

## БИНАРЛЫҚ ТЕРМИНДЕРДІҢ ФОРМАЛДЫҚ-ҚҰРЫЛЫМДЫҚ МОДЕЛЬДЕРІ

\*Абдикарим Н.<sup>1</sup>, Ысқақ Б.Д.<sup>2</sup>, Махсуд О.Ж.<sup>3</sup>

<sup>1</sup>ф.ғ.к, Ш. Шаяхметов атындағы «Тіл-Қазына» ұлттық ғылыми-практикалық орталығының жетекші ғылыми қызметкері, Астана, Қазақстан  
e-mail: [nursana@yandex.ru](mailto:nursana@yandex.ru)

<sup>2</sup>ф.ғ.к, Ш. Шаяхметов атындағы «Тіл-Қазына» ұлттық ғылыми-практикалық орталығының Терминология басқармасының басшысы, Астана, Қазақстан  
e-mail: [bakytgul.iskakova@mail.ru](mailto:bakytgul.iskakova@mail.ru)

<sup>3</sup>п.ғ.м., Ш. Шаяхметов атындағы «Тіл-Қазына» ұлттық ғылыми-практикалық орталығының ғылыми қызметкері, Астана, Қазақстан  
e-mail: [orazgul-astana@mail.ru](mailto:orazgul-astana@mail.ru)

**Аңдатпа.** Мақалада IT-технологиялардың, компьютерлік лингвистиканың мүмкіндіктерін терминологияда пайдалану үшін тілдік материалды құрылымы бойынша компьютер түсінегіндей етіп формалдаудың қажеттілігі сөз болады.

Мақаланың мақсаты – бинарлық терминдік тіркестердің сөз табы, синтаксистік тіркестерге ұйысудағы морфологиялық тұлғаларын есепке алып, олардың формалдық-құрылымдық модельдерін анықтау.

Қазақ тілін формалдаудың тарихына қысқаша шолу жасалады. Терминдік тіркестер лексика-синтаксистік категория болғанымен, сөздердің өзара қатынасқа түсе отырып, белгілі бір ұғымның атауы болуына морфология да елеулі қызмет атқаратындығы; оның тілдік бірліктерді формалдық-құрылымдық жағынан модельдеу барысында айқын көрініс табатындығы баяндалады.

Зерттеменің негізгі нәтижесі мен талдаулары ретінде мыналарды атауға болады: «Жасанды интеллект» саласының екікомпонентті терминдік тіркестерінің морфосинтаксистік құрылымы есепке алынып, оларды формалдауға қажетті шартты белгілер таңдалған; аталған саланың барлық терминдері EXCEL кестесіне енгізіліп, оларға сәйкес келетін модельдермен таңбаланған; терминдік тіркестерді парадигмалық-диструбивтік әдіспен талдау барысында олардың өнімді модельдері іріктеліп, типтендірілген және негізгі грамматикалық сипаттары көрсетіліп, мысалдар келтірілген.

Бір ғана саланың терминдері зерттеу материалы ретінде алынуы сөз таптарының термин жасаудағы белсенділігін, тіркесімділік қабілетін, сипатталатын терминдердің лексикасының қамтылу толықтығын айқындауға және т.б. нақты мәліметтерге қол жеткізуге мүмкіндік берген.

Авторлардың пікірінше, осы зерттеменің негізінде басқа да салалардың терминдік тіркестерін модельдеуге, жиі қолданылатын типтерін анықтауға және оларды арнайы мәтіндерден автоматты түрде танитын лексика-синтаксистік шаблон әзірлеуге болады.

Мақаланың практикалық маңыздылығы терминдік тіркестердің парадигмалық түрленуін талдауға, жүйелеуге және қазақ тілінің заңдылығына сай келмейтін терминдік тіркестерді анықтауға септігін тигізеді. Сонымен қатар зерттеу барысында мәлім болған фактілерге сүйене отырып, терминдер орфографиясы бойынша кейбір даулы мәселелерді шешуге және шеттілдік терминдердің аударылуына оң ықпал етуге болатындығы нақты мысалдар арқылы көрсетілген.

**Тірек сөздер:** термин, бинарлық терминдер, корпустық лингвистика, тег, Лейпциг ережесі, морфосинтаксистік модельдер, шартты белгілер, терминдерді автоматты тану, лексика-синтаксистік шаблон

### Негізгі ережелер

Қазіргі заманауи ІТ технологиялар мүмкіндіктерін терминология саласында мәліметтер базасынан ақпарат іздестіруге қажетті тіркестерді жүйелеу, сөздіктер мен тезаурустар әзірлеу, терминдерді концептуалдық талдау, автоматты түрде референция жасау, терминдер онтологиясын кеңейту, тілдік корпус ресурстарын пайдалану және т.б. мақсатында пайдалануға болады. Сонымен қатар құрылымдалған терминдердің, нақтырақ айтқанда, терминдік тіркестердің типтік модельдерін анықтаудың да терминология үшін өзіндік маңызы бар. Мәселе бірнеше сөздің тіркесуі (кейбір мәліметтерге сүйенсек, қазақ тілінде терминдік тіркестің құрамындағы сөз 7-ге дейін жетеді) арқылы жасалған күрделі терминдік тіркестер, соның ішінде екіқұрамды терминдік тіркестер жайында болмақ. Жалпы терминдік тіркестердің қазақ тіл білімінде ғана емес, әлемдік лингвистикада табиғаты толық ашылып, лингвистикалық мәртебесі айқындалмаған, терминдік атауы әлі де болса тұрақталмаған. Мысалы, қазақ терминологтары мұндай құрылымдарды *күрделі терминдер* (Ә. Қайдар) *сөз тіркесі түріндегі терминдер* (Ө. Айтбаев), *тіркесті терминдер*, *терминологиялық сөз тіркесі* (Ш. Құрманбайұлы), *терминологиялық тіркес* (С. Әлісжанов), *сөз тіркесі арқылы жасалған терминдер* (Ә. Әміров) және т.б. атап жүр.

