Valoda: nozīme un forma / Language: Meaning and Form 12, 2021, 7–30.
https://doi.org/10.22364/vnf.12.01

Grammatical patterns in the corpus-driven "Lexical Database of Lithuanian"

„Lietuviešu valodas lietošanas leksikona" gramatisko modeļu korpusa pētījums

Agnė Bielinskienė, Jolanta Kovalevskaitė, Erika Rimkutė

Vytautas Magnus University
V. Putvinskio str. 23–216, LT-44243, Kaunas, Lithuania
E-mail: *agne.bielinskiene@vdu.lt, jolanta.kovalevskaite@vdu.lt, erika.rimkute@vdu.lt*

This paper describes the grammatical patterning of two parts of speech – nouns and adjectives – included in the corpus-driven "Lexical Database of Lithuanian" as a foreign language. The lexical database is a lexicographic application of the Lithuanian Pedagogic Corpus (approx. 620.000 tokens) which was used to develop headword lists and to collect word usage information in the form of corpus patterns. In this project, we adopted a partially automated inductive procedure of Corpus Pattern Analysis for 207 verbs, 386 nouns, 87 adjectives, and 41 adverbs. The detected corpus patterns reflect different meanings of the headword. Each pattern presents information on grammatical, semantic, and lexical levels. Manually selected examples illustrate all pattern components.
In this paper, 673 patterns with nouns and 99 patterns with adjectives will be analysed discussing their syntactic behaviour in detail and providing some comments on lexis-grammar interface. The majority of patterns with nouns and adjectives are minimal patterns which include only the closest syntactical partners. This result is influenced by different procedures used to describe patterns with nouns, adjectives, and adverbs and patterns with verbs. Due to rich grammatical information, there are several similar patterns with one main (usually the most frequent) type and its variants. Pattern variants show that the grammatical characteristics of a specific word usage are rather individual.

Keywords: lexical database; Lithuanian language; "Lithuanian Pedagogic Corpus"; Corpus Pattern Analysis; pattern.

1. Introduction

Within the framework of the project "Lithuanian Academic Scheme for International Cooperation in Baltic Studies" (http://baltnexus.lt/en/baltic-studies-project) (2018–2020), a new "Lexical Database of Lithuanian as a Foreign Language" (hereinafter, database) is being prepared (more on the database see

Kovalevskaitė et al. 2020a). The target group of the database are A2–B2 (according to the "Common European Framework of Reference for Languages") learners of Lithuanian. The material for the database has been extracted from the subcorpus of written texts of the "Lithuanian Pedagogic Corpus" (Kovalevskaitė, Rimkutė 2020a), which consists of Lithuanian language textbooks and other learner-relevant authentic Lithuanian language material grouped according to different levels. Until now, there have been no similar lexicographical resources for Lithuanian. Both the corpus and the database are freely available for learners, teachers, and coursebook writers in the portal https://kalbu.vdu.lt/ since 2021.

Insights from the contextual and functional theory of meaning and the approach of corpus linguistics were adopted in database development. According to Sinclair (1991), not isolated words, but rather words in their contextual patterns are the true bearers of meaning. The main purpose of the database is to provide learners and teachers with appropriate material on language use. We chose to present information about word meaning and usage in the form of patterns by adapting one of the techniques of corpus-driven methods – the Corpus Pattern Analysis (Hanks 2008) to the Lithuanian language (see Section 2).

This article aims to describe the grammatical features of headwords stored in the database. Although grammatical information is only one of the three levels in word usage description (semantic information and the most typical collocates are also provided, see Section 3), it can show the differences in meaning related patterning and exemplify pattern variation of the same word class. Due to the large amount of the data, this article will focus on the grammatical patterns with nouns and adjectives, whereas patterns with verbs and adverbs are going to be described in further research.

2. Adaptation of Corpus Pattern Analysis

For the description of word usage, we used a slightly modified CPA approach (Corpus Pattern Analysis (CPA) (Hanks 2008; 2012)). CPA is based on the contextual and functional theory of meaning, where the meaning of a word is associated with a specific lexical and grammatical environment. "The Pattern Dictionary of English Verbs" (http://pdev.org.uk/) applied the CPA technique to describe the English verb usage: information in the dictionary is presented as a set of patterns for each verb entry. A pattern is a syntagmatic structure with semantic values for arguments (i.e., semantic types are indicated for each argument). For example, one of the three patterns for the verb ATTEND is presented in the "Pattern Dictionary of English Verbs" as follows:

(1) [[Human]] **attends** ([[Event]] or [[Location]])
Implicature: [[Human]] is present at [[Event]] or [[Location]]
*Example: The 18-year-olds are both **attending** the optional competition class at the college*

The entry consists of a pattern with its semantic values ([Human], [Event], [Location]), an implicature which explains the meaning, and a usage example. A semantic type [Event] is a generalization of a lexical set *meeting, conference, ceremony, seminar, dinner, reception*, etc.

Although CPA has mostly been applied to verbs, Hanks (2008, 117) observes that the method can also be adapted to nouns. However, research on the application of the CPA for the latter part of speech is still scarce (e.g., Hanks et al. 2019). When applying the CPA method to other languages and projects, researchers follow a rather limited set of guidelines. We explain our description strategy for nouns, adjectives, and adverbs in Section 2.2.

2.1. Recognition of patterns

For the recognition and definition of meaning-related patterning, we applied both automatic ("SketchGrammar" to detect grammatical patterns) and manual procedures (to identify semantic types and meaning and select examples). The database includes pattern descriptions of headwords which occur 100 or more times in the "Lithuanian Pedagogic Corpus", specifically, its subcorpus of written texts (approx. 620.000 tokens). For a more detailed description of pattern recognition procedure, see Kovalevskaitė et al. (2020b, 247–251).

First, "SketchGrammar" generated information on word patterning was analysed. This semi-automatic procedure helped to obtain grammatical and lexical information about the node-words from concordances (see Kovalevskaitė et al. 2020a, 248 for more details on "SketchGrammar" for Lithuanian). As "SketchGrammar" recognizes syntagmatic relations consisting of only two words, quite often linguists had to manually analyse longer structures to identify cases in which several elements function as one pattern component (adverbial, subject, object, or attribute). For example, a phrase *gero būdo žmogus* 'an easy-going person' must be described as an attributive pattern consisting of two components: a complex attribute [AtrA BŪDAS_sg.gen] and a modifier [Mod], which is realized semantically by a semantic type [human] and lexically by a collocate *žmogus* 'a person' (on the multilevel pattern structure, see Section 2.2.).

Next, we identified grammatical patterns, analysed the lexical sets of collocates and manually assigned them to semantic types. In this respect, our approach was like the one applied in the Tecling project in which usage patterns were "identified on the basis of syntactic functions and semantic types" (Giacomini et al. 2019, 495). On the other hand, in the Italian as a Foreign Language dictionary project, first, collocations of a node-word were collected and then concordance lines referring to these collocations were extracted for pattern identification (Giacomini et al. 2019, 495–496). Our pattern recognition approach turned out to be rather challenging for linguists: they had to deal simultaneously with high grammatical, lexical, and semantic data complexity. Thus, it might be assumed that the Italian as a Foreign Language approach could be practically easier and less time consuming.

One of the most challenging tasks was to assign the semantic types to lexical sets and determine the appropriate degree of generalization. In CPA, semantic types refer to a shallow semantic ontology. We chose to manually assign the semantic type when generalizing semantically similar lexical sets (groups of collocates) and used a bottom-up approach for noun description, like the one of the Italian as a Foreign Language dictionary project, i.e., semantic types were assigned during the pattern identification process (Giacomini et al. 2019). In the case of verbs and adjectives, we used three predefined semantic types. However, in further database development, more than three verb categories might be used, depending on the noun analysed. While analysing the noun *way,* Hanks et al. (2019, 252) observe that "any generalization about the syntactic and semantic properties of the governing verbs will have to be made for each and every pattern separately".

Our approach can be described as experimental in its nature. We see the patterns provided in the lexical database as an intermediate stage of analysis between the concordance of the word and its entry in the learner dictionary. Some parts of the CPA procedure (for example, pattern sorting) were not included. We maintain that the user will be able to judge the pattern frequency from the number of examples provided with that pattern.

2.2. Description of patterns

As Lithuanian is a highly inflected language, more detailed grammatical information might be especially relevant for the target audience of the database. Accordingly, in the adapted CPA procedure, we decided to use a **multilevel description of a pattern**, where grammatical (gramForm), semantic (semForm) and lexical (collocates) realizations of a node-word pattern are provided:

(2) "gramForm": [ARBATA] [su AtrN_ins] / [TEA] [with AtrN_ins]
"semForm": [Mod] [maistas] / [Mod] [food]
"collocates": [ARBATA] [su citrina] / [TEA] [with lemon]

The first level presents the information on the grammatical collocability and inflections, the second level shows semantic types related to grammatical categories, the third level lists the collocates of semantic types, and the fourth level (absent in the example (2)) shows examples from the corpus. The manually selected examples illustrate all pattern components, i.e., grammatical, semantical, and lexical features.

As shown in example (2), the node-word is capitalized and separate components in the levels "gramForm", "semForm" and "collocates" are given in square brackets. The variability of elements in the pattern is indicated with a vertical slash "|" (meaning 'either – or'). For example, in the pattern [Pred] [CENTRAS_acc]|[CENTRAS_ins], the object is expressed in the accusative or instrumental case. Variability criterion is particularly relevant for the syntactic functions of Object, Adverb and Predicate, as it is often necessary to include several adverbs, objects, or types of predicates in the same pattern. The node-word might

syntactically be a head of the phrase or a dependent element, e.g., the analysed node (marked as X) might occur in different positions as a modifier or a modified word: [AtrA] [**X**]: *jaunas* ***žmogus*** 'young **person**' and [**X**_gen] [Mod]: ***žmogaus*** *gyvenimas* '**person's** life'. Further on, the analysed node presented in Lithuanian examples and their translations into English is given in bold.

