Formal representation of the syntactic environment and semantic features of words

(SYNTEXT – a Web based system designed for frame lexicon)

Svetla Koeva[1], Emil Doychev[2] and Georgi Cholakov[3]

[1] Department for Computational Linguistics
IBL, BAS
52 Shipchenski prohod Blvd.
Sofia 1113, Bulgaria
svetla@ibl.bas.bg

[2, 3] e-Commerce Laboratory
University of Plovdiv
236 Bulgaria Blv., room 430
Plovdiv 4000, Bulgaria
{emil.doychev,georgi.cholakov}@ecl.pu.acad.bg

Abstract

SynText is a web-based system designed for adding, editing and validation of language data for the formal representation of syntactic frames. The system is implemented for the purposes of the Bulgarian frame lexicon but it may serve for the development of other language resources because its design is to the great extends language and theory independent.

1. Introduction

SynText (Syntactic lexicon Tool) is designed for adding, editing and validation of syntactic frames. The preliminary stage was the determination of a uniform theoretical model for the formal representation of the syntactic frames which itself presupposed the architecture of the corresponding software tool for data processing. In this regard several theoretical models directed to the verb semantics, predicate – argument structure and verbal alternations were took into consideration – among most prominent investigations in this field can be mentioned case theory (Fillmore 68), Levin's verb classes (Levin 93), FrameNet database (Collin & al. 98; Fillmore & Atkins 98), and many others (Dowty 96; Grimshow 90. Jackenndoff 90), etc.

It was supposed that some language features that are not handled by the preliminary accepted framework could be encountered. That is why options for adding new parameters for language description are provided, as well as for modification of the already chosen ones. This assumption practically made to the great extent the SynText system language independent as well as theory independent. As a result the system can be easily remobilized in order to be used for languages with different grammars or for one and the same language with different purposes.

The paper itself consists of (a) a description of the functionalities, architecture, technology and implementation of the web based system for editing, correction and verification of syntactic lexicon data, and (b) a presentation of the theoretical model for description of the Bulgarian frame lexicon and its implementation.

2. The system SynText

SynText allows the developers of the frame lexicon to work independently from each other and to use different operating systems (i.e. Windows or Linux). The SynText application has the following major characteristics:

- **Web application** – minimum requirements for the client machine, facile administration and support;

- **Multi-user ability** – many authors could work simultaneously on one and the same data base; the system supports user rights, driven by user roles, and provides authentication and special guest access;

- **Theory independent (partially)** – the theory dependent parameters and their values can be easily changed, if necessary, thus the verification of the theoretical hypothesis could be obtained;

- **Dynamic content** – the system allows fast and easy administration of the linguistic markers from the authorized person, therefore it is fully configurable and customizable;

- **Informational** – different checks up are enabled: to recall all units from one and the same type; to recall all units that satisfy particular criterion; to recall units that possess equal features; etc;

- **Multilanguage interface** – the SynText user interface is designed to allow easy change of the interface language (currently the system has only Bulgarian user interface);

- **Language independent content (partially)** – data from different languages could be added (in this version of the application not simultaneously); for this purpose the new language dependent features can be incorporated into the system framework;

- **Uniform** – the input data are unified by the frame of the current model;

- **Open** – the system is based on open source technologies, with open architecture and written in pure Java – so it can be deployed on different platforms.

Below briefly are viewed the two major aspects of the SynText implementation - its architecture and technology.

2.1. The architecture

From the architectural point of view the SynText system is a three–layered application (Figure 1).

