

Демак, тилнинг асосий вазифаси тафаккурни лисоний белгилар орқали вокелантиришгина эмас, балки уни шакллантириш, юзага келтириш ҳамдир. Тафаккурнинг асосий вазифаси эса борлиқни, ундаги нарса-ходисаларни билиш бўлиб, бу жараён тил ёрдамида амалга ошади.

### АДАБИЁТ:

- [1] Ирискулов М. Тилшуносликка кириш. Дорилфунун ва педагогика институтлари талабаларига қўлланма. – Тошкент: Ўқитувчи, 1992. – 256 б.  
[2] Махмудов Н. Ўзбек тили – халқ руҳининг кўзгуси / <https://uz24.uz/>  
[3] Семасиология / Ўзбек тили лексикологияси. – Тошкент: Фан, 1981. С.356.  
[4] Шмелев А.Г. Введение в экспериментальная психосемантику: теоретико-методологические основания и психодиагностические возможности. – М., 1983. С.412.  
[5] Глухов В.П. Основы психолингвистики. – М.: Высшая школа, 2005. С. 351.

*МРНТИ 16.31.21*

### Г. Тлегенова<sup>1</sup>, Е.ББесіров<sup>2</sup>

<sup>1</sup>А.Байтұрсынұлы атындағы Тіл білімі институты Қолданбалы лингвистика бөлімінің кіші ғылыми қызметкері, докторант  
Алматы, Қазақстан

<sup>2</sup>А.Байтұрсынұлы атындағы Тіл білімі институты Қолданбалы лингвистика бөлімінің кіші ғылыми қызметкері, магистрант  
Алматы қаласы, Қазақстан

### ЖАРТЫЛАЙ АВТОМАТТЫ МЕТАБЕЛГІЛЕНІМ ЖАСАУДЫҢ ПРАКТИКАЛЫҚ ТҰСТАРЫ

**Аннотация.** Метабелгіленім – ғылыми-зерттеу жұмыстарында белгілі бір кезеңге, стильге, авторға, тақырыпқа т.б. қатысты материалдар жинаудың таптырмас дереккөзі. Метабелгіленім ішкорпусты құру үшін корпуста қолданушыға мәтіндерді іздеуге, таңдауға мүмкіндік береді. Мақалада ұлттық корпустың құрамындағы мәтіндерді металингвистикалық белгілеу мәселесі қарастырылған. Ұсынылған қазақ мәтіндерінің метабелгіленім жүйелері сипатталған.

**Тірек сөздер:** корпус, метабелгіленім, метамәтіндік параметр, аннотацияланған корпус

### Г. Тлегенова<sup>1</sup>, Е.ББесіров<sup>2</sup>

<sup>1</sup>Младший научный сотрудник отдела прикладной лингвистики Института языкознания им. А. Байтұрсынова, докторант,  
Алматы, Казахстан

<sup>2</sup>Младший научный сотрудник отдела прикладной лингвистики Института языкознания им. А. Байтұрсынова, магистрант  
Алматы, Казахстан

## ПРАКТИЧЕСКИЕ АСПЕКТЫ ПОЛУАВТОМАТИЧЕСКОЙ МЕТАРАЗМЕТКИ

**Аннотация.** Метаразметка – незаменимый источник для сбора материала исследования, относящегося к определенному периоду, стилю, автору, теме и т.д. Метаразметка позволяет пользователю искать и выбирать тексты при создании собственного подкорпуса. В статье рассматривается проблема металингвистической разметки текстов при составлении национальных корпусов. Описываются системы метаразметки, применяемые для казахских текстов.

**Ключевые слова:** корпус, метаразметка, метатекстовый параметр, аннотированный корпус.

G. Tlegenova<sup>1</sup>, Y.B.Bessirov<sup>2</sup>

<sup>1</sup>Junior Researcher, Department of Applied Linguistics,  
Institute of Linguistics named after A. Baitursynov, Doctoral student,  
Almaty, Kazakhstan

<sup>2</sup>Junior Researcher, Department of Applied Linguistics,  
Institute of Linguistics named after A. Baitursynov, Master's degree student  
Almaty, Kazakhstan

## PRACTICAL ASPECTS OF SEMI-AUTOMATIC META-MARKING

**Annotation.** Meta-marking is an indispensable source for collecting research material related to a specific period, style, author, topic, etc. Meta-marking allows the users to search for and select texts when creating their own subcorpus. The article deals with the problem of metalinguistic marking of texts when compiling national corpora. Meta-marking systems used for Kazakh texts are described.