When describing grammatical patterns, **morphological categories** were marked using the Leipzig Glossing Rules, whereas **syntactic categories** were taken from the syntactically annotated Lithuanian corpus ALKSNIS (https://clarin.vdu.lt/xmlui/handle/20.500.11821/21). Syntactic categories were marked in accordance with the international abbreviations listed in Table 1. The English terms for syntactic categories and syntactic relations are based on Ambrazas (2006, 475; 486; 494; 463; 466; 468; 489; 493; 590).

Subject [Sub]	**Object [Obj]**	**AtrA**	**AtrN**	**Pred**	**Pred+N/ Pred+V**	**Adv**	**Mod**	**AtrQ**
+ case if not nominative	+ case information	attributes in concord	attributes not in concord: can be expressed as a preposition+ case	simple predicate	compound nominal or verbal predicate	adverbial	modifier	partitive genitive of quantity

Table 1. Syntactic and morphological categories used in patterns

If a certain usage was identified in the corpus, its specific grammatical features were reflected in the grammatical pattern (such as case, number, verb form, time, finite/non-finite form, degree, etc.). For example, in the pattern [X] [AtrN_inf]: ***vieta*** *pavalgyti* '**place** to eat', there is an additional marking indicating that this attribute not in concord is always used in infinitive form (*inf*); the pattern [X_sup] [Mod]: ***būtiniausius*** *darbus* 'the **most essential** works' shows that the adjective *būtinas* 'essential' is likely to be used in the superlative degree in the pattern with modifier. Due to these features, patterns with unique marking appeared: together with the main pattern they form one group and are labelled as pattern variants.

Patterns may reflect structures with fixed and variable lexical units. One of the patterns of a node-word *įprotis* 'habit' is [neturintis|neturinti žalingų X_pl.gen] [Mod] in which the [AtrN] complex component [neturintis|neturinti žalingų X_pl.gen] includes fixed lexis, e.g., *neturintis žalingų* ***įpročių*** *vyras* 'a man without bad **habits**'. In some cases, variable lexical units of the pattern are be indicated as, for example, various numerals which can occur with fixed lexis and are marked as x: [x|x milijonas|x tūkstantis X_pl.gen] [Mod]: *150* ***eurų*** *baudą* 'a fine of 150 **euro**'; *18 tūkst.* ***eurų*** *lengvata* 'a relief of 18 000 **euro**'.

In this lexicographic project, we use a broader understanding of grammatical collocability (Labutis 2002), accordingly, some components are included in the patterns even if from the perspective of valency they would be treated as non-obligatory. As a result, this increases the number and variability of grammatical patterns as, for example, in the case of verb combinations with various adverbs.

The collocability description of nouns, adjectives, and adverbs is different from verb description procedure. The usage patterns with nouns, adjectives, and adverbs include information about components which are directly related to the described node. Thus, components that are governed by the verb (such as subject, adverbial and objects) do not appear in the grammatical patterns with nouns, adjectives, and adverbs. For example, in the phrase *mąstyti AIŠKIAI apie Obj_acc*, the adverb *aiškiai* 'clearly' is related to the verb *mąstyti* 'to think', however, as the object [apie Obj_acc] is a verb-dependant component, it will not be included in the verbal pattern of the adverb *aiškiai* 'clearly': [AIŠKIAI] [Pred]. In the pattern description of the verb MĄSTYTI 'to think', all components (adverbial and the object) would be included as they are directly related to the node. As a result of this description procedure, the majority of patterns with nouns, adjectives and adverbs are minimal patterns which include only the closest syntactical partners.

In actual usage, one minimal pattern may be a constituent part of another minimal pattern. For example, the minimal attribute patterns with a noun may be integrated in the verbal pattern of the same noun. In the attribute pattern with the noun [AtrA] [ARBATA], the attribute [AtrA] may include adjectives indicating classificational characteristics (e.g., *juodoji* 'black', *žalioji* 'green', *šalta* 'cold') or physical properties (e.g., *karšta* 'hot', *kvapni* 'fragrant'), still, all these attributes may be included in the realisations of the verbal pattern with the noun ARBATA 'TEA' where *arbata* 'tea' is an object. The verbal pattern is [Pred] [ARBATA_gen]|[su ARBATA_ins], e.g.:

(3) – *Ko norėsit? – Žalios arbatos*. '– What would you like? – Green tea.'

(4) *Galite paruošti šaltos arbatos ir kavos?* 'Could you make cold tea and coffee?'

(5) *Vakare pakaitinkite kojas, dažnai gerkite karštos arbatos su medumi*. 'In the evening, keep your feet warm and drink hot tea with honey regularly.'

In example (5), the noun *arbata* 'tea' has two attributes – this attributive usage is recorded in two distinct patterns: [AtrA] [ARBATA] and [ARBATA] [su AtrN_ins].

Such a description of nouns, adjectives and adverb patterns allows us to show that different noun meanings can be identified in two-word combinations. The presented procedure was also chosen because of the practical reasons: if the attributive pattern is described separately, the structure of the verbal pattern is less complex. For example, the minimal pattern [AtrA] [X]: *geras žmogus* 'good person' is not included in the verbal pattern [Sub] [Pred+X]: *tu esi / geras* ***žmogus*** / 'you are/a good **person**', but it is implied (and shown by examples) that the noun in the compound nominal predicate can be used with the attribute.

Some minimal patterns with nouns may not accurately reflect the actual usage: in [AtrA] [LAIKAS] pattern, combinations in which the noun *laikas* 'time' is used with adjectives function as adverbial of time. To exemplify: ***ilgą laiką*** *buvo vadinamas...* '**for a long time** was named...'; *atidėti* ***ilgesniam laikui*** 'to postpone **for a longer time**', etc. In such cases, there is no need to present the attributive model [AtrA] [LAIKAS].

Overall, this section showed different procedures used to describe patterns with nouns, adjectives, and adverbs and patterns with verbs. It is also important to mention that other researchers include more distant syntactic partners in noun models, claiming that the aim of word usage description is "to present as detailed picture of usage as possible" (Marcinkevičienė 2000, 76). Hanks et al. (2019) are prone to tackle the usage complexities in both noun and verb descriptions: i.e., they maintain that when describing nouns in verbal patterns, elements that are governed by the noun (determiners and modifiers) should also be specified. If Hanks et al. (2019) position that syntactically subordinate parts of the noun phrases should be described in the verbal pattern had been adopted in the present project, verbal patterns with nouns would have included more elements governed by the noun and there would have been no need for minimal patterns. In this description, attributive patterns with nouns or adjectives (the most frequent type of minimal patterns) are very common. As a result, not all recognized patterns are distinct from each other concerning their grammatical level.

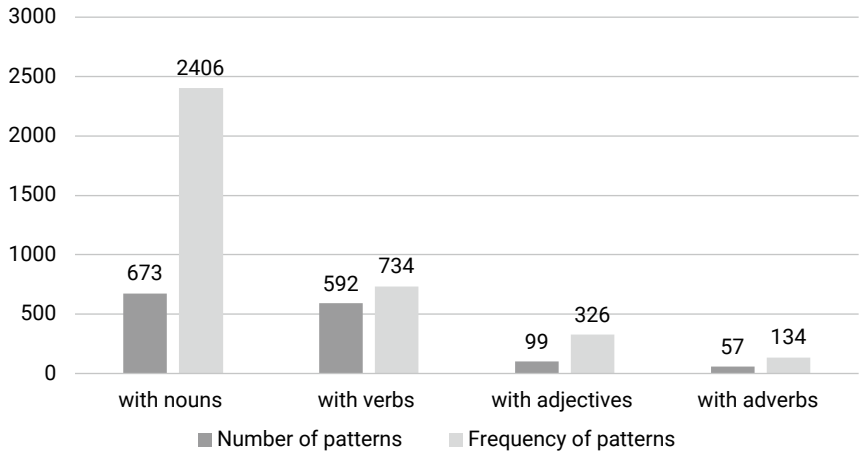
3. Grammatical patterns with nouns and adjectives

3.1. Number and frequency of grammatical patterns

The database includes entries for four-word classes (207 verbs, 386 nouns, 87 adjectives, and 41 adverbs) and presents corpus-driven information on their usage (Kovalevskaitė et al. 2020a). This section will give an overview of the identified grammatical patterns with nouns, verbs, adjectives, and adverbs. The number and frequency information of the grammatical patterns of these parts of speech (including verbs and adverbs, although they are not in the focus of the current work) is given in Table 1.

The analysis of the amount of pattern types shows that patterns with nouns (673) and verbs (592) prevail, whereas there are less patterns with adjectives (99) and adverbs (57). The frequency of patterns was calculated by considering all use cases of a particular pattern in the lexical entries: a pattern might apply to the concrete lexeme and to the several meanings of the same lexeme. For example, three different meanings of the adjective *geras* 'good' have the same pattern [GERAS] [Mod], where adjective functions as an attribute in concord.