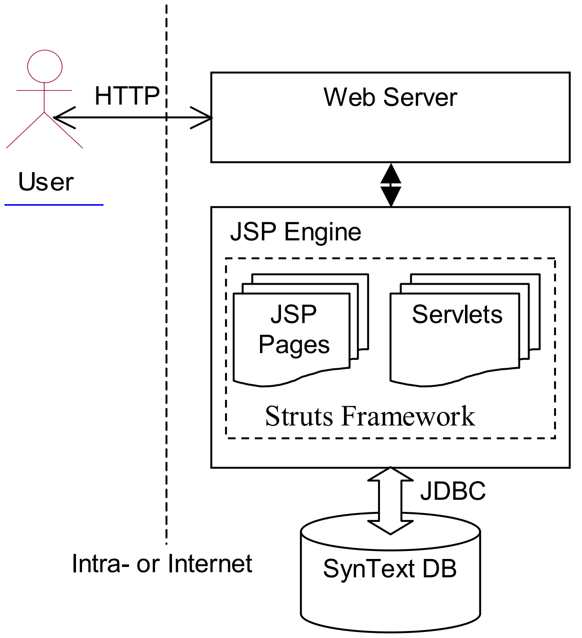


Figure 1: SynText Architecture

All user requests are handled by the Web Server, which is integrated with a JSP Engine for dynamic content generation by the Syntext business logic. The connection between the JSP Engine and the Database is performed by a JDBC driver.

2.2. The technology

The application business logic is primarily stored in the JSP Engine as JSP pages, Servlets and Java Beans. From the technological point of view the SynText application is a Model–View–Controller (MVC) application. This means that the functionality is partitioned into three interacting components – the **Model**, the **View**, and the **Controller** (Brown 01; Duffey 01). Each component maps to three main implementation technologies – beans, JSP, and Servlets.

The SynText is based on the Jakarta Struts Framework (http://jakarta.apache.org/struts) which implements the so called MVC Model 2 architecture. The Model 2 web application architecture uses a Servlet as a request dispatcher, a JavaBean that contains data for the request, and a JSP that presents data view to the user. The UML component diagram in the following Figure represents the Model 2 approach:

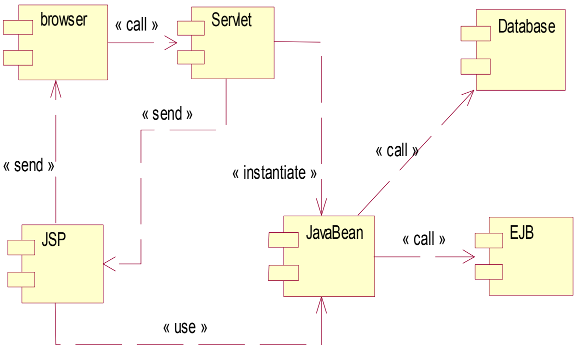


Figure 2: MVC

The Model 2 architecture utilizes a Servlet to receive the request. The Servlet delegates the collection of data for the request to a JavaBean. The JavaBean collects the data needed to satisfy the request by making calls to enterprise components like EJBs and databases and when it is finished collecting the needed data returns control back to the Servlet. The Servlet then forwards the request to the JSP which constructs the HTML response using the data from the JavaBean and its own HTML code. After construction the response is sent to the browser for display.

2.3. The Implementation

Currently the SynText implementation is based on open source technologies and products:

- Web Server – Apache (also tested on MS IIS 4+);

- JSP Engine – Jakarta Tomcat 4 (Specifications: Servlet 2.3, JSP 1.2);

- Database – MySQL 4.1 (with MySQL Connector/J 3.0 JDBC Driver);

- Jakarta Struts Framework 1.1;

- JDK 1.4.

3. Terminological assumptions

The state of the arts (concrete or abstract) to which simple sentences refer are (to some extend) constant in time and language independent. Thus a given predicate is always connected by exact semantic relations with respect to defined number of arguments. In other words, the semantic description of the predicate - argument structures is independent from the natural languages – the number of arguments and their specific semantic relations with the predicate are constant and language independent and thus inter-language differences appear at lexical and syntactical levels (Koeva 04).

We will start with the short terminological specifications because it is well known that in the field of the formal description of syntactic environments are many differences in the terms used.

We call *argumentness* the property of the predicate to incorporate a specific number of variables that correspond to the arguments and their syntactical positions in the sentence (Koeva 98a). For terminological convenience we accept the distinction between subject and complements.

We call a *syntactic lexicon entry* a unit consisting of a target title word, its explanatory definition, its categorial and morphosyntactic characteristics and its syntactic frame.

The *target word* is usually a verb, but in some cases could be a noun, an adjective or a preposition.

The *syntactic frame* is a set of possible syntactic structures associated with the target word.

Under *syntactic structure* we understand the number and combinatorial realizations of syntactic slots and their possible explication in the sentence, defined with syntactic categories and selective restrictions

Syntactic slots are positions of arguments (subject and complements).

The *selective restrictions* are viewed as values that a given candidate for the realization of syntactical phrase must satisfy in a particular position.

