LEXICAL COMPONENT OF THE SECOND CONSTITUENT OF MODAL VERB CONSTRUCTIONS IN THE TEXTS OF SCIENTIFIC-TECHNICAL COMMUNICATION

M.V. Tsinova, PhD (Odessa)

The paper describes the verbs as second constituents of the modal constructions "modal verb / verbs to have, to be in modal meaning + passive infinitive", "modal verb / verbs to have, to be in modal meaning + active infinitive" that function in the text corpuses of scientific-technical communication sublanguages "Automotive", "Heat engineering" and "Electrical engineering" and dependence of their lexical and grammatical peculiarities upon quantitative values.

Keywords: active and passive forms, discourse, frequency of usage, scientific stratification layer of lexicon, text corpus.

Цинова М.В. Лексична складова другого конституенту модальних дієслівних конструкцій в текстах науково-технічної комунікації. Стаття описує дієслова-другі конституенти модальних дієслівних конструкцій "модальне дієслово / дієслова *to have, to be* у модальному значенні + інфінітив пасив", "модальне дієслово / дієслова *to have, to be* у модальному значенні + інфінітив актив", які функціонують у текстових корпусах підмов науково-технічної комунікації "Автомобілебудування", "Теплотехніка" та "Електротехніка", залежність їх лексичних та граматичних особливостей від кількісних значень.

Ключові слова: активна та пасивна форми, науковий дискурс, стратифікаційний шар лексики, текстовий корпус, частота використання.

Циновая М.В. Лексическая составляющая второго конституента модальных глагольных конструкций в текстах научно-технической коммуникации. Статья описывает глаголы-вторые конституенты модальных глагольных конструкций "модальный глагол / глаголы *to have, to be* в модальном значении + инфинитив пассив", "модальный глагол / глаголы *to have, to be* в модальном значении + инфинитив актив", которые функционируют в текстовых корпусах подъязыков научно-технической коммуникации "Автомобилестроение", "Теплотехника" и "Электротехника", зависимость их лексических и грамматических особенностей от количественных значений.

Ключевые слова: активная и пассивная форм, научный дискурс, стратификационный слой лексики, текстовый корпус, частота использования.

Linguistic literature presents sufficiently detailed description of various aspects of modal verb constructions. If the attention of researchers was focused on the general theory of style in the late XX century, when modal verbs were considered from the point of view of style identifying scientific discourse markers and methods of linguistic statistics were widely implemented [6; 7], then nowadays, not exactly modal verbs are analyzed firstly, but a consolidated study of modality is occurring, and, secondly, syntactic, semantic and pragmatic side of modality is developed from the perspective of comparative linguistics [9; 11–13; 16]. The only work available in the literature on modern linguistics, where a statistical analysis for description of the infinitive forms in the modal verb constructions was referred to, was one by A.O. Churyumova [8]. However, there were neither theses nor articles found based on text corpus of scientific and technical communication. That is why in the present article we would like to analyze modal verb constructions and their constituents, which function in the scientific functional style.

The material for this article is based on the text body of three sublanguages, being part of the scientific discourse and serving the following fundamental fields of science and technology – "Automotive", "Heat Engineering" and "Electrical Engineering". Their text bodies were formed on the basis of 300 thousand language units. The sources from which the information was extracted were the texts from scientific journals published in the United Kingdom and the United States: IEEE Transactions on Power Apparatus and Systems; Power Engineering; Power; Automotive News; Combustion; Control and Optimization; Machine Design; Machinery and Production Engineering; Automotive Engineer.

The object of the study was identified as units of so-called "small syntax" [5, p. 5] – modal verb constructions (MVC), and, in particular, modal verb constructions functioning in the texts of scientific and technical communication. Out of the bulk of all constructions with modal verbs, which were identified as a result of the continuous survey of text samplings in mentioned already subject areas, the subject of the analysis is considered to be the most frequently used ones – "modal verb / verbs to have, to be a modal value + infinitive passive" and "modal verb / verbs to have, to be a modal value + infinitive active", where the first constituent – the actual modal verb – is implemented in the following selected units: can / could, may / might, must, to be + inf, to have + inf.

