Peculiarities of Generation of Semantics of Natural Language Speech by Helping Unlimited and Context-Dependent Grammar

Vasyl Lytvyn[0000-0002-9676-0180]1, Solomiya Kubinska[0000-0003-3201-635X]2, Andriy Berko[0000-0001-6756-5661]3, Tetiana Shestakevych[0000-0002-4898-6927]4, Lyubomyr Demkiv[0000-0002-2802-3461]5, Yuriy Shcherbyna[6]

[1-5]Lviv Polytechnic National University, Lviv, Ukraine
[6]Ivan Franko National University of Lviv, Lviv, Ukraine

Vasyl.V.Lytvyn@lpnu.ua[1], kubinskasm@gmail.com[2], Andrii.Y.Berko@lpnu.ua[3], Tetiana.V.Shestakevych@lpnu.ua[4], demkivl@gmail.com[5], Shcherbyna@franko.lviv.ua[6]

Abstract. The article deals with the use of generative grammars in linguistic modeling. Description of sentence syntax modeling is used to automate the analysis and synthesis of natural-language texts.

Keywords. Sentence structure, generative grammar, computer linguistic system.

1 Introduction

At the present stage of development, the need to develop common and specialized linguistic systems is forcing the use of applied and computer linguistics in the field of information technology [3-4, 7-14, 16, 22, 25-28, 32-33, 35, 43, 50-51, 53-55, 66]. Development of mathematical models of speech to provide computer linguistic systems enables the implementation of tasks of applied linguistics, such as analysis/synthesis of oral/written text content, description/indexing of text content, translation of texts, creation of lexicographic databases, etc. [7-8, 10-15, 25-28, 32-33, 43, 50-51, 53-55, 66]. Linguistic analysis of textual content consists of several sequential processes i.e. grapheme, morphological, syntactic, and semantic analysis [2-5, 9-10, 16-23, 28-30, 34]. For each of these stages, the corresponding models and algorithms are created [2-5, 9-10, 15-21, 44-49, 56-64]. An effective tool of linguistic modeling at the syntactic and semantic level of language is the main part of combinatorial linguistics - the theory of generative grammar, the beginning of which is based on the works of the American linguist N. Chomsky [16-20, 45-48, 56-64]. He used the technique of formal analysis of the grammatical structure of phrases to distinguish syntactic structure (constituents) as the basic scheme of a phrase, regardless of its meaning [16]. The ideas of N. Chomsky were developed by Soviet linguist A. V. Gladkyy [18 -

20], applying the concept of dependency trees and systems of components for modeling syntactic language level [30, 50]. He proposed a method of syntax modeling using syntactic groups that distinguish word components as units of constructing a dependency tree, such a representation combines the advantages of the method of direct constituents and dependency trees [4-5, 50].

Advantages of modelling using generative grammars are the ability to describe not only the syntactic level of language (rules for the formation of sentences from word forms) [12-14, 18-20, 48, 51], but also morpheme (rules for the formation of word forms from morphs) [9-10, 23, 38-41, 55], and semantic (rules of formation of meaningful sentences and texts) [2, 29-30, 47, 49, 56]. It is used for automation of word-change / word-formation processes, rubric or definition of keywords and formation of text content digest [7-8, 10-11, 23-28, 43, 66]. For example, when using automatic morphological synthesis, a computer linguistic system generates the necessary word forms based on the requirements for word forms and morpheme databases [32].

2 Connection of the Highlighted Issue to Important Scientific and Practical Work

Based on the importance of providing automatic processing of textual content in modern information tools (for example, information retrieval systems, machine translation, semantic, statistical, optical and acoustic analysis and synthesis of language, automated editing, extraction of knowledge of textual content, abstracting and annotating , indexing of textual content, educational and didactic, management of linguistic corpora, tools for compiling dictionaries of various types etc.), experts are intensively looking for new models, methods of their description and methods of automatic processing of textual content [7-8, 10-11, 23-28, 43, 66]. One such method is to develop general principles for the construction of lexicographic systems of syntactic type and to build on these principles these systems of textual content processing for specific languages [7-8, 10-11, 23-28, 43, 66].

Studies of linguists in the field of morphology, morphonology, structural linguistics have identified different structures for describing word forms [7-8, 10-15, 25-28, 32-33, 43, 50-51, 53-55, 66]. With the beginning of the development of the theory of generative grammars, linguists have focused not only on the description of ready word forms, but also on the processes of their synthesis [38]. Functional studies of linguists in Ukrainian are fruitful [2-3, 6, 12-14, 22-24, 28, 30, 32-33, 39, 42-43, 51-55], in particular, theoretical problems of morphological description [9-10, 23, 38-41, 55], questions of classification of morphemic and word-forming structure of derivatives of Ukrainian language [23], regularities of combinatorics of affixes [23, 55], modeling of word-formation mechanism of modern Ukrainian language in dictionaries of integral type [23, 38- 39], the principles of the internal organization of the word [23, 55], as well as the questions of the structural organization of distinct verbs and suffixes of eminative value [23, 38, 55], problems of word-forming motivation in the process of derivatives formation [23], regularities of realization of morphonological phenomena in Ukrainian word-formation [23, 39, 55], morphonological modifications in the process of verb-word change [23, 38, 55], morphonological processes in word-translation and word-formation of adjectives of modern Ukrainian literary language [23, 38-39], analysis and processing of textual content [7-8, 10-14, 22, 25-28, 30, 32-33, 43, 55, 66] etc.

Such a dynamic approach of modern linguistics in the analysis of the morphological level of the language, with the researcher's focus on the development of morphological rules, makes it possible to effectively apply the results of theoretical research in practice - for the construction of computer linguistic systems for processing textual content for various purposes [7-8, 67-74]. One of the first attempts to apply the theory of generative grammars to linguistic modeling is A.V. Gladkyy and I.A. Melchuk [18-20]. Developments by N. Chomsky [16-20, 45-48, 56-64] and A.V. [18-20], the study of M. Gross and A. Lanten [21], A.V. Anisimova [2-3], Yu.D. Apresian [4-5], N.C. Bilgaeva [9], I.A. Volkova and T.V. Rudenko [15], E.I. Bolshakova, ES Klyshynskyi, DV Lande, A.A. Noskova, OV Peskova, and E.V. Yagunova [10-11, 26-27], A.S. Gerasimova [17], B.K. Martynenko [29], A.E. Pentus and M.R. Pentus [34], E.V. Popova [35], V.S. Fomicheva [44] are applicable to the development of such textual content processing tools as information retrieval systems, machine translation systems, textual content annotation, morphological, syntactic and semantic analysis of textual content, educational and didactic textual content processing systems, to special systems of linguistic content etc. [7-8, 10-14, 22, 25-28, 30, 32-33, 43, 55, 66].