4. Approach adopted

We organized the language information in linguistic modules, called language parameters. For example linguistic modules could be explanatory definitions, selective restrictions, etc. Some parameters require free filling of the data, others require an option to be chosen from predefined list of values. Such values could be personal verb – impersonal verb; animate – non-animate, etc. The number and type of parameters as well as the lists of values can be reorganized easily according to the language specific features as well as to the particular tasks performed.

Between some parameters strong dependencies can be fixed, while other parameters are considered independent. For example personal verb will require a subject slot, impersonal – not. The dependencies can also be easily redefined, if necessary.

5. The organization of the frame lexicon

Bulgarian frame lexicon contains information concerning the syntactical environments of lexical units, their semantic combinability, as well as the possible formation of diatheses. The lexicon may include units of all parts of speech, but its main body consists of verbs only. The structure of a lexicon entry consists of the following units:

5.1. Title word

Each target title word in the frame lexicon is a lexical unit with unique meaning. The application of the system allows a word to be chosen from a previously drawn list, a new word to be entered, and a choice of words to be extracted from the list by defining a criterion – a single initial letter or a sequence of initial letters (for each of above options an information is available whether the word has been already processed and what definitions have been entered).

Each word in the list can be itself a link to the corpora driven examples (if such are available) supporting decision for meaning distribution of polysemous words.

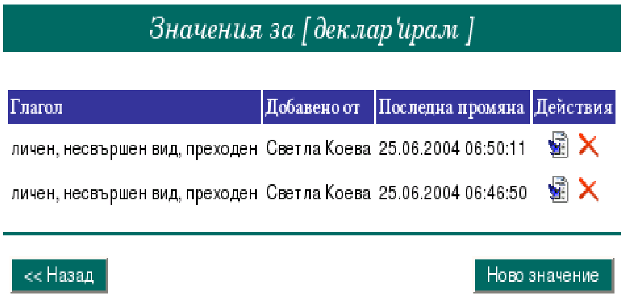
The list of the Bulgarian verbs included in the frame lexicon consists of the most frequent verbs extracted from the bank of Bulgarian structured corpus with 40 000 000 words. We work out with the frequency analysis of the verbal base forms, because we can provide POS disambiguation but we still could not automatically count the most frequent meanings.

The author can complete the following actions operating with the list of target words: adding new lexicon entries, editing existing lexicon entries and deleting target words with no associated lexicon entries. The special feature included is that each author can operate only with those lexicon entries which he/she has added personally.

5.2. Classification of the target word

The preliminary classification of a target word is necessary because the grammatical and morphosyntactical features determine the possible syntactic structures associated with a given word. The first main constituent of the grammatical characteristics is the part of speech information – it was already stipulated that currently work is done mainly on verbs. We have distinguished three groups of grammatical subclasses of Bulgarian verbs depending on the subject – personal, impersonal and third personal singular and plural (Koeva 98b; 04).

Twelve subclasses of personal verbs with a full paradigm of the categories of person and number (by taking into account the features of terminative and durative type, transitivity and intransitivity, constant reflexivity and constant reciprocity) were specified.

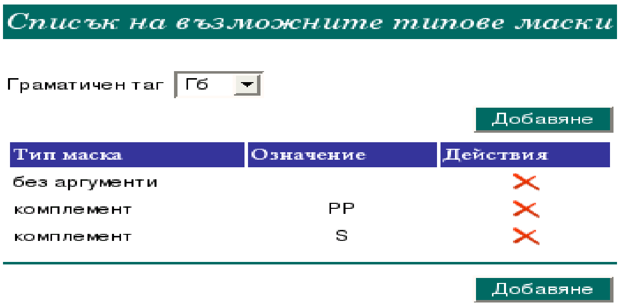


Значения за [деклар'ирам]

Глагол	Добавено от	Последна промяна	Действия
личен, несвършен вид, преходен	Светла Коева	25.06.2004 06:50:11	✕
личен, несвършен вид, преходен	Светла Коева	25.06.2004 06:46:50	✕

Figure 3. Classification of a target word

The verbs with a third person singular and plural subject (and also impersonal verbs) are divided into ten groups depending on the obligatory accusative and dative clitics that they incorporate; while taking into account the features of terminative and durative type, transitivity, intransitivity and constant reflexivity. The verbs of this group are intransitive with the exception of the accusativa tantum verbs, which are formally transitive, usually have no complements with the exception of the obligatory clitics, and have an inanimate subject. The enumerated language specific features have been chosen because they determine the formation of syntactic frames corresponding with Bulgarian verbs. It was already mention that the flexibility of the system allows the easy alternation of the relevant features and the definition of the valuable relations between them.



