhibit the same morphosyntactic feature – they are not marked for definiteness, as they are predominantly used with short forms.

However, some peculiar examples can be found in Swedish, e.g.:

(123) Swed. *Varje god pjäs innehåller flera akter.*
'Each good play contains several acts.'
(124) *Vi skrattar åt varje minst**a** lill**a** sak.*
'We laugh at every smallest.DEF tiny.DEF thing.'
(125) *varje hans gärning*
'his every deed' → literally 'every his deed'[113]
(126) *varje först**a** entusiastiskt försök*
'every first.DEF enthusiastic attempt'

Example (123) demonstrates the expected use of short adjectival forms after the distributive *every*. However, examples (124) – (126) show the universal interpretation of *every* because the following attributive adjectives are marked for definiteness. SAG explains the use of long adjectival forms in (125) as a lexicalised link between the distributive pronoun *every* and the superlative *smallest* (note that both adjectives agree in definiteness marking); (125) as an outdated use (in contemporary Swedish a possessive would be placed first); and (126) as a possible, but rare, use (note that only an ordinative is marked for definiteness, the second attribute is not). This is reflected in the changed structure of the Swedish NP and the changed status of various determiners due to the parallel universal and distributive readings of *every*.

[113] Examples (125) and (126) are taken from the Swedish Academy Grammar (SAG Vol. 2, 385).

Summing up, one could say that universally quantifying adjectives are used with their short forms in inherently definite NPs due to the semantics of totality or uniqueness.

4.4.6. Summary of the findings

The 30 adjectives in Table 9 have been divided into 4 major groups:

Table 10. Groups of adjectives based on Table 9 – summary of findings

No	Groups of adjectives	
1.	**Adjective-like adjectives** that exhibit 'proper' adjective-like behaviour (this group also includes the so-called absolute and mid-class modifiers, which have some irregular features), e.g., *švarus* 'clean'; absolutives like *optimalus* 'optimal'; and mid-class modifiers like *būdingas* 'typical, characteristic'	
2.	Adjectives that do not function like adjectives, viz., **displaced modifiers or adverbiatives** where sentence-level modification is downgraded to nominal level, e.g., *dažnas* 'frequent', *aiškus* 'evident'	
3.	**Verb-like adjectives** that have complex, verb-like argument structures, e.g., *vertas* 'worthy', *reikalingas* 'needed'	
4.	Adjectives that function as **quasi-determiners**, both definite and indefinite, viz., that establish minor categories of their own due to additional functions outweighing the traditional modifying function. This group includes several sub-groups:	
	4.1.	**similatives, dissimilatives**, **variatives** e.g., *panašus* 'similar', s*kirtingas* 'different, unlike', *įvairus* 'various, varied, miscellaneous'
	4.2.	**particularising attributives**, e.g., *konkretus* 'concrete, particular, specific'
	4.3.	**possessives**, e.g., *privatus* 'private, own, personal'
	4.4.	**quantifiers**, which include 4 different sub-categories: approximatives, e.g., *menkas* 'meagre, insignificant', markers of dispersed quantification, e.g., *atskiras* 'separate, individual, distinct', ordinatives, e.g., *galutinis* 'final, ultimate, terminal' and universal quantifiers, e.g., *visiškas* 'complete, total, full, absolute'

These groups have been expanded to include a few semantically similar members, which have also been checked in the CCLL for the proportion of long vs short forms. Almost all of them, with the exceptions of *specialus* 'special' and *ypatingas* 'special, peculiar, particular', do not differ from other members in their respective groups.

All the adjectives analysed in the chapter have multiple translations into English due to their fluid meanings. In Table 8 and Table 9, I have attempted to show as many of their alternative translations as possible. Yet it was entirely possible to find a common denominator (a prototypical member) for all the various groups created in the analysis. Some of the adjectives were analysed on the basis of just some of their meanings, e.g., *ryškus* 'bright' belongs to the group of approximatives due to its metaphoric, more abstract meaning, approaching that of an intensifier, as in *ryškus skirtumas* 'glaring difference = significant difference', which follows an attested path of grammaticalisation consisting in the concrete lexical meaning (brightness) being abandoned and drifting towards the more abstract meaning of a degree modifier (intensifier), ultimately landing in the group of displaced modifiers (adverbiatives).

Many of the analysed adjectives lack scalarity and gradation, which is one of the defining features of qualitative adjectives, and consequently cannot be used with degree modifiers. This is because some of them are non-scalar in that they themselves represent the end-points of the scale, e.g., *visiškas* 'total' – **visiškesnis.* Others, while allowing gradation, e.g., *gausus* 'abundant'– *gausesnis* – *gausiausias*, disallow the use of degree modifiers as this would be ungrammatical, e.g., **visai gausus* 'quite abundant', **labai tolesnis* 'very further', **pakankamai nuosavas* 'sufficiently own'.

Another common feature is the fact that several adjectives could represent more than one category, e.g., *individualus* 'individual' could be a distributive quantifier or a specificity indicator, and even be synonymous with the absolutive *unikalus* 'unique'; *tūlas* could be a quantifier or a variative; *dažnas* could be an adverbiative or a quantifier, etc.

4.5. Conclusions

There is no one single reason why an adjective in Lithuanian shows an absence of long forms. Some adjectives do not assume long forms because of semantic-pragmatic reasons, e.g., the group of adjective-like adjectives, many of which denote properties of such unstable, impermanent, or temporary nature that they cannot establish *ad hoc* categories like long adjectival forms do. Other adjectives lack long forms because they do not denote properties,

but serve as quasi-determiners to express qualification, similarity, variation, specificity, possession, etc.

A group that stands out for a different reason is the so-called displaced modifiers because they, unlike traditional attributes, encode sentence-level modifications (like adverbials) on the level of a noun phrase. One could say that their locus and the scope of their function do not match. They therefore do not command the features typical of attributive qualitative adjectives.

Verb-like adjectives that exhibit complex verb-like argument structures unsurprisingly resemble predicates and subsequently, because of their verb-like behaviour, conceal their attributive properties and consistently appear in short forms only. This group includes a rather large number of adjectives.

The fourth and most distinct group is that of quasi-determiners (both definite and indefinite). It has been argued that attributive adjectives in Lithuanian differ from true determinatives in that they perform two functions – they modify and determine simultaneously (Trakymaitė 2018). It has also been argued that NPs containing attributive long adjectives form *ad hoc* categories (Holvoet, Sprauniené 2012); and by assuming the morphological definiteness marker an adjective loses its ability to be gradable (scalar), e.g., *balti* 'white.NONDEF' can become *baltesni* 'whiter' but *balt**ieji*** 'white.DEF' cannot. An *ad hoc* category is established on the basis of a prominent property, e.g., *white.* Many of the adjectives exhibiting the absence of long forms denote undisclosed properties (variatives, similatives), properties of undefined values (approximatives) or rather, not properties but the structure of the category itself (specificity markers, dispersed quantifiers). Therefore, they seem to be losing their attributive adjectival properties, including the opposition of long and short forms, and seem to be functioning as determinatives more than modifiers. Some of the quasi-determiners render NPs indefinite, e.g., approximatives, variatives, dispersed quantifiers, specificity markers, and consequently show almost a total absence of long forms.

Other quasi-determiners render NPs definite by virtue of their inherent semantic definiteness, e.g., ordinatives and universal quantifiers. Yet, because of the nature of their relational character they, as opposed to other qualitative adjectives, rather resemble relational adjectives that cannot acquire long forms.

In Swedish, all of these quasi-determiners (comparative pronouns like *annan* 'other', s*amma* 'same', *sådan* 'such', *likadan* 'similar/alike'; perspective pronouns like *ena* 'one', *höger* 'right', *vänster* 'left', *norra* 'northern'; focusing pronouns *själv* 'self', *egen* 'own', *enda* 'sole/one'; ordinative pronouns like *nästa* 'next', *första* 'first', *sista* 'last') are considered

to be relational pronouns because they, in various ways, relate the referent to others with regard to properties or with regard to identity (SAG 2, 236). The borderline between these relational pronouns and adjectives is, in many cases, undefined. The focus is not on any property denoted by the pronoun/adjective, but on the relation between referents. The same can be attributed to Lithuanian quasi-determiners.

As a next step in trying to determine a broader picture of the uses of Lithuanian long and short adjectives, it would be interesting to compile an alternative frequency list of Lithuanian adjectives that are mostly used with long forms and analyse how they map onto the definiteness marking system.

An interesting question briefly touched upon in this section is why certain NPs can be considered taxonomic, e.g., *šalt**oji** kava* 'the cold.DEF coffee' referring to *frappe* style coffee is an established term, while **šalt**asis** *maistas* 'the cold.DEF meals' cannot. As demonstrated in examples (66) – (69), there are numerous cases of frequently used NPs, e.g., *būding**ieji** bruožai* 'typical.DEF features' that I would not classify as established taxonomies. A further and deeper study of the established taxonomy of NPs would provide interesting material and insight into the overall ability to establish categories and classifications of the Lithuanian language.

In addition, the analysis of the features typical of the attributive adjectives, such as ordering restrictions and scopal implications (identified by Belk and briefly mentioned in 3.2) affecting interpretation and linear ordering of attributes, is yet another field waiting to be researched.

Finally, the issue of short forms used in established terminology, e.g., *balta* vs *juoda duona* 'white vs black bread' (the latter being bread made from rye flour), *geltonas sūris* 'yellow cheese' referring to a type of fermented cheese, *mobilus* and not *mobil**usis** 'mobile.DEF' as an established nominalisation for a mobile phone, needs to be researched and analysed.

Concluding, I would like to add that data-driven research has led me from Lithuanian asymmetry of uses of long versus short adjectival forms to cross-linguistically established minor categories of quasi-determiners, independently established and described in, e.g., CGEL or SAG. I hope that this chapter ill serve as yet another puzzle piece in solving the adjectival (in)definiteness marking in Lithuanian and other languages.

5. ADJECTIVES PREDOMINANTLY AND FREQUENTLY USED WITH LONG FORMS

5.1. Background

As mentioned in section 4.1.2, in order to understand (and gather some quantitative data on) the extent of the use of LF versus SF of adjectives, a list of 111 adjectives able to assume LFs was created based on the data presented in the FrD (Appendix A). For these 111 adjectives, the distribution of their LF and SF was calculated based on the counts displayed under each individual entry in FrD (see Appendix C for a detailed description and data). While section 4 focused on those showing an absence of uses with LFs, this section will take a closer look at those that are predominantly used with LF.

Even though several studies of the Lithuanian ADM have been published in recent years, few tangible and quantified data are available on the distribution of uses of LFs and SFs. One particular adjective, *pastar**asis*** 'the latter, recent, the aforementioned', has been recognised as one almost solely used with its LFs (Sprauniené 2011, 93). A question that this chapter tries to answer is: are there more? When looking at the distribution of LFs and SFs, researchers have focused on different linguistics contexts and explanations (Parenti 1995, Kamandulytė, Tuškevičiūtė 2008). Šereikaitė in her recent study of Lithuanian adjectives links the distribution of LFs and SFs to the concepts of strong (based on familiarity) *versus* weak (based on uniqueness) definiteness (Šereikaitė 2019). In a study of the structure of attributive and predicative utterances containing adjectives (Kamandulytė-Merfeldienė, Balčiūnienė 2016, 129-130), it is claimed that the attributive adjectives of interest to us here are more frequent (76%) in official speech situations as compared with spontaneous non-official speech[114] (46%). Following this line of thought we could say that long adjectival forms in their attributive function would be less likely to appear in spontaneous spoken language. In her study of the use of LF in the corpus of the spoken language, Sprauniené (2011, 115) points out that the majority of uses of LFs (86%) appear in the official registers of the spoken language, viz., media, academic and semi-public communication (e.g., less familiar conversations amongst colleagues, communication in service-providing situations and similar), in situations that resemble the "prepared speech" type

[114] The actual term used by the authors is *spontaninė neoficiali kalba* (Kalmandulytė-Merfeldienė, Balčiūnienė 2016, 129-130) as opposed to *oficiali kalba* and *viešoji kalba* 'the public speech'.

of texts. Moreover, 58% of these LF uses were taxonomic; hence the factors determining the use of LF are not only associated with the communication situation, but also with the type of definiteness denoted by them.

Summarising previous research, one could say that expectations of finding LFs in data of the spoken language should be fairly low. Additionally, the LF used in nominalisations, e.g., *balt**ieji*** 'the white.DEF [ones]' meaning the white pawns in the game of chess, make up less than 0.5% of all the adjectives in the above-mentioned study by Kamandulytė-Merfeldienė and Balčiūnienė (Kamandulytė-Merfeldienė, Balčiūnienė 2016, 130, footnote (4)). This will be of importance in the data analysis ahead.

With this in mind, we set out to search for linguistic evidence of adjectives predominantly used with the long (definite) forms in contemporary Lithuanian.