Терминдік тіркестер еркін сөз тіркестерінің үлгісіне тәуелді болғанымен, грамматикалық емес, белгілі бір ғылым саласының тұжырымдалған ұғымының формуласы іспеттес, тұтасқан күйінде бір ғана лексикалық бірлік болып табылатын функционалдық қызметтегі *лексика-синтаксистік категория* екені белгілі. Оларды ғылыми мәтіндер мен ғылыми дискурста қайталанатын әрі туындатуға болатын құрылымдалған жүйе тұрғысынан қарастыруға болады. Мұндай жағдайда терминдік тіркестердің синтаксистік қана емес, морфологиялық белгілерінің де есепке алынуы шарт. Соңғысы сөз таптарының топтастырылуымен тығыз байланысты әрі сөздің жалпы категориалды (лексикалық және грамматикалық мағыналардың бірлесуінен пайда болған) мағынасы, синтаксистік белгілері бойынша ажыратылады. Ал синтаксистік белгісі бойынша терминдердің сөйлем мүшесі болу қызметіне емес, олардың тіркесу қабілетіне баса назар аударылуы керек. Сөз тіркестері көбінесе морфологиялық тұлғалар арқылы жүзеге асырылатындықтан, терминдік тіркестерді морфосинтаксистік тұрғыдан қарастыруға болады. Бұл синтаксистік тәсілмен жасалған терминдік тіркестердің құрамындағы сыңарлардың жалпы категориалды мағынасын, сөз таптарымен байланысын, оны қамтамасыз ететін морфологиялық тұлғалар т.б. элементтерді есепке алып түгендеуді және оларды құрайтын элементтерді шартты белгілермен тегтеуді қажет етеді.

Синтаксистік тұрақты тіркестер ретіндегі терминдік тіркестер корпустық лингвистикада «коллакация» деген терминмен аталады. Лексикалық

бірліктердің тіркесімділігін анықтаудың ең қарапайым әдісі – статистикалық зерттеулер нәтижесі бойынша сөз тіркестерінің жиілік тізімін құру болса, қазіргі кезде коллакацияларды статистикалық аппараттың көмегімен танытын жалпыға қолжетімді сервистер де бар: мәтінді автоматты өңдеу сайтынан биграммды іздеу сервисі; корпус-менеджер CQP; Sketch Engine жүйесі және т.б. Мәліметтерді анықтаудың мұндай әдісімен қатар лексика-синтаксистік шаблондар арқылы да іздеп табатын әдіс қолданылады. «Лексико-синтаксический шаблон определяется как «структурный образец языковой конструкции, который отображает ее лексические и поверхностно-синтаксические свойства» [6] Басқаша айтқанда, лексика-синтаксистік шаблон дегеніміз тіркесімдегі көптеген лексемалардың маңызды грамматикалық сипатын көрсететін тілдік конструкциялардың моделі немесе құрылымдық үлгісі. Осындай үлгілерді алдын ала әзірлеп, терминдік тіркестердің мүмкін болатын модельдерін анықтау, оны компьютер «оқи алатындай» болу үшін арнайы шартты белгілер қажет болады. Алғашқы болып қазақ тілінің деректерін формалдау, статистикалық өңдеумен айналысқан ғалымдарымыздың еңбектерінде мынадай шартты белгілер қолданылған: З – зат есім, Г – етістік, С – сын есім, У – үстеу, А – сан есім, Е – есімдік, О – одағай, К – көмекші есімдер, Л – еліктеуіш, Д – модаль сөз. [2, 22 б.].

Типологиялық лингвистиканың және компаративистиканың дамуының нәтижесінде тілші ғалымдарға кез келген тілдің материалын «оқып, тануға» мүмкіндік беретін, тіл фактілерін грамматикалық категориялары бойынша қысқартып тегтеуге арналған халықаралық Лейпциг ережесі (барлығы 11 ереже) немесе морфологиялық глостәу [7] тәртібі әзірленгені белгілі. Корпустық лингвистика да осы Ережені басшылыққа алады. Аталған Ереже тілдік бірліктерді сипаттау стандартын ғана белгілейтіндіктен, оны әр тіл өзінің заңдылықтарына ыңғайластырып, оған барынша жуықтастырып жасақтауға тырысады. Мысалы, орыс тілінің терминдік тіркестерінің лексика-синтаксистік шаблондарында септік тұлғалары сандармен ( $N_1$  (именительный падеж),  $N_2$  (родительный падеж), ... т.б.) белгіленген. Қазақ лингвистикасында арнайы бекітілген шартты белгілер жүйесі жоқ болғандықтан, біз төл тіліміздің Ұлттық корпустарында, түркі тілдерінің корпустарында қолданылып жүрген және жоғарыда аталған Ережеге сүйене отырып, терминдік тіркестердің құрылымын сипаттап шықтық. Шартты белгілермен таңбалау барысында грамматикалық тұлғаларды барынша айқын көрсету үшін біршама нақтылықтар енгіздік: сөз таптары бас әріптермен, ал грамматикалық көрсеткіштер кіші әріптермен қысқартылып, латын қаріптерімен белгіленді.

### **Кіріспе**

Тілдік бірліктерді құрылымына қарай талдау формалдық грамматика мектебі өкілдерінің еңбектерінен бастау алып, америкалық дескриптивтер тілдік элементтерді *дистрибутивтік талдау және тікелей құрастырушылары бойынша талдау, синтаксистік трансформация* әдістерін қалыптастырды; бұл үрдіс арқылы машинаның адамды түсінуін қамтамасыз ететін қолданбалы лингвистика мәселелерінің шешімі табылғаны белгілі.

Тікелей құрастырушылары бойынша талдауда тілдік жүйенің бірліктері бір-бірімен сан түрлі әрі күрделі қатынастармен байланысқан деген уәж басшылыққа алынады. Ал трансформациялық әдіс бойынша синтаксистік жүйелердің бірі түпкі ядролық, басқасы туынды бірнеше кіші жүйелерге бөлінеді, кез келген синтаксистік тип белгілі бір ядролық типтердің түрленген комбинациясы болып табылады. Алдыңғы екі әдіс тілдік материалдарды бақылау арқылы тілдік модельдерді анықтау, олардың типтері, жасалу жолдарын түсіндіруді мақсат етті. Ал барлық тілдерге ортақ модель ұсынуды көздеген туындатушы грамматика өкілдерінің пікірінше: «... грамматика дегеніміз белгілі бір тілдегі сөйлемнің дұрыс («белгіленім жасалған») сөйлемдерін туындататын өзіндік механизм. Сол тілді тұтынушылар лингвистер жасаған туындатушы модельдердің дұрыстығын тексеріп отырады» [1, 161 б.]. Туындатушы грамматика өкілдері берілген тілдік жүйенің элементтерінен туындау, орналасу және синтезделу картинасын көруді көздейді. Осы әдістердің қай қайсысы да табиғи тілді автоматты өңдеу, машиналық аударма жасау, ауызша сөйлеу тілін тану, ақпараттық іздеу және т.б. көптеген прагматикалық сұраныстарды қанағаттандыруға пайдасын тигізді