**Keywords:** corpus, meta-marking, metatext parameter, annotated corpus.

Тіл білімінде соңғы жылдары компьютерлік технологиялардың көмегімен автоматтандыру мәселесі қолға алынып келеді. Ендеше ақпараттық технологияларды үйлесімді пайдалану қазіргі таңда өзекті мәселелердің біріне айналып отыр. Ғылымда, білім беру жүйесінде компьютерлерді пайдаланудың маңызы зор. Өйткені компьютер адам қызметінің барлық саласында еңбек тиімділігін арттыру құралына айналды.

Корпустық лингвистика қазақ тіліне еніп отырғанына көп уақыт бола қоймаса да, өз кезегінде практикалық-теориялық маңыздылығы арту үстінде. Көптеген жаңа терминдер еніп, оларға анықтама беріліп, заман сұранысына сай жаңа бағыттың қарқынды дамуына көңіл бөлініп отыр.

Корпус құруда метабелгіленім қоюдың маңыздылығы өте қажетті. Айталық

Қазақ тіл біліміндегі корпусты қолдану барысында ең әуелі таңдама мәтіндердің параметрлерін анықтап алуымыз қажет. Мысал ретінде, көркем әдебиет, оның ішіндегі әңгіме не роман, әңгіменің жазылған жылы, жазған адамның ер не әйел екендігі (жынысы), яғни алып отырған мәтін туралы жалпы ақпараттың бірізді формамен, математика тілімен айтқанда алгоритмі құрылуы тиіс. Осы алгоритмнің негізінде программа құрылып компьютердің көмегімен жүзеге асады. Метабелгіленім қою мәтін туралы ақпаратты алуда және біршама мәселелерді шешуде көмегін тигізеді.

«**Метаразметка текстов** является существенной частью поискового аппарата КорпусА. Под метаразметкой понимается приписывание тексту атрибутов, характеризующих обстоятельства его создания, автора, тематику, жанровые особенности и др. Метаразметка необходима прежде всего для того, чтобы исследователь мог составлять по своему желанию произвольные выборки текстов с заданными внешними параметрами: например, тексты мемуарного характера, тексты, написанные мужчинами, тексты, написанные авторами, родившимися между 1940 и 1960 гг., тексты автобиографий, тексты проповедей, тексты романов и повестей, и т. д., и т. п.» [1].

Метабелгіленімдер – ғылыми-зерттеу жұмыстарында белгілі бір кезеңге, стильге, авторға, тақырыпқа т.б. қатысты материалдар жинаудың таптырмас дереккөзі десек те болады.

Метабелгіленімдер қойылған корпустардан зерттеушілер өзіне қажетті стиль, кезең, автор т.б. ақпараттар бойынша мәліметтерді тез тауып алуына мүмкіндік алады. Ал бұл жетістік қазіргі дамыған еліміздің ғылыми-зерттеушілік әлеуетін, ғылым мен білімді жан-жақты дамытатын бірден-бір күш екендігі сөзсіз. Корпустың қандай мәтінмен жұмыс жасасақ та ең алдымен жинақталған мәтін туралы ақпарат беретін метамәтіндік параметрлер нақтылануы қажет. Метамәтіндік белгіленімдер корпустың қай түріне болмасын енгізілуге тиісті маңызды ақпараттар болып табылады

Жартылай метабелгіленім енгізу бағдарламасы 23 параметрден тұрады:

1) Автордың аты-жөні. Бұл параметрде автордың толық аты-жөні, болмаған жағдайда бос қалдырылады.

2) Құрастырушы. Ұжымдық, жеке құрастырушылардың аты-жөні беріледі.

3) Мәтін аты. Бұл жерде мәтіннің атауы беріледі. (Мәтін оқулық ішінен алынған тақырыптан алынса, тақырып атауы жазылады. Ал оқулық атауы дереккөзде көрсетіледі).