3 Analysis of scientific results

Grammar has the following important in the linguistic aspect. We will interpret terminal symbols as wordforms (some natural language), auxiliary symbols - as syntactic categories (for example, V is verb, S is noun, A is adjective, $\tilde{V}$ is group of verbs, $\tilde{S}$ is noun group), initial symbol – as (sentences), and terminal strings, that are deduced - as correct sentences of the given language. Then the derivation of a sentence is naturally interpreted as its syntactic structure, presented in terms of direct constituents. Let's explain the examples.

1. Весела посмішка твого сина наповнює мене безмежним щастям [12-14, 51].
2. In seinem bedeutendsten Werk zeigt er die bunte welt des ukrainischen Dorfes in ihrem einmaligen Reiz [31].

Constructing a grammar G_1 that will generate phrases of the corresponding language (Ukrainian, English or German), syntactically same and very simple [1, 6, 10-14, 22, 24, 28, 31, 36-42, 49, 51-55, 65] . Let us write out only the scheme of this grammar; its terminal symbols are the word forms of the corresponding language, and the auxiliary dictionary contains the above syntactic categories. The symbols of these categories are provided with indexes corresponding to their morphological features, for example $S_{ж,од,р}$. The initial character is indicated by R.

Diagram of grammar G_1. The contents of the designations used G_1 are explained after the table 1.

Table 1. Rules for the formulation of sentences in Ukrainian

№	The name of the rule	Rule
I.	Choice of structure R	$R \to \# \tilde{S}_{x,y,н,w} \tilde{V}_{y,тепер,w} \#$.
II.	Deploying a Nominal Group	1) $\tilde{S}_{x,y,z,3} \to \tilde{S}_{x,y,z,3} \tilde{S}_{x',y',p,w}$; 2) $\tilde{S}_{x,y,z,3} \to A_{x,y,z} \tilde{S}_{x,y,z,3}$; 3) $K_1 \tilde{S}_{x,y,z,w} K_2 \to K_1 S^{займ}_{x,y,z,w} K_2$, де K_1 – a symbol other than a symbol $A_{x,y,z}$, a K_2 – a character other than an index character $z' = p$. Symbols K_1 and K_2 are contextual constraints. The substantive meaning of their introduction to this rule is that the principal member of a nominal group should not be implemented by a personal pronoun if it is preceded by a definition expressed by an agreed adjective, or if it is followed by a nominal group in the generic case, for example, the impossibility of a *new self* or *tenderness*. In poetic language, such combinations suggest; 4) $\tilde{S}_{x,y,z,3} \to S_{x,y,z}$.
III.	Deployment verbal groups	1) $\tilde{V}_{y,тепер,w} \to V_{y,тепер,w} \tilde{S}_{x',y',зн,w'} \tilde{S}_{x'',y'',ор,w''}$; 2) $\tilde{V}_{y,тепер,w} \to V_{y,тепер,w} \tilde{S}_{x',y',ор,w'} \tilde{S}_{x'',y'',зн,w''}$; 3) $\tilde{V}_{y,тепер,w} \to V_{y,тепер,w} \tilde{S}_{x',y',зн,w'}$; 4) $\tilde{V}_{y,тепер,w} \to V_{y,тепер,w} \tilde{S}_{x',y',ор,w'}$.
IV.	Realization syntactic categories word forms	1) $S_{ч,y,z} \to син_{y,z}, \dots$; 2) $S_{ж,y,z} \to посмішка_{y,z}, \dots$; 3) $S_{сер,y,z} \to щастя_{y,z}, \dots$ 4) $S^{займ}_{x,од,z,1} \to я_z$; 5) $S^{займ}_{x,од,z,2} \to ти_z$; 6) $A_{x,y,z} \to веселий_{x,y,z}, безмежний_{x,y,z}, мій_{x,y,z}, твій_{x,y,z}, \dots$; 7) $V_{y,тепер,w} \to наповнити_{y,тепер,w}, \dots$

Each line in this diagram is not a single rule, but a abbreviation entry of several rules. Thus, line II.1 forms 648 rules for forming phrases in Ukrainian [6, 12-14, 24, 42, 51, 52, 55]:

$\tilde{S}_{ч,од,н,3} \to \tilde{S}_{ч,од,н,3} \tilde{S}_{ч,од,р,1}$; $\tilde{S}_{ч,од,р,3} \to \tilde{S}_{ч,од,р,3} \tilde{S}_{ч,од,р,1}$; …; $\tilde{S}_{сер,мн,м,3} \to \tilde{S}_{сер,мн,м,3} \tilde{S}_{сер,мн,род,3}$, where the abbreviation *ч* is masculine, *од* is singular, *н* is a nominative distinguisher, *р* is a generic distinguisher, *сер* is an average genus, *мн* is a plurality, *м* is a ablative, 1 is the first person of the noun, 3 is the third person of the noun. The same method of reduction is used in the following examples. However, for the sake of simplicity, we will refer to the lines of such abbreviated entries as rules (Table 2). Rule IV does not take into account the matching A of the noun of the creature S in the accusative case. Designation: $\#$ - the symbol of the boundary of the sentence, which is terminal (in the text, the left border is capitalized in the first letter of the word, and the right - in a dot); x, y, z, w - characteristics of word forms corresponding to the type, number, case, person, for example, $веселий_{ж,од,н} = весела$. An example of a grammar G_1 derivation for generating a Ukrainian sentence is shown in fig. 1.