Қазақ тіл білімінде тілдік материалдарды математикалық әдіспен өңдеу «... был произведен Х.К. Жубановым вместе с сотрудниками программно-методического сектора Наркомпроса республики в период создания нового казахского латинизированного алфавита и выработки орфографических правил. Ученый сознательно придерживался принципа количественного подхода к исследованию порядка построения синтаксических величин, позиционных отношений слов ... и т.д.» [2, 13 б.] Одан кейін сандық-статистикалық әдісті пайдаланған ғалымдарымыз қазақ тілі буындарының типтерін (С. Кеңесбаев), ең көп жұмсалатын синтаксистік қатынастарды (М. Балақаев), қазақ фонетикасы мен фонологиясын (Ж. Аралбаев) осы бағытта зерттеп, айтулы нәтижелерге қол жеткізді. 1980 жылдары ҚазССР ҒА Тіл білімі институтында «Автоматизация лингвистических работ и статистика речи» деп аталатын топ құрылып, электрондық есептеуіш машинаның (ЭВМ) көмегімен құман-қазақ тілдерінің, Абай тілінің, М. Әуезовтің 20 томдық шығармалар жинағының, ертегі мәтінінің, кейбір оқулықтардың және т.б. стильдік мәтіндердің жиілік сөздігін жарыққа шығарды. Мұндай сөздіктер негізінен қазақ тілін үйренушілер үшін лексикалық минимум белгілеуге пайдаланылды. Сол кездегі еңбектерде құрылымдық лингвистиканың негізгі нысанасына синхрондық тұрғыдан тілдің ішкі бірлігін зерттеу, лингвистикалық талдауларды формалдау, оның құрылымдық қабаттары (стратификация) мен тілдік деңгей ұғымдары т.б. қамтылған. Зерттеушілер структурализмнің түрлі бағыттарының әдіснамалық негіздерін сыни тұрғыдан талдай келе, оның әдістерінде тілдік бірліктерді нақты және дәлдікпен сипаттау, тілдік фактілерді жүйелеу тәрізді көптеген жаңашылдық пен құндылық бар екендігін атап өтеді [3, 5]. Демек қазақ ғалымдары есептеуіш техника мүмкіндігін пайдаланып, формалды-құрылымдық зерттеулер жүргізуден қалыс қалмады дегуге болады. Оның маңыздылығы жайында А.Жұбанов былай дейді: «...

формалды грамматика қазақ тілін терең игеруге бағыттталып, жалпы филологиялық дәстүрлерге негізделіп, оның мүмкін болатын барлық формалары туралы ақпараттар есепке алынған жағдайда, қазақ сөздері туралы энциклопедиялық ақпараттар мен нақты лингвистикалық (лексика-семантикалық, сөзжасамдық, тұлғаландырушы және т.б.) модельдерін құруға болады» [4, 51 б.]. Зерттеуіміздің нысанасына ілініп отырған терминдер тіркесіндегі «... бағыныңқы сыңарлар толықтауыштық та, анықтауыштық та мәнге ие емес, тек қана негізгі тірек сөздің мағлұматын толықтырып, ақпараттық жағынан айқындай түсуге қызмет етеді [5, 19 б.]. Осы жағынан алғанда, терминдік тіркестерді дәстүрлі құрылымдық синтаксис тұрғысынан емес, формалдық грамматика тұрғысынан зерделеу маңызды. Қазіргі IT-технологиялар, компьютер лингвистикасы, оның бір саласы ретіндегі корпуслық лингвистиканың дамуы бұрынғыдан да үлкен мүмкіндіктерге жол ашты. Әлем лингвистикасында терминдік тіркестерді мәтіннен іздеп табу, анықтауды теориялық және практикалық аспектіде қарастыратын көптеген еңбектер жарияланып, зерттеу нәтижелері қолданысқа енгізіліп жатыр. Егер біз қазақ тіл білімінің де заманауи тенденциялар көшінен қалмауын қадағалап, үлесімізді қосқымыз келсе, мәтінді автоматты тануды жүзеге асыруға септігін тигізетін амал-тәсілдерін, жолдарын қарастыруымыз керек.

### **Материалдар мен әдістердің сипаттамасы**

ҚР Үкіметінің жанындағы Республикалық терминология комиссиясының 2021 жылғы 30 сәуірдегі I отырысында және аталған жылдың 3 қыркүйегіндегі II отырысында Назарбаев университетінің ұсынысымен бекітілген жасанды интеллект саласының 933 термині зерттеу материалы ретінде пайдаланылды. Мақалада көтерілген мәселені жан-жақты ашу үшін осы бағыттағы қазақ тіл білімінің зерттеулеріне шолу жасалып, анализ, синтез әдістері пайдаланылды; терминдік тіркестердің типтік модельдерін жүйелеу үшін топтастыру, ал қазақ тілінің табиғатына жат кейбір тіркестердің дұрыстығына көз жеткізу үшін орыс тілімен салғастыру әдісі; терминдік тіркестер жасалатын сөз табы мен грамматикалық формаларды есепке алу мақсатында статистикалық және деривациялық талдау; оларды формалдық-құрылымдық жағынан модельдеу, формалдау, тікелей құрастырушыларға бөлу және ұқсас тіркестерді пайдаланып, ұсыныстар беру үшін парадигмалық-дистрибутивтік талдау әдістері қолданылды.

### **Нәтижелер**

Нысанаға ілінген саланың барлық терминдері шартты белгілермен таңбаланып, EXCEL кестесіне енгізілді және олардың сандық мөлшері анықталды. Мысалы, төменде екі зат есімнен құралған, бағыныңқысы *-лық* типіндегі жұрнақпен тұлғаланған,  $Natr+N$  моделіне (қатыстық сын есімнің зат есіммен қабыса байланысуы) сәйкес келетін атрибутивтік қатынастағы терминдік тіркестердің жалпы саны 71 екендігін EXCEL кестесі көрсетті.