In the previous descriptions of the research results devoted to the functioning of the modal verb constructions in scientific and technical discourse [4] all cases of MVC usage, belonging to already mentioned types and forms as well as aspectual and voice forms of notional verbs as the second constituents of MVC, were specified, counted and presented. This paper goes on with the description of the second constituents of modal verb constructions, but its goals are: 1) identification of the possible relationship between the formal and lexical features of notional verbs, included in MVC, and the role played by the quantitative values in the formation of such relationship; 2) a description of lexical connection between notional verbs and the first constituents.

To achieve the goals set it was necessary to perform the following tasks:
– to make a list of notional verbs included in the MVC and functioning in text samples with the greatest frequency and determine their grammatical forms;
– to select a group of notional verbs, combined with the maximum number of modal verbs;
– to distribute the allocated notional verbs due to the stratification layers.

In their research the authors used the following methods: method of "quantitative" mathematics for summing up, comparison and generalization of data; statistical method of rank correlation required for the bundle of the verbal lexemes due to the stratificational layers; expert evaluation, which involved into the research expert-professionals in the technical areas of knowledge, which make the material for the article, to confirm or refine the data obtained by the method of rank correlation.

40 most frequently used notional verbs as second constituents of MVC were sorted out for the purpose of research. They appeared to be common for the three fundamental areas of science and technology considered and were used in the text corpora with the absolute frequency of more than 10 (including verbs with lower frequency would make the results of the study unrepresentative). Then a list was compiled where they were represented in order of diminishing frequency.

Based on the objectives of the article, it was necessary to solve the problem of belonging of the verbs-second constituents of MVC to a particular startificational layer. As it has already been mentioned, two methods were used, i.e. a method of rank correlation and survey of specialists. Taking into consideration the probabilistic nature of a rank correlation method, the introduction of an expert evaluation gave an opportunity to get more correct results and avoid rough estimates. For the sake of rank comparison of allocated notional verbs, on the one hand probabilistic-statistical models of technical sublanguages "Automotive", "Heat Engineering" and "Electrical Engineering" were used, and within them the rank correlation between the verbal lexemes also took place (in this case the type of such

Spearman Rank Correlation formula was selected $r_s =$ 1 $-$ 6 $(\Sigma d^2 + T_a + T_b) / N(N^2 - 1)$, and, on the other hand – the frequency common literary dictionary by Thorndike and Lordge dictionary [15] (here we chose formula $r_s = 1 - 6 \Sigma d^2 / N(N^2 - 1)$. With a small ranks difference between units of frequency lists of technical text corpora and common literary dictionary, and, according to the experts, lack of terminological characteristic for contents of all four frequency dictionaries they were assigned to the layer of commonly used vocabulary (e.g. use, have, make, do, etc.). If the difference between the ranks of the units of frequency lists in "Automotive", "Heat Engineering", "Electrical Engineering" sublanguages and Thorndike and Lordge dictionary is significant enough, it is not so high between verbs of technical dictionaries, and, in addition, expert judgment is positive with the respect to the degree of the terminological colouring increase in units of Thorndike and Lordge dictionary, but not in the three fundamental areas of science and technology, these lexemes are defined as belonging to general scientific field (e.g. provide, determine, operate, etc.). In the process of compilation of the terminological vocabulary layer we only resorted to the experts' evaluation of notional verbs analysis, as those units were identified as belonging to the terminological area of science and technology with which the experts are familiar professionally.