Minimal patterns are especially frequent. For example, there are approx. 1000 occurrences of attributive patterns with nouns. However, some patterns apply to the usage of only one word (a certain patterning of a node-word is recognized as a pattern if it occurs in concordances no less than 3 times. Thus, even if



Picture 1. Number of pattern types and their total frequency

a certain pattern applies only to one word (or meaning), it occurs in the corpus more than one time). As the "Lithuanian Pedagogic Corpus" is relatively small (620.000 tokens), unique patterns amount to one-third or one-fourth in each group of all analysed node-words, specifically, there are 23% patterns with nouns, 25% with adjectives, 30% with adverbs, and even 74% with verbs (compare the relation between the number of pattern types and their total frequency in Picture 1). Thus, unique patterns with verbs are less frequent and describe the patterning of one verb. This outcome results from the different procedures used to describe patterns with nouns, adjectives, and adverbs in comparison to patterns with verbs, as explained in Section 2.2.

3.2. Types of grammatical patterns

This section presents grammatical patterns (further on, patterns) with nouns (Section 3.2.1.), and adjectives (Section 3.2.2.). First, the main (generalizing) pattern (presented in bold in all tables) is given, followed by the details on its variants. The observed trend of the main pattern frequency shows that in nearly all cases, the main patterns prevail, whereas their variants are either less frequent or occur only once. The data in the tables will illustrate only the patterns discussed in this article, whereas the database contains more of them.

3.2.1. Grammatical patterns with nouns

There are 673 nominal patterns in total. As most often the patterns coincide with word combinations, they are classified syntactically and semantically. We distinguished two pattern types depending on the head word of the phrase: (1) nominal patterns

(see 3.2.1.1.) and (2) verbal patterns (see 3.2.1.2.). According to their semantic function, nominal patterns were grouped into: (1) attributive (see 3.2.1.1.1.), (2) objective (see 3.2.1.1.2.), (3) partitive genitive of quantity (AtrQ) (see 3.2.1.1.3.), (4) adverbial (see 3.2.1.1.4.) and (5) patterns with other functions (see 3.2.1.1.5.). In verbal patterns, nouns can function as subjects, as part of compound predicates, objects, and adverbials.

3.2.1.1. Nominal patterns

The patterns with nouns constitute a rather large category (1156). Predominantly (97.15%), these are attributive patterns (1123), while less than 3% are objective patterns, patterns of partitive genitive of quantity, nominal patterns with adverbials and others. If we compare nominal (3.2.1.1.) and verbal patterns (3.2.1.2.), their total frequency is similar (1123 and 1239, respectively). However, nominal patterns have fewer types and higher frequency, whereas verbal pattern types vary greatly, but are less frequent.

3.2.1.1.1. Attributive nominal patterns

The patterns of this type are very common (1123). The analysed nouns in such patterns can stand as (1) modifiers (Mod) with attributes in concord (AtrA) or not in concord (AtrN) or as (2) attributes not in concord (see Table 2).

In this and other tables, the main pattern(s) is/are presented in bold, the frequency of the pattern is given in brackets and examples illustrating the pattern are given in italics. In the examples, the node-word (X) is given in bold. After the main pattern(s), its/their variants are listed, with information presented in the same way as for the main pattern: pattern is followed by frequency numbers and examples. Some examples are given in head form, i.e., singular nominative case. If examples occur in other than singular nominative case in the corpus, they are provided in their original form.

Although there are fewer variants of the 1st and 2nd pattern types, they occur with the higher frequency (see Table 2). On the contrary, there is a higher variation of the 3rd type patterns, but they are not as common. In addition to the main syntactic function, information on the number and case, or only on the number if the word is uncountable or used in plural form, is provided (see lines 1.1, 1.2, and 2.1 in Table 2).

In the patterns of the 3rd type, the function of the attribute not in concord is performed by a complex unit. A complex attribute can include an adjective in concord [AtrA] with a noun or an adjective not in concord [AtrN] with a noun, e.g.: [AtrA X_gen] [Mod]; [AtrN X_gen] [Mod] (see Table 2, line 3). Depending on the node-word, pattern variants include information on number and case. The described nouns *amžius* ‘age’, *metai* ‘year’*, klasė* ‘category’ (see Table 2, line 3.1) are used in genitive case (most often singular) and in constructions with adjectives or numerals express various features: *įvairaus* ***amžiaus*** *vaikams* ‘for children of all **ages**’, *turistinės* ***klasės*** *viešbutyje* ‘in the economy **class** hotel’. Other variants of the pattern (see line 3.2) are characterized by nouns denoting a measure

1. [AtrN] [X] (315): *karo* ***metai*** ‘war **years**’; **[X_gen] [Mod]** (195): ***žmogaus*** *gyvenimas* ‘**person’s** life’
1.1. [AtrN] [X_sg] (21): *upės* ***vanduo*** ‘river **water**’; [AtrN] [X_pl] (19): *mano* ***pinigai*** ‘my **money**’; [AtrN] [X_loc] (2): *namo* ***viduje*** ‘**inside** the house’
1.2. [X_sg.gen] [Mod] (57): ***šeimos*** *gyvenimas* ‘**family** life’; [X_pl.gen] [Mod] (25): ***lietuvių*** *kalba* ‘**Lithuanian** language’; [X_sg] [Mod] (1): ***ekskursijos*** *kaina* ‘price of **excursion**’
2. [AtrA] [X] (308): *jaunas* ***žmogus*** ‘young **person**’
2.1. [AtrA] [X_pl] (15): *dideli* ***pinigai*** ‘big **money**’; [AtrA] [X_sg] (15): *šaltas* ***vanduo*** ‘cold **water**’; [AtrA]|[AtrN] [X] (1): *valgomoji* ***druska***|*jūros* ***druska*** ‘table **salt**’|‘sea **salt**’
3. [AtrA X_gen] [Mod] (1): *raudonos* ***spalvos*** *drabužius* ‘red **colour** clothes’; **[AtrN X_gen] [Mod]** (1): *lietuvių* ***kalbos*** *žodyno* ‘Lithuanian **language** dictionary’
3.1. [AtrA X_sg.gen] [Mod] (5): *įvairaus* ***amžiaus*** *vaikams* ‘for children of all **ages**’; [AtrN X_sg.gen] [Mod] (1): *plento* ***taurės*** *etapas* ‘highway **cup** stage’; [AtrA|AtrN X_sg.gen] [Mod] (1): *turistinės* ***klasės*** *viešbutyje* ‘in the economy **class** hotel’; [AtrA X_gen] [Mod] (1): *praėjusio* ***amžiaus*** *(pirmojoje) pusėje*|*XIX* ***amžiaus*** *pabaigoje* ‘in the (first half) of the last **century**’|‘in the end of the 19th **century**’
3.2. [x X_gen] [Mod] (9): *x* ***metų*** *olimpiada* ‘the Olympics of **year** x’; [x metrų|centimetrų X_gen]|[iki x metrų|centimetrų X_gen išaugti_ptcp.act] [Mod] (1): *2-6 m* ***aukščio*** *krūmai* ‘bushes 2-5 m **tall**’; [x X_pl.gen noun_sg.gen] [Mod] (1): *20* ***metrų*** *ilgio lazdos* ‘sticks of 20 **meters**’
3.3. [Mod] [su X_ins] (12): *padažas su* ***vynu*** ‘sauce with **wine**’; [Mod] [su X_sg.ins] (2): *kibiras su* ***vandeniu*** ‘bucket of **water**’; [Mod] [be X_gen] (2): *važiavimas be* ***bilieto*** ‘driving without a **ticket**’; [Mod] [iš X_gen] (2): *svečiai iš* ***užsienio*** ‘guests from **abroad**’; [Mod] [į X_acc] (2): *bilietas į* ***filmą*** ‘ticket to **film**’; [Mod] [iki x X_gen]|[nuo x iki x X_gen] (1): *vaikams iki 12* ***metų*** (‘for children up to 12 years’); [Mod] [iš viso X_gen] (1): *sportininkai iš viso* ***pasaulio*** ‘sportsmen from all over the **world**’; [Mod] [iš x X_sg.gen] (1): *vairuotojas iš 6* ***buto*** ‘driver from **flat** 6’
3.4. [X] [su AtrN_ins] (16): ***kambarys*** *su dušu* ‘**room** with shower’; [X] [apie AtrN_acc] (5): ***istoriją*** *apie tiltą* ‘**story** about the bridge’; [X] [su AtrN_ins] (5): ***vakarienei*** *su bulvėmis* ‘for the **supper** with potatoes’; [X] [į AtrN_acc] (4): ***teisė*** *į privatumą* ‘**right** to privacy’; [X] [iš AtrN_gen] (3): ***studentas*** *iš Vilniaus* ‘**student** from Vilnius’; [X] [po atviru dangumi] (1): ***muziejus*** *po atviru dangumi* ‘open-air **museum**’
3.5. [X] [apie AtrN_acc]|[dėl AtrN_gen] (1): ***klausimų*** *apie (plaukų) priežiūrą*|***klausimų*** *dėl (pajamų) stabilumo* ‘**questions** about (hair) care’/‘**questions** about (income) stability’; [X] [į AtrN_acc]|[per AtrN_acc]|[po AtrN_acc] (1): ***kelionė*** *į Havajus*|***kelionė*** *per Lenkiją*|***kelionės*** *po Prancūziją* ‘**trip** to Hawaii’|‘**trip** via Poland’|‘**trips** in France’
3.6. [X] [AtrN_inf] (6): ***vieta*** *pavalgyti* ‘a place to eat’; [X_sg] [AtrN_inf] (2): ***noras*** *bendrauti* ‘a wish to communicate’
3.7. [Mod] [X_pl.dat] (1): *filmai* ***suaugusiems*** ‘**films** for grown ups’ [Mod] [X_loc] (1): *nakvynė* ***viešbutyje*** ‘night at the **hotel**’; [X] [AtrN_dat] (3): ***meilė*** *sportui* ‘**love** for sports’; [X_pl] [AtrN_ins] (1): ***paslaugų*** *internetu* ‘internet **services**’.