№	Термин	Саны	Белгілеу
1	аскестік ажыратыны	1	Natr+N
2	аскестік сөздері	1	Natr+N
3	саясилық ақпарат	1	Natr+N
4	саясилық интерпретация	1	Natr+N
5	саясилық регрессия	1	Natr+N
6	психологиялық функция	1	Natr+N
7	векторлық объективтілік	1	Natr+N
8	тригонометриялық жұп	1	Natr+N
9	кадрлық бағдар	1	Natr+N
10	матрицалық диагональ	1	Natr+N
11	геометриялық композиция	1	Natr+N
12	аппараттық элемент	1	Natr+N
13	функционалдық оператор	1	Natr+N
14	аппараттық құралдар	1	Natr+N
15	гармоникалық жетек	1	Natr+N
16	гармоникалық жетек	1	Natr+N
17	интерпретациялық сөз	1	Natr+N
18	кадрлық функция / көрсеткіш	1	Natr+N
19	кадрлық тереңдік	1	Natr+N
20	матрицалық диагональ	1	Natr+N
21	доғалық элемент	1	Natr+N
22	доғалық элемент	1	Natr+N
23	доғалық элемент	1	Natr+N
24	доғалық элемент	1	Natr+N
25	доғалық элемент	1	Natr+N
26	доғалық элемент	1	Natr+N
27	доғалық элемент	1	Natr+N
28	доғалық элемент	1	Natr+N
29	доғалық элемент	1	Natr+N
30	доғалық элемент	1	Natr+N
31	доғалық элемент	1	Natr+N
32	доғалық элемент	1	Natr+N
33	доғалық элемент	1	Natr+N
34	доғалық элемент	1	Natr+N
35	доғалық элемент	1	Natr+N
36	доғалық элемент	1	Natr+N
37	доғалық элемент	1	Natr+N
38	доғалық элемент	1	Natr+N
39	доғалық элемент	1	Natr+N
40	доғалық элемент	1	Natr+N
41	доғалық элемент	1	Natr+N
42	доғалық элемент	1	Natr+N
43	доғалық элемент	1	Natr+N
44	доғалық элемент	1	Natr+N
45	доғалық элемент	1	Natr+N
46	доғалық элемент	1	Natr+N
47	доғалық элемент	1	Natr+N
48	доғалық элемент	1	Natr+N
49	доғалық элемент	1	Natr+N
50	доғалық элемент	1	Natr+N
51	доғалық элемент	1	Natr+N
52	доғалық элемент	1	Natr+N
53	доғалық элемент	1	Natr+N
54	доғалық элемент	1	Natr+N
55	доғалық элемент	1	Natr+N
56	доғалық элемент	1	Natr+N
57	доғалық элемент	1	Natr+N
58	доғалық элемент	1	Natr+N
59	доғалық элемент	1	Natr+N
60	доғалық элемент	1	Natr+N
61	доғалық элемент	1	Natr+N
62	доғалық элемент	1	Natr+N
63	доғалық элемент	1	Natr+N
64	доғалық элемент	1	Natr+N
65	доғалық элемент	1	Natr+N
66	доғалық элемент	1	Natr+N
67	доғалық элемент	1	Natr+N
68	доғалық элемент	1	Natr+N
69	доғалық элемент	1	Natr+N
70	доғалық элемент	1	Natr+N
71	доғалық элемент	1	Natr+N
72	доғалық элемент	1	Natr+N
73	доғалық элемент	1	Natr+N
74	доғалық элемент	1	Natr+N
75	доғалық элемент	1	Natr+N
76	доғалық элемент	1	Natr+N
77	доғалық элемент	1	Natr+N
78	доғалық элемент	1	Natr+N
79	доғалық элемент	1	Natr+N
80	доғалық элемент	1	Natr+N
81	доғалық элемент	1	Natr+N
82	доғалық элемент	1	Natr+N
83	доғалық элемент	1	Natr+N
84	доғалық элемент	1	Natr+N
85	доғалық элемент	1	Natr+N
86	доғалық элемент	1	Natr+N
87	доғалық элемент	1	Natr+N
88	доғалық элемент	1	Natr+N
89	доғалық элемент	1	Natr+N
90	доғалық элемент	1	Natr+N
91	доғалық элемент	1	Natr+N
92	доғалық элемент	1	Natr+N
93	доғалық элемент	1	Natr+N
94	доғалық элемент	1	Natr+N
95	доғалық элемент	1	Natr+N
96	доғалық элемент	1	Natr+N
97	доғалық элемент	1	Natr+N
98	доғалық элемент	1	Natr+N
99	доғалық элемент	1	Natr+N
100	доғалық элемент	1	Natr+N

Сурет 1 – Natr+N типіндегі бинарлық терминдік тіркестердің Excel кестесіндегі көрінісі

Терминдердің формальдық-құрылымдық сипатын анықтауға негіз болып отырған жасанды интеллект саласының 933 терминінің 353 термині бір сөзден жасалған жалаң термин болса, 335 термин екі сыңарлы, 224 термин үш сыңарлы, 16 термин төрт сыңарлы, ал 5 термин бес сыңарлы екендігі анықталды. Яғни жалаң терминдерге қарағанда, толық мағыналы 2-5 сөзге дейін тіркескен терминдер көп (353<580).

Бір сөзден жасалған жалаң терминдердің 153-і төл сөздерімізден немесе материалдық калька зат есімнен, 22-сі – қимыл есімінен, 22-сі – зат есімнен, субстантивтенген сын есімнен (өткізгіштік, көбейтіндік, ондық, оналтылық), 5-еуі – біріккен сөзден, 3-еуі – иелік формадағы зат есімнен (асинхронды, фракталды) және есімше, үстеу, көптік жалғаулы зат есімнен жасалған жалаң терминдер 1-1-ден кездеседі. Мұндағы бір назар аударатын мәселе – *-дық* және *-ды* тұлғалы сын есімдердің субстантивтік қызмет атқаруы, яғни осы формалар арқылы бұл сөздер атауыштық мәнге ие болған.