Списък на възможните типове маски

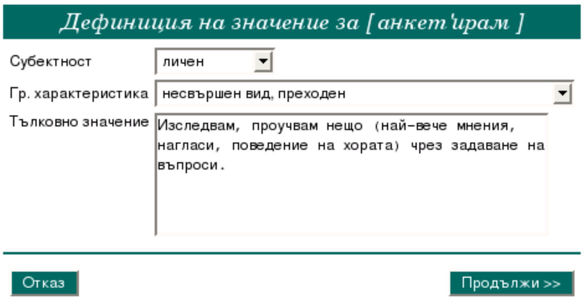
Граматичен таг: Гб

Тип маска	Означение	Действия
без аргументи		✕
комплемент	PP	✕
комплемент	S	✕

Figure 4. Admissible syntactic frames according to a given verb type

5.3. Explanatory definition of the target word

The particular meanings of the polysemy word are in separate lexicon entries, regardless of whether they have different environments or not and are entered in text boxes. It is recommended the number of arguments to be evident from the meaning definition. It can be seen that the already chosen or inserted information is kept to the final validation of the lexicon entry.



Дефиниция на значение за [анкет'ирам]

Субектност: личен

Гр. характеристика: несвършен вид, преходен

Тълковно значение: Изследвам, проучвам нещо (най-вече мнения, нагласи, поведение на хората) чрез задаване на въпроси.

Figure 5. Explanatory definition to a target word

5.4. Syntactic frames

5.4.1. Specification of syntactic phrases

The syntactical phrases that can be candidates for arguments in Bulgarian are: NP (noun phrase), PP (preposition phrase), AdvP (adverb phrase), AP (adjective phrase), S (sentence), SC (small clause). The permissible combinations for particular verb classes may be explicitly listed, but for convenience the application is constructed in such a way that the developers each time select the candidates for arguments by marking them in check boxes.

Thus each verb with a unique meaning is attributed with its obligatory environment, understood as obligatory syntactical positions in the sentence, and not as obligatory explicitness. For a single verb with a unique meaning there might be more than one combination of obligatory environments – i.e. syntactic structures.

Each personal verb incorporates an argument – a noun phrase (NP_0) or a sentence (S_0) that are realized as the subject in the sentence. The subject may be not explicitly stated – with personal verbs the values for person and number of the omitted pronoun subject are contained in the verb inflexion.

The types of argument structures related to the subject of Bulgarian verbs can be characterized as follows:

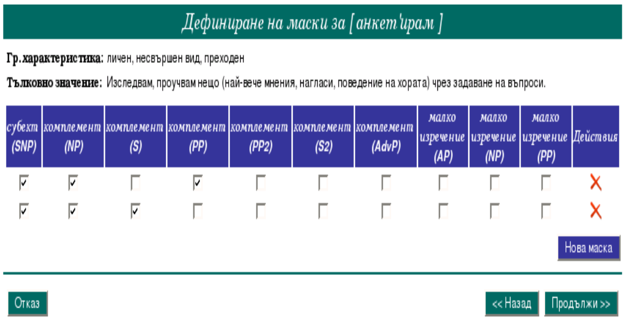
- With explicitly or implicitly expressed subject with a full paradigm of the category of person;
- With explicitly or implicitly expressed third-person subject;
- With no subjective argument.

The types of Bulgarian argument structures, concerning the complements of Bulgarian verbs, can be classified as follows:

- With a single NP complement;
- With a NP complement and a S complement;

- With a NP complement and PP complements, regardless of their number;
- With a NP complement, PP complements, regardless of their number, and an S complement;
- With PP complements, regardless of their number;
- With PP complements, regardless of their number, and a S complement;
- With a S complement;
- With an AdvP predicate modifier;
- With a SC (small clause) NP argument;
- With a SC (small clause) PP argument;
- With a SC (small clause) AP argument;
- With no complements.