Table 2. Attributive nominal patterns

describing a specific quantity: height, length, year, etc. They also include a flexible lexical item which denotes a certain measure and is marked as x.

Postpositionally used complex attributes not in concord can be seen as a distinct variant of the 3rd type pattern. In the example [X] [su AtrN_ins] *kambarys* ***su dušu*** 'room with **shower**', the attribute not in concord is expressed by a prepositional construction. The majority of the attributes are formed with prepositions *su, ant, apie, be, dėl, iki, į, iš, nuo, po, prie* 'with, on, about, without, due, till, to, from, under, at' and accompanying genitive, accusative or instrumental case forms. The node word can be in both positions of the pattern (see Table 2 and compare 3.3 and 3.4). There are some pattern variants, where complex attribute not in concord is constructed with several prepositions (see Table 2, line 3.5). The occurrence of the attribute not in concord (see Table 2, line 3.6) in infinitive form and nominal cases (dative, instrumental, locative) (see Table 2, line 3.7) has been detected only in several patterns.

3.2.1.1.2. Objective nominal patterns

Although objective nominal patterns are a rather rare nominal pattern type, they are more frequent in verbal combinations (see 3.2.1.2). Typical objective nominal patterns include prepositional combinations of cases (genitive, accusative, instrumental) with prepositions *apie, su, dėl, tarp* 'about, with, due to, between' (see Table 3, lines 4 and 4.1). We maintain that prepositional constructions function as object in the nominal pattern when used with nouns denoting a person or an animal, for example, *nuomonė apie* ***moteris*** 'opinion about **women**', *bendravimas su* ***gyvūnais*** 'communication with **animals**'.

4. [Mod] [prep Obj_] [X] [apie Obj_acc] (2): ***informacija*** *apie šalį* '**information** about the country'; [Mod] [apie X_acc] (1): *nuomonė apie* ***moteris*** 'opinion about **women**' [Mod] [su X_pl.ins] (1): *bendravimas su* ***gyvūnais*** 'communication with **animals**'
4.1. [X] [su Obj_ins] [apie Obj_acc] (1): ***pokalbiui*** *su darbdaviu* ('for **conversation** with employer'); ***pokalbis*** *apie kiną* '**conversation** about film'; [X] [su Obj_ins] [dėl Adv_gen] (1): ***kova*** *su burlentininkais* '**fight** with windsurfers'
5. [X_sg] [Obj_dat] (1): ***paminklą*** *kunigaikščiui Gediminui* '**monument** for the duke Gediminas'

Table 3. Objective nominal patterns (the numbering is continued)

3.2.1.1.3. Nominal patterns with partitive genitive of quantity (AtrQ)

This type of nominal pattern expresses a certain quantity (see Table 4). The main quantity indicating patterns are [Mod] [X_gen] and [Adv] [X_pl.gen], where the noun in genitive position has the [AtrQ] function – partitive genitive of quantity (see Table 4, line 6). The node-word can be present in both positions of the pattern.

6. [Mod] [X_gen] (1): *puokštę **gėlių*** ‘a bouquet of **flowers**’; **[Adv] [X_pl.gen]** (1): *300 **metrų*** ‘300 **metres**’
6.1. [daugelis] [X_pl.gen] (1): *daugelis (šio ženklo) **atstovų*** ‘many **representatives** of this sign’; [gabalėlis] [X_sg.gen] (1): *gabalėlis **pyrago*** ‘a piece of **cake**’; [kilogramas] [X_pl.gen] (1): *kilogramas **obuolių*** ‘a kilogram of **apples**’; [Mod] [X_sg.gen] (1): *5 valandą **ryto*** ‘at 5 o‘clock in the **morning**’; [pusė stiklinės]|[x šaukštų]|[truputį] [X_sg.gen] (1): *truputį **sviesto*** ‘a little bit of **butter**’
7. [X] [AtrQ] (1): ***dalis** gatvių* ‘**part** of streets’
8. [x X] [Mod] (1): ***54 minutės** (šiaurės) platumos* ‘**54 minutes** of (north) latitude’
8.1. [x X_nom|x X_gen] [Mod_sg.gen] (1): *antra **valanda** nakties* ‘two **o‘clock** at night’; [x X|x X_pl.gen] [Mod_sg.gen] (1): *17 **laipsnių** šilumos* ‘17 **degrees**’

Table 4. Nominal patterns with partitive genitive of quantity (the numbering is continued)

The quantity modifier is expressed by more specific nouns denoting the quantity (e. g., *kilogramas* ‘kilogram’), more abstract quantity nouns (e.g., *dalis* ‘part’, *daugelis* ‘majority’, *gabalėlis* ‘piece’, *truputis* ‘little’), or numerals (e.g., ***300** metrų* ‘**300** metres’) (see Table 4, line 6.1). Some nouns realize the function of quantity modifier in constructions with numerals (e.g., ***54 minutės** (šiaurės) platumos* ‘**54 minutes** of (north) latitude’; *antra **valanda** nakties* ‘two **o‘clock** at night’).

3.2.1.1.4. Nominal patterns with adverbials

The patterns with adverbials form a small group of nominal patterns and are more frequent with verbal adverbial patterns (see 3.2.1.2.3).

9. [Mod] [Adv] [Mod] [už X_sg.acc] (1): *žuvusiųjų už **laisvę*** ‘deceased for **freedom’**; [X_pl] [Adv_inf]|[noun_dat+inf] (1): ***sąlygos** treniruotis|**sąlygos** grybams augti* ‘**conditions** to exercise’|‘**conditions** for mushrooms to grow’; [X] [Adv_dat] (1): ***įrankiai** desertui* ‘**cutlery** for dessert’; [X] [ant Adv_gen] (1): ***žiedas** ant piršto* ‘**ring** on the finger’

Table 5. Nominal patterns with adverbials (the numbering is continued)

As is seen in Table 5, the two adverbial patterns denote purpose and aim: [Mod] [už X_sg.acc]; [X_pl] [Adv_inf]|[noun_dat+inf]. They are formed either with a preposition or an infinitive construction; the purpose is expressed by the dative [X] [Adv_dat] and the place is indicated by a preposition [X] [ant Adv_gen].

3.2.1.1.5. Other nominal patterns

Some nouns cannot be assigned to any of the categories above and form individual patterns. For example, some of the analysed nouns are constituents of parentheses:

[pron_gen X_sg.ins]: *mano* ***nuomone*** 'in my **opinion**'; [mano X_sg.ins]: *mano* ***galva*** 'to my **mind**'. In other cases, a noun can be a part of: (a) complex subject [AtrA] [X+noun_pl.gen]: *didžioji* ***dauguma*** *gyventojų* 'the **majority** of inhabitants'; (b) time adverbial [AtrA+[AtrN]+X]: *pirmojoje metų* ***pusėje*** 'in the first **half** of the year'; (c) complex attribute [Num_card+pagal X_sg.acc] [Mod]: *trečias pagal* ***dydį*** *miestas* 'the third town in **size**'; [Mod]|[Pred] [aplink X_acc]: *kelionę aplink* ***pasaulį***|*plaukti aplink* ***pasaulį*** 'journey around **the world**|swim around **the world**'.

3.2.1.2. Verbal patterns

These patterns include verbs in the predicate function, whereas the analysed nouns can be used in subject and compound nominal predicates (see 3.2.1.2.1), object (see 3.2.1.2.2) or adverbial (see 3.2.1.2.3) functions. Table 6 presents the frequency of the verbal patterns according to their syntactic function (the number of patterns is not given in Table 6). The total frequency of verbal patterns (1239) resembles that of the nominal patterns (1156).

Syntactic noun function	**Frequency**	**Percentage**
Object	383	30.91
Subject	460	37.13
Compound nominal predicate	86	6.94
Adverbials	310	25.02
Place	164	13.24
Time	62	5.00
Manner	42	3.39
Aim	19	1.53
Quantity	14	1.13
Cause	8	0.65
Condition	1	0.08
In sum	**1239**	**100.00**

Table 6. Syntactic functions of verbal patterns with nouns

3.2.1.2.1. Subjective verbal patterns

As is seen in Table 6, the patterns in which the analysed nouns function as a subject (460) are especially frequent. A detailed information of the main patterns and their variants is presented in Table 7.