Тіркесті терминдердің көпшілігі N+N моделімен (зат есім+зат есім) жасалған субстантивтік тіркестер екендігі анықталды. Екікомпонентті терминдік тіркестердің тірек сөз табы – зат есім және қимыл есімі, ал бағыныңқы сыңарда зат есім, адъективтенген зат есім, сын есім, есімше, үстеу жұмсалады. Төмендегі кестеден бинарлық терминдердің формальдық-құрылымдық сипатын көрсететін ең өнімді (3-тен жоғары қайталанған) модельдердің статистикасын мысалдарымен қоса көруге болады:

Кесте 1. Бинарлық терминдердің модельдері және жиілігі туралы мәлімет

№	Типтік моделдері	Саны	Мысалдар
1	Natr+N	80	гармоникалық жетек
	Npf+N	32	тізбекті доғалақ, шектеулі кіріс
	FUSWatr+N	6	уақытаралық камера
	Natr+INF	10	векторлық талдау
2	NgenØ+Npp3	92	қоспа моделі
	Npl,genØ+Npp	14	деректер жиыны, құралдар панелі
	Ngen+Npp3	4	шығыстың басқарылуы

	N+N	8	планшет компьютер, іздеуіш робот
3	Nacc+INF	37	батчты нормалау, қозғалысты жоспарлау
	Npl,acc+INF	9	белгілерді айрықшалау
	NaccØ+INF	3	тойтарма тойтару, монета лақтыру
4	ADJ+N	33	жасырын папка
5	ADV+N	19	сирек матрица
	ADV+INF	38	қатаң бақылау, жылдам прототиптеу
6	PCP+N, PCPnf+N, PCP+Npl	14	бағытталған жарық, анықталмаған жүйе, қиылыспайтын жиындар, күтілетін пайдалылық
	PCP+INF	3	тұйықталған басқару
7	CVB+INF	3	тереңдете оқыту, мөлшерлеп оқыту

Жоғарыдағы кестеден көріп отырғанымыздай, екі компонентті терминдердің басым бөлігі – NgenØ+Npp3 моделіне сәйкес келетін түркі изафетінің 1-ші заңдылығы бойынша жасалған 92 термин; одан кейін Natr+N моделіне сәйкес келетін атрибутивтік қатынастағы терминдер – 80 термин; 3-ші орында ADV+N, ADV+ INF модельдеріне сәйкес келетін 57 термин анықталды, олар үстеудің зат есіммен немесе қимыл есімімен қабыса байланысуы арқылы жасалған; ал Nacc+INF моделіне сәйкес келетін табыс септігіндегі зат есімнің қимыл есіміне немесе күрделі етістікке меңгерілуі арқылы жасалған терминдердің саны – 37. Бұл модельдерді тек қана тізіп көрсетіп қоймай, оларды басыңқы және бағыныңқы сыңарларының тұлғаларын есепке ала отырып, топтастырдық.

### Талқылау

Жасанды интеллект саласының бинарлық терминдері бойынша ең өнімді тәсіл ретінде 7 түрлі типтік модель анықталды және олардың морфосинтаксистік тұлғалары төмендегідей (Кесте 1-дің түсіндірмесі):

1) Natr+N, Npf+N, FUSWatr+N, Natr+INF модельдерінің типтік ерекшелігі мынадай: қатыстық формадағы (-лық (atr)), иелік формадағы (-лы (pf)) зат есімдер немесе біріккен зат есім, зат есіммен, кейде қимыл есімімен қабыса байланысқан;

2) N+N, NgenØ+Npp3 (Npl,genØ+Npp3), Ngen+Npp3 модельдеріне тән типтік ерекшелік – түркі изафетінің үлгісінде жасалуы: бағыныңқы-басыңқы сыңарларда зат есім түбір тұлғаларымен қабысу немесе зат есімдердің ілік септігінің ашық, жасырын түрінде матасуы арқылы жасалған; мұндай терминдік тіркестердің түркі изафетінің II заңдылығы бойынша жасалғандары I, III заңдылығына қарағанда әлдеқайда көп;

3) Nacc+INF, NaccØ+INF, Npl,acc+INF, NaccØ+CMPL, Nacc+CMPL модельдері зат есімнің жекеше, көпше түріне табыс септігінің жалғауы ашық және жасырын жалғанып, қимыл есіміне немесе күрделі етістікке меңгерілуі арқылы жасалғандығы тағы бір типтік қатарды құрайды;

4) ADJ+N, ADJ+ INF модельдерінің бағыныңқы компоненті сын есім, басыңқысы зат есім немесе қимыл есімімен жасалып, қабыса байланысқан;

5) ADV+N, ADV+INF модельдері үстеулердің зат есім немесе қимыл есіммен қабыса байланысуы арқылы жасалған;

6) PCP+N, PCPnf+N, PCP+Npl, PCP+INF модельдері адъективтенген есімшелердің (бұрынғы және ауыспалы өткен шақтық) болымды немесе болымсыз формасы зат есімнің жекеше, көпше түрімен, кейде қимыл есімімен қабыса байланысқан;

7) CVB+INF моделі онша өнімді болмағанымен, көсемшенің қимыл есімімен тіркесіп келуі арқылы жасалған терминдік тіркестер де кездеседі.

Осылайша терминдік тіркестердің формалдық құрылымдарын шартты белгілермен көрсетіп, типтік түрлерін анықтаудың заманауи зерттеулер мен мәтінді автоматты өңдеу үшін теориялық әрі практикалық маңызы бар. «Химиялық формулалар химияда қандай рөл атқарса, синтаксистік құрылымдарды бейнелейтін формальды және абстракт құралдар, грамматикада дәл сондай рөл атқарады» [8, 64 б.]. Мұндай құрылымдардың әрбір элементі есепке алынуы үшін сөз болып отырған тіркестердің грамматикалық қатынастағы қызметі, сөз таптарының категориалды мағынасына қарай шартты белгілері нақтылануы шарт. Қазақ тілінің Ұлттық корпусында функционалды қосымшалар ретінде салыстыру (CF (-дай/дей)), ұқсату (AF (-ша/ше)), иелік (PF-лы/лі), субстантив (SF (-лық/лік), атрибут анықтауыш (ATR -лық/лік)) формалары көрсетіліп, тиісті шартты белгілерге [9] басыбайлы етілген. Біз де зерттеу барысында грамматикалық тұлғасы бірдей болғанымен, мағынасы басқа *-лық/лік* жұрнақтарын екі түрлі шартты белгілермен (sf, atr) көрсеттік. Осылайша саралап көрсету өзіндік мәнге ие. Мысалы, төмендегі тіркестерде табыс септігі мен иелік формадағы функционалдық жұрнағының кейбір графикалық бейнесі (-ды/ді, -ты/ті) бірдей болғанымен, грамматикалық мағынасы басқа; осындай түсініксіздікті болдырмас үшін табыс септігін – ACC (acc), иелік формасын – PF (pf) түріндегі шартты белгілермен таңбаладық: Npl, **acc**+INF (сәулелерді трассалау), Natr+N**pf**+Npl (механикалық роликті бұрандалар) және т.б.. Нәтижесінде Npf+N моделі бойынша 32, ал Nacc+INF моделіне сәйкес келетін 29 терминдік тіркес анықталды. Біріншісі «бар» деген мағынаны білдіретін функционалдық жұрнақпен тұлғаланған зат есім мен зат есімнің тіркесі, мысалы, *терминалды ахуал, зиянды бағдарлама, бұрамалы түйрегіш* және т.б.; екіншісі табыс септігіндегі зат есімнің қимыл есіміне меңгеріле байланысуы арқылы жасалған тіркестер: *сөзді тану, бейнені қалпына келтіру, онтологияны оқыту* және т.б. Сонымен қатар формалдық тұлғасы бойынша Npf+INF немесе Nacc+INF құрылымдық моделінің нақты қайсысына сәйкестендіруге болатындығы қиындық тудыратын 8 терминдік тіркес анықталды. Бірақ *афферективті есептеу, инкрементті оқыту, бейімді басқару, визуалды бағдарламалау, стандартты ауытқу* тәрізді тіркестердің қазақ терминжасамына сәйкестігі күмән тудырады. Өйткені қазақ тілінде Npf+N моделіне сәйкес келетін тіркестер болғанымен, Npf+INF моделі, яғни «бар» мағыналы сын есім жұрнағымен тұлғаланған есім сөздің қимыл есіміне сабақтаса байланысуы қазақ тілінің табиғатына жат. Сондықтан бұл тіркестерді орысша баламасымен салыстырып көрдік: афферективные