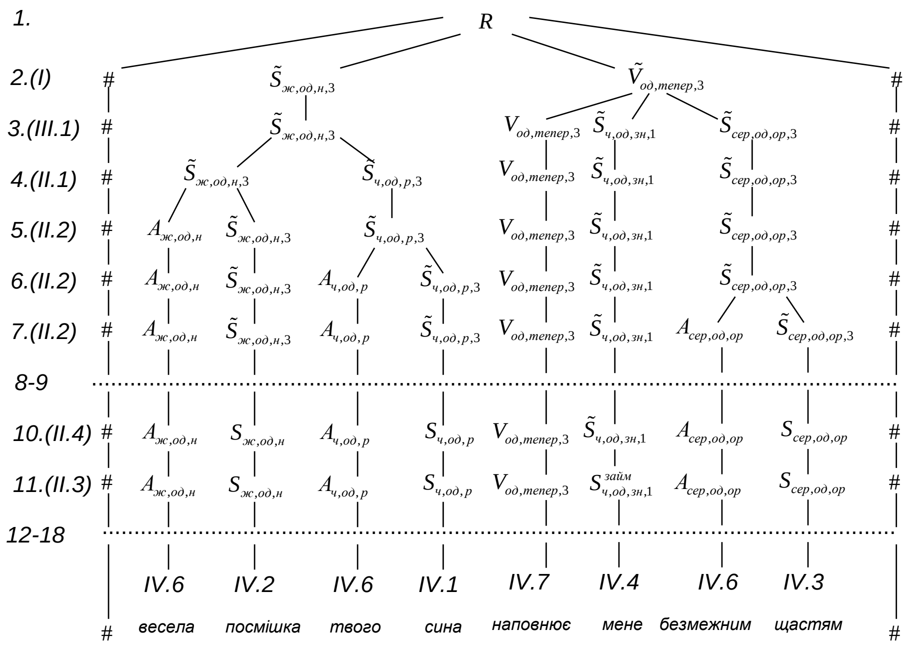

Fig. 1. Example of phrase structure grammar (type 0) for generating a Ukrainian sentence

Grammar can give rise to other phrases (which are not necessarily meaningful), for example: *я наповнюю тебе щастям*, *веселе безмежне моє щастя наповнює тобою твоїй посмішці щастя твоєї посмішки щастя*, etc. Grammar G_1 generates infinitely many different phrases (as opposed to grammar G_0), since its composition includes the so-called cyclic rules (II.1 and II.2). The peculiarity of such a rule is that the result of its application contains the occurrence of its left, so that it can always be applied to its own result, which leads to an infinite number of phrases: yes, along with a group of *весела посмішка* you can get a *весела весела посмішка*, further *весела*

весела весела посмішка, etc., so, the adjective *весела* can be repeated as many times as you like. This raises the question of the infinite number of phrases in natural language, in relation to which we note that at any given moment the number of words of any natural language is finite. In addition, the maximum length of phrases occurring in a language is practically limited: it is unlikely that people use phrases more than, say, 1000 words. It follows that the number of phrases in natural language should be finite. But it is not possible to specify the longest phrase: whichever phrase is not suggested, we can always extend it by adding, for example, another homogeneous term or a contract sentence of which. In natural language, there are fundamental possibilities for constructing as many long phrases as possible, that is, potentially real phrases of any length, although in practice large phrases do not use them. This potential unlimited length of phrases cannot be ignored by formal grammarians, since their task is to model the fundamental possibilities of natural language. If the lengths of grammar-generated phrases are infinite, then the set of all these phrases is infinite.

For German, record II.1 forms 288 rules: $\tilde{S}_{m,einz,nom,3} \to \tilde{S}_{m,einz,nom,3}\ \tilde{S}_{m,einz,gen,1}$; $\tilde{S}_{m,einz,gen,3} \to \tilde{S}_{m,einz,gen,3}\ \tilde{S}_{m,einz,gen,1}$; …; $\tilde{S}_{n,plur,akk,3} \to \tilde{S}_{n,plur,akk,3}\ \tilde{S}_{n,plur,akk,3}$, where m is Maskulinum (masculine), *einz* is Einzahl (singular), *nom* is Nominativ (nominal), *gen* is Genitiv (genetive), *n* is Neutrum (neuter), *akk* is Akkusativ (accusative), *plur* is Plural (plural), 1 is erste Person Substantiv (first person of the noun), 3 is dritte Person Substantiv (third person of the noun) [12-13, 31, 36-37, 41, 49]. In the table 2 sets out the rules for formulating sentences in German.

Table 2. Rules for the formulation of sentences in German

№	The name of the rule	Rule
I.	Choice of structure R	1) $R \to \#\tilde{S}_{x,y,nom,w}\ \tilde{V}_{y,pr,w}\ \#$; 2) $R \to \#\tilde{S}_{x,y,dat,w}\ \tilde{V}_{y,pr,w}\ \#$.
II.	Deploying a Nominal Group	1) $\tilde{S}_{x,y,z,3} \to \tilde{S}_{x,y,z,3}\ \tilde{S}_{x',y',gen,w}$; 2) $\tilde{S}_{x,y,z,3} \to A_{x,y,z}\ \tilde{S}_{x,y,z,3}$; 3) $K_1\tilde{S}_{x,y,z,w}\ K_2 \to K_1 S^{pron}_{x,y,z,w} K_2$, де K_1 — a symbol other than a symbol, and K_2 — a character other than an index $z' = gen$. Symbols are contextual constraints. The K_1 and K_2 substantive meaning of their introduction to this rule is that the principal member of the nominative group should not be implemented by a personal pronoun if it is preceded by a definition expressed by an agreed adjective, or if it is followed by іменна група в a generic case, such as the impossibility of neu ich (new) or er Zärtlichkeit (he of tenderness). In poetic language, such combinations suggest; 4) $\tilde{S}_{x,y,z,3} \to S_{x,y,z}$;