The analysed noun most frequently is in the subject position. We distinguish several main patterns here due to the differences in predicate structure and word order: [Pred] [X]; [X] [Pred+V]; [X] [Pred+N]. If the noun (subject) is used only in singular or plural form, this morphological information in added to the main pattern, and, accordingly, the number of variants increases. The variants of

1. [X] [Pred+N] (96): ***žmonės*** *buvo linksmi* '**people** were happy'; **[Pred] [X]** (9): *puošia* ***ženklas*** '**sign** decorates'; **[X] [Pred+V]** (2): ***tikslas*** *yra suvaidinti* 'the **aim** is to act'
1.1. [X_sg] [Pred] (9): ***vanduo*** *bėga* '**water** is running'; [X]|[X_pl.gen] [Pred] (1): ***grybai*** *turi*|*yra* ***grybų*** '**mushrooms** have'|'there are **mushrooms**'
1.2. [Pred]|[Pred+V] [X] (1): ***autobusai*** *vyksta*|***autobusas*** *pradės vežti* '**buses** go'|'**buses** will start to go'; [X]|[X_dat] [Pred+V]|[Pred] (1): ***Dvyniai*** *pamilsta*|***Dvyniai*** *gali susidurti*|***Dvyniams*** *reikia pasiruošti* '**Gemini** fall in love'|'**Gemini** can meet with'|'**Gemini** need to prepare'
1.3. [X] [Pred+N] (96): ***žmonės*** *buvo linksmi* '**people** were happy'; [X_pl] [Pred+N] (6): ***valgiai*** *yra svarbūs* '**meals** are important'; [X] [Pred+V+N] (6): ***dovanos*** *turėtų būti simbolinės* '**presents** should be symbolic'; [X] [Pred+N]|[Pred+V+N] (3): ***gyvenimas*** *bus aktyvus*|***gyvenimas*** *gali būti įvairus* '**life** will be active'|'**life** can be various'; [X_sg] [Pred+N] (3): *(mirties)* ***priežastis*** *yra skendimai* '(death) **cause** is drowning'
2. [Sub] [Pred+X] (40): *esu (jautrus)* ***žmogus*** 'I am a (sensitive) **person**'
2.1. [Sub] [Pred+X_sg] (4): *inteligentas jau buvo aukštesnė* ***klasė*** 'an intelligent was already a higher **class**'; [Sub] [Pred+X_nom] (2): *jis yra* ***studentas*** 'he is a **student**'; [Sub] [Pred+X_sg.ins] (2): *televizija tapo...* ***būdu*** 'television became... **a way'**; [Sub] [Pred+V+X] (2): *mes galime būti* ***draugai*** 'we can be **friends**'; [Sub]|[Sub_inf] [Pred+X] (1): *meilė yra (tikras)* ***dalykas***|*gauti bilietus... yra...* ***dalykas*** 'love is a (real) **thing**'|'to get tickets... is... a(n)… **thing**'*;* [Sub] [Pred+N+X_nom] (1): *buvo įsteigtas (kultūros)* ***centras*** 'a (cultural) **centre** was established'
2.2. [Sub] [Pred+AtrA X_sg.gen] (4): *esu linksmo* ***būdo*** 'I have a joyful **character**'; [Sub] [Pred+AtrA X_sg.gen sudėjimo] (1): *esu normalaus* ***kūno*** *sudėjimo* 'I have a normal **physique**'; [Sub] [Pred+AtrA|AtrN X_sg.gen] (1): *būna garbingo* ***amžiaus*** 'are of honourable **age**'; [Sub] [Pred+[AtrN]+X] (1): *tai yra pačių futbolininkų* ***reikalas*** 'is a **matter** of the football players themselves'
2.3. [Sub] [Pred+x X_pl.gen] (1): *pokalbiai būna trijų* ***tipų*** 'conversations are of three **types**'; [Sub] [Pred+x X_sg.gen noun_sg.gen] (1): *autobusas yra aštuonių* ***metrų*** *pločio* 'the bus is eight **metre** wide'

Table 7. Verbal patterns with noun as a subject and with compound nominal predicate

the main pattern are also related with the predicate composition: the pattern can include a simple predicate: [X_sg] [Pred]: *vanduo bėga* '**water** is running' (see Table 7, line 1.1) or a compound verbal predicate (e.g., in a combination with a simple predicate): [Pred]\|[Pred+V] [X]: ***autobusai*** *vyksta*\|*autobusas pradės vežti* '**buses** go'\|'**buses** will start to go'. Information on number and case can also be included as in the following impersonal pattern where subject is in dative [X_dat] [Pred+V]: ***Dvyniams*** *reikia pasiruošti* '**Gemini** need to prepare' (see Table 7, line 1.2). Other patterns have a compound nominal predicate (Pred+N) and (or) a compound predicate with verb and noun (Pred+V+N) (see Table 7, line 1.3).

Quite often, a noun is a part of the compound nominal predicate. For example, consider the pattern [Sub] [Pred+X] (see Table 7, line 2). The analysed word stands as the last element of the compound nominal predicate. The number

of pattern variants increases due to the indicated information on number and (or) case. The usage has shown that complex attributes (including the analysed noun) can be a constituent part of the compound nominal predicate (see Table 7, lines 2.2 and 2.3): *esu linksmo* ***būdo*** 'I have a joyful **character**'; *būna garbingo* ***amžiaus*** 'are of honourable **age**' (see 3.2.1.1.1 for more information on complex attributes).

3.2.1.2.2. Objective verbal patterns

Many verbal patterns with the noun in the object function occur only once, therefore, we will discuss this type in general, without providing all use cases. The analysed noun in object position may be used in different cases: accusative: [Pred] [X_acc] (57): *rasiu* ***vietą*** 'will find a **place**'; genitive: [Pred]|[Pred+V] [X_pl.gen] (6): *ėmėsi* ***priemonių***|*galite imtis* ***priemonių*** 'took **action**|you may take **action**'; instrumental: [Pred] [X_ins] (12): *mostelėjo* ***ranka*** 'waved **a hand**'.

Other objective patterns are formed with prepositions: *su*, *nuo*, *už*, *prie*, *pas*, *apie*, *į*, *iš*, *be*, *dėl* 'with, from, behind, near, at, about, to, from, without, due'. To give an example: [Pred] [su X_ins] (6): *susikalbėti* ***su dievais*** 'to talk **to gods**'; [Pred] [nuo X_gen] (1): ***nuo vėjo*** *lingavo* 'swayed **in the wind**'; [Pred] [už X_pl.acc] (1): *perka* ***už pinigus*** 'buy **for money'**; [Pred] [už X_pl.gen] (1): *neužsikabina* ***už augalų*** 'do not hang **on plants**'; [Pred]|[Pred+V] [X_acc]|[X_dat]|[su X_ins]|[už X_gen] (1): *norite ištekėti* ***už vyriškio*** 'want to marry **to a man**'. Pattern variants occur due to different predicate structure, case and preposition agreement, the added elements, and included additional information on the grammatical category of number.

3.2.1.2.3. Verbal patterns with adverbials

The patterns with nouns as adverbials make up approximately 25% (310) of all verbal pattern occurrences. Adverbials can refer to place, time, manner, aim, quantity, cause, or condition (see Table 6). Due to the large amount of data (unique patterns) and the fact that a number of patterns occur only once, we will present only selected examples. Some verbal patterns given as examples in this section have been shortened for clarity.

Nouns in **place** indicating adverbials are expressed in locative [Pred] [X_loc] (16): *esu (miesto)* ***centre*** 'I am in the town **centre**'; ***gyvenime*** *būna* 'happens in **life**'; instrumental case: [Pred] [X_ins] (12): ***veidu*** *pasruvo* 'flooded the **face**'; *ėjo* ***keliu*** 'went along the **road**'; and prepositions: *į, ant, po, nuo, per, iki, palei, už, tarp, pas, vidury, virš, iš, arčiau, arti, link, netoli, pro* 'to, on, under, from, through, until, near, behind, between, at, in the middle, above, from, nearer, near, towards, close, through'. Examples of patterns with the most common prepositions in this category are as follows: [Pred] [į X_acc] (23): *pasinerti* ***į*** *(fantastikos)* ***pasaulį*** 'dive **into** the (fictional) **world**'; [Pred] [ant X_gen] (1): ***ant lapo*** *užsirašyti* 'to write down **on the paper**'; [Pred]|[Pred+V] [po X_acc] (1): *vaikščiojo* ***po senamiestį*** 'walked **along the old town**'; [Pred] [po X_ins]|[ant X_gen] (1): *padėdavo* ***po staltiese*** 'put **under the cloth**'.

Time is expressed by a noun in accusative: [Pred]||[Pred+V] [X_acc] (29): *prinoksta (rugpjūčio)* ***mėnesį*** ‘ripens in **the month** of (August)’; dative: ***vakarienei*** *davė* ‘gave for **supper**’; instrumental case: [Pred]||[Pred+V] [X_pl.ins] (1): ***vakarais*** *mėgstu skaityti* ‘**in the evenings**, I like to read’. Examples of patterns with the most common prepositions are as follows, to mention just a few: [Pred]||[Pred+V] [per X_acc] (1): ***per*** *(žydėjimo)* ***laikotarpį*** *išskiria* ‘**during the period** of blossoming release’; [Pred]||[Pred+V] [po X_gen] (1): ***po ligos*** *praradęs* ‘lost **after illness**’; [Pred] [iki X_gen] (1): *dirbau* ***iki nakties*** ‘worked **till night**’.

Manner is expressed by a noun in instrumental case, e.g.: [Pred] [į X_ acc]||[X_pl.ins] (1): *(lygiomis)* ***dalimis*** *sumaišykite* ‘mix in equal **parts**’, as well as noun+preposition constructions: *be, pagal, per, po , iš, su, už* ‘without, according to, during, under, from, with, behind’): [Pred] [be X_pl.gen] (1): ***be problemų*** *susidorosite* ‘you will deal with it **without problems**’: [Pred] [pagal X_pl.acc] (2): ***pagal galimybes*** *atsižvelkite* ‘take into consideration **according to the possibilities**’.

Aim is signalled by a noun in genitive [Pred]||[Pred+V] [X_sg.dat] (1): *kviečia* ***vakarienės*** ‘invites for **supper**’; dative: [Pred] [X_dat] (10): *ruoštis* ***šventei*** ‘prepare for **celebration**’; instrumental case [Pred] [AtrA X_pl.ins] (1): *darbo* ***reikalais*** *važiuojate* ‘travel for work **purposes**’. Moreover, aim can be expressed by a construction with a noun in dative and infinitive: [Pred] [X_pl.dat+inf] (1): *naudojami* ***produktams*** *dažyti* ‘used for **product** painting’.