Table 1 below gives quantitative indicators of realization of the notional verbs in the active (A) and passive (P) forms in text samples, and their belonging to the particular stratificational layer – commonly used (CU), general scientific (GS), terminological (T):

Voice forms in the list are distributed in the following way:

– passive forms of the verbs as second MVC constituents make up the majority – twenty one verbs are predominantly used in the passive form, and according to the calculation it is 61.56% of the total number of word-forms in the list), e.g. can be found, must be used;

– only seven verbs (provide, cause, take, meet, lead, require, include) make 15,5% of the list given in table

1 and are more likely to occur in the active form, e.g. could cause, may require;

– five verbs (achieve, design, consider, derive, remove) make 9,3% and are used only in the passive form, e.g. could be designed, may be considered;

– five verbs (operate, produce, increase, eliminate, select) make 8% and are equally or almost equally used in the active and passive forms as a part of the MVC, e.g. can be operated, have to select;

– two verbs (occur, have – 6,5%) have never been found in the passive form, e.g. may have, can occur. The fact that two last verbs do not have passive forms is connected with the specifics of their semantics; they can participate in the sentence when the agent is specified [14, p. 270].

The fact that the most part of the language units of notional verbs were found in the text samples in the passive voice can be explained by its content and task-orientation of statements, i.e. extralinguistic factors. However, as to content-informational plan of the texts, it implies the action which is fulfilled not only by a human, but also by a machine, a turbine or an electric grid, which are equipped with automatic control, that is why the share of notional verbs used in the active form, is significant enough.

According to lexical markers the list is composed of eighteen commonly used verbs and twenty general scientific ones and two terms.

Data of Table 1 allow us to trace the possible interdependence of grammatical and lexical features of the notional verbs. Lexemes of different stratificational layers, which are used in various voice forms, are studied in the work. The analysis shows:

– 311 commonly used notional verbs are used in the active voice and 201 verbs are used in the passive voice, which is expressed as a percentage ratio 60% : 40% respectively;

– in general scientific layer 116 verbs are used in the active voice and 249 ones in the passive (the percentage as 31% : 69% respectively);

– in terminological layer 8 verbs are used in the active voice and 25 in the passive (percentage ratio as 24% : 76% respectively).

Table 1

№	Main Verb, Stratification Layer	Active Forms F	Passive Forms F	Total Frequency F*
1	use, CU	15	87	102
2	make, CU	16	39	55
3	have, CU	37	-	37
4	provide, GS	23	9	32
5	cause, CU	25	3	28
6	see, CU	12	16	28
7	apply, CU	1	26	27
8	take, CU	16	11	27
9	determine, GS	5	21	26
10	find, CU	5	21	26
11	reduce, GS	10	13	23
12	obtain, GS	1	21	22
13	achieve, GS	-	21	21
14	expect, CU	6	15	21
15	meet, CU	14	6	20
16	occur, CU	20	-	20
17	design, GS	-	19	19
18	operate, GS	9	10	19
19	do, CU	7	11	18
20	draw, GS	3	15	18
21	consider, CU	-	17	17
22	control, T	2	15	17
23	produce, GS	8	9	17
24	calculate, GS	1	15	16
25	handle, GS	7	9	16
26	increase, GS	8	8	16
27	represent, GS	4	12	16
28	write, CU	2	14	16
29	lead, GS	12	2	14
30	show, CU	5	9	14
31	derive, GS	-	13	13
32	eliminate, GS	7	6	13
33	program(me), T	3	10	13
34	remove, GS	-	13	13
35	add, GS	2	10	12
36	generate, GS	5	7	12
37	require, CU	8	4	12
38	select, GS	6	2	12
39	build, CU	4	7	11
40	include, CU	8	3	11
	Total amount:	317	553	870

The following peculiar feature of the scientific texts is obvious: the form of the active voice prevails in less terminologically coloured lexics, but the higher the level of the terminological colouring is, the higher the proportion of the passive voice forms is. Thus, the results of this phase of the study show that there is a kind of relationship between the type of voice form and lexical content of notional verbs-second constituent of modal verb constructions.

If the previous stage of the study describes the interdependence between a notional verb and its lexical meaning, then this part will determine the possible lexical relationship of notional verbs-second constituents with the first constituents of MVC – can / could, may / might, must, to be + inf, to have + inf.

For this purpose 20 notional verbs that function in the texts of scientific communication with the maximum number of modal verbs were selected. The results of the analysis were summarized in Table 2.