№	The name of the rule	Rule
		5) $\tilde{S}_{x,y,z,3} \to \tilde{S}_{x,y,z,3}\ \tilde{S}_{x',y',akk,w}$; 6) $\tilde{S}_{x,y,z,3} \to \tilde{S}_{x,y,z,3}\ \tilde{S}_{x',y',dat,w}$
III.	Deployment verbal groups	1) $\tilde{V}_{y,pr,w} \to V_{y,pr,w}\ \tilde{S}_{x',y',akk,w'}\ \tilde{S}_{x'',y'',gen,w''}$; 2) $\tilde{V}_{y,pr,w} \to V_{y,pr,w}\ \tilde{S}_{x',y',gen,w'}\ \tilde{S}_{x'',y'',akk,w''}$; 3) $\tilde{V}_{y,pr,w} \to V_{y,pr,w}\ \tilde{S}_{x',y',akk,w'}$; 4) $\tilde{V}_{y,pr,w} \to V_{y,pr,w}\ \tilde{S}_{x',y',akk,w'}\ \tilde{S}_{x'',y'',gen,w''}\tilde{S}_{x''',y''',dat,w'''}$; 5) $\tilde{V}_{y,pr,w} \to V_{y,pr,w}\ \tilde{S}_{x',y',akk,w'}\ \tilde{S}_{x'',y'',dat,w''}$; 6) $\tilde{V}_{y,pr,w} \to V_{y,pr,w}\ \tilde{S}_{x',y',dat,w'}\ \tilde{S}_{x'',y'',akk,w''}$; 7) $\tilde{V}_{y,pr,w} \to V_{y,pr,w}\ \tilde{S}_{x',y',dat,w'}\ \tilde{S}_{x'',y'',akk,w''}\tilde{S}_{x''',y''',dat,w'''}$; 8) $\tilde{V}_{y,pr,w} \to V_{y,pr,w}\ \tilde{S}_{x',y',dat,w'}\ \tilde{S}_{x'',y'',dat,w''}\tilde{S}_{x''',y''',akk,w'''}$; 9) $\tilde{V}_{y,pr,w} \to V_{y,pr,w}\ \tilde{S}_{x',y',akk,w'}\ \tilde{S}_{x'',y'',dat,w''}\tilde{S}_{x''',y''',dat,w'''}$; 10) $\tilde{V}_{y,pr,w} \to V_{y,pr,w}\ \tilde{S}_{x',y',nom,w'}$;
IV.	Realization syntactic categories word forms	1) $S_{m,y,z} \to Reiz_{y,z},\dots$ (де *der Reiz – attractiveness*); 2) $S_{f,y,z} \to Welt_{y,z},\dots$ (де *die Welt – world*); 3) $S_{n,y,z} \to Werk_{y,z}, Dorf_{y,z},\dots$ (where das Werk is work, das Dorf is a village); 4) $S^{pron}_{x,einz,z,1} \to er_z$ (де *er – he*); 5) $S^{pron}_{x,plur,z,3} \to sie_z$ (де *sie* – they); 6) $A_{x,y,z} \to sein_{x,y,z}, ihr_{x,y,z}, bedeutend_{x,y,z}, bunt_{x,y,z}, ukrainisch_{x,y,z}, einmaling_{x,y,z},\dots$ (where *sein – my, ihr - theirs, bedeutend - important, bunt - colorful, ukrainian - ukrainian, einmalig - unique*); 7) $V_{y,pr,w} \to zeigen_{y,pr,w},\dots$ (де *zeigen – show*).

An example of grammar output for generating a German sentence is shown in fig. 2.

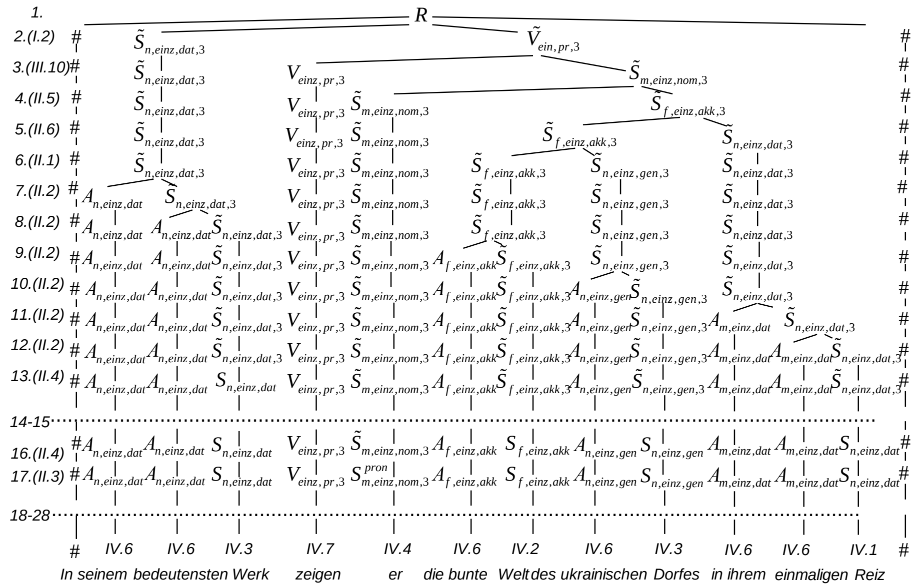

Fig. 2. An example of a phrase structure grammar (type 0) for generating a German sentence

Let's return to fig. 1, where each step of the output consists in the deployment of one of the characters of the previous chain (for example, in the transition from 2 to 3 chains the symbol $\tilde{V}_{од,тепер,3}$ is expanded in 3 characters - $V_{од,тепер,3}$ $\tilde{S}_{ч,од,зн,1}$ $\tilde{S}_{сер,од,ор,3}$), or in its replacement by another (for example, in the transition from 10 to 11, $\tilde{S}_{ч,од,зн,1}$ the symbol is replaced by $S^{займ}_{ч,од,зн,1}$) , the other characters are overwritten unchanged (wildcard rules). Expandable, replaceable, or rewritable characters are *ancestors*, and characters that we get from deployment, replacement, or rewriting are their *descendants* (descendants of descendants are also descendants). We connect the ancestors with the lineages with their immediate descendants. Then we get the component tree, or the syntactic structure of the phrase in terms of immediate components (type 1). To illustrate this, let more clearly eliminate the scheme in Fig. 1 are all descendant characters that rewrite without modification (for example, $\tilde{S}_{ч,од,зн,1}$ in chains 4-10) and combine the same steps 4-7 (Fig. 3).

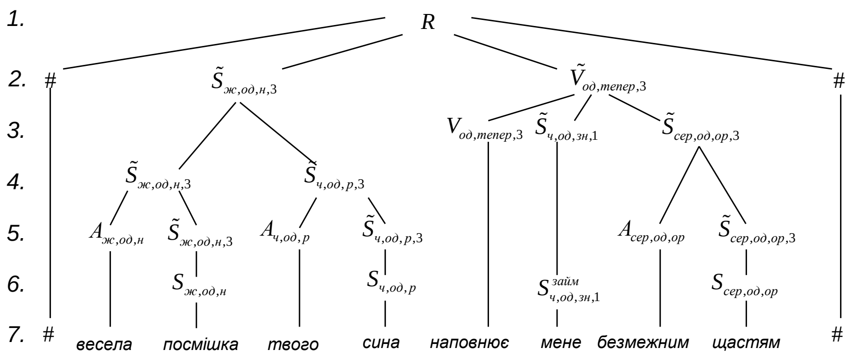

Fig. 3. Example of context-dependent grammar of immediate components (type 1)

Context-free grammars are a special case of grammars of direct components (an example of using context in G_1 is rule II.3). Their value is due to the following circumstances:

- rejection of the context (there is exactly one character in the left side of the rule) makes the grammar structure easier and easier to study;
- although in natural languages the replacement of one unit by another is permissible in certain contexts, it is advisable to investigate the possibility of describing the language without distracting from that fact.