Quantity is shown by a noun in genitive [Pred] [x X_pl.gen] (1): *turi 206 000* ***gyventojų*** ‘has 206 000 **inhabitants**’; accusative: [Pred] [X_sg.acc]||[visą X_sg.acc]||[x X_pl.acc]||[porą X_pl.gen] (1): *(visą)* ***savaitę*** *buvau* ‘stayed for the (whole) **week**’; *sirgau dvi* ***savaites*** ‘was ill for two **weeks**’; *šaunu būtų porą* ***savaičių*** *per savo atostogas pagyventi pas kokius malonius žmones* ‘it would be nice to stay with some pleasant people for a couple of **weeks** during holidays’. Usually, in verbal patterns with the adverbial of quantity, an exact amount (marked by x) is indicated, besides, these patterns often include fixed lexis of quantity (*visą, porą*) ‘all, a couple’.

Cause is expressed by a construction with noun in the instrumental case [Pred] [AtrN+X_sg.ins] (1): *sveikinu gimtadienio* ***proga*** ‘I congratulate you on **the occasion** of your birthday’ or prepositions: *dėl, be, per* ‘because, without, through’: [Pred+V] [dėl X_gen] (1): ***dėl dokumento*** *turite kreiptis* ‘**for document** apply’. **Condition** is expressed by the instrumental case of the analysed noun. Interestingly, this noun is a part of a complex structure: [AtrA+X_ins] [Pred]||[Pred+V] (1): *tokiu* ***atveju*** *reikės pateikti* ‘**in this case**, you will need to provide’; *geriausiu* ***atveju*** *baigsis* ‘**at best**, it will end’, where a part [AtrA+X_ins] of the pattern is realised by several collocations *tokiu atveju* ‘on the occasion of’, *geriausiu atveju* ‘at best’. The mentioned pattern with the adverbial of cause [Pred] [AtrN+X_sg.ins] also illustrates complex adverbial structure in which several nouns (referring to types of occasions) for the attribute not in concord can be selected: *gimtadienio* ***proga*** ‘on **the occasion** of your birthday’, *sukaktuvių* ***proga*** ‘on the **occasion** of the anniversary’. Complex structures including fixed lexis are

described in other adverbials, to mention just a few: [Pred]|[Pred+V] [X_sg.gen metu] (1): ***vasaros*** *metu paskatino* ‘in **summer**-time encouraged to’; [Pred] [visas_ acc X_acc] (1): *visą* ***gyvenimą*** *dirbau* ‘worked all my life’.

3.2.1.3. Patterns with clauses, parentheses and forms of address

These patterns are only typical for selected nouns. They stand as constituent elements of composite sentences with a conjunction signalling the beginning of subordinate clause. In some cases, the word analysed is in the subject position and has a predicate (see Table 8, line 1.1). In other cases, the word is a part of the predicate (see Table 8, line 1.2). Usually, the nouns are abstract (*klausimas, dalykas, reikalas* ‘question, thing, issue’) and behave like shell nouns in colligation (Drūlienė 2014).

1. Patterns as constituents of clauses, parentheses and forms of address
1.1. [Pred] [X], ar (1): *iškilo* ***klausimas****, ar* ‘a **question** has risen whether’; [Pred] [X], kad|jog (1): *neatmetama* ***galimybė****, kad* ‘a **possibility** that… is not rejected’; [Pred] [X_pl.nom], kad (1): *sklinda* ***kalbos****, kad* ‘**rumour** goes that’; [Pred+[tik]+X_sg.gen klausimas], kada (1): *yra tik* ***laiko*** *klausimas, kada* ‘it is only a question of **time** when’; [X_sg.nom] [Pred+tas], kad (1): ***reikalas*** *tas, kad* ‘the **thing** is that’
1.2. [Pred+AtrA X_sg.gen], kad (1): *yra tokio* ***dydžio****, kad* ‘is of the **size** that’; [Pred+X], dėl kurios (1): *yra ir kita* ***priežastis****, dėl kurios* ‘there is another **reason** because of which’; [Pred+X], kad (1): ***tiesa****, kad* ‘the **truth** is that’; [Pred+X_pl.gen], kuris (1): *yra* ***dalykų****, kuriuos* ‘there are **things** which’; [su X_ins], kad (1): *su* ***sąlyga****, kad* ‘on **condition** that’
1.3. [X_sg.acc sakant], [Pred] (1): ***tiesą*** *sakant,* ‘to tell the **truth**’; [X_voc]|[X_ voc+noun_prop], [Pred] (1): ***ponia|ponia*** *Laima,* ‘**Madam|Madam** Laima’; jei [Pred+X], [Pred_imp] (1): *jei yra* ***galimybė****, išmokite* ‘if there is a **possibility**, learn’

Table 8. Patterns with clauses, parentheses, and forms of address

In some cases, patterns function as parentheses (***tiesą*** *sakant* ‘to tell **the truth**’) or forms of address: ***ponia*** *Laima* ‘**Madam** Laima’ (see Table 8, line 1.3). Fixed lexical and grammatical structure is common to these patterns: [su X_ins], kad: *su* ***sąlyga****, kad* ‘on **condition** that’; jei [Pred+X], [Pred_imp]: *jei yra* ***galimybė****, išmokite* ‘if there is a **possibility**, learn’.

3.2.2. Grammatical patterns with adjectives

The patterns, in which adjectives function as attributes, are especially common (see Table 9). Less frequent patterns include adjectival constructions with adverbs (see Table 10) or adjective as a part of compound nominal predicate (see Table 11).

In pattern [X] [Mod], the noun stands as a modified word [Mod] and adjective has the role of the attribute in concord. Due to additional grammatical features

(definiteness, degree, gender (see Table 9, line 1.1)), the pattern [X] [Mod] has many variants. For example, the pronominal form (definiteness) marked in this pattern signifies a quality adjective, whereas a combination of words will most probably be a term: *raudonieji serbentai* 'red currant', *smulkieji gamintojai* small producers'.

1. [X] [Mod] (110): ***geras*** *darbas* '**good** job'*;* ***nauja*** *galimybė* '**new** opportunity'
1.1. [X_def] [Mod] (6): ***senoji*** *karta* 'the **old** generation'; ***raudonieji*** *serbentai* '**red** currant'; [X]|[X_def] [Mod] (3): ***mažas*** *miestelis*|***mažoji*** *salė* '**small** town|**small** hall'; ***smulkus*** *ūkininkas*|***smulkieji*** *gamintojai* '**small** farmer|**small** producers'*;* [X]|[X_sup] [Mod] (2): ***populiari*** *vieta;* ***artimiausiu*** *metu* '**popular** place; in the **nearest** time'*;* [X_f] [Mod] (1): ***liekna*** *mergina* ('**slim** girl'); [X_com]|[X_sup]|[X_pos] [Mod] (1): ***aukštos*** *kokybės* 'of **high** quality'; ***aukščiausios*** *kokybės* 'of **the highest** quality'; ***aukštesnė*** *klasė* '**the higher** class'
2. [X_gen noun_gen] [Mod]: ***ramaus*** *būdo moteris* 'woman of a **calm** character'
2.1. [X_sg.gen noun_sg.gen] [Mod] (2): ***laisvo*** *stiliaus imtynininkas* '**free** style wrestler'; [noun_gen+X] [Mod] (2): *pasitikėjimo* ***vertas*** *gėrimas* 'a trust**worthy** drink'; [X_pl.gen akių] [Mod] (1): ***mėlynų*** *akių vyriškis* 'a **blue** eyed man'; [X_sg.gen naudojimo] [Mod] (1): ***asmeninio*** *naudojimo lėkštes* 'plates of **personal** use'; [X+noun_gen]|[noun_gen+X] [Mod] (1): ***pilname*** *dūmų kambaryje* 'in a room **full** of smoke'

Table 9. Nominal attributive patterns

The second category includes adjectives as attributes not in concord. Typically, an attribute not in concord consists of two parts: an adjective and a noun in genitive cases (see the respective noun patterns described in 3.2.1.1.1). Grammatical features of this structure practically do not vary: plural vs. singular; adjective governing the case (e.g., *vertas* 'worth'); word order can vary (e.g., *pilnas* 'full'). On the other hand, depending on the described adjective, the pattern can include fixed lexis – in this way, grammatical patterns reflect collocations: *atviro būdo* 'of open character'*, asmeninio naudojimo* 'of personal use'*, nuostabaus grožio* 'of wonderful beauty'. In some once occurring patterns, adjectives are part of a collocation and mark (1) time [nuo X|X_sup laikų] [Pred]: *yra nuo* ***seniausių*** *laikų* 'is from **the oldest** times'; [X_sg.ins metu] [Pred]: ***paskutiniu*** *metu padidėjo* ('**lately** has increased') (2) place [X] [visame pasaulyje]: ***garsus*** *visame pasaulyje* '**famous** in the whole world'; (3) attribute [vienas+X_sup.pl.gen] [Mod]: *vieni* ***garsiausių*** *(Brazilijos) atlikėjų* 'some of **the most famous** (Brazilian) performers'; [X+[Mod]+Adv_loc]: ***vieninteliame*** *pasaulyje muziejuje* 'in **the only** museum of the world'.

In rather frequent adjectival patterns [Adv] [X], the analysed adjective is characterised by an adverb: *visiškai* ***naujas*** 'completely **new**'; *labai* ***mažas*** 'very **little**'. Here, variation occurs due to definiteness and degree of the pattern constituents: [Adv]\|[Part] [X_def]\|[X_comp]\|[X_sup]: *gerokai* ***didesnė***\|*dar* ***didesnis*** 'much **bigger**'\|'even **bigger**' (see Table 10). As syntactic adjective functions may vary in these models, they were not semantically classified.