Table 2

№	Main Verbs, used in Active (A) and Passive (P) Forms	Modal Verbs										Total Frequency F*
		can/could+		may/might+		must+		to be+		to have +		
		A	P	A	P	A	P	A	P	A	P	
1	use	5	62	4	22	1	1	4	1	1	1	102
2	make	8	21	4	2	2	7	-	5	2	4	55
3	have	12	-	18	-	5	-	1	-	1	-	37
4	provide	20	5	1	-	-	4	2	-	-	-	32
5	apply	1	17	-	6	-	1	-	-	-	2	27
6	take	6	3	5	2	2	6	-	-	3	-	27
7	determine	4	18	-	-	-	2	-	1	1	-	26
8	find	2	13	3	3	-	4	-	-	-	1	26
9	reduce	10	7	-	2	-	2	-	-	-	2	23
10	meet	6	2	-	-	7	1	-	2	1	1	20
11	design	-	9	-	1	-	8	-	-	-	1	19
12	operate	7	5	-	2	2	1	-	-	-	2	19
13	do	3	6	2	2	1	3	-	-	1	-	18
14	draw	3	12	-	2	-	-	-	-	-	1	18
15	consider	-	7	-	4	-	4	-	1	-	1	17
16	control	1	12	-	1	1	2	-	-	-	-	17
17	calculate	1	10	-	2	-	1	-	1	-	1	16
18	increase	3	5	4	1	1	1	-	-	-	1	16
19	remove	-	8	-	1	-	2	-	-	-	2	13
20	add	1	4	-	-	1	5	-	-	-	1	12

If we consider 20 selected verbs from the point of view of their lexical meanings, we can see that 9 out of them belong to the commonly used layer (use, make, have, apply, take, find, meet, do, consider), 10 verbs belong to the general scientific layer (provide, determine, reduce, design, draw, operate, calculate, increase, remove, add) and one verb – to terms (control).

The first units to be analyzed were the ones which belong to commonly used layer of the language. According to the results of contextual analysis the notional commonly used verbs can be found in the descriptions of the preliminary or preparatory actions and procedures (e.g. The same procedure can be applied to fault-analysis procedure as well … [1]). The

quantitative calculations within this stratification layer reveal the following:

– in their vast majority (53% of language units) they are combined with the most frequently used modal verbs can / could, thus confirming the ability or skill to perform an appropriate future action or procedure by the objects described in the texts of scientific communication;

– the same verbs which occur in the text packages in combination with the modal verbs may / might show a much lower activity (only 23% of all the language units are used), which can be explained by may / might semantics, as in the texts of scientific communication authors rarely describe the preparation for any action or procedure, expressing doubt and uncertainty about the possibility of this action performed. However, as it goes from the list, two verbs of this group 'use' and 'have' show considerable frequency in combination with this modal verb – 26 and 18, respectively. Here they are used in the meaning of "authorization". But, on the whole, the results of the calculations showed that the notional commonly used verbs join the modal verbs 'may / might' more than twice as little;

– with modal verbs 'must', 'to be + inf', 'to have + inf' commonly used verbs are used with the total absolute frequencies of 45 (14% of all the language units of the layer), 14 (which is 4% of all the language units of the layer), 19 (6% of all the language units of the layer) respectively. As we see from Table 2 verbs 'apply, take, find' are not combined with the modal verb 'to be + inf' at all, as the research of contexts of "Automotive", "Heat Engineering" and "Electrical Engineering" has shown that situationally there can not be any kind of a preliminary agreement for the fulfilment of such actions in the texts of scientific communication.

The second place is occupied by ten notional verbs-second MVC constituents included in the general scientific stratificational layer, such as "provide, determine, reduce, design, draw, operate, calculate, increase, remove, add". Though as to the amount of different words they outnumber one unit of the commonly used layer of lexics, but the total frequency of language units of the considered lexical layer is

almost twice as little, that is 330 and 188 respectively. It is believed that such distribution on stratification layers is characteristic of the text of scientific and technical communication, as noted by many researchers [1; 2; 6; 9]. They indicate the actions necessary for work, calculations, changes in the parameters and rules of operation of devices and other objects of technical activity (e.g. In this step an intermediate vector can be calculated … [1]). The results of calculation show:

– the same thing as to the previous units of a stratification layer happens to the verbs as second constituents of MVC in general scientific lexics, where they have the highest frequency of usage with a modal verb 'can / could' (132 units of the layer, i.e. 68% of all general scientific lexemes), thus expressing the objective possibility of the fulfilment of the action they specify;

– the total absolute frequencis and the proportions of notional verbs which go together with the modal verb 'may / might' are significantly different in this lexical layer (16 words, and respectively 8.5% of all the volume of lexemes of the layer). It means that if the units of commonly used lexical layer showed one hundred percent, though not very high frequency, compatibility with 'may / might', two verbs 'determine' and 'add' are not generally used with a modal verb 'may / might' at all, and the others have a very low frequency of usage (from 1 to 5);

– much higher figures are shown as to notional verbs which go with 'must' (30 words, and respectively 16%). This fact reveals the obligatory nature and necessity character of the actions denoted by these verbs in a general scientific layer, points at necessity of specification, calculation or changes of the objects parameters;

– the comparison of 'must', 'to be + in'f and 'to have + inf' gives the same picture as with the verbs of commonly used layer, but with smaller quantities, as the language units frequencies in general scientific layer are lower themselves. There are only two verbs 'provide' and 'calculate' to go with 'to be + inf', i.e. the efficiency of this modal verb is 1.5% (3 cases of occurrences); modal verb 'to have + inf' has a share of four times higher than that of 'to be + inf' – 6%

(11 occurrences) in the general bulk of ten general scientific verbs considered, as it participates in the construction of revealing the meaning of obligation which is characteristic of the texts in scientific and technical discourse.

The only verb which was included in the terminological layer – control – also has the highest numerical value when it goes with 'can / could' (13 language units, i.e. 76% terminological layer); completely irrelevant frequency with 'may / might' (1 language units, 6% lexemes of this layer) and the median frequency with 'must' (3 lexemes, 17% units of this layer). There is not any usage of the verb 'control' with 'to be+inf' and 'to have+inf', which shows that the authors of scientific works do not modify syntax constructions with modal verbs to convey the meaning of necessity with terminological units.

Statistical data obtained for the determination of a possible lexical relationship between the two parts of the MVC (a modal verb and a notional verb) with the transition from less terminological lexics to the terms showed a gradual and significant increase of the amount of notional verbs with the modal verb 'can / could' at the expense of suspension of functioning with 'may / might' and 'to be + inf', the terms even with 'have + inf'. This shows that the lexical component of notional verbs responds to the frequency of their usage with the first constituents of the MVC, which can be considered to be a definite trend in the texts of scientific communication.

The results of the study of three text corpora which belong to different fundamental areas of science and technology "Automotive", "Heat Engineering" and "Electrical Engineering" give us an opportunity to come to the following general conclusions about the features of the modal verb constructions functioning in these areas:

1) the statistical analysis determined that there is a particular relationship between the grammatical forms and lexical meanings of notional verbs-second constituents of MVC in a text corpora of "Automotive", "Heat Engineering " and "Electrical Engineering": the less terminological the verb is, the more it is used in the form of active voice. Thus in comparison with the other layers of vocabulary verbs-terms have the maximum number of units in the form of passive voice;

2) the attempt to find the lexical relationship between the first and second constituents of MVC has shown that this relationship can be traced quite clearly in calculating the usage of notional verbs as second constituents of MVC with modal verbs in different stratification layers. This is evident in the increasing of quantity of notional verbs usages with 'can / could' at the expense (at the same time) of diminishing the number of verbs with 'may / might', 'to be + inf' and 'have to + inf', which depends on the degree of terminological colouring of a language unit.

Further research is planned to describe the same small units of syntax, but with other constituents – adjectives functioning in the constructions such as "modal verb + be + adjective", their lexical and grammatical features.