This differentiates between contextual use cases and non-contextual cases. Of particular interest is the study of situations where context is meaningful, but formally taken into account by context-free rules, that is, they are not considered as context (new categories are introduced into the grammar). Thus, in grammar G_1 context-dependent rule II.3 is eliminated (alg. 1).

Algorithm 3. Conversion of context-dependent grammar into context-free grammar.

Step 1. New characters $\tilde{S}'_{x,y,z}$ are introduced into the auxiliary dictionary to interpret non-noun groups, as opposed to characters $\tilde{S}_{x,y,z}$ that indicate arbitrary noun groups.

Step 2. Regulation II.3 is replaced by two new rules: $\tilde{S}_{x,y,z,w} \to S^{займ}_{x,y,w}$ and $\tilde{S}_{x,y,z,3} \to \tilde{S}'_{x,y,z}$.

Step 3. In rules II.1, II.2 and II.4, all occurrences of characters $\tilde{S}_{x,y,z,3}$ are replaced by characters $\tilde{S}'_{x,y,z,w}$. When you deploy an arbitrary namespace $\tilde{S}_{x,y,z,w}$ to a structure $A+S$ or $S+S_p$ to make sure that a personal pronoun such as *I, you, you, he, she, it,* that cannot have definitions (*A or* S_p: *новий я* or *ми посмішки*) appears in the design header. This is determined in different ways.

1. Personal pronouns are considered as personal class nouns - $S^{займ}$ and are treated as noun groups $(\tilde{S})$ alongside ordinary nouns. It is only allowed to move from a noun group $\tilde{S}$ to $S^{займ}$ the condition that it has not previously separated from itself to the left or right (rules II.1 and II.2), that is, if there is no adjective to the left of the symbol $\tilde{S}$ and no noun group to the right in the case. This condition is taken into account in Regulation II.3.
2. The pronoun is considered a special class of nouns, but along with the category arbitrary noun group $\tilde{S}$ enter the category of proper noun (non-noun) group $\tilde{S}'$, and the symbol $\tilde{S}$ during the output - before its deployment - must be replaced by either a symbol $S^{займ}$ (which can not expand further), or to the symbol $\tilde{S}'$ (which is expanded in the usual way); A and S_p only appear with $\tilde{S}'$, but $\tilde{S}'$ cannot become a pronoun.
3. Pronouns do not consider nouns and use a symbol *M* for them. Then, most grammar rules of G_1 are duplicated, for example, with rule I they introduce rule I': $R \to M_{од,наз,w}\tilde{V}_{од,тепер,w}$; along with Regulation III.3 - Regulation III.3 ': $\tilde{V}_{y,тепер,w} \to V_{y,тепер,w}M_{x',y',зн,w'}$ etc. The received grammar will be context-free.

The example above shows that in natural languages situations where context-dependent phenomena are described and as context-independent are possible, that is, in terms of context-free grammars. This complicates the description, introducing new categories and rules. Not every context-dependent is replaced by equivalent context-free grammar. It is known that there are direct component languages that are not context-free languages, such languages $a^n b^n a^n$ $(aba, aabbaa, \ldots)$ or $a^n b^n c^n$. Almost all examples of non-context-free languages are abstract in nature and have no interpretation in natural languages.

However, for any context-free grammar you can construct an equivalent binary context-free grammar. For example, the context-free grammar shown in Fig. 2, converted into binary by replacing rules III.1 and III.2 with the following new rules:

III. 1') $\tilde{V}_{y,тепер,w} \to \tilde{V}^1_{y,тепер,w}\tilde{S}_{x'',y'',ор,w'''}$; III. 1'') $\tilde{V}^1_{y,тепер,w} \to V_{y,тепер,w}\tilde{S}_{x',y',зн,w'}$;

III. 2') $\tilde{V}_{y,тепер,w} \to \tilde{V}^2_{y,тепер,w}\tilde{S}_{x'',y'',зн,w'''}$; III. 2'') $\tilde{V}^2_{y,тепер,w} \to V_{y,тепер,w}\tilde{S}_{x',y',ор,w'}$.

It is still necessary to replace rule I by rule I ': $R \to \tilde{S}_{x,y,н,w}\tilde{V}_{y,тепер,w}$; thus eliminating boundary characters (generally, in context-free grammar, boundary symbols are not formally required, whereas in grammars of direct constituents having context-dependent rules, boundary characters may be required as context (rule II.3 in G_1). There is a restriction (not more than two characters in the right-hand side of the rules) can be overlaid on arbitrary grammars of direct components, formulating it somewhat differently: each rule has the form of $Z_1CZ_2 \to Z_1WZ_2$, where W of one or two characters. For any grammar of direct constituents, one can construct an equivalent binary grammar of direct constituents (Fig. 4).

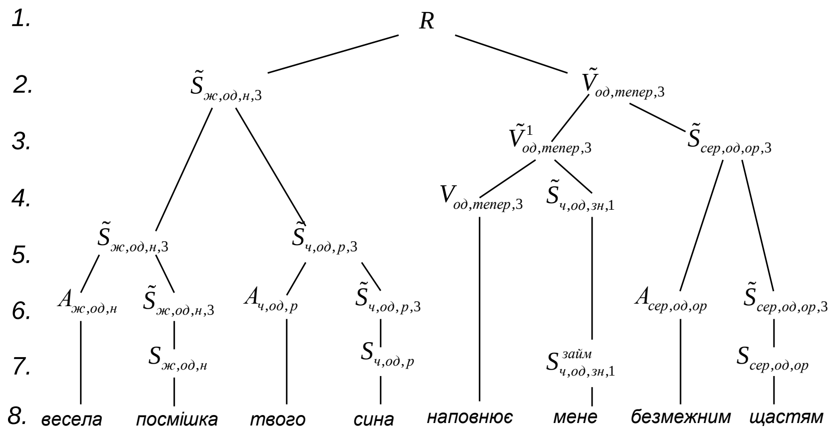

Fig. 4. Example of context-free grammar (type 2)

The peculiarity of regular grammars is a specific form of output. Let's build for example regular grammar G_2, that is, the generation of sentences of the type *Весела посмішка наповнює безмежним щастям* (simplified version of the sentence in Fig. 1). Diagram of grammar G_2.