вычисления – *аффективті есептеу*; двунаправленный поиск – *қосбағытты іздеу*, пограничные вычисления – *шектесімді есептеу* және т.б. Көріп отырғанымыздай, мұнда бағыныңқы сыңар иелік формадағы *-ды* жұрнақтары (PF (pf)) емес, субстантивтік формадағы *-дық* жұрнақтары (SF (sf)) немесе басқа тұлғалар екендігі белгілі болды. Бұл тіркестер *добыча данных* – *деректерді алу, распознавание лица* – *бетті тану* сияқты табыс жалғаулы емес. Олай болса, басыңқысы қимыл есімі, бағыныңқы сыңары табыс септігінен болмаған жағдайда оны *бейімді басқару (адаптивное управление)* емес, *бейімдеп басқару* ретінде аударылғаны дұрыс. Өйткені тілімізде бағыныңқысы көсемшемен тұлғаланған *мөлшерлеп оқыту (пристрелочное обучение)* деген терминдік тіркестер бар. Сондай-ақ бағыныңқысы иелік формаға ұқсас мынадай тіркестер де кездеседі: *тоғыспалы тексеру* (перекрестная проверка); мұны *тоғыстыра тексеру, тоғыстырып тексеру* ретінде; *стандартное отклонение* – *стандартты ауытқу* емес, *стандарттық ауытқу* ретінде аударған дұрыс деген ойдамыз. Сонымен, Npf+INF немесе Nacc+INF моделінің қайсысына сәйкестігін олардың құрамындағы тіркестердің категориалды мағынасын саралап, осыған ұқсас құрылымдарға дистрибуциялық талдау жасай келе, жоғарыда аталған 8 терминнің Nsf+INF, CVB+INF, Natr+Nsf модельдеріне сәйкес келетіндігін (субстантивтің немесе көсемшенің қимыл есіміне бағынатындығын, атрибутивтің субстантивке бағынатындығын) көруге болады. Яғни тіркестердің бағыныңқы сыңарларындағы ұқсас элементтердің грамматикалық тұлғаларын ажыратып белгілеудің әрі олардың басыңқысының өзіне тәуелді элементті меңгеру қабілетінің тіліміздің заңдылығына сәйкестігін саралау шеттілдік терминдерді игеру барысында ескерілуі тиіс. Терминдік тіркестерді модельдеу барысында өте аз қолданылатын модельдер де кездеседі. Олар, бір жағынан, өнімсіз модельдер, екінші жағынан, қазақ тілінің тіркесімділік заңдылығына сәйкес келетіндігі күмән тудырады. Мысалы, Ndat+INF және Npl,ins+ N модельдері бойынша 1-1 тіркес (ұсыныстарға оқыту, символдармен операциялар) анықталды. Бұл терминдердің орысша нұсқасы – *обучение представлениям, операции над символами*, біздің ойымызша, *ұсыныстар бойынша оқыту, символдар операциясы* ретінде аударылғаны дұрыс. Орыс тілінде септеулік шылаулар да, түркі изафеті де жоқ екендігін ескерсек, басқа тілдің ықпалымен жасанды тіркестердің қатарын көбейтудің маңызы жоқ.

### **Қорытынды**

Жасанды интеллект саласының терминдік тіркестерінің сыңарларының грамматикалық бейнесі анықталып, екіқұрамды терминдік тіркестердің морфосинтаксистік модельдері жайында мәліметтер алынды: мұндай тіркестердің жасалуына қатысатын сөз таптары мен грамматикалық формалары шектеулі; олардың бағыныңқы сыңарларында зат есім, сын есім, үстеу, өткен шақ есімшесінің (-ған, -атын тұлғалы) болымды, болымсыз формасы, көсемше (-п, а тұлғалы), басыңқы сыңарында зат есім мен қимыл есімі; ал грамматикалық тұлғалардан атау, ілік және табыс септіктерінің ашық,

жасырын түрлері, көптік жалғауы және субстантив, иелік формадағы функционалдық жұрнақтар жиі қолданылады.

Терминдік тіркестерді формалдық-құрылымдық жағынан әрбір элементіне маңыз бере отырып талдау арқылы грамматикалық тұлғалары бірдей болғанымен, семантикасы басқа тұлғалар шартты белгілердің көмегімен ажыратылып таңбаланды, бұл тілді автоматты өңдеу кезінде омонимияның алдын алуға көмектеседі.

Бинарлық терминдердің өнімді және өнімсіз модельдері анықталды; өнімді модельдердің негізінде шетілдік терминдерді аударуда қазақ терминжасамының заңдылығына сәйкестігін тексеруге болады; ал өнімсіз модельдерің (көбінесе 1 ғана модель болған жағдайда) қазақ тілінің тіркесімдік заңдылығына қайшы келетіндігі немесе сәтсіз аударылған терминдік тіркестер екендігі анықталды; олар нақты мысалдармен пайымдалып, ықтимал нұсқалары берілді; бұл болашақта осындай олқылықтарды болдырмауға септігін тигізеді.