3. [Adv] [X] (51): *visiškai* ***naujas*** ‘completely **new**’*; labai* ***svarbus*** ‘very **important**’
3.1. [Adv] [Part] [X] (2): *tik labai* ***geras*** ‘only very **good**’; *dar labai* ***mažas*** ‘still very **little**’; [Adv]||[Part] [X] (2): *labai* ***didelis*** ‘very **big**’; *gana* ***didelis***|*tik* ***didelis*** ‘rather **big**’|‘only **big**’*;* [Adv]||[Part] [X_def]||[X_comp]||[X_sup] (2): *gerokai* ***didesnė***|*dar* ***didesnis*** ‘much **bigger**’|‘even **bigger**’

Table 10. Adjectival patterns (the numbering is continued)

In comparison with nominal attributive patterns (Table 9) and adjectival patterns (Table 10), verbal patterns have a higher variation (see Table 11).

The most typical is the verbal use of adjectives when they are a part of compound nominal predicate. In this type, the main pattern [Sub] [Pred+X] is the most frequent (see Table 11, line 7). It is important to note that grammatical patterns depicted in Table 11 include more elements than in the minimal patterns of the adjective: as an adjective is a part of the compound nominal predicate, such patterns include elements characterizing the collocability of the verb.

As is seen in Table 11, in pattern types 4, 5, and 6, an adjective is of neuter gender and functions as a part of a compound nominal predicate (the verb *būti* ‘be’ is often implied): *visada (yra)* ***malonu*** ‘(is) always **pleasant**’*, buvo* ***gausu*** *uogų* ‘there were **plenty** of berries’, *(yra)* ***smagu****, kad...* ‘(is) **great** that…’. The constructions are also impersonal. Patterns with subject in dative are also impersonal (see Table 11, line 10). Adjectives fall into two semantic categories: physical (e.g., *karšta, šalta, pilna, gausu* ‘hot, cold, full, plenty’) and evaluative (e.g., *malonu, ramu, svarbu, lengva* ‘pleasant, calm, important, easy’).

Type 4 patterns with adverbs [Adv] [Pred+X_n] may vary because of word order differences or adverbial components (e.g., [Adv_acc]||[Adv] [Adv_loc] [Pred+X_n]). When an adjective is a constituent part of a compound nominal predicate which governs the case, the object is included in the pattern (see Table 11, line 5). In this type, the variants of the main model differ in their predicate structure, which can have two ([Pred+X_n] [Obj]) or three elements as, for example, auxiliary verb+neuter gender adjective+infinitive [Pred+X_n+V] [Obj_acc]: *(yra)* ***lengva*** *palaikyti gerą formą* ‘**easy** to keep shape’. When the adjective in the pattern expresses fulness/plentiness, it is in genitive case: ***pilna*** *naftos* ‘**plenty** of oil’*,* ***gausu*** *uogų* ‘**plenty** of berries’. As the number of syntactic partners in verbal patterns depends on the predicate, the variants of this pattern may include adverbials: [Adv] [Pred+X_n+V] [Obj_acc] (1): *tada* ***būtina*** *išsaugoti pagarbą* ‘then it is **necessary** to keep respect’. Type 6 patterns may include the beginning of the relative clause: [Pred+X_n], kad (***malonu****, kad...* ‘**it is pleasant** that’).

In types 7, 8, and 9 (see Table 11), an adjective, as a part of a compound nominal predicate, is used with a subject [Sub] [Pred+X]*: mokytojai buvo* ***geri*** ‘teachers were **good**’; *namas yra* ***didelis*** ‘house is **big**’. These patterns are the most frequent as they represent a personal syntactic structure with a subject. One variant of the 7th type is a three-word complex predicate, which can be used instead of the two-word predicate: [Sub] [Pred+X]||[Pred+V+X]: *pasiūlymai būna*

4. [Adv] [Pred+X_n] (6): *visada **malonu*** ‘always a **pleasure**’; *nėra labai **svarbu*** ‘it is not very **important**’
4.1. [Pred+X_n] [Adv] (1): *visada būdavo **karšta*** ‘would always be **hot**’; [Pred+[Adv]+X_n] (1): *buvo tikrai **smagu*** ‘it was real **fun**’; [Adv_acc]|[Adv] [Adv_loc] [Pred+X_n] (1): *žiemą būna labai **šalta*** ‘it is **cold** in winter’; *dabar Estijoje labai **šalta*** ‘it is very **cold** in Estonia’; *šiandien **šalta*** ‘it is **cold** today’
5. [Pred+X_n] [Obj_pl.gen] (1): *buvo **gausu** uogų* ‘there were **plenty** of berries’
5.1. [Adv]|[Adv_loc] [Pred+X_n] [Obj_gen] (1): *jūroje yra **pilna** naftos* ‘there is **plenty** of oil in the sea’; [Pred+X_n+V] [Obj_acc] (1): ***lengva** palaikyti gerą formą* ‘it is **easy** to keep a shape’; [Pred+X_n+V] [Obj_gen]|[Obj_acc] (1): ***svarbu** nepamiršti saugumo taisyklių* ‘it is **important** not to forget security rules’; [Adv] [Pred+X_n+V] [Obj_acc] (1): *tada **būtina** išsaugoti pagarbą* ‘then it is **necessary** to keep respect’
6. [Pred+X_n], kad (3): ***malonu**, kad* ‘it is **pleasant** that’; ***keista**, kad* ‘it is **strange** that’
6.1. [Pred+X_n], kad|jog (2): ***panašu**, kad|jog* ‘it is **likely** that’; [Pred+X_n+V], kad (1): ***svarbu** žinoti, kad* ‘it is **important** to know that’
7. [Sub] [Pred+X] (43): *mokytojai buvo **geri*** ‘the teachers were **good**’; *namas yra **didelis*** ‘the house is **big**’
7.1. [Sub] [Pred+X_f] (1): *esu **liekna*** ‘I am **slim**’; [Sub] [Pred+X]|[Pred+V+X] (5): *pasiūlymai būna **paprasti**|mergina turi būti **paprasta*** ‘suggestions can be **simple**’|‘a girl has to be **simple**’; [Sub] [Pred+X_def|X_comp|X_sup] (2): *greitis yra **didesnis*** ‘speed is **higher**’; [Sub] [Pred+verb_ptcp.pass.prs+X] (1): *erškėtuogės valgomos **šviežios*** ‘rosehips are eaten **fresh**’
8. [Sub] [Pred+X] [Obj_dat] (1): *ežerai **tinkami** ir karosams* ‘lakes also **suitable** for carps’
8.1. [Sub] [Pred+X] [Obj_gen] (2): *lentynos **pilnos** knygų* ‘shelves **full** of books’; *šuo **vertas** geresnio maisto* ‘a dog is **worth** better food’; [Sub] [Pred+X] [Obj_ins]|[dėl Obj_gen] (1): *restoranas buvo **garsus** dėl pačios skaniausios juodosios arbatos* ‘the restaurant was **famous** for the tastiest black tea’
9. [Sub] [Adv] [Pred+X] (2): *aš jau **senas*** ‘I am already **old**’; *vaikai jau **suaugę*** ‘the children are **grown up**’
9.1. [Sub] [Pred+[Adv]+X] (2): *miltai tikrai **švieži*** ‘flour is really **fresh**’; *vaisiai būna gana **smulkūs*** ‘fruits tend to be rather **small**’; *pertraukos labai **trumpos*** ‘breaks are very **short**’; [Sub] [Pred+[Adv]+X_comp] (1): *mes daug **mažesni*** ‘we are much **smaller**’; [Sub] [Pred+[Adv]+X]|[Pred+X] [su Obj_ins] (1): *susiformavimas buvo glaudžiai **susijęs** su judėjimu* ‘formation was closely **related** to movement’; [Sub] [Adv_ins]|[Adv]|[Adv_acc] [Adv_loc] [Pred+X] (1): *šiuo metu vis dar **populiarios** asmenukės* ‘at present, selfies are still **popular**’
10. [Sub_dat] [Pred+X_n+V_inf] (2): *gyventojams yra **malonu** gauti* ‘inhabitants are very **pleased** to receive’
10.1. [Sub_dat] [Pred+X_n], kad (1): *jam buvo labai **įdomu**, kad* ‘it was very **interesting** to him that’

Table 11. Verbal patterns with adjective as part of the compound nominal predicate (the numbering is continued)

paprasti|*mergina turi būti* ***paprasta*** ‘suggestions are **simple**|girl has to be **simple**’. In addition, type 7 includes information on degree: [Sub] [Pred+X_def|X_comp|X_sup]: *greitis yra* ***didesnis****, gyventojų skaičius yra* ***didžiausias*** ‘speed is **higher**, the number of inhabitants is **the biggest**’. In some patterns, a predicate is used not only with subjects (see Table 11, line 8), but also with adverbials (see Table 11, type 9). As in type 5, in the pattern of type 8 the object is necessary because of the analysed adjective: the genitive case is governed specifically by adjectives such as *tinkamas, pilnas, vertas, gausus* ‘appropriate, full, worthy, plenty’.