1) $R \to \tilde{S}_{x,y,н,w}$

2) $S_{x,y,z} \to весела_{x,y,z}S_{x,y,z}$

3) $S_{x,y,z} \to безмежний_{x,y,z}S_{x,y,z}$

4) $S_{ж,y,н} \to посмішка_{ж,y,н}V_{y,3}$

5) $S_{сер,y,ор} \to щастя_{сер,y,ор}V_{y,3}$

6) $S_{ж,y,н} \to посмішка_{ж,y,н}$

7) $S_{сер,y,ор} \to щастя_{сер,y,ор}$

8) $V_{y,3} \to наповнити_{y,3}S_{x,y',ор}$

The given sentence will have the following interference in this grammar: there are

R

(1) $S_{ж,од,н}$

(2) *весела* $S_{ж,од,н}$

(4) *весела посмішка* $V_{од,3}$

(8) *весела посмішка наповнює* $S_{сер,од,зн}$

(3) *весела посмішка наповнює безмежним* $S_{сер,од,зн}$

(7) *весела посмішка наповнює безмежним щастям.*

Each intermediate chain contains exactly one auxiliary symbol in the last position. The sentence is generated from left to right: at each step, a specific wordform is displayed, followed by an auxiliary symbol indicating which construction should follow this wordform. A wordform is then issued that begins or is contained in this construction, followed by an auxiliary symbol of the next construction, and so on. Regular grammar implies that it follows a given word form, with the prediction depth being one adjacent symbol; each regular choice is entirely conditioned by one previous choice. It should be noted that it is impossible to derive a natural representation of the structure of the immediate components of that sentence (as it was done for context-dependent and context-free grammar) from the derivation of a sentence in regular grammar. That is, regular grammars give some structure to constituents, however, these constituents are usually purely formal in nature and are not amenable to natural interpretation (Fig. 5).

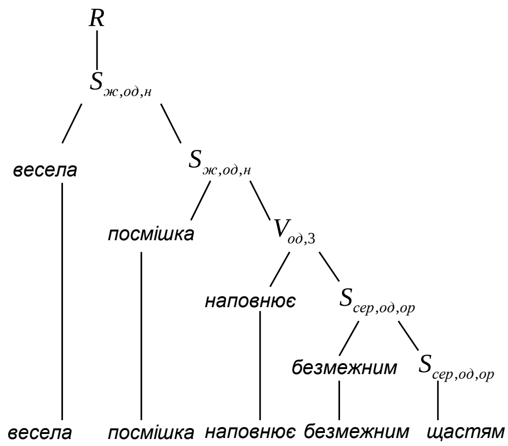

Fig. 5. Example of regular grammar (type 3)

Incorrectly split the sentence into two components - *весела*, etc., as well as attributing the categories to the resulting component. In the sentence *Посмішка наповнює мене щастям* the result would be even worse: the combination of *мене щастям* is a component. This sentence is not generated by grammar G_2, but it is easy to supplement it (by introducing two rules). The interpretation of output in regular grammars is pointless. They use a different interpretation of regular output: as a sequence of predictions and their implementations. There are context-free languages that are not generated by regular grammars. An example is the language that consists of view chains $a^n b^n$.

4 Conclusions

The article discusses known methods and approaches of addressing the automatic processing of textual content and highlights the shortcomings and benefits of existing approaches and results in the syntactic aspects of computational linguistics. The general conceptual principles of modeling of word-exchange processes in the formation of text arrays on the example of Ukrainian and German sentences were formed, then, by proposing syntactic models and word-classifications of the lexical composition of Ukrainian and German sentences, the lexicographic rules of syntactic type for automated processing were developed. The application of the technique allows to achieve higher reliability indicators in comparison with the known analogues, as well as demonstrates high efficiency in applied applications in the construction of new information technologies of lexicography and the study of the word-exchange effects of natural languages. The work has practical value because the proposed models and rules make it possible to effectively organize the process of creating lexicographic systems for processing syntactic textual content.

References

1. Angliyskaya grammatika v dostupnom izlozhenii. In: http://real-english.ru/crash/lesson3.htm, last accessed 2019/11/21.
2. Anisimov, A.V., Marchenko, O.O., Nykonenko, A.O.: Alhorytmichna model′ asotsiatyvno-semantychnoho kontekstnoho analizu tekstiv pryrodnoyu movoyu. In: Probl. Prohramuv, 2-3, 379-384. (2008)
3. Anisimov, A.V.: Komp'yuternaya lingvistika dlya vsekh: mify, algoritmy, yazyk. Dumka, Kyiv, Ukraine (1991)
4. Apresyan, Y.D.: Idei i metody sovremennoy strukturnoy lingvistiki. Prosveshcheniye, Moscow (1966)
5. Apresyan, Y.D.: Neposredstvenno sostavlyayushchikh metod. In: http://tapemark.narod.ru/les/332a.html, last accessed 2019/11/21.
6. Bahmut, A.Y.: Poryadok sliv. In: Ukrayins′ka mova: Entsykl, 3, 675-676. (2007)
7. Vysotska, V., Fernandes, V.B., Emmerich, M.: Web content support method in electronic business systems. In: CEUR Workshop Proceedings, Vol-2136, 20-41 (2018)
8. Lytvyn, V., Vysotska, V.: Designing architecture of electronic content commerce system. In: Computer Science and Information Technologies, Proc. of the X-th Int. Conf. CSIT’2015, 115-119 (2015)
9. Bil'gayeva, N.T.: Teoriya algoritmov, formal'nykh yazykov, grammatik i avtomatov. VSGTU, Ulan-Ude (2000)
10. Bolshakova, Y.I., Klyshinskiy, E.S., Lande, D.V., Noskov, A.A., Peskova, O.V., Yagunova, Ye.V.: Avtomaticheskaya obrabotka tekstov na yestestvennom yazyke i komp'yuternaya lingvistika. MIEM, Moscow (2011)
11. Lytvyn, V., Vysotska, V., Burov, Y., Veres, O., Rishnyak, I.: The Contextual Search Method Based on Domain Thesaurus. In: Advances in Intelligent Systems and Computing, 689, 310-319 (2018)
12. Su, J., Vysotska, V., Sachenko, A., Lytvyn, V., Burov, Y.: Information resources processing using linguistic analysis of textual content. In: Intelligent Data Acquisition and Advanced Computing Systems Technology and Applications, Romania, 573-578, (2017)
13. Lytvyn, V., Sharonova, N., Hamon, T., Vysotska, V., Grabar, N., Kowalska-Styczen, A.: Computational linguistics and intelligent systems. In: CEUR Workshop Proceedings, Vol-2136 (2018)
14. Lytvyn, V., Vysotska, V., Dosyn, D., Burov, Y.: Method for ontology content and structure optimization, provided by a weighted conceptual graph, Webology, 15(2), pp. 66-85 (2018)