5.2. Data and method

Based on the list of 111 most frequently used adjectives that take LFs (see Appendix A), the percentage of distributions between LFs and SFs were calculated. Subsequently, the normal distribution curve was fitted to the data (see Appendix C) and p-values were calculated to identify the statistically significant instances of counts of LFs. With a relaxed p-value of 0.10, a list of 9 adjectives was compiled (Appendix D) that included the following: *didis* 'great, famous, sublime', *pastaras* 'the latter', *vyresnis* 'older, senior, superior', *šventas* 'holy, sacred, saint, blessed, sacrosanct', *ankstyvas* 'early, precocious, premature', *viešas* 'public, open' and *gimtas* 'native, inborn, innate', which according to the statistics display a particular behaviour with regard to usage, viz., they predominantly appear with the LFs.

Table 11. List of FrD adjectives with p-value < 0.10

No	Adjective	Translation	% of LFs in FrD	p-value (< 0.10)
1	pastaras	the latter/the recent	100.00	p = 0.0000
2	didis	great/famous/sublime	90.96	p = 0.0000
3	gimtas	native/born/innate	89.16	p = 0.0000
4	vyresnis	older/senior/superior	66.53	p = 0.0023

No	Adjective	Translation	% of LFs in FrD	p-value (< 0.10)
5	viešas	public/open	56.84	p = 0.0099
6	šventas	holy/sacred/saint/blessed/sacrosanct	48.82	p = 0.0282
7	ankstyvas	early/precocious/premature	48.04	p = 0.0310
8	senas	old/ancient	43.85	p = 0.0499
9	grynas	pure/net/neat/clear/clean	41.84	p = 0.0617

Knowing that attributive long adjectival forms are less likely to appear in spontaneous speech situations, an additional data test was applied. The mentioned 9 adjectives were tested in the Sketch Engine family corpus of Lithuanian-language texts collected from the Internet, Lithuanian Web 2014 or ltTenTen14, which reflect speech situations (web-based texts, internet commentary and similar[115]) more resembling that of spontaneous non-official speech. The hypothesis was that if the adjectives listed above maintain their statistically significant high counts of LFs, they truly are the ones predominantly used with LFs in both written (as attested by the Frequency Dictionary) and spoken language.

The complete paradigms of both long and short forms (including those typical of colloquial usage, i.e., the locative without the final vowel *-e* (in the tables "Loc.-e"), the feminine '-om' case, which could be used for both plural dative and instrumental (in the tables "-om") were compiled, as illustrated below in Table 12, Table 13, Table 14 and Table 15 with the paradigm of *šventas* 'holy, sacred, saint, blessed, sacrosanct':

Table 12. *šventas* – MASCULINE, NONDEF (SF) paradigm

Sg M		Instances	**Pl M**		Instances
Nom.	*šventas*	8396	**Nom.**	*šventi*	2151
Gen.	*švento*	4706	**Gen.**	*šventų*	3569
Dat.	*šventam*	523	**Dat. +s**	*šventiems*	174

[115] See https://www.sketchengine.eu/lttenten-lithuanian-corpus/ for more information. Also, https://www.sketchengine.eu/documentation/tenten-corpora/ for more on the TenTen family of corpora.

Sg M		Instances	**Pl M**		Instances
			Dat. -s	*šventiem*	7
Acc.	*šventą*	4833	**Acc.**	*šventus*	1133
Inst.	*šventu*	1701	**Inst.**	*šventais*	929
Loc. +e	*šventame*	334	**Loc. +e**	*šventuose*	161
Loc. -e	[*šventam*]	= DAT.SG.M	**Loc. -e**	*šventuos*	1
Total:		**20493**	**Total**		**8125**
Total of SF of *šventas*					**28618**

Table 13. *šventa* – FEMININE, NONDEF (SF) paradigm

Sg F		Instances	**Pl F**			Instances
Nom.	*šventa**	7734	**Nom.**	[*šventos*]		= GEN.SG.F
Gen.	*šventos*	4002	**Gen.**	[*šventų*]		= GEN.PL.M
Dat.	*šventai***	3568	**Dat.**	*šventoms*		151
Acc.	[*šventą*]	= ACC.SG.M	**Acc.**	[*šventas*]		= NOM.SG.M
Inst.	[*šventa*]	= NOM.SG.F	**Inst.**	*šventomis*		793
Loc. +e	*šventoje*	659	**Loc. +e**	*šventose*		429
Loc. -e	*šventoj*	65	**Loc. -e**	[*šventos*]		= GEN.SG.F
'-om'	n/a	only PL.F	**'-om'**	*šventom****		94
Total		**16028**	**Total**			**1467**
Total of SF of *šventas*						**17495**

The total count of the short forms of the adjective *šventas*, both masculine and feminine, is **46113**. The following instances of homonyms were evident in the corpus, but could not be filtered manually due to high counts:

***** *šveñta* (ADJ, NEUTER) ≠ *šventà* (ADJ, NOM.SG.F)

(127) *Pasakiau* *ir* *šventa!*
say.1SG.PST and sacred.NEUT.
'Because I said so full stop!'

****** *šveñtai* (ADJ, DAT.SG.F) ≠ *šventaĩ* (ADVERB)

(128) *Tuo* *šventai* *tikiu* *iki* *šiol.*
this.INST.SG.M sacredly believe.1SG.PRS until now
'I totally believe in that'

******* *šventom* = *šventoms* (ADJ, DAT.PL.F) = *šventomis* (ADJ, INST.PL.F)

In colloquial use, this form could function as a substitute to the two above-mentioned plural cases, dative and instrumental. Also, the feminine paradigm of plural definite forms is used with it, as is demonstrated in Table 15.

Table 14. *šventasis* – MASCULINE, DEF (LF) paradigm

Sg M		Instances	**Pl M**		Instances
Nom.	*šventasis*	13787	**Nom.**	*šventieji*	4358
Gen.	*šventojo*	14714	**Gen.**	*šventųjų*	11671
Dat.	*šventajam*	1401	**Dat. +s**	*šventiesiems*	660
			Dat. -s	*šventiesiem*	3
Acc.	*šventąjį*	4779	**Acc.**	*šventuosius*	2208
Inst.	*šventuoju*	3385	**Inst.**	*šventaisiais*	1446
Loc. +e	*šventajame*	1916	**Loc. +e**	*šventuosiuose*	165
Loc. -e	[*šventajam*]	= DAT.SG.M	**Loc. -e**	*šventuosiuos**	0
Total:		**39982**	**Total**		**20511**
Total of LF of *šventasis*					**60493**

* The colloquial locative (Loc.-e) occurred once in the corpus, but as was the case with many colloquial locatives of examined adjectives, it was a typo[116], viz., the incorrect ending *-uosiuos* was used for plural accusative.

Table 15. *šventoji* – FEMININE, DEF (LF) paradigm

Sg F		Instances	**Pl F**		Instances
Nom.	*šventoji**	8890	**Nom.**	[*šventosios*]	= GEN.SG.F
Gen.	*šventosios*	21419	**Gen.**	[*šventųjų*]	= GEN.PL.M
Dat.	*šventajai*	1283	**Dat.**	*šventosioms*	47
Acc.	*šventąją*	5563	**Acc.**	*šventąsias*	701
Inst.	*šventąja*	2004	**Inst.**	*šventosiomis*	89
Loc. +e	*šventojoje*	4065	**Loc. +e**	*šventosiose*	173
Loc. -e	*šventojoj*	90	**Loc. -e**	[*šventosios*]	= GEN.SG.F
'-om'	n/a	only PL.F	**'-om'**	*šventosiom*	4
Total		**43314**	**Total**		**1014**
Total of SF of *šventas*					**44328**

[116] It is difficult to say whether it was a typo or simply inability to select the right ending. In my opinion, it could be both as double diphthongs could be difficult to both correctly select and/or type. See Appendix D for all the detailed data from the ltTenTen14 corpus.

* The corpus contained several instances of the toponym *Švent**oji***, a Lithuanian seaside resort. These were not eliminated from the data count as the form originates in the nominalisation of *švent**oji*** 'holy.NOM.SG.F.DEF'.

The total count of the long adjectival forms of the adjective *švent**asis***, both masculine and feminine, is **104821**. The distribution of SF versus LF is 30.55% versus 69.45% respectively, which means that well above 50% of all forms of this adjective encountered in the web corpus of modern Lithuanian are long (definite). For a detailed analysis of all the 9 adjectives in the ltTenTen14 corpus see Appendix D.

Obviously, the data acquired is not absolute, as a few technical challenges needed to be addressed. Firstly, there was the issue of homonymy. As shown in Table 13 and the comments below, a few adjectives similarly to *šventas* have homonymous forms "hidden" in NOM.SG.F (homonymous with the neuter gender forms of respective adjectives) and in DAT.SG.F (homonymous with adverbs). Other adjectives have some additional instances of homonymy (see, e.g., *didis* in Appendix D). Since it was physically impossible to fine-tune data because of the high numbers and lack of annotation, I chose not to remove the counts of homonymous forms from the short form paradigms, but rather to keep them in the calculations. Since the goal was to identify the percentage of LFs used, the logic behind including the counts of SF homonyms that could not be dealt with manually due to high counts was that it would potentially decrease the percentage of LFs (as the total count of short ones would increase, the total count of the long ones would automatically decrease, increasing the chance of them being removed from the list of the atypically behaving adjectives, viz., predominantly appearing with LFs against expectations). Therefore, those adjectives that still showed a very high percentage of LFs would be of significance.

Secondly, the instances of nominalisations and terminology-like uses (cases of taxonomic definiteness) had to be addressed. In this study, I have chosen to keep them in my calculations for the following reasons:

- As has been established, all long adjectival forms always encode definiteness both on the level of individual reference and on that of generic use (taxonomy) (for a more detailed discussion, see 4.2.2).
- In case of nominalisations, the characterising feature of which is the notion of maximal inclusivity, they could be considered elliptical structures with omitted nouns, as the property denoted by the adjective is key to establishing a category and key to identifying referents as belonging to this category, e.g., *šventieji* 'the saints' → ALL holy persons who have died and have been declared saints by

relevant authorities. The property of sanctity is essential to the extent that it can substitute the noun in an NP and function as a category denominator on its own.

Thirdly, I believe that the issue with regard to the syntactic position of adjectives with LF must be addressed, as obviously we are only interested in attributive and not in predicative uses of adjectives. It must be noted that they do not appear predicatively (in cases of both primary and secondary predications, see Section 4.3.2 on this matter):

(129) **Kariai buvo švent**ieji**.* versus *Kariai buvo šventi.*
'The warriors were holy.DEF.' 'The warriors were holy.NONDEF.'

They can only appear predicatively if they are part of established taxonomic NPs as in (128) *šventieji kariai*, or if they themselves are nominalisations (129), which as explained above, could be interpreted as elliptical taxonomic NPs.

(130) *Jie buvo švent**ieji** kariai.*
'They were holy.DEF warriors.'
(131) *Jie buvo švent**ieji**.*
'They were holy.DEF. → 'They were saints.'

Therefore, I do not think that the analysis of the syntactic function of the LF would be of relevance here. The same can be said about SF, as even if they can and do appear predicatively, removing them from the pool of instances of SFs would only weaken the test of my hypothesis following the same logical pattern as in the case of homonymy explained above.

5.3. Findings

After the initially established data (the list of adjectives in Table 11) was run in the ltTenTen14 corpus, the following distribution of LFs *versus* SFs for the 9 chosen adjectives has been established:

Table 16. Adjectives ranked highest to lowest based on the share of LFs in ltTenTen14

No	Adjective	Translation	% of LFs in ltTenTen14	% difference between preceding and succeeding adjectives
1	pastaras	the latter/the recent	99.88	n/a
2	didis	great/famous/sublime	92.75	7.13
3	gimtas	native/born/innate	83.47	9.28
4	šventas	holy/sacred/saint/blessed/taboo	69.45	14.02
5	viešas	public/open	67.33	2.12
6	ankstyvas	early/precocious/premature	42.57	24.76
7	vyresnis	older/senior/superior	41.88	0.69
8	senas	old/ancient	39.89	1.99
9	grynas	pure/net/neat/clear/clean	36.79	3.10

It has been established, as demonstrated, that 5 adjectives (No 1 to 5 above) are predominantly used with their long (definite) forms. *Didis*, *gimtas*, *šventas*, and *viešas*, along with the anaphoric *pastarasis*, form a core of Lithuanian adjectives used with long forms. Another 4 adjectives remain below the 50% limit when it comes to their appearance with LFs.

The adjectives seem to have split into two distinct groups marked by the significant increase in the percentage difference (24.76) between *viešas* and *ankstyvas*. More thorough statistical data would be needed to explain this increase. This, however, will not be addressed here. Let us recall that that all the 9 adjectives in FrD[117], which is based on representative samples of the written language, have p-values less than 0.10, with adjective *grynas* having the highest one of p = 0.0617, which still is closer to the traditionally recognised one of p = 0.05. Moreover, as mentioned, the list could be further expanded to include additional 8 adjectives if the p-value is further relaxed to p = 0.25 (see Appendix C). This represents tangible quantified data on the uses of predominantly LFs.