For the concrete practical needs for the developing of the lexicon we created masks with combinations of complements and subjects that are possible for the different types of verb subclasses.



Дефиниране на маски за [анкет'ирам]

Гр.характеристика: личен, несвършен вид, преходен

Тълковно значение: Изследвам, проучвам нещо (най-вече мнения, нагласи, поведение на хората) чрез задаване на въпроси.

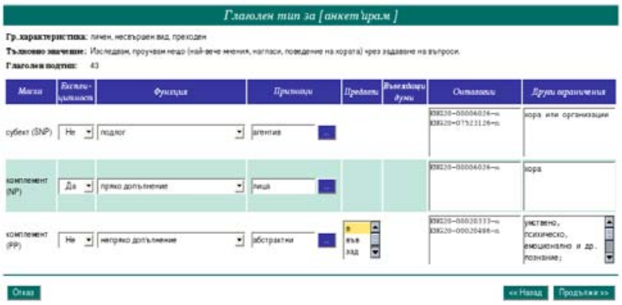
субект (SNP)	комплемент (NP)	комплемент (S)	комплемент (PP)	комплемент (PP2)	комплемент (S2)	комплемент (AdvP)	малко изречение (AP)	малко изречение (NP)	малко изречение (PP)	Действия
☑	☑	☐	☑	☐	☐	☐	☐	☐	☐	X
☑	☑	☑	☐	☐	☐	☐	☐	☐	☐	X

Figure 6. Syntactic slots selection

For example, the masks for personal verbs includes the following components – NP subject, NP, S, PP, PP, S, AdvP, SC; for third personal verbs – NP subject, S subject, NP, S, PP; and for impersonal – no complements, NP, S, PP.

5.4.2. Information concerning each argument separately

All options are not predefined and can alternate from task to task or language to language – here the Bulgarian frame lexicon is exemplified.



Глаголен тип за [анкет'ирам]

Гр.характеристика: личен, несвършен вид, преходен

Тълковно значение: Изследвам, проучвам нещо (най-вече мнения, нагласи, поведение на хората) чрез задаване на въпроси.

Глаголен подтип: 43

Маски	Експли-цитност	Функция	Признаци	Предлози	Въвеждащи думи	Онтология	Други ограничения
субект (SNP)	Не	подлог	агентив			ENG20-00006026-n ENG20-07523126-n	хора или организации
комплемент (NP)	Да	пряко допълнение	лица			ENG20-00006026-n	хора
комплемент (PP)	Не	непряко допълнение	абстрактни	в във зад		ENG20-00020333-n ENG20-00020486-n	умствено, психическо, емоционално и др. познание;

Figure 7. Information in syntactic structure

Every syntactic structure includes information about the phrases' explicitness (check box), syntactic function (list box) belonging to a given slot, selection restriction (tree structure), specification for prepositions (list) and complementizers (list), specification for the top most entry in the Bulgarian WordNet and other specific comments if necessary. The last two specifications are entered in text boxes.

- **Explicitness of the phrase**

The phrases that express the arguments may be obligatory explicit (in rare cases in Bulgarian) or non–explicit, which means that the potential possibility for a syntactical realization of the phrase exists, but its explicitness is not mandatory because it is understood from the context in a broader sense (verb morphology, preceding text, extralinguistic information, etc.)

- **Syntactical function**

The syntactical functions (names of syntactical positions taken from traditional grammar) are subject, direct object, indirect object, adverbial, subjective clause, objective clause, adverbial clause, and small clause. Free relatives are not encoded in the syntactic structures as it is accepted that each NP can be described with a free relative.

- **Semantic features (selective restrictions)**

The arguments are realized in syntactic structures with concrete words that may be compatible or not with the meaning of the verb. We call selectivity the semantic restrictions to a given argument in a certain context. Due to the fact that selective restrictions act between a concrete predicate and the arguments that belong to it, they can be different for each separate case.

The most general semantic classification distinguishes among abstract and concrete nonus. On their part, concrete nouns can be animate or inanimate. Animate nouns may be classified as persons and non–persons, persons as agents or experiencers. The Figure 8 presents a convenient method for classification of the selective restrictions with nouns as the over–line restriction implicates the corresponding under–line one. The developers of the lexicon can choose a feature from the tree and the selected feature includes all features that it dominates over and inherits the characteristics only of the features that dominate it. Thus, if the feature of *person* has been selected, it shall include the features of *agent*, and *experiencer*, and inherits the features of *animate* and *concrete*. More that one feature can pertain to an argument and the selected features are conjunctive or disjunctive.