15. Volkova, I.A., Rudenko ,T.V.: Formal'nyye grammatiki i yazyki. Elementy teorii translyatsii. In: Izdatel'skiy otdel fakul'teta vychisli-tel'noy matematiki i kibernetiki MGU im. M.V.Lomonosova, (1999)
16. Hakman, O.V.: Heneratyvno-transformatsiyna linhvistyka N. Khoms′koho yak vyrazhennya yoho linhvistychnoyi filosofiyi. In: Mul′tyversum. Filosofs′kyy al′ma-nakh, 45, 98-114. (2005)
17. Gerasimov, A.S.: Lektsii po teorii formal'nykh yazykov. In: http://gasteach.narod.ru/au/tfl/tfl01.pdf, last accessed 2019/11/21.
18. Gladkyy, A.V.: Sintaksicheskiye struktury yestestvennogo yazyka v avtomatizirovan-nykh sistemakh obshcheniya. Nauka, Moscow (1985)
19. Gladkyy, A.V. Mel'chuk, I. A.: Elementy matematicheskoy lingvistiki. In: Nauka, Moscow (1969)
20. Gladkyy, A.V.: Formal'nyye grammatiki i yazyki. In: Nauka, Moscow (1973)
21. Gross, M., Lanten, A.: Teoriya formal'nykh Grammatik. Mir, Moscow (1971)
22. Darchuk, N.P.: Komp'yuterna linhvistyka (avtomatychne opratsyuvannya tekstu). In: Kyyivs'kyy universytet VPTS, Kyiv, Ukraine (2008)
23. Demeshko, I.: Typolohiya morfonolohichnykh modeley u viddiyeslivnomu slovotvorenni suchasnoyi ukrayins'koyi movy. In: Zbirnyk naukovykh prats' Linhvistychni studiyi, 19, 162-167. (2009)
24. Zubkov, M.: Ukrayins'ka mova: Universal'nyy dovidnyk. In: VD Shkola, Kyiv, Ukraine (2004)
25. Gozhyj, A., Kalinina, I., Vysotska, V., Gozhyj, V.: The method of web-resources management under conditions of uncertainty based on fuzzy logic, 2018 IEEE 13th International Scientific and Technical Conference on Computer Sciences and Information Technologies, CSIT 2018 – Proceedings 1, 343-346 (2018)
26. Gozhyj, A., Vysotska, V., Yevseyeva, I., Kalinina, I., Gozhyj, V.: Web Resources Management Method Based on Intelligent Technologies, Advances in Intelligent Systems and Computing, 871, 206-221 (2019)
27. Vysotska, V., Chyrun, L.: Analysis features of information resources processing. In: Computer Science and Information Technologies, Proc. of the Int. Conf. CSIT, 124-128 (2015)
28. Lyubchenko, T.P.: Leksykohrafichni systemy hramatychnoho typu ta yikh zastosuvannya v za-sobakh avtomatyzovanoho opratsyuvannya movy. In: Avtoref. dys. kand. tekhn. nauk: spets. 10.02.21, Kyiv, Ukraine (2011)
29. Martynenko, B.K.: Yazyki i translyatsii: Ucheb. Posobiye. In: Izd. 2-ye, ispr. i dop. – SPb.: Izd-vo S.-Peterb. un-ta, (2008)
30. Marchenko, O.O.: Alhorytmy semantychnoho analizu pryrodnomovnykh tekstiv. In: Avtoref. dys. na zdobuttya nauk. stupenya kand. fiz.-mat. nauk: spets. 01.05.01. Kyiv, Ukraine (2005)
31. Noskov, S.A.: Samouchitel' nemetskogo yazyka. Deutsch für sie. Nauka, Kyiv, Ukraine (1999)
32. Partyko, Z.V.: Prykladna i komp'yuterna linhvistyka. In: Afisha, Lviv, Ukraine (2008)
33. Vysotska, V.: Linguistic Analysis of Textual Commercial Content for Information Resources Processing. In: Modern Problems of Radio Engineering, Telecommunications and Computer Science, TCSET'2016, 709–713 (2016)
34. Pentus, A.Y., Pentus, M.R.: Teoriya formal'nykh yazykov: Uchebnoye posobiye. In: Izd-vo TSPI pri mekhaniko-matematicheskom f-te MGU, Moscow (2004)
35. Popov, E.V.: Obshcheniye s EVM na yestestvennom yazyke. Nauka, Moscow (1982)
36. Postnikova, O.M.: Nimets′ka mova. Rozmovni temy: leksyka, teksty, dialohy, vpravy. T. 1, A.S.K, Kyiv, Ukraine (2001)