In an attempt to analyse and group the findings, I have chosen to look at all 18 of them:

[117] Frequency Dictionary of the Written Lithuanian Language (Utka 2009): http://donelaitis.vdu.lt/publikacijos/Dazninis_zodynas.pdf

Table 17. Adjectives with the relaxed p-value ≤ 0.25 ranked according to p-values (smallest to highest)

No	Adjective	Translation	% of LFs in FrD	% of LFs in ltTenTen14	p-values
1	pastaras	the latter/recent	100	99.88	0.0000
2	didis	great/famous/sublime	90.96	92.75	0.0000
3	gimtas	native/born/innate	89.16	83.47	0.0000
4	vyresnis	older/senior/superior	66.53	41.88	0.0023
5	viešas	public/open	56.84	67.33	0.0099
6	šventas	holy/sacred/saint/blessed/sacrosanct	48.82	69.45	0.0282
7	ankstyvas	early/precocious/premature	48.04	42.57	0.0310
8	senas	old/ancient	43.85	39.89	0.0499
9	grynas	pure/net/neat/clear/clean	41.84	36.79	0.0617
10	tikras	true/real/proper	36.44	-	0.1043
11	idealus	ideal/perfect	34.12	-	0.1281
12	aukštas	tall/high	30.78	-	0.1685
13	garsus	famous/prominent/known/loud	29.20	-	0.1903
14	juodas	black	29.39	-	0.1876
15	laisvas	free/liberal	29.83	-	0.1814
16	amžinas	eternal/perpetual/everlasting/timeless	29.69	-	0.1833
17	jaunas	young/youthful/adolescent	26.06	-	0.2382
18	bendras	common/joint/shared/general	29.19	-	0.2526

It is challenging to group the findings in clearly defined semantic or functional groups. I will try to summarise and provide some tentative groups[118].

Anaphora and Deixis

*Pastar**asis*** in its meaning 'the latter' is clearly anaphoric. In its meaning 'recent.DEF', alongside *vyresn**ysis*** and *ankstyv**asis***, it is deictic. *Vyresnis* also has a semantic feature of [+animacy] and refers to someone/something that is older than X; and *ankstyvas*, which is anaphoric in that it implies some relative temporal relations, viz., someone/something is earlier/premature than X and that that someone/something could be considered *vėlyvas* 'late, belated, tardy'.

Totality

The adjective *bendras* in its meaning 'joint/shared', e.g., is clearly quantifying, e.g., *bendr**oji** rinka* 'the single.DEF market' speaking of the EU internal market. It could also be interpreted as such in its meaning 'common/general', e.g., *bendr**ieji** planai* 'general.DEF plans' as opposed to *special**ieji** planai* 'particular.DEF plans' speaking of municipal/urban developments. The meanings 'general' and 'common' seem to have correlation with a universal quantifier *all*, viz., *common → common to all*, as in *bendr**asis** Vilniaus planas* 'general.DEF plan of Vilnius' → *the plan of the whole city, not only of certain parts*.

Intensity and scalarity

Tikras and *grynas* both denote intensity, degree and possibly quantification, e.g., *gryn**asis** pelnas* 'net.DEF profit' ≈ *tik pelnas, vienas pelnas* 'sole profit, profit after all the necessary subtractions', *tikr**oji** vertybė* 'true.DEF virtue', *gryn**oji** vilna* 'pure.DEF wool'. One could also say that these two denote proximity to a prototype, scalarity of some sort in that they denote a scalar endpoint, e.g., on the scale describing the purity of wool, one end is occupied by *gryn**oji** vilna* as opposed to not so pure wool. Also, *idealus*, which similarly to *optimalus* (see Section 4.4.1 for analysis) is an absolute adjective denoting a scalar endpoint and therefore non-gradable[119].

[118] This is not an established classification but rather an attempt to reflect on the statistical evidence in search of common denominator. More thorough studies of the subject are needed to make linguistic claims.

[119] The superlative forms are attested: *idealiausias*, but there is no scalar difference between *idealus* and *ideliausias*. The comparative *idealesnis* is also attested but has a very low frequency and is seldom used.

Adjectives that represent culture-bound concepts

These with the exception of *pastaras* include the 4 adjectives predominantly used with LFs *didis* [+animacy, human], *gimtas* which is inherently definite and unique, *viešas* and *šventas*. In addition to these, I would add *garsus* [+animacy, human], *amžinas* and *laisvas* [+animacy]: *did**ysis** kunigaikštis* 'grand.DEF duke, *gimt**oji** šalis* 'native.DEF country', *viešasis transportas* 'public.DEF transport', *švent**oji** giraitė* 'holy.DEF grove', *gars**ioji** frazė* 'famous.DEF phrase', *amžin**asis** gyvenimas* 'eternal.DEF life', *laisv**oji** rinka* 'free.DEF market'. All these describe results of human activity and culture rather than properties existing independently of human activity like e.g., colours that are perceived through a physical process of seeing (visual perception).

Typologically frequent adjectives describing social and cultural hierarchies

This group would include the antonym pair *senas* and *jaunas*, as well as *aukštas* and *juodas*. The typological underlying motivation for certain semantic types of adjectives is presented and discussed in Section 2.3. Undeniably, these adjectives can be used as simple attributes in their core meanings to describe all types of nouns, e.g., *sena sofa* 'an old couch', *jaunas rašytojas* 'a young writer', *aukštas stulpas* 'a high pillar', *juodas katinas* 'a black cat', all denoting properties at the core of their semantic frames. However, they can also be used to establish categories describing relations and positions in social and cultural hierarchies, e.g. *aukšt**oji** kultūra* 'high.DEF culture', *jaun**oji** karta* 'young.DEF generation', *sen**asis** Vilnius* 'old.DEF Vilnius' not in the meaning of age, but rather in contrast with other parts of the urban development of the city (restrictive attribute), similarly *Juod**oji** Vokė* lit.'black.DEF Vokė', a toponym, a settlement and *Balt**oji** Vokė* lit. 'white.DEF Vokė', *juod**oji** magija* 'black.DEF magic', etc. In most of these uses, they represent cases of taxonomic reference. These adjectives are in some way fundamental to our conceptualisation of reality, which is reflected in their typological universality. That is why it is perhaps natural that they more than others are put to use in categorising and organising specific domains of experience where they acquire metaphorical senses. That is why they are prone to establishing taxonomic categories.

5.4. Conclusions of section 5

In an attempt to quantify the distribution of long and short adjectival forms in contemporary Lithuanian, a distinct group of adjectives (p-value < 0.05) that are predominantly used with LFs has been identified based on statistical data

tested on the corpora of both written and web-based language use, which in genre is closer to the spoken language. Their definiteness (except for *pastaras*) mostly reflects socially and/or culturally determined identifiability.

Others, in the bigger group, with a relaxed p-value $\leq 0{,}25$, show a frequent use with LFs. This group contains several typologically attested universal adjectives that seem to be fundamental to human categorisation and conceptualisation of reality in both their direct and metaphorical meanings.

6. CONCLUSIONS

The aim of this dissertation was to analyse and describe the Lithuanian adjectival definiteness marking, to place it in a typological context and hence to deepen the understanding of it, and to contribute to cross-linguistic research into this phenomenon. Lithuanian with the sole morphologically grammaticalised marker of definiteness, viz., a dedicated set of endings of the long form adjectival paradigm, represents a prototypical case study for the research of ADM as a cross-linguistic feature. Much of the thinking and inspiration came from the studies of other European languages that also have the grammatical feature of the ADM in their inventories. Quite a few terms and analysis models have their roots in the studies of continental Scandinavian, Dutch and some Slavic languages. Most data (both Lithuanian and that of other languages) has been collected from the corpora of contemporary languages reflecting the actual usage.

In the data analysis of sections 3, 4 and 5, three different synchronic cross-sections of Lithuanian adjectives were presented, revealing the linear structure of the definite Lithuanian NP, as well as providing some tangible data on the actual distribution of long and short adjectival paradigms in the contemporary language, identifying some groups of Lithuanian adjectives that do not take long adjectival forms and those that are predominantly used in their long forms, and hence also contributing to the descriptive classification of the Lithuanian adjectives.

It has been shown that ADM is a morphosyntactic feature with implications for both the category of definiteness and the syntax of the noun phrase.

It has been shown that ADM directly correlates with the syntactic structure of the NP whereby the feature [+DEF] may be and often is encoded in several loci with different degrees of impact. It also has been shown that the different loci of [+DEF] correlate with weak *vs* strong types of definiteness, where NP constituents in the left periphery encode discourse-bound definiteness while constituents in the right (closer to the head) denote identifiability-based definiteness.

It has been demonstrated that there is no one reason why some Lithuanian prenominal adjectival modifiers are predominantly used with short forms. This is due to the fact that:

- the properties they denote are not suitable for establishing categories;
- they denote not properties, but something else, e.g. (dis)similarity, quantification, variations, etc.

I have also found that a small group of Lithuanian adjectives is predominantly used in their long forms. Definiteness expressed by these adjectives mostly reflects socially and/or culturally determined identifiability.

Trying to apply the lessons learned to the descriptive Lithuanian grammar, the following questions arise requiring further study and research:

- There is a need for a re-think of traditional word classes to accommodate some aspects directly resulting from the ADM, i.e., should quasi-determiners be acknowledged; should quantifiers be acknowledged; should indefiniteness markers be acknowledged?
- There is a need for a more comprehensive classification of Lithuanian adjectives on a par with those included in the 2004 work edited by R.M.W. Dixon and Alexandra Y. Aikhenvald entitled *Adjective Classes: A Cross-linguistic Typology*. Such a study could substantially widen our understanding and description of this part-of-speech.
- There is a need for recognition and research on cross-linguistic minor categories such as, similatives, dissimilatives and variatives (supported by the morphological argument → *toks* 'such', *kitoks* 'of another kind, different', *visoks* 'of all kinds, of all sorts').

With regards to the adjectival definiteness marking itself, even though many aspects of the ADM have been covered, many still remain to be researched from a typological perspective, including, amongst others, the following:

- The ADM and the typology of proper names: uses like *sen**ieji** Kazlauskai* 'old.DEF Kazlauskas [family]', *įžym**usis** Čiurlionis* 'famous.DEF Čiurlionis', *Nauj**ieji** Trakai* 'new.DEF Trakai [toponym]', and similar;
- The ADM in vocative constructions: *Brang**usis**!* 'dear.DEF', *Gerbiam**ieji** svečiai!* 'honoured.DEF guests', *Brangūs susirinkus**ieji**!* 'dear.DEF assembled [here]'.
- A typological comparison of the ADM in the two Baltic languages, Lithuanian and Latvian, would be of importance.

Concluding, I would like to hope that this dissertation will spark more interest in cross-linguistic studies of the adjectival definiteness marking.

REFERENCES

Ambrazas, Vytautas (ed.) et al. 2006, *Lithuanian Grammar*, Second revised edition, Vilnius: Baltos lankos.

Belk, Zoë C. E. 2017, *Attributes of Attribution*, Doctoral dissertation, University College London.

Börjars, Kerstin 1994, Swedish double determination in a European typological perspective, *Nordic Journal of Linguistics*, 17, 219–252.

Corbett, Greville G. 2004, The Russian Adjective: A Pervasive yet Elusive Category, in Dixon, R.M.W. & Alexandra Y. Aikhenvald (eds.), *Adjective Classes: A Cross-linguistic Typology*, Oxford: Oxford University Press, 199–222.

Croft, William 2004, *Typology and Universals*, Cambridge: Cambridge University Press.

Cuzzolin, Pierluigi & Christian Lehman 2004, Comparison and Gradation, in Gerd Booij et al. (eds.), *Morphologie. Ein internationales Handbuch zur Flexion und Wortbildung* 2, Berlin: Mouton de Gruyter, Berlin, 1212–1220.

Dahl, Östen & Maria Koptjevskaja-Tamm (eds.) 2001, *The Circum-Baltic Languages*, Volume 2: Grammar and Typology, John Benjamins Publishing Company.

Dahl, Östen 2004, Definite articles in Scandinavian: Competing grammaticalization processes in standard and non-standard varieties, in Bernd Kortmann (ed.), *Dialectology Meets Typology: Dialect Grammar from a Cross-Linguistic Perspective*, Berlin, New York: Mouton de Gruyter, 148–180.

Dahl, Östen 2010, *Grammaticalization in the North: Noun Phrase Morphosyntax in Scandinavian Vernaculars*, Stockholm: Stockholm University.

Delsing, Lars-Olof, 1993, *The Internal Structure of Noun Phrases in the Scandinavian Languages – A Comparative Study*, Doctoral dissertation, Lund: University of Lund.

Diessel, Holger, 1999, *Demonstratives: Form, Function, and Grammaticalization*, Amsterdam/Philadephia: John Benjamins Publishing Company.

Dixon, Robert M.W 1982, *Where have all the adjectives gone?*, Berlin: Walter de Gruyter & Co.

Dixon, Robert M.W. 2004, Adjective Classes in Typological Perspective, in Robert M.W. Dixon & Alexandra Y. Aikhenvald (eds.),

Adjective Classes: A Cross-linguistic Typology, Oxford: Oxford University Press, 1–49.