REFERENCES

1. Alekseev P.M. Some problems in the theory and practice of statistical lexicography / P.M. Alekseev // Text Statics. – Minsk, 1969. – Volume 1. – P. 12–37. 2 . Alekseev P.M. Frequency English-Russian dictionary of newspaper lexics / P.M.Alekseev, L.A.Turygina. – M . : Voenizdat, 1974. – 260 p. 3. Belyaeva Z.F. Comparative analysis of the modal verb combinations in the American scientific and technical literature : diss. ... Candidate of Philological Science : 10.02.04 / Z.F. Belyaeva. – Omsk, 1973. – 191 p. 4. Borisenko T.I. The Peculiarities of model verb constructions functioning in technical sublanguages / T.I. Borisenko, M.V. Koshuba, E.V. Mardarenko, M.V. Tsinovaya // Notes on Romano-Germanic Philology. – Odesa: ONU after I.I.Mechnikov. – № 1(32). – 2014. – P. 25–34. 5. Budagov R.A.To what extent "text linguistics" is linguistics? / R.A. Budagov // NDVSh. Philology. – 1979. – № 2. – P. 13–19. 6. Budkova S.S. Lexicographic description of English terminology of radiation and plasma technologies [Text]: diss. on scientific. Ph.D. degree. Philology. Sciences: special. 10.02.04 "Germanic languages" / S.S. Budkova. – Omsk, 2012. – 24 p. 7. Buntina T.A. Relationship of lexical-semantic and grammatical characteristics of the verb in functional style : dis. on scientific. Ph.D. degree. Philology. Sciences : special. 10.02.04 "Germanic languages" / T.A. Buntina. – K., 1979. – 26 p. 8. Churyumova A.O.

Modal constructions with the ptogressive infinitive in modern English [Electronic resource] / A.O. Churyumova // Humanities research. – 2013. – № 4. – Available at : http://human.snauka.ru/2013/04/2653 (date accessed: 07/31/2014) 9. Grebneva O.V. Modal necessity in legislative texts of the European Union: dis. on scientific. Ph.D. degree. Philology. Sciences : special. 10.02.20 "Comparative-historical, typological and comparative linguistics" / O.V. Grebneva. – Kazan, 2004. – 20 p. 10. Dyachenko G.F. The semantics of the verb in English technical sublanguages : diss. Candidate. Philology. Sciences. : 10.02.04 / G.F. Dyachenko. – Odessa, 1980. – 198 p. 11. Olenchuk O.G. Language means of problematic modality representation : dis. on scientific Ph.D. degree. Philology. Sciences: special. 10.02.19 "Theory of Language" / O.G. Olenchuk. – Izhevsk, 2012. – 22 p. 12. Muhometzyanova Y. Functional-semantic aspect of statements with probability meaning : dis. on scientific. Ph.D. degree. Philology. Sciences : special. 10.02.19 "Theory of Language" / Yu.V. Muhometzyanova. – Tver, 2007. – 20 p. 13. Samoilova Yu.A. Double and triple combinations of modal verbs in the language of Lowland Scotland (Scotts) / Yu.A. Samoylova // News RSPU. AI Herzen. – 2008. – Issue № 73-1. – P. 400–404. 14. Smirnitsky A.I. Morphology of English / A.I. Smirnitsky. – M. : Foreign. languages publishing house??, 1959. – 439 p. 15. Thorndike E., Lordge J. The Teacher's Word Book of 30 000 words / E. Thorndike, J. Lordge. – New York : Columbia University Teacher's college press, 1968. – 274 p. 16. Umatova M.B. Comparative analysis of modal verbs in English and Uzbek [Text] / M.B. Umatova // Young scientist. – 2014. – № 7. – P. 641–642.

LIST OF SOURCES

1. Robert A.M. van Amerongen A Rank-Oriented Setup for the Compensation Algorithm / Robert A.M. van Amerongen. – EEE Transactions on Power System, 1990. – Vol. 5. – N 1. – P. 283. 2. Glielmo L. Automotive: Control Systems in Millions of Copies / L. Glielmo, Ken Butts, Carlos Canudas-de-Wit, Ilya Kolmanovsky. – Transactions on Automatic Control [Electronic Resource]. – Available at: IoCT-Part1-03Automotive Control.pdf