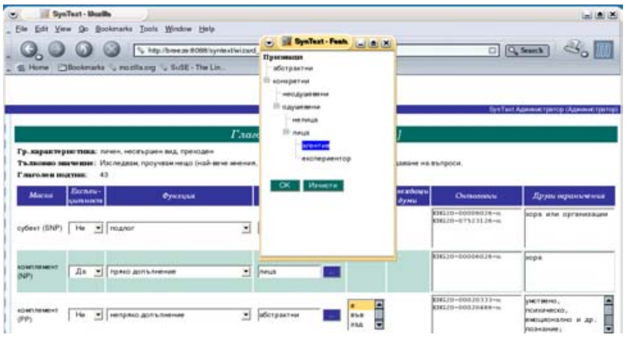


Figure 8. Selective restrictions' tree

• Specification of the prepositions for preposition groups

The concrete prepositions for a given argument expressed with a prepositional phrase are to be selected from a list box – it is permissible more than one preposition to be selected.

• Specification of the complementizers

The types of subordinate clauses depend on the method of conjunction – interrogative, relative, conjunctional, thus the respective complementizers are to be selected from a list box (more than one choice is permissible).

▪ Specification for the top most WordNet synset

In some cases the selective restrictions do not provide all necessary restrictive information. Besides the general cases there may also be cases where concrete restrictions are required, as for example *liquid, food*, etc. That is why we think reasonable to include the link to the top most synonymous set (or the conjunction of top most synonymous sets) taken from the Bulgarian WordNet (Koeva & al. 04). The top most set should dominate all appropriate sets for a given syntactic slot i.e. *liquid* is a hypernym of *water, milk, liquor,* etc.

• Other necessary comments

The developer has the opportunity to enter concrete restrictions, for example additional selective restrictions, notes about the prepositions, control information for equi and raising verbs, surface order information, etc.

• Illustrative examples

At least five illustrative examples usually taken from our corpora have to be presented – if there is more then one arguments combination, for each combination an example has to be provided.

6. Main priorities

The main priorities of the organization of the lexicon presented are as follows:

6.1. Linguistic

The theoretical investigation and its implementation combine the lexicon with grammar rules for the verbal transformations (diatheses).

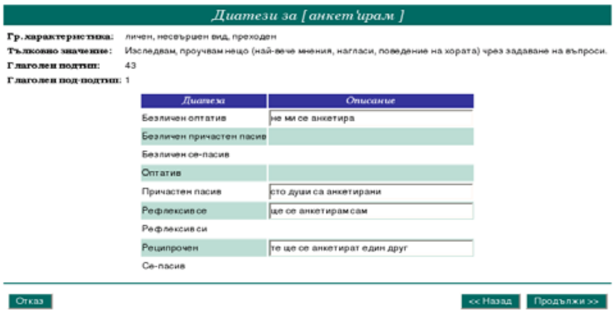
Diatheses are the transformations where the number of arguments and/or their syntactic realizations change, but the basic meaning of the verb remains. The following verb diatheses can be observed in the Bulgarian language: Real reflexives; Real reciprocals; Real optatives; Impersonal optatives; Real participle passives; Impersonal participle passives; Real se–passives; Impersonal se–passives (Koeva 98b; 04).

The Bulgarian verbs with a full paradigm of the category of person can be classified as productive as regards to transformations. The verbs from the second group (with a limited paradigm of the category of person – third person singular and plural) are productive as regards to the impersonal diathesis. The verbs from the third group (the impersonal verbs) are not productive for transformations.

The values of the following three parameters are of significance for the formation of diatheses: argumentness, selectiveness, and perfectiveness.

The number and type of the complements, as well as the presence or lack of a subject, act as determinative conditions for the formation of all types of diatheses. Even though the list of the minimum required selective restrictions that guarantee the correctness of each possible sentence has not been specified, the features that play a part in the formation of the transformations in are: experiencer, agent, person, animate, non–animate.

With the term perfectiveness we denote the way the action signified by the predicate runs. The following general tendency can be observed: the durative type is productive, while the terminative type is unproductive for diatheses.