Dixon, Robert M.W. & Alexandra Y. Aikhenvald (eds.) 2004, *Adjective Classes: A Cross-linguistic Typology*, Oxford: Oxford University Press.

Dryer, Matthew S. & Haspelmath, Martin (eds.) 2013, The World Atlas of Language Structures Online, Leipzig: Max Planck Institute for Evolutionary Anthropology. Available online at http://wals.info.

Ekerot, Lars-Johan 2011, *Ordföljd, tempus, bestämdhet*, Malmö: Gleerups.

Faarlund, Jan, Svein Lie & Kjell Ivar Vannebo (eds.) 1997, *Norsk referansegrammatikk*, Oslo: Universitetsforlaget.

Gelumbeckaitė, Jolanta 2020, Pronominierte Nominalformen in der Wolfenbütteler Postille (1573–1574), *Baltistica* 55, 49–82.

Gillon, Carrie & Solveiga Armoskaite 2015, The Illusion of NP/DP divide: Evidence from Lithuanian, *Linguistic Variation*, Volume 15, Issue 1, Jan 2015, p. 69–115.

Golden, Anne, Kirsti Mac Donald & Else Ryen (eds.) 1997, *Norsk som fremmedspråk: Grammatikk* (3. utgave) Oslo: Universitetsforlaget.

Hansen, Björn 2004, The Life Cycle of a Definiteness Marker: The Long and Short Form of the Adjective in Old Church Slavonic, Russian, Serbian and Croatian. *Zbornik Matice srpske za filologiju ilingvistiku* 47. 1–2, 51–74.

Hansen, Erik & Lars Heltoft 2010, *Grammatik over det Danske Sprog*, Århus: Syddansk Universitetsforlag.

Haspelmath, Martin 1997, *Indefinite Pronouns*, Oxford: Oxford University Press.

Haspelmath, Martin 1999, Explaining article-possessor complementarity: economic motivation in noun phrase syntax, *Language* 75.2, 227–243.

Haspelmath, Martin 2001a, The European linguistic area: Standard Average European, in Martin Haspelmath (ed.), *Language typology and language universals. (Handbücher zur Sprach- und Kommunikationswissenschaft)*, Berlin: de Gruyter, 1492-1510.

Haspelmath, Martin 2001b, Word classes and parts of speech, *International Encyclopedia of the Social & Behavioral Sciences*, 16538–16545.

Haspelmath, Martin 2010, Comparative concepts and descriptive categories in crosslinguistic studies. *Language* 86.3, 663–687.

Haspelmath, Martin 2018, How comparative concepts and descriptive linguistic categories are different, in Olmen, Daniël, Tanja Mortelmans & Frank Brisard (eds.), *Aspects of Linguistic Variation*, Berlin, Boston: De Gruyter Mouton, 83-114.

Haspelmath, Martin 2020, General linguistics must be based on universals (or nonconventional aspects of language), online pre-print version available here (2021-11-02):
https://www.researchgate.net/publication/341293285_General_linguistics_must_be_based_on_universals_or_nonconventional_aspects_of_language

Hawkins, John A. 1978, *Definiteness and Indefiniteness: A Study in Reference and Grammaticality Prediction*. London: Croom Helm.

Heltoft, Lars 2010, Paradigmatic structure and reanalysis: from NPs to DPs in Scandinavian. *Acta Linguistica Hafniensia*, Vol. 42, No 1, 11-25.

Holvoet, Axel & Rolandas Mikulskas (eds.) 2006, *Daiktavardinio junginio tyrimai*, Vilnius: Lietuvių kalbos institutas.

Holvoet, Axel & Aurelija Tamulionienė 2006, Apibrėžtumo kategorija, *Daiktavardinio junginio tyrimai*, Vilnius: Lietuvių kalbos institutas, 11– 32.

Holvoet, Axel 2009, *Bendrosios sintaksės pagrindai*, Vilnius: Vilniaus universitetas & Asociacija *Academia Salensis*.

Holvoet, Axel & Birutė Spraunienė 2012, Towards a semantic map for definite articles in the Baltic, *Baltic Linguistics* 3 (2012), 65–99.

Holvoet, Axel & Birutė Spraunienė 2014, Ad hoc taxonomies: A Baltic parallel to the Scandinavian absolute positives, in Ērika Sausverde & Ieva Steponavičiūtė (eds.), *Fun and Puzzles in Modern Scandinavian Studies*, *Scandinavistica Vilnensis 9*, Vilnius University, 47–62.

Huddleston, Rodney & Geoffrey K. Pullum (eds.) 2002, *The Cambridge Grammar of the English Language*, Cambridge: Cambridge University Press.

Judžentis, Artūras 2012, *Lietuvių kalbos gramatinės kategorijos*, Vilnius: Vilniaus universiteto leidykla.

Julien, Marit 2005, *Nominal Phrases from a Scandinavian Perspective*, Amsterdam: John Benjamins Publishing Company.

Kamandulytė, Laura & Marija Tuškevičiūtė 2008, Būdvardžio vartojimo skirtumai sakytinės kalbos registruose, *Darbai ir dienos*, 50, p. 91–108.

Kamandulytė-Merfeldienė, Laura & Ingrida Balčiūnienė 2016, Atributinių ir predikatinių junginių su būdvardžiais dažnumas ir struktūra sakytinėje kalboje, *Lituanistica* 2016 (2), 127-137.

Kibort, Anna 2008, "Definiteness." Grammatical Features. 25 January 2008. http://www.grammaticalfeatures.net/features/definiteness.html.

Kibort, Anna & Greville G. Corbett 2008, "Number." Grammatical Features, 25 January 2008,
http://www.grammaticalfeatures.net/features/number.html.

Kniūkšta, Pranas 1976, *Priesagos* -inis *būdvardžiai: daryba, reikšmės, gramatiniai sinonimai*, Vilnius: Mokslas.

Koptjevskaja-Tamm, Maria 2003, *A woman of sin, a man of duty,* and *a hell of a mess*: Non-determiner genitives in Swedish, in Frans Plank (ed.), *Noun Phrase Structure in the Languages of Europe*, Berlin, New York: Mouton de Gruyter, 515-558.

Koptjevskaja-Tamm, Maria & Bernhard Wälchli 2001, The Circum-Baltic languages: An areal typological approach, in Dahl, Östen & Maria Koptjevskaja-Tamm (eds.), *The Circum-Baltic Languages,* Volume 2: Grammar and Typology, John Benjamins Publishing Company, 615–750.

Kortmann, Bernd (ed.) 2004, *Dialectology Meets Typology: Dialect Grammar from a Cross-Linguistic Perspective*, Berlin: Walter de Gruyter.

Kortmann, Bernd & Johan van der Auwera (eds.) 2011, *The Languages and Linguistics of Europe*, De Gruyter Mouton.

Kozhanov, Kiril 2010, Notes on the use of Lithuanian indefinite pronouns, *Baltic Linguistics* 1, 79-110.

Langacker, Ronald W. 2008, *Cognitive Grammar: A Basic Introduction*, Oxford: Oxford University Press.

Lohndal, Terje 2007, On the structure and development of nominal phrases in Norwegian, in Stark, Elisabeth, Elisabeth Leiss & Werner Abraham (eds.), *Nominal Determination: Typology, context constrains and historical emergence*, Amsterdam/Philadelphia: John Benjamins Publishing Company, 287–310.

Lohrmann, Susanne 2011, A unified structure for Scandinavian DPs, in Sleeman, Petra & Harry Perridon (eds.), *The Noun Phrase in Romance and Germanic: Structure, Variation, and Change*, in Amsterdam: John Benjamins Publishing Company, 111– 125.

Loman, Bengt 1956, Om relationen mellan ordföljd och betydelse hos framställda attributiva bestämningar till substantiviska huvudord, *ANF* 71, 218–244.

Lyons, Christopher 2003, *Definiteness*, Cambridge: Cambridge University Press.

Mauri, Caterina 2014, What do connectives and plurals have in common? The linguistic expression of ad hoc categories, in Blochowiak, Joanna, Stephanie Durrlemann-Tame, Cristina Grisot & Christopher

Laenzlinger (eds.) 2014, *Linguistic papers dedicated to Jacques Moeschler*, Genève: University of Geneva Publication.

Mel'čuk, Igor' A. 1998, *Kurs obščej morfologii*. Tom 2. Moskva: Progress, Jazyki russkoj kul'tury; Vena: Wiener Slavistischer Almanach, 29–31.

Mickūnaitė-Griškevičienė, Aurelija 2004, *Aktualioji skaida lietuvių ir norvegų kalbose*, Humanitarinių mokslų daktaro distertacija, Vilniaus universitetas.

Mikulskas, Rolandas 2006, Apibrėžiamųjų būdvardžių aprašo perspektyva. *Daiktavardinio junginio tyrimai*, Vilnius: Lietuvių kalbos institutas, 33-65.

Olmen, Daniël, Tanja Mortelmans & Frank Brisard (eds.) 2018, *Aspects of Linguistic Variation*, Berlin, Boston: De Gruyter Mouton, 83–114.

Otrębski, Jan 1968, Die Entstehung des bestimmten Adjektivs im Slavischen und Baltischen, in Zeitschrift für vergleichende Sprachforschung 82. Bd., 1. H. (1968), 122-128. Göttingen: Vandenhoeck & Ruprecht.

Parenti, Alessandro 1995, Observations on the Use of the Definite Adjective in Bretkūnas' Translation of the New Testament. *Res Baltic*, 77–87.

Paulauskienė, Aldona 1994, *Lietuvių kalbos morfologija: Paskaitos lituanistams*, Vilnius: Mokslo ir enciklopedijų leidykla.

Perridon, Harry 1989. *Reference, Definiteness and the Noun Phrase in Swedish*. Doctoral Dissertation. Amsterdam: University of Amsterdam.

Plank, Frans (ed.) 2003, *Noun Phrase Structure in the Languages of Europe*. Berlin, New York: Mouton de Gruyter.

Riessler, Michael 2016, *Adjective attribution* (Studies in Diversity Linguistics 2), Berlin: Language Science Press.

Rosinas, Albertas 1975, Ar baltų **i-, *io-* resp. *ī, *iā-* kamienai buvo reliatyviniai?, *Baltistica* 11.2, 165–170.

Rosinas, Albertas 1976, Baltų kalbų parodomųjų įvardžių funkcijų ir reikšmių klausimu. *Baltistica* 12.2, 150–155.

Rosinas, Albertas 1981, Kelios pastabos parodomųjų ir nežymimųjų įvardžių evoliucijos klausimu, *Baltistica* XVII (1), 16-27.

Rosinas, Albertas 1996, *Lietuvių bendrinės kalbos įvardžiai*. Vilnius: Mokslo ir enciklopedijų leidykla.

Rosinas, Albertas 2009, *Baltų kalbų įvardžių semantinė ir morfologinė struktūra. Sinchronija ir diachronija*. Vilnius: Mokslo ir enciklopedijų leidybos institutas.

Rutkowski, Paweł 2002, Noun/Pronoun Asymmetries: Evidence of Support of the DP Hypothesis in Polish. *Jezykslovlje* 3.1-2, 159–170.

Rutkowski, Paweł and Ljiljana Progovac 2006, Classifying Adjectives and Noun Movement in Lithuanian, in *Minimalist views on language design: Proceedings of the 8th Seoul Conference on Generative Grammar*, Seoul, 265-277.

Schwarz, Florian 2009, "Two Types of Definites in Natural Language", *Open Access Dissertations*. 122.
https://scholarworks.umass.edu/open_access_dissertations/122.

Schwarz, Florian 2013, Two types of definites cross-linguistically, *Language and Linguistic Compass* 7/10, 534-559.

Schwarz, Florian 2019, Weak vs. strong definite articles: Meaning and form across languages, in Aguilar-Guevara, Ana, Julia Pozas Loyo & Violeta Vázquez-Rojas Maldonado (eds.). 2019, *Definiteness across languages* (Studies in Diversity Linguistics 25), Berlin: Language Science Press.

Sleeman, Petra & Harry Perridon (eds.) 2011, *The Noun Phrase in Romance and Germanic: Structure, Variation, and Change*, Amsterdam: John Benjamins Publishing Company.

Sommer, Florian 2018, The historical morphology of definiteness in Baltic, *Indo-European Linguistics* 6, 152-200.

Spraunienė, Birutė 2008a, Apibrėžtumo kategorija ir jos tyrimai lietuvių kalbotyroje, *Lietuvių kalba* 2, 1-12.

Spraunienė, Birutė 2008b, Paprastųjų ir įvardžiuotinių būdvardžių opozicija lietuvių kalboje kaip apibrėžtumo sistema, *Acta Linguistica Lithuanica LIX* (2008), 109-139.

Spraunienė, Birutė 2011, *Apibrėžtumo žymėjimas lietuvių kalboje lyginant su danų ir kitomis artikelinėmis kalbomis*, Humanitarinių mokslų daktaro distertacija, Vilniaus universitetas.