11. **[Pred+V] [noun_dat X_ins noun_ins]** (1): *Jums* ***patogiu*** *laiku galite peržiūrėti* ‘You can watch at a **convenient** time’; [Pred+X_comp] [nei x metai]|[už Obj_acc] (1): *esate* ***jaunesnis*** *nei 26 metai*|*būtų* ***jaunesnė*** *už mane* ‘you are 26 years **younger**’|‘she would be **younger** than me’; [Pred_imp+X] [Pred_imp] (1): *būk* ***geras****, paskubėk* ‘be **kind**, hurry up’; [Pred_imp+X], [Pred_imp] (1): *būkite* ***maloni****, po to užeikite* ‘be so **kind** to come in after that’; [X_n], [Pred] (1): ***aišku****, tu juokauji* ‘it is **clear** that you are joking’; [X_voc+noun_prop], [Pred] (1): ***Miela*** *Laima, rašau tau* ‘**Dear** Laima, I am writing to you’

Table 12. Once occurring patterns with adjectives (the numbering is continued)

Finally, in Table 12, we provide once occurring patterns reflecting the usage of a specific adjective. The examples include sayings, cliches of a specific genre, and various constructions (with fixed and variable elements).

4. Conclusions

The presented database has appr. 3700 one-word and multi-word entries. More than 700 entries describe one-word lexical units and their patterns: 207 verbs, 386 nouns, 87 adjectives and 41 adverbs. The majority of the patterns describe nouns (673) and verbs (592) followed by adjectives (99) and adverbs (57). Pattern types with verbs are numerous, but of lower frequency. As a contrast, the frequency of patterns with nouns is higher but their variation is lower (however, this ratio is different in nominal and verbal patterns). This result is influenced by different procedures used to describe patterns with nouns, adjectives, and adverbs and patterns with verbs. As patterns are rich with grammatical information, there are several similar patterns with one main (usually the most frequent) pattern and its variants. Pattern variants show that the grammatical characteristics describing the usage of a specific word are rather individual.

The analysis has shown the typical usage of the two parts of speech – nouns and adjectives. Although the grammatical information presents only one pattern description level, some lexical differences of the same part of speech can also be seen. For example, the nouns *truputis* ‘a little’, *daugelis* ‘many’ are used as adverbs

due to their semantics; the adjectives *pilnas* 'full', *vertas*, 'worthy' differently from other adjectives are predominantly used in verbal patterns.

The representation of the usage pattern in the lexical database consists of three levels. Although these levels are interconnected with each other, it is still complicated to analyse all of them at once. In this article, we discussed the grammatical level of the patterns. It has been shown that there are a lot of similar grammatical patterns in the noun and adjective groups. Nevertheless, most of these grammatically similar patterns are disambiguated on the semantic and lexical levels where semantic types and/or collocates are provided.

Although the focus was on the grammatical patterns, the analysis does not allow to show that distinct word meanings may have different usage patterns. For example, the entry of the noun *tiesa* 'truth' describes the three meanings of the word. 7 patterns describe the 1st meaning, whereas the 2nd and the 3rd meanings have only one pattern each. When used in its 1st and 2nd meanings, the noun has a singular form, whereas in the 3rd meaning, the noun is used in plural: *amžinos tiesos* 'eternal truths'. In further research, it would be interesting to analyse the 'local grammar' (i.e., usage patterns) of different nouns, adjectives, verbs, and adverbs and see to what extent different word meanings are related to different patterning.

Acknowledgement

We would like to thank our colleague from the Centre of Computational Linguistics, Loïc Boizou, for helping to process the data, Jurgita Vaičenonienė, for translating the article into English, and Kristina Vaisvalavičienė for the translation of the summary into Latvian.

Abbreviations

acc	accusative
comp	comparative degree
dat	dative
def	definiteness
f	feminine gender
gen	genitive
imp	imperative mood
inf	infinitive
ins	instrumental
loc	locative
nom	nominative
pl	plural
pos	positive
prop	proper noun
sg	singular
sup	superlative degree
voc	vocative

References

1. Ambrazas, Vytautas (ed.). 2006. *Lithuanian Grammar*. Vilnius: Baltos lankos.
2. Boizou, Loic, Kovalevskaitė, Jolanta, Rimkutė, Erika. 2020. Lithuanian Pedagogic Corpus: correlations between linguistic features and text complexity. *The Ninth Conference Human Language Technologies – the Baltic Perspective.* Utka, Andrius, Vaičenonienė, Jurgita, Kovalevskaitė, Jolanta, Kalinauskaitė, Danguolė (eds.). *Frontiers in Artificial Intelligence and Applications.* 328. Amsterdam, Berlin, Tokyo, Washington, DC: IOS Press, 233–240. DOI: 10.3233/FAIA200628
3. *Common European Framework of Reference for Languages: Learning, teaching, assessment*. 2020. Strasbourg: Council of Europe Publishing. Available at: www.coe.int/lang-cefr
4. Drūlienė, Viltė. 2014. *Lietuvių kalbos apibendrinamieji daiktavardžiai: leksinė reikšmė ir pragmatinės funkcijos.* Kaunas: Vytauto Didžiojo universiteto leidykla.
5. Giacomini, Laura, DiMuccio-Failla, Paolo. 2019. Investigating semi-automatic procedures in Pattern-Based Lexicography. *Proceedings of eLex2019*. Sintra, Portugal. Brno: Lexical Computing CZ, 490–505.
6. Hanks, Patrick. 2008. Lexical patterns: from Hornby to Hunston and beyond. *Proceedings of the 13th EURALEX International Congress*. Barcelona: Seria Activitats 20, 89–129. Available at: https://www.euralex.org/elx_proceedings/Euralex2008/004_Euralex_2008_Plenary_Patrick%20Hanks_Lexical%20Patterns_from%20Hornby%20to%20Huston%20and%20Beyond.pdf
7. Hanks, Patrick. 2012. How people use words to make meanings: semantic types meet valencies. *Input, Process and Product: Developments in Teaching and Language corpora.* Boulton, Alex, Thomas, James (eds.). Brno: Masaryk University Press, 52–67.
8. Hanks, Patrick, Može, Sara. 2019. The way to analyse 'way': A case study in word-specific local grammar. *International Journal of Lexicography.* 32(3), 247–269. DOI: https://doi.org/10.1093/ijl/ecz005
9. Kovalevskaitė, Jolanta et al. 2020a. The first corpus-driven Lexical Database of Lithuanian as L2. *The Ninth Conference Human Language Technologies – The Baltic Perspective.* Utka, Andrius, Vaičenonienė, Jurgita, Kovalevskaitė, Jolanta, Kalinauskaitė, Danguolė (eds.). *Frontiers in Artificial Intelligence and Applications.* 328. Amsterdam, Berlin, Tokyo, Washington, DC: IOS Press, 245–252. DOI: 10.3233/FAIA200630
10. Kovalevskaitė, Jolanta, Rimkutė, Erika. 2020b. Mokomasis lietuvių kalbos tekstynas: naujas išteklius lietuvių kalbos besimokantiesiems. *Darnioji daugiakalbystė.* 17, 201–233. DOI: https://doi.org/10.2478/sm-2020-0019
11. Labutis, Vitas. 2002. *Lietuvių kalbos sintaksė*. Vilnius: Vilniaus universiteto leidykla.
12. Marcinkevičienė, Rūta. 2000. Patterns of word usage in corpus linguistics. *Kalbotyra.* 49(3), 71–80.
13. Sinclair, John. 1991. *Corpus, Concordance, Collocation*. Oxford: Oxford University Press. DOI: https://doi.org/10.1177/003368829302400207

Kopsavilkums

Šajā rakstā tiek raksturotas divu vārdšķiru – substantīvu un adjektīvu – gramatiskās lietošanas īpašības. Datu avots – leksikas datubāze „Lietuvių kalbos vartosenos leksikonas" 'Lietuviešu valodas lietošanas leksikons', kura pamatā ir „Lietuviešu valodas mācību korpuss" (ap 620 tūkst. vārdu). Korpuss tika izmantots šķirkļu saraksta izveidei un informācijas par vārdu lietošanu apkopošanai. Piemērojot adaptētu un daļēji automatizētu lietošanas modeļu analīzi, šajā projektā modeļu apraksti sagatavoti 207 verbiem, 386 substantīviem, 87 adjektīviem un 41 adverbam. Lietošanas modeļi tiek sasaistīti ar atšķirīgām aprakstāmā vārda nozīmēm. Lietošanas modelī tiek izdalīti trīs līmeņi: gramatiskais, semantiskais un leksiskais. Manuāli atlasītos atlasītos piemēros tiek atspoguļoti visi komponenti, kas tika piefiksēti modelī.

Šajā rakstā tiek analizēti 673 lietošanas modeļi ar substantīviem un 99 lietošanas modeļi ar adjektīviem: tiek raksturotas to sintaktiskās īpašības, skaidrotas leksikas un gramatikas saiknes. Modeļu ar substantīviem un adjektīviem lielāko daļu sastāda minimālie modeļi, kuros tiek iekļauti vistuvākie sintaksiskie partneri. Tādu rezultātu ietekmēja atšķirīgi vārdšķiru aprakstīšanas principi – viena veida principi tika piemēroti verbiem, citādi – substantīviem, adjektīviem un adverbiem. Lietošanas modeļi aptver plašu gramatisko informāciju, tādēļ līdzīgu modeļu grupā tiek izdalīts pamatmodelis (parasti tas, kurš ir visbiežāk sastopams) un tā varianti. Variantu daudzveidība parāda, ka ir cieša saikne starp gramatiskām īpašībām un individuālu vārda lietojumu.

Atslēgvārdi: leksikas bāze; lietuviešu valoda; „Lietuviešu valodas mācību korpuss"; lietošanas modeļu analīze; lietošanas modelis.