Қазақ тіліндегі терминдерді автоматты өңдеуге қажетті модельдер анықталды; ғылым тілінің корпусын жасау үшін морфологиялық тегтеуге қажетті шартты белгілердің қатары (сөз болып отырған саланың терминдеріне қатысты) айқындалды (Қараңыз: қысқартылған шартты белгілер). Мұндай белгілерді қазақ тілінің деректерін заманауи компаративистикалық зерттеулерде халықаралық стандарттаға сай глостау практикасында пайдалануға болады.

Қазақ терминжасамында белсенді қолданылатын бірқатар функционалдық жұрнақтар анықталғанымен, олар қазақ тілінің корпустарында арнайы белгімен берілмеген. Мысалы, іс-әрекеттің орындаушысын білдіретін, зат есімдердің қабысуы арқылы, яғни түркі изафетінің I заңы бойынша жасалған *іздеуіш робот, болжауыш модельдер, шектеуші көлем* және т.б. тіркестер жиі кездеседі. Татар тілі корпусында мұндай тұлғалар осы шақ есімшесінің *-учы, -уче* тұлғасы – PCP\_PR, башқұрт тілінде – SUP ретінде белгіленген. Және де осындай формаларды ажырату *болжағыш, болжауыш, болжаушы* сөздеріндегі мәндес жұрнақтарды нақты тану үшін де қажет. Бұл қазақ тілінің терминжасамында жиі қолданылатын жұрнақтарды да (-малы, -ғыш және т.б.) есепке алып, түркі тілдерінің тәжірибелерін ескере отырып, тегтеуді жетілдірудің қажеттілігін көрсетеді. Сондай-ақ терминдік тіркестердің құрамында күрделі етістіктер, лексикалық біртұтас бірліктер де кездеседі, оларды да шартты белгілермен таңбалаудың жолдарын іздестіру керек.

#### Қысқартылған шартты белгілер

N – зат есім

ADJ – сын есім

ADV – үстеу

INF – тұйық етістік

PCP – есімше

Ø – нөлдік тұлға

CVB – көсемше

gen (GEN) – ілік септік

dat (DAT) – барыс септік

acc (ACC) – табыс септік

atr (ATR) – атрибут анықтауыш форма

pf (PF) – тәуелдік форма

pp3 (PP3) – III жақ тәуелдік жалғауы

pl (PL) – көптік жалғауы

### ӘДЕБИЕТ

- [1] Кондрашов Н.А. История лингвистических учений: Учеб. Пособие для студентов пед. ин-тов по спец. №2010 «Рус.яз и лит.». – М.: Просвещение, 1979. – 224 с.
- [2] Джубанов А.Х. Квантативная структура казахского текста (опыт лингвистического анализа на ЭВМ). – Алма-Ата: Наука, 1987. – 147 с.
- [3] Аралбаев Ж.А. Вокализм казахского языка (Очерки по экспериментальной фонетике и фонологии), Алма-Ата: Наука, 1970. – 179 с.
- [4] Жубанов А.К. Основные принципы формализации содержания казахского текста. – Алматы, 2002. – 250 с.
- [5] Русская грамматика Том 2. Синтаксис. / Ред. Шведова Н.А., Арутюнова Н.Д., Бондарко А.В. и др. – М.: Наука, 1980. – 717 с.
- [6] Захаров В.П., Хохлова М.В. Автоматическое выявление терминологических словосочетаний. – 2013. – URL: <https://www.academia.edu/8587168> (дата обращения: 13.01.2023).
- [7] The Leipzig Glossing Rules: Conventions for interlinear morpheme-by-morpheme glosses. – URL: <https://www.eva.mpg.de/lingua/pdf/Glossing-Rules.pdf> (accessed: 12.10.2021).
- [8] Тестелец Я.Г. Введение в общий синтаксис.: учеб. пособие. – РГГУ, 2001. – 805 с.
- [9] Қазақ тілінің ұлттық корпус. – URL: <https://qazcorpus.kz/about/4/> (жүктелген күні: 22.02.2023).

### REFERENCES

- [1] Kondrashov N.A. (History of linguistic doctrines). Istoria lingvistsicheskikh ýcheni: Ýчеб. Posobie dlá stýdentov ped. in-tov po spes. №2010 «Rýs.iáz 1 lit.». – М.: Prosveshenie, 1979. – 224 s. [in Rus.]
- [2] Djýbanov A.H. Kvantativnaja strýktýra kazahskogo teksta : opyt lingvistsicheskogo analiza na EVM (Quantative structure of the Kazakh text: experience of linguistic analysis on a computer). – Alma-Ata: Naýka, 1987. – 147 s. [in Rus.]
- [3] Aralbaev J.A. Vokalizm kazahskogo iazyka: Oчерkы po eksperimentalnoi fonetike 1 fonologii (Vocalism of the Kazakh language (Essays on experimental phonetics and phonology). – Alma-Ata: Naýka, 1970. – 179 s. [in Rus.]
- [4] Jýbanov A.K. Osnovnye prinsýpy formalizatsii soderjaniá kazahskogo teksta (Basic principles of formalization of the content of the Kazakh text). – Almaty, 2002. – 250 s. [in Rus.]
- [5] Rýsskaja gramatika. Том 2. Sintaksis (Russian Grammar Volume 2. Syntax). / Red. Shvedova N.A., Arýtúnova N.D., Bondarko A.V. 1 dr. – М.: Naýka, 1980. – 717 s. [in Rus.]
- [6] Zaharov V.P., Hohlova M.V. Avtomaticheskoe vyjavlenie terminologicheskikh slovsochetani. – 2013 (Automatic detection of terminological phrases). – URL: <https://www.academia.edu/8587168> (accessed: 13.01.2023). [in Rus.]
- [7] The Leipzig Glossing Rules: Conventions for interlinear morpheme-by-morpheme glosses. – URL: <https://www.eva.mpg.de/lingua/pdf/Glossing-Rules.pdf> (accessed: 12.10.2021).
- [8] Testeles Ia.G. Vvedenie v obshni sintaksis.: ýчеб. Posobie (Introduction to General Syntax: Proc. Allowance). – RGGÝ, 2001. – 805 s. [In Rus.]
- [9] Qazaq tiliniñ últtyq korpus. – URL: <https://qazcorpus.kz/about/4/> (accessed: 22.02.2023). [In Kaz.]