Stark, Elisabeth, Elisabeth Leiss & Werner Abraham (eds.) 2007, *Nominal Determination: Typology, context constrains and historical emergence*, Amsterdam/Philadelphia: John Benjamins Publishing Company

Šereikaitė, Milena 2019, Strong vs. weak definites: Evidence from Lithuanian adjectives, in Aguilar-Guevara, Ana, Julia Pozas Loyo & Violeta Vázquez-Rojas Maldonado (eds.) 2019, *Definiteness across languages* (Studies in Diversity Linguistics 25), Berlin: Language Science Press.

Tekorienė, Dalia 1987 (rec.), Albertas Rosinas, Lietuvių bendrinės kalbos įvardžių semantinė struktūra, 1984, *Baltistica*, T.23 (1), 86-92.

Teleman, Ulf, Staffan Helleberg & Erik Andersson (eds.) 1999, *Svenska Akademiens grammatik*. Stockholm: Svenska akademien, Nordstedts ordbook.

Trakymaitė, Ringailė 2018, Determination and modification: Topology of prenominal attributes in Lithuanian, *Kalbotyra* 71, 84-133. Vilnius: Vilnius University press.

Trost, Pavel. 1966. O složeném adjektivu baltoslovanském, in Zdeňka Havránková & Jaroslav Popela (eds.), Prof. Dr. Antoniń u Dostálovi, DrSc., k 60. Narozeninám (Bulletin Ústavu ruského jazyka a literatury 10), Praha: Universita Karlova, 35–38.

Tumėnas, Stasys 1988, Artroido (įvardinio artikelio) reiškiniai šiaurinėse lietuvių kalbos tarmėse, *Kalbotyra* 39.1, 90–97.

Ulvydas, Kazimieras (ed.) 1965, *Lietuvių kalbos gramatika*, Vilnius: Mintis.

Utka, Andrius 2009, *Dažninis rašytinės lietuvių kalbos žodynas 1 milijono žodžių morfologiškai anotuoto tekstyno pagrindu*. Kaunas: Vytauto Didžiojo universitetas.

Vaičiulytė-Semėnienė, Loreta 2006, Morfosintaksinis posesyvumo žymėjimas: vidinė ir išorinė posesyvumo raiška, *Daiktavardinio junginio tyrimai*, Vilnius: Lietuvių kalbos institutas, 151-178.

Vaitkutė, Ieva 2017, *Apibrėžtumo rodiklių nebuvimo kompensavimas vertimuose iš švedų kalbos į lietuvių*, Humanitarinių mokslų bakalauro darbas, Vilniaus universitetas.

Vaitkutė, Ieva 2019, *Lietuvių kalbos įvardis „*toks*" similiatyvų tipologijos kontekste*, Humanitarinių mokslų magistro darbas, Vilniaus universitetas.

Valeckienė, Adelė 1957, Dabartinės lietuvių kalbos įvardžiuotinių būdvardžių vartojimas, *Literatūra ir kalba* 2, 161–328.

Valeckienė, Adelė 1986, Apibrėžtumo / neapibrėžtumo kategorija ir pirminė įvardžiuotinių būdvardžių reikšmė, *Lietuvių kalbotyros klausimai* 25, 168–189.

Valeckienė, Adelė 1998, *Funkcinė lietuvių kalbos gramatik*a, Vilnius: Mokslo ir enciklopedijų leidybos institutas.

Van de Velde, Freek 2011, Anaphoric adjectives becoming determiners, in Petra Sleeman & Harry Perridon (eds.) 2011, *The Noun Phrase in Romance and Germanic: Structure, Variation, and Change*, Amsterdam: John Benjamins Publishing Company, 111-125.

van der Auwera, Johan & Evie Coussé 2016, *Such* and *sådan* – the same but different. *Nordic Journal of English Studies* 15 (3), 15-32.

van der Auwera, Johan & Kalyanamalini Sahoo 2018, *Such* similatives: a cross linguistic reconnaissance, *Language Sciences.* https://doi.org/10.1016/j.langsci.2018/12.002**.**

Vaskelaitė, Ramunė 2016, Įvardžiuotinių formų nevartojimas terminijoje ir jo keliamas iššūkis normai, *Bendrinė kalba* 89.

Wälchli, Bernhard 2011, The Circum-Baltic languages, in Bernd Kortmann & Johan van der Auwera (eds.), *The Languages and Linguistics of Europe*, De Gruyter Mouton, 325–340.

Zwicky, Arnold M. 1986, German Adjective Agreement in GPSG. *Linguistics* 24, 957–999.

Appendix A.

Comments:

- All items shaded in grey exhibit less that 2% of LF in FrD.
- All items removed from the data analysed contained more than 1% of LFs in CCLL. All items included in the analysis contained less than 1% of LFs in CCLL and the collocation analysis was run on all of them.
- For all items removed from the data analysis in this section, collocation list checks have not been made due to high counts of LF to be manually assessed. Also, all these items were intuitively assessed as exhibiting no strange behaviour patterns with regards to their appearance in LFs, e.g., their ability to establish an *ad hoc* category was deemed fully functional, e.g.:
    - *sunk**ieji** galvosūkiai* 'difficult.DEF puzzles' → *Iš pradžių jis išsprendė lengvus, o tada perėjo prie sunk**iųjų** galvosūkių*. 'To begin with, he solved easy [puzzles], and then he moved on to the difficult puzzles.'
    - *puik**ioji** lašiša* 'splendid salmon' → *Vakarienei šiandien – puik**ioji** norvegiška lašiša!* 'For dinner today, the splendid Norwegian salmon!'
    - *keist**ieji** radiniai*' strange.DEF findings' → *Mokslininkai nežinojo kaip apibūdinti keist**uosius** radinius.* 'Scientists did not know how to describe the strange findings.'
- The two adjectives in bold (*paprastas* and *individualus*) were included in the data analysis for the reasons explained in the table below and in the text of the chapter.

No	Rank in FrD	Lexeme	Translation	Total count of LFs in FrD	% of LFs	Comment
0	53.	didelis	big/large	n/a	n/a	Under this lexeme, paradigms of 2 adjectives are mixed up, viz., *didelis* and *didis*. Only comparative forms of *didelis* can assume LFs. Due to this, I have chosen not to include this in the data.
1	64.	naujas	new/novel	1349	17.87	
2	80.	svarbus	important	1157	3.03	
3	105.	geras	good/kind	941	5.10	
4	123.	įvairus	various/varied/miscellaneous/sundry	852	0.00	
5	141.	bendras	common/joint/shared	778	25.19	
6	156.	mažas	little/small	706	23.80	
7	172.	senas	old/ancient	650	43.85	
8	181.	aukštas	tall/high	627	30.78	

No	Rank in FrD	Lexeme	Translation	Total count of LFs in FrD	% of LFs	Comment
9	199.	sunkus	heavy/difficult/hard	587	1.53	It was removed from the data analysed in the dissertation, as in CCLL it showed 9.39% of LF. NB due to high numbers (total count of LFs 2594) it was impossible to run a collocation check, but intuitively it seems like an ordinary adjective with no unusual behaviour in relation to appearance with LF vs SF.
10	267.	atskiras	separate/individual/special/distinct/detached	460	0.43	
11	300.	reikalingas	needed/required/necessary/requisite	410	0.00	
12	310.	paskutinis	last/final	400	17.25	
13	334.	panašus	similar/like/alike/analogous/resemblant	376	0.00	
14	335.	tikras	true/real/proper	376	36.44	
15	336.	ilgas	long/lengthy	374	2.41	

No	Rank in FrD	Lexeme	Translation	Total count of LFs in FrD	% of LFs	Comment
16	342.	didis	great/famous/sublime	354	90.96	In FrD, presented count is 368. It mistakenly contains 14 forms of *didelis*. Hence the adjusted number of instances – 354.
17	349.	aiškus	clear/understandable/explicit/evident	359	0.00	
18	421.	konkretus	concrete/particular/specific	313	0.00	
19	427.	specialus	special/particular	309	10.03	
20	432.	jaunas	young/youthful/adolescent	307	26.06	
21	438.	pastaras	the latter/the recent	305	100.00	
22	474.	laisvas	free/liberal	285	29.83	
23	496.	būdingas	typical/characteristic/specific	276	0.73	
24	522.	**paprastas**	simple/ordinary/normal/average	265	6.42	LFs of this adjective are frequently used in terminology (botany, biology, medicine, etc.) corresponding to the use of Latin

No	Rank in FrD	Lexeme	Translation	Total count of LFs in FrD	% of LFs	Comment
						vulgaris or similar, e.g., *paprastieji spuogai* (Acne vulgaris). See 1.2 for reasons why it was included in the data analysis (CCLL adjusted data showed 0.72% of LF uses).
25	565.	skirtingas	different/unlike/separate/distinct/diverse	246	0.00	
26	592.	vyresnis	older/senior/superior	239	66.53	
27	594.	artimas	close/familiar/near	236	11.02	
28	603.	baltas	white/clean	234	18.80	
29	613.	gražus	beautiful/pretty/nice/lovely/picturesque	229	3.06	
30	618.	juodas	black	228	29.39	
31	673.	blogas	bad/evil/poor/wrong/ill	210	5.24	
32	675.	stiprus	strong/powerful/mighty	210	6.67	
33	685.	sudėtingas	complex/complicated/multiplex/elaborate	208	0.48	

No	Rank in FrD	Lexeme	Translation	Total count of LFs in FrD	% of LFs	Comment
34	688.	įdomus	interesting/exciting/entertaining	207	0.48	
35	700.	gyvas	live/alive/living/vivid/animate	203	16.75	
36	782.	platus	wide/broad/extensive/spacious	181	13.26	
37	808.	stambus	large/large-scale/bulky/hefty	177	4.52	
38	834.	tamsus	dark/overcast/gloomy/sombre	171	16.96	
39	844.	trumpas	short/brief/laconic	169	5.33	
40	847.	lengvas	easy/light/effortless	168	13.10	
41	849.	brangus	expensive/costly/precious/dear	167	9.58	
42	862.	svetimas	outlandish/strange/foreign	165	14.55	
43	889.	raudonas	red	159	18.87	
44	891.	šaltas	cold	159	3.77	
45	893.	**individualus**	individual	158	3.16	See 1.2 for reasons why this adjective was included in the analysis. After the collocation

No	Rank in FrD	Lexeme	Translation	Total count of LFs in FrD	% of LFs	Comment
						analysis of CCLL data, the LFs comprise 0.44% of all uses.
46	894.	lietuviškas	Lithuanian	158	17.09	
47	904.	vidutinis	average/medium/middle/moderate/normal	157	1.27	2 instances of LF attested.
48	913.	rimtas	serious/solid/sober/grave	156	0.64	1 instance of LF attested. It was removed from the data analysed, as in CCLL it showed 2.23% of LFs. Intuitively, it seems like an ordinary adjective with no unusual behaviour in relation to appearance with LF vs SF.
49	927.	atviras	open/overt/public/honest	154	18.18	
50	960.	sveikas	healthy/whole/intact/sound	150	5.33	
51	990.	smulkus	small/petty/fine	146	8.90	
52	1007.	realus	real/realistic/actual	143	9.09	
53	1050.	garsus	famous/prominent/known/loud	137	29.20	

No	Rank in FrD	Lexeme	Translation	Total count of LFs in FrD	% of LFs	Comment
54	1083.	neaiškus	unclear/uncertain/obscure/vague/indistinct	134	0.00	
55	1086.	šiltas	warm	134	2.24	
56	1091.	ypatingas	special/particular/distinct/especial	132	1.52	2 instances of LF attested. It was removed from the data analysed, as in CCLL it showed 5.80% of LFs.
57	1092.	naudingas	useful/beneficial/helpful	132	15.15	
58	1096.	puikus	great/excellent/splendid	132	1.52	2 instances of LF attested. It was removed from the data analysed, as in CCLL it showed 2.14% of LFs.
59	1106.	žemas	low/short/inferior	131	14.50	
60	1121.	amžinas	eternal/perpetual/everlasting/timeless	128	29.69	
61	1136.	šventas	holy/sacred/saint/blessed/taboo	127	48.82	
62	1165.	gilus	deep/abysmal/profound/thoughtful	123	2.44	
63	1179.	ramus	calm/quiet/peaceful/tranquil	122	0.82	1 instance of LF attested. It was removed from the data analysed, as in CCLL it showed 11.50% of LFs.