37. Postnikova, O.M.: Nimetsʹka mova. Rozmovni temy: leksyka, teksty, dialohy, vpravy. T. 2, A.S.K, Kyiv, Ukraine (2001)
38. Potapova, H.M.: Morfonolohiya viddiyeslivnoho slovotvorennya (na materiali slovotvirnykh hnizd z vershynamy - diyeslovamy ta viddiyeslivnykh slovotvirnykh zon). In: Dys. kand. nauk: 10.02.02, Ukraine (2008).
39. Rusachenko, N.P.: Morfonolohichni protsesy u slovozmini ta slovotvori staroukrayins'koyi movy druhoyi polovyny XVI – XVIII st. In: Avtoreferat dysertatsiyi na zdobuttya naukovoho stupenya kandydata filolohichnykh nauk, http://auteur.corneille-moliere.com/?p=history&m=corneille_moliere&l=rus, last accessed 2019/11/21.
40. Torosyan, O.M.: Funktsionalʹni kharakterystyky pryslivnykiv miry ta stupenya v suchasniy anhliysʹkiy movi. In: avtoref. dys. na zdobuttya nauk. stupenya kand. filol. na-uk, http://disser.com.ua/contents/6712.html, last accessed 2019/11/21.
41. Turysheva, O.O.: Porushennya ramkovoyi konstruktsiyi v suchasniy nimetsʹkiy movi: funktsionalʹnyy aspekt, normatyvnyy status. In: Avtoref. dys. kand. filol. nauk: spets. 10.02.04 (2012)
42. Ukrayins'kyy pravopys. In: In-t movoznavstva im. O.O. Potebni NAN Ukrayiny, In-t ukr. movy NAN Ukrayiny, Nauk. dumka, Kyiv, Ukraine (2007)
43. Vysotska, V., Hasko, R., Kuchkovskiy, V.: Process analysis in electronic content commerce system. In: Proceedings of the International Conference on Computer Sciences and Information Technologies, CSIT 2015, 120-123 (2015)
44. Fomichev, V.S.: Formal'nyye yazyki, grammatiki i avtomaty. In: http://www.proklondike.com/books/thproch/, last accessed 2019/11/21.
45. Chomsky, N.: O nekotornykh formal'nykh svoystvakh grammatik. In: Kiberneticheskiy sbornik, 5, 279-311. (1962)
46. Chomsky, N., Miller, G. A.: Formal'nyy analiz yestestvennykh yazykov. In: Kiberneticheskiy sbornik, 1, 231-290. (1965)
47. Chomsky, N.: Yazyk i myshleniye. In:Publikatsii OSiPL. Seriya monografiy, 2. (1972)
48. Chomsky, N.: Sintaksicheskiye struktury. In: Sbornik Novoye v lingvistike, 2, 412-527. (1962)
49. Chepurna, Z.V.: Transformatsiya poryadku sliv u prostomu rechenni pry perekladi z nimetsʹkoyi movy ukrayinsʹkoyu. In: Naukovi zapysky, 89(1), 232-236. (2010)
50. Sharov, S.A.: Sredstva komp'yuternogo predstavleniya lingvisticheskoy informatsii. In: http://www.ksu.ru/eng/science/ittc/vol000/002/, last accessed 2019/11/21.
51. Korobchinsky, M., Vysotska, V., Chyrun, L., Chyrun, L.: Peculiarities of Content Forming and Analysis in Internet Newspaper Covering Music News, In: Computer Science and Information Technologies, Proc. of the Int. Conf. CSIT, 52-57 (2017)
52. Kanishcheva, O., Vysotska, V., Chyrun, L., Gozhyj, A.: Method of Integration and Content Management of the Information Resources Network. In: Advances in Intelligent Systems and Computing, 689, Springer, 204-216 (2018)
53. Naum, O., Chyrun, L., Kanishcheva, O., Vysotska, V.: Intellectual System Design for Content Formation. In: Computer Science and Information Technologies, Proc. of the Int. Conf. CSIT, 131-138 (2017)
54. Vysotska, V., Lytvyn, V., Burov, Y., Gozhyj, A., Makara, S.: The consolidated information web-resource about pharmacy networks in city, CEUR Workshop Proceedings, 239-255 (2018)
55. Chyrun, L., Vysotska, V., Kis, I., Chyrun, L.: Content Analysis Method for Cut Formation of Human Psychological State, Proceedings of the 2018 IEEE 2nd International Conference on Data Stream Mining and Processing, DSMP 2018, 139-144 (2018)

56. Chomsky, N.: Three models for the description of language. In: I. R. E. Trans. PGIT 2, 113-124. (1956)
57. Chomsky, N.: On certain formal properties of grammars, Information and Control 2. In: A note on phrase structure grammars, Information and Control 2, 137-267, 393-395. (1959)
58. Chomsky, N.: On the notion «Rule of Grammar». In: Proc. Symp. Applied Math., 12. Amer. Math. Soc. (1961)
59. Chomsky, N.: Context-free grammars and pushdown storage. In: Quarterly Progress Reports, 65, Research Laboratory of Electronics, M. I. T. (1962)
60. Chomsky, N.: Formal properties of grammars. In: Handbook of Mathemati-Mathematical Psychology, 2, ch. 12, Wiley, 323-418. (1963)
61. Chomsky, N.: The logical basis for linguistic theory. In: Int. Cong. Linguists, (1962)
62. Chomsky, N., Miller, G. A.: Finite state languages. In: Information and Control 1, 91-112. (1958)
63. Chomsky, N., Miller, G. A.: Introduction to the formal analysis of natural languages. In: Handbook of Mathematical Psychology 2, ch. 12, Wiley, 269-322. (1963)
64. Chomsky, N., Schützenberger M.P.: The algebraic theory of context-free languages. In: Computer programming and formal systems, North-Holland, Amsterdam, 118–161 (1963)
65. English Verbs (Part 1) – Basic Terms. In: http://sites.google.com/site/englishgrammarguide/Home/english-verbs--part-1----basic-terms, last accessed 2019/11/21.
66. Vysotska, V., Chyrun, L., Chyrun, L.: Information Technology of Processing Information Resources in Electronic Content Commerce Systems. In: Computer Science and Information Technologies, CSIT'2016, 212-222 (2016)
67. Lytvyn, V., Vysotska, V., Uhryn, D., Hrendus, M., Naum, O.: Analysis of statistical methods for stable combinations determination of keywords identification. In: Eastern-European Journal of Enterprise Technologies, 2/2(92), 23-37 (2018)
68. Lytvyn, V., Vysotska, V., Pukach, P., Brodyak, O., Ugryn, D.: Development of a method for determining the keywords in the slavic language texts based on the technology of web mining. In: EasternEuropean Journal of Enterprise Technologies, 2(2-86), 4-12 (2017)
69. Su, J., Sachenko, A., Lytvyn, V., Vysotska, V., Dosyn, D.: Model of Touristic Information Resources Integration According to User Needs, 2018 IEEE 13th International Scientific and Technical Conference on Computer Sciences and Information Technologies, CSIT 2018 – Proceedings 2, 113-116 (2018)
70. Lytvyn, V., Vysotska, V., Dosyn, D., Lozynska, O., Oborska, O.: Methods of Building Intelligent Decision Support Systems Based on Adaptive Ontology. In: International Conference on Data Stream Mining and Processing, DSMP 2018, 145-150 (2018)
71. Burov, Y., Vysotska, V., Kravets, P. Ontological approach to plot analysis and modeling. CEUR Workshop Proceedings, Vol-2362, 22-31 (2019)
72. Chyrun, L., Kis, I., Vysotska, V., Chyrun, L.: Content monitoring method for cut formation of person psychological state in social scoring, 2018 IEEE 13th International Scientific and Technical Conference on Computer Sciences and Information Technologies, CSIT 2018, 106-112 (2018)
73. Lytvyn, V., Pukach, P., Bobyk, I., Vysotska, V.: The method of formation of the status of personality understanding based on the content analysis. In: Eastern-European Journal of Enterprise Technologies, 5/2(83), 4-12 (2016)
74. Lytvyn, V., Vysotska, V., Rzheuskyi, A.: Technology for the Psychological Portraits Formation of Social Networks Users for the IT Specialists Recruitment Based on Big Five, NLP and Big Data Analysis. In: CEUR Workshop Proceedings, Vol-2392, 147-171. (2019)