### ФОРМАЛЬНО-СТРУКТУРНЫЕ МОДЕЛИ БИНАРНЫХ ТЕРМИНОВ

\*Абдиқарим Н.<sup>1</sup>, Ысқақ Б.Д.<sup>2</sup>, Махсуд О.Ж.<sup>3</sup>

\*<sup>1</sup>к.ф.н., ведущий научный сотрудник Управления терминологии  
Национальный научно-практический центр им. Ш. Шаяхметова  
Астана, Казахстан, e-mail: [nursana@yandex.ru](mailto:nursana@yandex.ru)

<sup>2</sup>к.ф.н., руководитель Управления терминологии, Национальный научно-  
практический центр им. Ш. Шаяхметова  
Астана, Казахстан, e-mail: [bakytgul.iskakova@mail.ru](mailto:bakytgul.iskakova@mail.ru)

<sup>3</sup>м.п.н., научный сотрудник Управления терминологии, Национальный  
научно-практический центр им. Ш. Шаяхметова  
Астана, Казахстан, e-mail: [orazgul-astana@mail.ru](mailto:orazgul-astana@mail.ru)

**Аннотация.** В статье рассматривается вопрос о необходимости формализации языкового материала, чтобы компьютер мог понимать его структуру и давать возможность использования IT-технологий и компьютерной лингвистики в терминологии.

Цель статьи – определить формально-структурные модели бинарных терминосочетаний с учетом частей речи и их морфологических признаков в составе синтаксических словосочетаний.

В статье дан краткий обзор истории формализации казахского языка. Несмотря на то что терминосочетания являются лексико-синтаксической категорией, морфология также играет значительную роль в образовании названий определенных понятий, вступая в связь между словами. Также сообщается о том, что это отчетливо проявляется в процессе моделирования языковых единиц с формально-структурной точки зрения.

В качестве основных результатов и анализов исследования можно отметить следующее: учитывалась морфосинтаксическая структура двухкомпонентных терминосочетаний области «Искусственный интеллект», а также подобраны условные знаки, необходимые для их формализации. Все термины этого поля включены в таблицу EXCEL и помечены соответствующим им моделям. При анализе терминосочетаний парадигматико-дистрибутивным методом были отобраны, типизированы их продуктивные модели, показаны основные грамматические характеристики и приведены примеры.

Использование в качестве исследовательского материала терминов только одной области позволяли получить доступ к конкретной информации, которые помогают определить активность части речи в создании терминов, способность сочетания, полноту охвата лексики описываемых терминов и т.д.

По мнению авторов, на основе проведенного исследования возможно моделирование терминосочетаний других отраслей, определения наиболее часто употребляемых типов и разработка лексико-синтаксического шаблона, автоматически распознающего их в специальных текстах.

Практическая значимость статьи заключается в анализе, систематизации парадигматической трансформации терминосочетаний, в выявлении терминосочетаний, не соответствующих правомерности казахского языка. В то же время на основе фактов, которые стали известны в ходе исследования, на конкретных примерах показано, что можно решить некоторые спорные вопросы, касающиеся правописания терминов и оказать положительное влияние на перевод иноязычных терминов.

**Ключевые слова:** термин, бинарные термины, корпусная лингвистика, тег, правило Лейпцига, морфосинтаксические модели, условные знаки, автоматическое распознавание терминов, лексико-синтаксический шаблон

## **FORMAL STRUCTURAL MODELS OF BINARY TERMS**

\*Abdikarim N.<sup>1</sup>, Iskak B.D.<sup>2</sup>, Mahsud O.J.<sup>3</sup>

\*<sup>1</sup>Cand. of Phil. Sc., leading researcher of the Department of Terminology  
National scientific and practical center named after Sh. Shayakhmetov

Astana, Kazakhstan, e-mail: [nursana@yandex.ru](mailto:nursana@yandex.ru)

<sup>2</sup>Cand. of Phil.Sc., Head of the Department of Terminology, National Scientific and Practical Center named after Sh. Shayakhmetov

Astana, Kazakhstan, e-mail: [bakytgul.iskakova@mail.ru](mailto:bakytgul.iskakova@mail.ru)

<sup>3</sup>Master, Researcher, Department of Terminology, National Scientific and Practical Center named after V.I. Sh. Shayakhmetov

Astana, Kazakhstan, e-mail: [orazgul-astana@mail.ru](mailto:orazgul-astana@mail.ru)

**Abstract.** The article deals with the issue of the need to formalize the language material structure to be clear for the software, so this makes it possible to use IT technologies and computational linguistics in terminology.

The purpose of the article is to determine the formal structural models of binary term combinations, taking into account the parts of speech and their morphological features as part of syntactic phrases.

A brief review of the history of formalizing the Kazakh language is presented in the article. The term combinations are a lexico-syntactic category, morphology also plays a significant role in forming the name of a certain concept, entering into a connection between words; it is reported that this is clearly manifested in the process of modeling language units from a formal structural point of view.

As the main results and analyses of the study, the following can be noted: the morphosyntactic structure of two-component term combinations in Artificial Intelligence (AI) was taken into account, and the conventional signs necessary for their formalization were selected. All terms in this field are included in MS Excel and tagged with their respective models. When analyzing term combinations by the paradigmatic-distributive method, their productive models were selected, and typified, the main grammatical characteristics were shown, and examples were given in detail.

The use of terms of only one area as research material made it possible to gain access to specific information that helps to determine the activity of a part of speech in creating the terms, the ability to combine, the completeness of the coverage of the vocabulary of the terms described, etc.

The authors consider that on the basis of the study, it is possible to model term combinations in other industries, determine the most commonly used types, and develop a lexico-syntactic template that automatically recognizes them from special texts.

The practical significance of the article lies in the analysis, and systematization of the paradigmatic transformation of term combinations and the identification of term combinations that do not correspond to the legitimacy of the Kazakh language. At the same time, based on the facts that became known in the course of the study, specific examples show that it is possible to resolve some controversial issues regarding the spelling of terms and have a positive impact on the translation of foreign terms.

**Keywords:** term, binary terms, corpus linguistics, tag, Leipzig rule, morphosyntactic models, conventional signs, automatic term recognition, lexico-syntactic pattern

*Статья поступила 11.04.2023*