No	Rank in FrD	Lexeme	Translation	Total count of LFs in FrD	% of LFs	Comment
64	1182.	normalus	normal/regular/ordinary/average	121	0.00	
65	1204.	aktyvus	active/energetic/lively	118	7.63	
66	1214.	nemažas	considerable/not small	117	0.00	
67	1262.	siauras	narrow/tight	113	10.62	
68	1289.	savarankiškas	independent/autonomous/self-sufficient	111	0.00	It was removed from the data analysed, as in CCLL it showed 1.19 % of LFs.
69	1294.	vienodas	uniform/equal/same/homogeneous/like	111	0.00	
70	1296.	keistas	strange/odd/bizarre/weird	110	1.82	2 instances of LF attested. It was removed from the data analysed, as in CCLL it showed 1.72 % of LFs.
71	1301.	ryškus	bright/stark	110	0.91	1 instance of LF attested.
72	1328.	tuščias	empty	108	1.85	2 instances of LF attested. It was removed from the data analysed, as in CCLL it showed 2.15 % of LFs.
73	1341.	žalias	green	107	18.69	

No	Rank in FrD	Lexeme	Translation	Total count of LFs in FrD	% of LFs	Comment
74	1343.	karštas	hot	106	2.83	
75	1346.	vertas	worthy/worth/deserving/valuable	106	0.00	
76	1351.	oficialus	official	105	18.10	
77	1367.	patogus	comfortable/convenient/handy	104	0.00	This adjective in CCLL contained 0.20% of LFs. For unknown reasons, it has a very low count of LFs even though one can establish a category, e.g., *patogioji avalynė* 'comfortable.DEF footwear', *patogioji kelionių agentūra* 'convenient.DEF travel agent', *patogusis fotelis* 'comfortable.DEF armchair', etc. In this regard, it is somewhat similar to *įdomus*.
78	1377.	griežtas	strict/stringent/tight/austere	103	4.85	
79	1383.	ankstyvas	early/precocious/premature	102	48.04	
80	1389.	gausus	abundant/numerous/plentiful/ample/bountiful	102	0.00	

No	Rank in FrD	Lexeme	Translation	Total count of LFs in FrD	% of LFs	Comment
81	1401.	storas	thick/heavy/fat/corpulent	102	17.65	
82	1406.	lygus	equal/level/like/smooth/flat	101	0.99	1 instance of LF attested. In CCLL, this adjective contained 2018 instances of LFs. It was impossible to calculate the count of SF due to a very high number of homonyms (nouns like *lyga* 'league', *lygis* 'level', etc.). Yet, I deemed that 2018 is a high number allowing elimination of this adjective from the data analysed.
83	1417.	privatus	private/personal/proprietary/individual	100	1.00	1 instance of LF attested.
84	1419.	švarus	clean/pure/clear/fresh/immaculate	100	0.00	
85	1424.	ankstus	previous/early	99	14.14	Only comparative forms of this adjective assume LFs.
86	1446.	grynas	pure/net/neat/clear/clean	98	41.84	

No	Rank in FrD	Lexeme	Translation	Total count of LFs in FrD	% of LFs	Comment
87	1450.	pavojingas	dangerous/hazardous/serious/precarious	98	0.00	In CCLL, this adjective contained 1.48% of LFs. It was removed from the data analysed.
88	1459.	galutinis	final/ultimate/terminal	97	0.00	
89	1477.	menkas	meagre/insignificant/poor	96	0.00	
90	1488.	malonus	pleasant/enjoyable/kind/nice	95	2.1	2 instances of LF attested.
91	1501.	viešas	public/open	95	56.84	
92	1502.	baisus	terrible/horrible/awful/gruesome	94	3.19	
93	1519.	žymus	famous/eminent/significant/celebrated	94	5.32	
94	1536.	teisingas	right/righteous/just/correct/fair/truthful	93	0.00	In CCLL, this adjective contained 1.79% of LFs. It was removed from the data analysed.
95	1548.	optimalus	optimal/optimum/superb/top-notch	92	0.00	
96	1554.	tolesnis	further/subsequent/successive	92	1.09	1 instance of LF attested.

No	Rank in FrD	Lexeme	Translation	Total count of LFs in FrD	% of LFs	Comment
97	1558.	efektyvus	effective/efficient/valid	91	1.10	1 instance of LF attested. In CCLL, this adjective contained 1.65% of LFs. It was removed from the data analysed.
98	1570.	dažnas	frequent/habitual/periodic/repeated	90	0.00	
99	1579.	pilnas	full/whole/complete/plump	90	10.00	
100	1614.	sausas	dry/arid/droughty/dead	88	1.14	1 instance of LF attested. In CCLL, this adjective contained 17.74% of LFs. It was removed from the data analysed.
101	1625.	lankstus	flexible/supple/versatile	87	0.00	In CCLL, this adjective contained 2.76% of LFs. It was removed from the data analysed.
102	1628.	modernus	modern/up-to-date/contemporary	87	19.54	
103	1633.	silpnas	weak/fragile/frail/lax	87	10.34	
104	1638.	universalus	universal/versatile/all-around	87	12.64	

No	Rank in FrD	Lexeme	Translation	Total count of LFs in FrD	% of LFs	Comment
105	1644.	galingas	powerful/mighty/potent/strong	86	6.98	
106	1664.	idealus	ideal/perfect	85	34.12	
107	1675.	populiarus	popular	85	2.35	
108	1698.	reikšmingas	significant/meaningful/important/weighty	84	0.00	
109	1704.	visiškas	complete/total/full/absolute/superior/superb	84	0.00	
110	1706.	atsakingas	responsible/liable/accountable	83	0.00	In CCLL, this adjective contained 5.45% of LFs. It was removed from the data analysed.
111	1709.	gimtas	native	83	89.16	

Appendix B.

Collocations of *konkretus* 'concrete, specific, particular'

No	Collocation	Translation	Instances	
1	daiktavardžiai	nouns	22	22
2	muzika	music	11	11
3	linksniai	cases (as in a paradigm of noun cases)	10	10
4	klausimai	questions	7	7
5	dalykai	subjects	5	5
6	NOMINALISATION		5	
7	poezija	poetry	4	4
8	prasmė	meaning	3	3
9	abstrakcija	abstraction	2	2
10	dvasingumas	spirituality	2	2
11	ekonomika	economy	2	2

No	Collocation	Translation	Instances	
12	filosofija	philosophy	2	2
13	konkretika	specifics	2	2
14	materialistai	materialists	2	2
15	objektas	object	2	2
16	pavidalas	form/shape/guise	2	2
17	pažinimas	cognition	2	2
18	politika	politics	2	2
19	santykiai	relations	2	2
20	turinys	content	2	2
21	analogija	analogy	1	**86** the sum of terminology-like uses
22	apraiška	manifestation	1	
23	aspektas	aspect	1	
24	būtis	existence	1	

No	Collocation	Translation	Instances
25	forma	form	1
26	gėris	good, goodness, kindness	1
27	intelektas	intellect	1
28	išraiška	expression	1
29	kainos	prices	1
30	kalba	language	1
31	klasifikatoriai	classifiers	1
32	kontekstas (reikšmės)	context (of a meaning)	1
33	lygmuo	level	1
34	mainai	exchange	1
35	materializmas	materialism	1
36	mokslai	sciences/studies	1
37	pastoracija	pastoral care	1

No	Collocation	Translation	Instances
38	postūmis	push/impulse/stimulus	1
39	pozicija	position	1
40	programa	programme	1
41	pusė	side	1
42	raiška	expression/marking	1
43	rašymas	writing	1
44	reikšmė	meaning	1
45	šalis	country	1
46	sąvoka	notion	1
47	simbolis	symbol	1
48	substratas	substrate	1
49	sugebėjimas	ability	1
50	tarpsnis (laiko)	period (of time)	1
51	terpė	environment	1

No	Collocation	Translation	Instances
52	tikrovė	reality	1
53	tikslai	aims/goals	1
54	transcendencija	transcendence	1
55	tyrimas	research/investigation	1
56	žmogiškumas	humanity	1

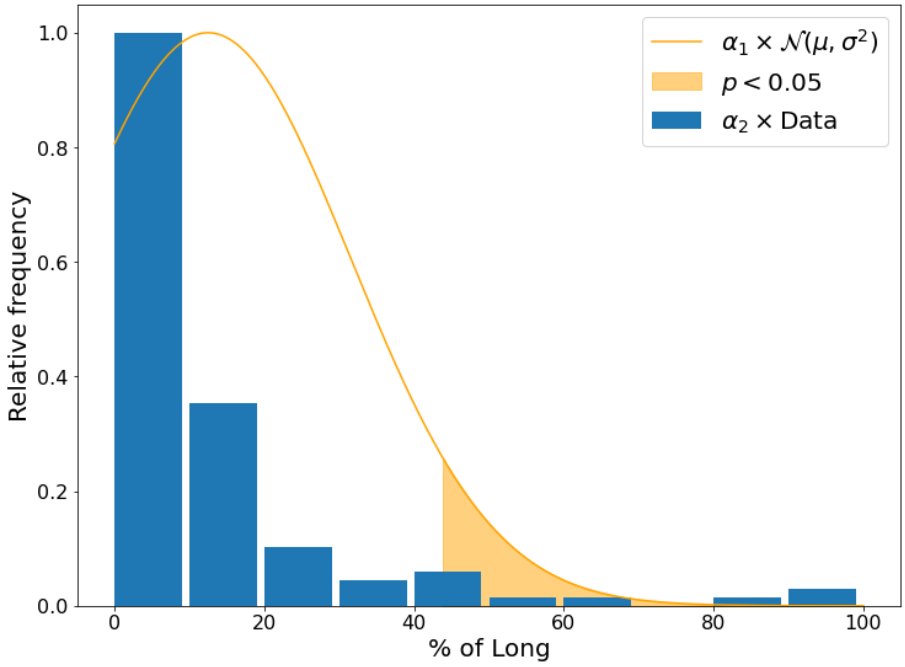
Total	Terms	Nominalisations	Other
127	86	5	**36**

Percentage:	**%**
Terminology	**68.00**
Nominalisations	**4.00**
Other uses	**28.00**

Appendix C.

Comments:

- The 18 highlighted adjectives (16.22% of 111 on the list) have 25% or more of LF as drawn from FrD.
- The mean of the LF percentage sample is μ = 12.5%, while the standard deviation is σ = 19.0%.
- A normal distribution $\mathcal{N}(\mu, \sigma^2)$ was fitted onto the data according to the mean and the standard deviation values provided above.
- To afford illustrative superimposing, the data with the normal distribution in the figure below, the scaling factors $\alpha_1 = 47.739$ and $\alpha_2 = 0.015$ were applied.



- Based on these two values, the p-values were calculated. The adjectives with p-value less than 0.05 (cut-off at 44%) are the following: *didis*, *pastaras*, *vyresnis*, *šventas*, *ankstyvas*, *viešas* and *gimtas* (p-values in bold in the Comment section below). If the p-value is relaxed[120] to 0.10 (cut-off at 36.9%), our list of adjectives is expanded with *senas* 'old' and *grynas* 'pure'. One could say that the fact that these adjectives are

[120] I am aware that there has been a lot of scientific debate whether it is acceptable to relax the p-value. I share the opinion that it is, depending on the field and type of research.

predominantly used with their LFs is statistically significant. If the p-value was relaxed to 0.25 (cut-off at 25%), it would include all the 18 highlighted adjectives.

- It must be noted that p-values have been calculated for all the 111 adjectives on the list. I have chosen not to include the additional ones in order not to burden the visual impact of the table. None of them were of significance (as can be observed from the percentage of LF in FrD in the table).

No	Rank in FrD	Lexeme	Translation	Total count of LFs in FrD	% of LFs	Comment
0	53.	didelis	big/large	n/a	n/a	Under this lexeme, paradigms of 2 adjectives are mixed up, viz., *didelis* and *didis*. Only comparative forms of *didelis* can assume LFs. Due to this, I have chosen not to include this in the data.
1	64.	naujas	new/novel	1349	17.87	
2	80.	svarbus	important	1157	3.03	
3	105.	geras	good/kind	941	5.10	
4	123.	įvairus	various/varied/miscellaneous/sundry	852	0.00	
5	141.	bendras	common/joint	778	**25.19**	p = 0.2526
6	156.	mažas	little/small	706	23.80	

No	Rank in FrD	Lexeme	Translation	Total count of LFs in FrD	% of LFs	Comment
7	172.	senas	old/ancient	650	43.85	**p = 0.0499**
8	181.	aukštas	tall/high	627	**30.78**	p = 0.1685
9	199.	sunkus	heavy/difficult/hard	587	1.53	
10	267.	atskiras	separate/individual/special/distinct/detached	460	0.43	
11	300.	reikalingas	needed/required/necessary/requisite	410	0.00	
12	310.	paskutinis	last/final	400	17.25	
13	334.	panašus	similar/like/alike/analogous/resemblant	376	0.00	
14	335.	tikras	true/real/proper	376	**36.44**	p = 0.1043
15	336.	ilgas	long/lengthy	374	2.41	

No	Rank in FrD	Lexeme	Translation	Total count of LFs in FrD	% of LFs	Comment
16	342.	didis	great/famous/sublime	354	90.96	In FrD, the presented count is 368. It mistakenly contains 14 forms of *didelis*. Hence the adjusted number of instances – 354. **p = 0.0000**
17	349.	aiškus	clear/understandable/explicit/evident	359	0.00	
18	421.	konkretus	concrete/particular/specific	313	0.00	
19	427.	specialus	special/particular	309	10.03	
20	432.	jaunas	young/youthful/adolescent	307	**26.06**	p = 0.2382
21	438.	pastaras	the latter/the recent	305	**100.00**	**p = 0.0000**
22	474.	laisvas	free/liberal	285	**29.83**	p = 0.1814
23	496.	būdingas	typical/characteristic/specific	276	0.73	
24	522.	paprastas	simple/ordinary/normal/average	265	6.42	