Диатези за [анкет'ирам]

Гр. характеристика: личен, несвършен вид, преходен
Тълковно значение: Изследвам, проучвам нещо (най-вече мнения, нагласи, поведение на хората) чрез задаване на въпроси.
Глаголен подтип: 43
Глаголен под-подтип: 1

Диатеза	Описание
Безличен оптатив	не ми се анкетира
Безличен причастен пасив	
Безличен се-пасив	
Оптатив	
Причастен пасив	сто души са анкетирани
Рефлексив се	ще се анкетирам сам
Рефлексив си	
Реципрочен	те ще се анкетират един друг
Се-пасив	

Figure 9. Generation of transformations

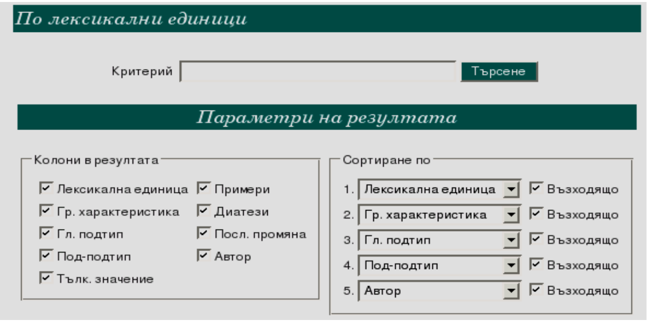
The correctness of the information encoded in the lexical entries can be easily verified. The combination of possible diatheses is unique for a given verb and thus it validates the input data in the lexicon. For example if the grammatical characteristics for a particular Bulgarian verb are **personal transitive (non–)perfective verb** and the syntactic structure is NP subject and NP complement, and the chosen features are agent – non–animate then the se–passive diathesis is permissible.

6.2. Flexibility and standardization

Flexibility of the system predetermines the broad scope of its usage in different tasks for different languages. But at the same time the uniform framework presupposes the standardization of the classified language data into the chosen paradigm.

6.3. Checks up

The system provides wide spectrum of checks up. This option might be used both for comparison and validation of the language data. Every parameter included in the system can be used as a selection criterion for the checks up definition – for instance, checks up by author or by target words, by verb types included, by similarities between syntactic structures, etc.



По лексикални единици

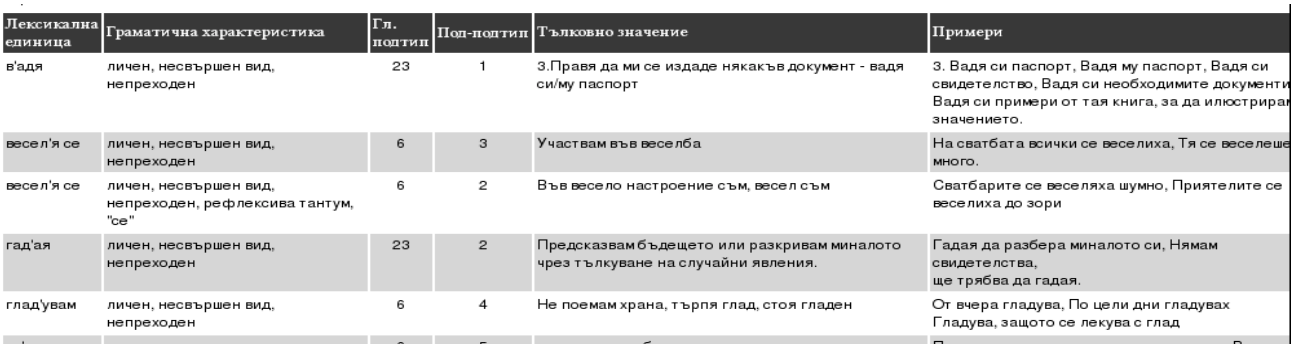
Критерий [ ] Търсене

Параметри на резултата

Колони в резултата
- ☑ Лексикална единица
- ☑ Гр. характеристика
- ☑ Гл. подтип
- ☑ Под-подтип
- ☑ Тълк. значение
- ☑ Примери
- ☑ Диатези
- ☑ Посл. промяна
- ☑ Автор