No	Rank in FrD	Lexeme	Translation	Total count of LFs in FrD	% of LFs	Comment
25	565.	skirtingas	different/unlike/separate/distinct/diverse	246	0.00	
26	592.	vyresnis	older/senior/superior	239	**66.53**	**p = 0.0023**
27	594.	artimas	close/familiar/near	236	11.02	
28	603.	baltas	white/clean	234	18.80	
29	613.	gražus	beautiful/pretty/nice/lovely/picturesque	229	3.06	
30	618.	juodas	black	228	**29.39**	p = 0.1876
31	673.	blogas	bad/evil/poor/wrong/ill	210	5.24	
32	675.	stiprus	strong/powerful/mighty	210	6.67	
33	685.	sudėtingas	complex/complicated/multiplex/elaborate	208	0.48	
34	688.	įdomus	interesting/exciting/entertaining	207	0.48	
35	700.	gyvas	live/alive/living/vivid/animate	203	16.75	
36	782.	platus	wide/broad/extensive/spacious	181	13.26	

No	Rank in FrD	Lexeme	Translation	Total count of LFs in FrD	% of LFs	Comment
37	808.	stambus	large/large-scale/bulky/hefty	177	4.52	
38	834.	tamsus	dark/overcast/gloomy/sombre	171	16.96	
39	844.	trumpas	short/brief/laconic	169	5.33	
40	847.	lengvas	easy/light/effortless	168	13.10	
41	849.	brangus	expensive/costly/precious/dear	167	9.58	
42	862.	svetimas	outlandish/strange/foreign	165	14.55	
43	889.	raudonas	red	159	18.87	
44	891.	šaltas	cold	159	3.77	
45	893.	individualus	individual	158	3.16	
46	894.	lietuviškas	Lithuanian	158	17.09	
47	904.	vidutinis	average/medium/middle/moderate/normal	157	1.27	2 instances of LF attested.
48	913.	rimtas	serious/solid/sober/grave	156	0.64	1 instance of LF attested.

No	Rank in FrD	Lexeme	Translation	Total count of LFs in FrD	% of LFs	Comment
49	927.	atviras	open/overt/public/honest	154	18.18	
50	960.	sveikas	healthy/whole/intact/sound	150	5.33	
51	990.	smulkus	small/petty/fine	146	8.90	
52	1007.	realus	real/realistic/actual	143	9.09	
53	1050.	garsus	famous/prominent/known/loud	137	**29.20**	p = 0.1903
54	1083.	neaiškus	unclear/uncertain/obscure/vague/indistinct	134	0.00	
55	1086.	šiltas	warm	134	2.24	
56	1091.	ypatingas	special/particular/distinct/especial	132	1.52	2 instances of LF attested.
57	1092.	naudingas	useful/beneficial/helpful	132	15.15	
58	1096.	puikus	great/excellent/splendid	132	1.52	2 instances of LF attested.
59	1106.	žemas	low/short/inferior	131	14.50	
60	1121.	amžinas	eternal/perpetual/everlasting/timeless	128	**29.69**	p = 0.1833

No	Rank in FrD	Lexeme	Translation	Total count of LFs in FrD	% of LFs	Comment
61	1136.	šventas	holy/sacred/saint/blessed/sacrosanct	127	**48.82**	**p = 0.0282**
62	1165.	gilus	deep/abysmal/profound/thoughtful	123	2.44	
63	1179.	ramus	calm/quiet/peaceful/tranquil	122	0.82	1 instance of LF attested.
64	1182.	normalus	normal/regular/ordinary/average	121	0.00	
65	1204.	aktyvus	active/energetic/lively	118	7.63	
66	1214.	nemažas	considerable/not small	117	0.00	
67	1262.	siauras	narrow/tight	113	10.62	
68	1289.	savarankiškas	independent/autonomous/self-sufficient	111	0.00	
69	1294.	vienodas	uniform/equal/same/homogeneous/like	111	0.00	
70	1296.	keistas	strange/odd/bizarre/weird	110	1.82	2 instances of LF attested.
71	1301.	ryškus	bright/stark	110	0.91	1 instance of LF attested.
72	1328.	tuščias	empty	108	1.85	2 instances of LF attested.

No	Rank in FrD	Lexeme	Translation	Total count of LFs in FrD	% of LFs	Comment
73	1341.	žalias	green	107	18.69	
74	1343.	karštas	hot	106	2.83	
75	1346.	vertas	worthy/worth/deserving/valuable	106	0.00	
76	1351.	oficialus	official	105	18.10	
77	1367.	patogus	comfortable/convenient/handy	104	0.00	
78	1377.	griežtas	strict/stringent/tight/austere	103	4.85	
79	1383.	ankstyvas	early/precocious/premature	102	48.04	**p = 0.0310**
80	1389.	gausus	abundant/numerous/plentiful/ample/bountiful	102	0.00	
81	1401.	storas	thick/heavy/fat/corpulent	102	17.65	
82	1406.	lygus	equal/level/like/smooth/flat	101	0.99	1 instance of LF attested.
83	1417.	privatus	private/personal/proprietary/individual	100	1.00	1 instance of LF attested.
84	1419.	švarus	clean/pure/clear/fresh/immaculate	100	0.00	

No	Rank in FrD	Lexeme	Translation	Total count of LFs in FrD	% of LFs	Comment
85	1424.	ankstus	previous/early	99	14.14	Only comparative forms of this adjective assume LFs.
86	1446.	grynas	pure/net/neat/clear/clean	98	**41.84**	**p = 0.0617**
87	1450.	pavojingas	dangerous/hazardous/serious/precarious	98	0.00	
88	1459.	galutinis	final/ultimate/terminal	97	0.00	
89	1477.	menkas	meagre/insignificant/poor	96	0.00	
90	1488.	malonus	pleasant/enjoyable/kind/nice	95	2.1	2 instances of LF attested.
91	1501.	viešas	public/open	95	**56.84**	**p = 0.0099**
92	1502.	baisus	terrible/horrible/awful/gruesome	94	3.19	
93	1519.	žymus	famous/eminent/significant/celebrated	94	5.32	
94	1536.	teisingas	right/righteous/just/correct/fair/truthful	93	0.00	
95	1548.	optimalus	optimal/optimum/superb/top-notch	92	0.00	

No	Rank in FrD	Lexeme	Translation	Total count of LFs in FrD	% of LFs	Comment
96	1554.	tolesnis	further/subsequent/successive	92	1.09	1 instance of LF attested.
97	1558.	efektyvus	effective/efficient/valid	91	1.10	1 instance of LF attested.
98	1570.	dažnas	frequent/habitual/periodic/repeated	90	0.00	
99	1579.	pilnas	full/whole/complete/plump	90	10.00	
100	1614.	sausas	dry/arid/droughty/dead	88	1.14	1 instance of LF attested.
101	1625.	lankstus	flexible/supple/versatile	87	0.00	
102	1628.	modernus	modern/up-to-date/contemporary	87	19.54	
103	1633.	silpnas	weak/fragile/frail/lax	87	10.34	
104	1638.	universalus	universal/versatile/all-around	87	12.64	
105	1644.	galingas	powerful/mighty/potent/strong	86	6.98	
106	1664.	idealus	ideal/perfect	85	**34.12**	p = 0.1281
107	1675.	populiarus	popular	85	2.35	

No	Rank in FrD	Lexeme	Translation	Total count of LFs in FrD	% of LFs	Comment
108	1698.	reikšmingas	significant/meaningful/important/weighty	84	0.00	
109	1704.	visiškas	complete/total/full/absolute/superior/superb	84	0.00	
110	1706.	atsakingas	responsible/liable/accountable	83	0.00	
111	1709.	gimtas	native/inborn/innate	83	**89.16**	**p = 0.0000**

Appendix D.

Comments:

- The 9 adjectives identified in Appendix A. with the relaxed p-value less than 0.10 were examined in the Sketch Engine family corpus *Lithuanian Web 2014* (ltTenTen14).
- The distribution of percentage of LF *versus* SH was calculated based on counts in the corpus.
- Some data manipulation was carried our (see Section 5.2 for a detailed explanation). Check the paradigms of individual adjectives below for raw data and explanations on data calculation.
- The findings are summarised below:

Table 1. Adjectives listed alphabetically

No	Alphabetically	Translation	% LF
1	ankstyvas	early/precocious	42.57
2	didis	great/famous/sublime	92.75
3	gimtas	native/inborn/innate	83.47
4	grynas	pure/net/neat/clear/clean	36.79
5	pastaras	the latter/the recent	99.88
6	senas	old/ancient	39.89
7	šventas	holy/sacred/saint/blessed/sacrosanct	69.45
8	viešas	public/open	67.33
9	vyresnis	older/senior/superior	41.88

Table 2. Adjectives listed according to their % of LF values (highest to lowest)

No	Highest to lowest	Translation	% LF	diff %LF
1	Pastaras	the latter/the recent	99.88	n/a
2	Didis	great/famous/sublime	92.75	7.13
3	gimtas	native/inborn/innate	83.47	9.28
4	šventas	holy/sacred/saint/blessed/sacrosanct	69.45	14.02
5	viešas	public/open	67.33	2.12
6	ankstyvas	early/precocious	42.57	**24.76**
7	vyresnis	older/senior/superior	41.88	0.69
8	senas	old/ancient	39.89	1.99
9	grynas	pure/net/neat/clear/clean	36.79	3.10

- As demonstrated in Table 2., the adjectives can be split into two groups: 1) those whose values are higher than 50% implying that their use with LFs exceeds their use with SFs; 2) those whose values are under 50% implying that they are more frequently used with SFs. Notably, there is a change in values with a difference of 24.76 between the adjectives *viešas* and *ankstyvas*.

Paradigms of individual adjectives created based on the data encountered in ltTenTen14:

PASTARAS	M			PASTARA	F			PASTARASIS	M			PASTAROJI	F		
Sg		**Pl**		**Sg**		**Pl**		**Sg**		**Pl**		**Sg**		**Pl**	
Nom	38	**Nom**	0	**Nom**	1	**Nom**	n/a	**Nom**	23755	**Nom**	16963	**Nom**	12999	**Nom**	n/a
Gen	15	**Gen**	10	**Gen**	5	**Gen**	n/a	**Gen**	12518	**Gen**	20523	**Gen**	9971	**Gen**	n/a
Dat	6	**Dat +s**	11	**Dat**	7	**Dat**	1	**Dat**	2871	**Dat +s**	2192	**Dat**	1404	**Dat**	388
		Dat -s	0							**Dat -s**	6				
Acc	40	**Acc**	21	**Acc**	n/a	**Acc**	n/a	**Acc**	13029	**Acc**	40586	**Acc**	3831	**Acc**	3904
Inst	25	**Inst**	41	**Inst**	n/a	**Inst**	10	**Inst**	55555	**Inst**	23205	**Inst**	516	**Inst**	4402
Loc	12	**Loc +e**	15	**Loc +e**	19	**Loc**	8	**Loc**	1610	**Loc +e**	466	**Loc +e**	1458	**Loc**	488
		Loc -e	1	**Loc -e**	4	**"om"**	6			**Loc -e**	2	**Loc -e**	24	**"osiom"**	152
	136		99		36		25		109338		103943		30203		9182

2 total count 46, adjusted to 2 as contains typos

TOTAL SF **296**

TOTAL LF **252666**

% LF 99.88299

DIDIS	M			DIDI	F			DIDYSIS	M			DIDŽIOJI	F		
Sg		**Pl**		**Sg**		**Pl**		**Sg**		**Pl**		**Sg**		**Pl**	
Nom	6159	**Nom**	264	**Nom**	2483	**Nom**	n/a	**Nom**	24662	**Nom**	13167	**Nom**	59287	**Nom**	n/a
Gen	1504	**Gen**	1645	**Gen**	1462	**Gen**	n/a	**Gen**	50826	**Gen**	27737	**Gen**	58625	**Gen**	n/a
Dat	662	**Dat +s**	131	**Dat**	4464	**Dat**	42	**Dat**	3082	**Dat +s**	1800	**Dat**	4242	**Dat**	1247
		Dat -s	2							**Dat +s**	8				
Acc	1531	**Acc**	852	**Acc**	1303	**Acc**	235	**Acc**	8769	**Acc**	5345	**Acc**	33450	**Acc**	4272
Inst	1203	**Inst**	352	**Inst**	2111	**Inst**	73	**Inst**	3954	**Inst**	1336	**Inst**	6381	**Inst**	1692
Loc	95	**Loc +e**	97	**Loc +e**	297	**Loc**	28	**Loc**	3459	**Loc +e**	7485	**Loc +e**	23447	**Loc**	1374
		Loc -e	6	**Loc -e**	52	**"om"**	15			**Loc -e**	8	**Loc -e**	175	**"osiom"**	63
	11154		3349		12120		393		94752		56886		185432		8648