Сортиране по
1. Лексикална единица ☑ Възходящо
2. Гр. характеристика ☑ Възходящо
3. Гл. подтип ☑ Възходящо
4. Под-подтип ☑ Възходящо
5. Автор ☑ Възходящо

Figure 10. Selection criteria for the checks up



Лексикална единица	Граматична характеристика	Гл. подтип	Под-подтип	Тълковно значение	Примери
в'адя	личен, несвършен вид, непреходен	23	1	3.Правя да ми се издаде някакъв документ - вадя си/му паспорт	3. Вадя си паспорт, Вадя му паспорт, Вадя си свидетелство, Вадя си необходимите документи Вадя си примери от тая книга, за да илюстрирам значението.
весел'я се	личен, несвършен вид, непреходен	6	3	Участвам във веселба	На сватбата всички се веселиха, Тя се веселеше много.
весел'я се	личен, несвършен вид, непреходен, рефлексива тантум, "се"	6	2	Във весело настроение съм, весел съм	Сватбарите се веселяха шумно, Приятелите се веселиха до зори
гад'ая	личен, несвършен вид, непреходен	23	2	Предсказвам бъдещето или разкривам миналото чрез тълкуване на случайни явления.	Гадая да разбера миналото си, Нямам свидетелства, ще трябва да гадая.
глад'увам	личен, несвършен вид, непреходен	6	4	Не поемам храна, търпя глад, стоя гладен	От вчера гладува, По цели дни гладувах Гладува, защото се лекува с глад

Figure 11. Bulgarian frame lexicon body

7. Conclusions

The proposed methodology for formal description of syntactic frames is to the great extends language independent and theory independent. Thus it can be easily used for the formal descriptions of languages different from Bulgarian.

The further functionality of the SynText system is XML import / export option of the lexicon data – this will allow convenient archive of the data base, as well as data exchange with other natural language processing tools using XML format.

References

Brown 01 S. Brown at all. 2001. *Professional JSP* 2nd Edition. Wrox Press, Birmingham.

Collin & al. 98 Collin F. Baker, Charles J. Fillmore, and John B. Lowe. 1998. *The Berkeley FrameNet project*. In “Proceedings of the 17th International Conference on Computational Linguistics (COLING/ACL– 98)”, pages 86–90, Montreal. ACL.

Dowty 96 D. R. Dowty. 1996 *On the semantic content of the notion thematic role*. In G. G. Chierchia, B. H. Partee, and R. Turnen eds., “Prpoperties, Tipes and Meaning”, Dordrecht: Kluwer.

Duffey 01 K. Duffey at all. 2001. *Professional JSP Site Design*. Wrox Press, Birmingham,

Fillmore 68 Charles J. Fillmore. 1968. *The case for case*. In E. Bach and R. Harms (eds.), Universals in Linguistic Theory. New York: Holt, Rinehart & Winston, 1968.

Fillmore & Atkins 98 Charles J. Fillmore and B.T.S. Atkins. 1998. *FrameNet and lexicographic relevance*. In “Proceedings of the First International Conference on Language Resources and Evaluation”, Granada, Spain, 28– 30 May 1998.

Grimshow 90 Jane Grimshow. 1990. *Argument Structure*, Cambridge USA: MIT Press.

Jackenndoff 90 R. Jackenndoff. 1990. *Semantic Structures*, Linguistic Inquiry Monograph 18, Cambridge: MIT Press.

Koeva 98a S. Koeva. 1998. *Argument structure, thematical relations and syntactic realization of arguments*. In “Language consciousness”, Sofia, 206– 230.

Koeva. 98b. S. Koeva 1998 *Reflexive, passive, optative, reciprocal and impersonal verbs in Bulgarian,* In. “Scientific studies of Plovdiv University”. 97– 112.

Koeva. 04 S. Koeva 2004 *Semantical ans syntactical descriptions of Bulgarian deatheses*, in: Bulgarian linguistics IV, Sofia: Axademic Press. 2004, 182-231.

Koeva at al. 04 S. Koeva, Tinko Tinchev, and Stoyan Mihov *Bulgarian Wordnet-Structure and Validation*, in: Romanian Journal of Information Science and Technology, Volume 7, No. 1-2, 61-78.

Levin 93 Beth Levin. 1993. *English Verb Classes and Verb Alternations: A Preliminary Investigation*. University of Chicago Press.