4464	contains homonyms, as *didžiai* → dat.sg.f ≠ adverb
8	total count 33, adjusted to 8 of true loc., as contains homonyms with the verb 'didžiuotis' 1sg.fut + typos

TOTAL SF **27016**

TOTAL LF **345718**

% LF 92.75194

GIMTAS	M			GIMTA	F			GIMTASIS	M			GIMTOJI	F		
Sg		**Pl**		**Sg**		**Pl**		**Sg**		**Pl**		**Sg**		**Pl**	
Nom	200	**Nom**	5471	**Nom**	237	**Nom**	n/a	**Nom**	1891	**Nom**	379	**Nom**	5015	**Nom**	n/a
Gen	292	**Gen**	1291	**Gen**	315	**Gen**	n/a	**Gen**	4779	**Gen**	1881	**Gen**	8756	**Gen**	n/a
Dat	130	**Dat +s**	11	**Dat**	50	**Dat**	2	**Dat**	1162	**Dat +s**	62	**Dat**	929	**Dat**	27
		Dat -s	0							**Dat -s**	0				
Acc	584	**Acc**	160	**Acc**	n/a	**Acc**	n/a	**Acc**	4126	**Acc**	1263	**Acc**	5063	**Acc**	736
Inst	172	**Inst**	18	**Inst**	n/a	**Inst**	112	**Inst**	422	**Inst**	120	**Inst**	4823	**Inst**	210
Loc	143	**Loc +e**	93	**Loc +e**	264	**Loc**	38	**Loc**	3385	**Loc +e**	527	**Loc +e**	2783	**Loc**	251
		Loc -e	6	**Loc -e**	42	**"om"**	3			**Loc -e**	1	**Loc -e**	65	**"osiom"**	1
	1521		7050		908		155		15765		4233		27434		1225

5471	contains homonyms with verb 'gimti' inf.
237	contains homonyms with *gimta* adj.neutr
50	contains homonyms with verb 'gimti' 2sg.subv
1	total count 4, adjusted to 1 as contains typos

TOTAL SF **9634**

TOTAL LF **48657**

% LF 83.47258

ŠVENTAS	M			ŠVENTA	F			ŠVENTASIS	M			ŠVENTOJI	F		
Sg		**Pl**		**Sg**		**Pl**		**Sg**		**Pl**		**Sg**		**Pl**	
Nom	8396	**Nom**	2151	**Nom**	7734	**Nom**	n/a	**Nom**	13787	**Nom**	4358	**Nom**	8890	**Nom**	n/a
Gen	4706	**Gen**	3569	**Gen**	4002	**Gen**	n/a	**Gen**	14714	**Gen**	11671	**Gen**	21419	**Gen**	n/a
Dat	523	**Dat +s**	174	**Dat**	3568	**Dat**	151	**Dat**	1401	**Dat +s**	660	**Dat**	1283	**Dat**	47
		Dat -s	7							**Dat -s**	3				
Acc	4833	**Acc**	1133	**Acc**	n/a	**Acc**	n/a	**Acc**	4779	**Acc**	2208	**Acc**	5563	**Acc**	701
Inst	1701	**Inst**	929	**Inst**	n/a	**Inst**	793	**Inst**	3385	**Inst**	1446	**Inst**	2004	**Inst**	89
Loc	334	**Loc +e**	161	**Loc +e**	659	**Loc**	429	**Loc**	1916	**Loc +e**	165	**Loc +e**	4065	**Loc**	173
		Loc -e	1	**Loc -e**	65	**"om"**	94			**Loc -e**	0	**Loc -e**	90	**"osiom"**	4
	20493		8125		16028		1467		39982		20511		43314		1014

7734	contains homonyms with *šventa* adj.neutr
3568	contains many homonyms, as *šventai* → dat.sg.f ≠ adverb
0	total count 1, adjusted to 0 as contains a typo

TOTAL SF **46113**

TOTAL LF **104821**

% LF 69.44824

VIEŠAS	M			VIEŠA	F			VIEŠASIS	M			VIEŠOJI	F		
Sg		**Pl**		**Sg**		**Pl**		**Sg**		**Pl**		**Sg**		**Pl**	
Nom	9643	**Nom**	4365	**Nom**	5775	**Nom**	n/a	**Nom**	8886	**Nom**	5513	**Nom**	13427	**Nom**	n/a
Gen	8264	**Gen**	4206	**Gen**	3588	**Gen**	n/a	**Gen**	53230	**Gen**	54124	**Gen**	51493	**Gen**	n/a
Dat	1253	**Dat +s**	391	**Dat**	58601	**Dat**	243	**Dat**	3076	**Dat +s**	2039	**Dat**	3658	**Dat**	1280
		Dat -s	3							**Dat -s**	4				
Acc	7842	**Acc**	2154	**Acc**	n/a	**Acc**	n/a	**Acc**	9081	**Acc**	8317	**Acc**	8338	**Acc**	3670
Inst	1221	**Inst**	705	**Inst**	n/a	**Inst**	300	**Inst**	5073	**Inst**	1795	**Inst**	1572	**Inst**	722
Loc	3370	**Loc +e**	1450	**Loc +e**	7298	**Loc**	5362	**Loc**	6466	**Loc +e**	1828	**Loc +e**	13383	**Loc**	3374
		Loc -e	1	**Loc -e**	312	**"om"**	11			**Loc -e**	2	**Loc -e**	84	**"osiom"**	5
	31593		13275		75574		5916		85812		73622		91955		9051

4365	contains homonyms with verb 'viešėti' 3.prs
5775	contains homonyms with *vieša* adj.neutr
58601	contains homonyms, as *viešai*→ dat.sg.f ≠ adverb
2	total count 22, adjusted to 2 as contains typos

TOTAL SF **126358**

TOTAL LF **260440**

% LF 67.3323

ANKSTYVAS	M			**ANKSTYVA**	F			**ANKSTYVASIS**	M			**ANKSTYVOJI**	F		
Sg		**Pl**		**Sg**		**Pl**		**Sg**		**Pl**		**Sg**		**Pl**	
Nom	3823	**Nom**	506	**Nom**	2038	**Nom**	n/a	**Nom**	924	**Nom**	1427	**Nom**	1368	**Nom**	n/a
Gen	4605	**Gen**	878	**Gen**	3265	**Gen**	n/a	**Gen**	3021	**Gen**	2664	**Gen**	3606	**Gen**	n/a
Dat	305	**Dat +s**	48	**Dat**	255	**Dat**	45	**Dat**	220	**Dat +s**	123	**Dat**	256	**Dat**	61
		Dat -s	0							**Dat -s**	0				
Acc	6374	**Acc**	319	**Acc**	n/a	**Acc**	n/a	**Acc**	477	**Acc**	595	**Acc**	702	**Acc**	326
Inst	349	**Inst**	356	**Inst**	n/a	**Inst**	83	**Inst**	827	**Inst**	979	**Inst**	176	**Inst**	90
Loc	678	**Loc +e**	92	**Loc +e**	2722	**Loc**	346	**Loc**	623	**Loc +e**	285	**Loc +e**	1131	**Loc**	255
		Loc -e	0	**Loc -e**	80	**"om"**	3			**Loc -e**	0	**Loc -e**	6	**"osiom"**	0
	16134		2199		8360		477		6092		6073		7245		732

2038	contains homonyms with *ankstyva* adj.neutr
255	contains homonyms, as *ankstyvai* → dat.sg.f ≠ adverb

TOTAL SF **27170**

TOTAL LF **20142**

% LF 42.57271

VYRESNIS	M			VYRESNĖ	F			VYRESNYSIS	M			VYRESNIOJI	F		
Sg		**Pl**		**Sg**		**Pl**		**Sg**		**Pl**		**Sg**		**Pl**	
Nom	5835	**Nom**	10161	**Nom**	2164	**Nom**	n/a	**Nom**	10437	**Nom**	5933	**Nom**	6852	**Nom**	n/a
Gen	17654	**Gen**	8417	**Gen**	3355	**Gen**	n/a	**Gen**	4178	**Gen**	8483	**Gen**	2432	**Gen**	n/a
Dat	1140	**Dat +s**	5844	**Dat**	355	**Dat**	592	**Dat**	1052	**Dat +s**	1130	**Dat**	663	**Dat**	62
		Dat -s	31							**Dat -s**	1				
Acc	999	**Acc**	2014	**Acc**	517	**Acc**	474	**Acc**	894	**Acc**	1130	**Acc**	538	**Acc**	57
Inst	1051	**Inst**	1392	**Inst**	657	**Inst**	141	**Inst**	1684	**Inst**	610	**Inst**	637	**Inst**	24
Loc	1432	**Loc +e**	55	**Loc +e**	70	**Loc**	662	**Loc**	20	**Loc +e**	2	**Loc +e**	21	**Loc**	11
		Loc -e	0	**Loc -e**	1	**"ėm"**	16			**Loc -e**	0	**Loc -e**	2	**"osiom"**	1
	28111		27914		7119		1869		18265		17289		11145		155

TOTAL SF **65013**

TOTAL LF **46854**

% LF	**41.88367**

SENAS	M			SENA	F			SENASIS	M			SENOJI	F		
Sg		**Pl**		**Sg**		**Pl**		**Sg**		**Pl**		**Sg**		**Pl**	
Nom	31616	**Nom**	14974	**Nom**	18442	**Nom**	n/a	**Nom**	9482	**Nom**	9227	**Nom**	10446	**Nom**	n/a
Gen	25474	**Gen**	28872	**Gen**	15411	**Gen**	n/a	**Gen**	18873	**Gen**	19130	**Gen**	23507	**Gen**	n/a
Dat	1893	**Dat +s**	2073	**Dat**	15203	**Dat**	415	**Dat**	956	**Dat +s**	839	**Dat**	1027	**Dat**	393
		Dat -s	47							**Dat -s**	1				
Acc	17969	**Acc**	12405	**Acc**	n/a	**Acc**	n/a	**Acc**	6364	**Acc**	5577	**Acc**	6570	**Acc**	4545
Inst	4592	**Inst**	6745	**Inst**	n/a	**Inst**	1491	**Inst**	2373	**Inst**	1953	**Inst**	2348	**Inst**	818
Loc	2142	**Loc +e**	1851	**Loc +e**	1389	**Loc**	1281	**Loc**	4608	**Loc +e**	1226	**Loc +e**	3303	**Loc**	2081
		Loc -e	7	**Loc -e**	156	**"om"**	247			**Loc -e**	1	**Loc -e**	94	**"osiom"**	5
	83686		66974		50445		3434		42656		37954		47295		7842

25474	contains homonyms with verb 'senti' 3.pst
14974	contains homonyms with noun 'senis' voc.sg.m
18442	contains homonyms with *sena* adj.neutr
15203	contains homonyms, as *senai*→ dat.sg.f ≠ adverb
247	contains homonyms with verb 'senti' 1pl.pst
1	total count 13, adjusted to 1 due to typos

TOTAL SF **204539**

TOTAL LF **135747**

% LF 39.89203

GRYNAS	M			GRYNA	F			GRYNASIS	M			GRYNOJI	F		
Sg		**Pl**		**Sg**		**Pl**		**Sg**		**Pl**		**Sg**		**Pl**	
Nom	8690	**Nom**	1048	**Nom**	3666	**Nom**	n/a	**Nom**	3278	**Nom**	1365	**Nom**	904	**Nom**	n/a
Gen	4486	**Gen**	1667	**Gen**	1475	**Gen**	n/a	**Gen**	3137	**Gen**	9054	**Gen**	1493	**Gen**	n/a
Dat	220	**Dat +s**	40	**Dat**	13283	**Dat**	12	**Dat**	85	**Dat +s**	123	**Dat**	42	**Dat**	36
		Dat -s	1							**Dat -s**	0				
Acc	2876	**Acc**	695	**Acc**	n/a	**Acc**	n/a	**Acc**	692	**Acc**	1712	**Acc**	302	**Acc**	173
Inst	2017	**Inst**	3187	**Inst**	n/a	**Inst**	103	**Inst**	252	**Inst**	4241	**Inst**	262	**Inst**	45
Loc	3177	**Loc +e**	81	**Loc +e**	58	**Loc**	14	**Loc**	19	**Loc +e**	11	**Loc +e**	16	**Loc**	2
		Loc -e	0	**Loc -e**	3	**"om"**	13			**Loc -e**	0	**Loc -e**	0	**"osiom"**	0
	21466		6719		18485		142		7463		16506		3019		256

3666	contains homonyms with *gryna* adj.neutr
13283	contains homonyms, as *grynai* dat.sg.f ≠ adverb

TOTAL SF **46812**

TOTAL LF **27244**

% LF 36.78838

Vilnius University Press
9 Saulėtekio Ave., Building III,
LT-10222 Vilnius, Lithuania
Email: info@leidykla.vu.lt,
www.leidykla.vu.lt
Print run 15