4) Мерзімі (жазылған уақыты). Мұндай белгіленімдер автордың шығарманы жазу барысында мәтіннің соңында қалдырған мәліметтерден алынады.

5) Белгілеуші. Бұл жерге метабелгіленімді жасаған адамның аты-жөні жазылады.

6) Аудармашының аты-жөні. Мәтін аударма болған жағдайда аудармашының аты-жөні көрсетіледі.

7) Автордың жынысы. Автордың жынысы мәтін авторы біреу болған жағдайда көрсетіледі; яғни ер не әйел деп біреуі қойылады. Ұжымдық мәтін авторлары әр түрлі жынысты болған жағдайда «ер/әйел» деген белгіленім қойылады.

8) Автордың туған уақыты. Автордың туған жылы, күні туралы нақты мәліметтер көрсетіледі. Ұжымдық авторлар болған жағдайда бос қалдырылады.

9) Серия аты. Мәтін телеарна, радиодан алынған (ауызша, жазбаша) болса, бағдарлама сериясы жазылады.

10) Аудитория жасы. Мектепке дейінгі, бастауыш мектеп жасы, жасөспірімдік кезең, жастар әдебиеті, бейтарап.

11) Субкорпус. Газет мәтіндері, ауызша, мультимедиялық, көркем мәтіндер, поэтикалық мәтіндер, оқыту, ғылыми-техникалық корпус, ісқағаз мәтіндері.

12) Мәтін хронотопы. Мәтінде суреттелген оқиғамен шамамен алғандағы мерзімі, өткен орны көрсетіледі.

13) Стиль. Көсемсөз (публицистикалық), ғылыми-көпшілік, ғылыми стиль, аралас, көркемсөз, ісқағаз, сөйлеу стильдері көрсетіледі.

14) Мәтін типі. Белгілі бір стильдің типтері көрсетіледі.

15) Сөзформа саны. Жеке файлда тұрған әрбір мәтінді ашқанда компьютер экранының төменгі жағында сол мәтіндегі сөз саны шығып тұрады, осы сан метабелгіленімге енгізіледі.

16) Таралым типі. Мәтіннің қай жерден алынғаны жазылады (газет, журнал, ғаламтор-ресурс, кітап, мультимедиа)

17) Мәтін формасы. Мәтіндер жазбаша мәтіндер болса, «жазбаша», егер видео/аудио жазбалар болған жағдайда «ауызша» деген белгіленім қойылады.

18) Графика. Мәтіннің қандай графикада жазылғандығын көрсетеді (кирил жазуы, латын жазуы, төте жазуы).

19) Мәтін тақырыбы. Тақырып белгілі бір салаға байланысты қойылады. Тақырып көркем әдебиетке қойылмайды.

20) Мәтін көзі. Бұл жерде мәтін алынған дереккөзі көрсетіледі (баспа, көшірме (сканер), интернет-ресурс, электронды көшірме, қолмен терілген, қағаз көшірме).

21) Корпусқа енгізілу уақыты. Мәтіннің корпусқа енгізілген мерзімі жазылады.

22) Жарық көрген уақыты. Кейбір мәтіндер (шығарма, монография, оқулық т.б.) жазылған уақытында жарық көрсе, кейбірі кейін өңделіп қайта басылады. Бұл жерде кітаптың жарыққа шыққан, яғни қайта басылып шыққан уақыты көрсетіледі.

23) Дереккөзі. Мәтіннің дереккөзі ретінде ол туралы толық ақпарат жазылады (жері, баспа, жылы).

Жоғарыда аталған метабелгіленім параметрлері бойынша корпусқа енгізілетін мәтіндерге белгіленім осы жартылай автоматты метабелгіленім қою бағдарламасы арқылы жүзеге асады (1-сурет).

1-сурет – 23 параметр енгізілген Метабелгіленім бағдарламасы

Демек, корпустың қандай түрін алсақ та, қандай мәтінмен жұмыс жасасақ та ең алдымен жинақталған мәтін туралы ақпарат беретін метамәтіндік параметрлер компьютерлік базаға салынуы қажет. Сондықтан корпус құрастыруда метабелгіленімдер (шығарма атауы, авторы, жылы, шыққан орны, стилі, жанры, көлемі т.б.) жасаудың әдіс-тәсілдерін, түрлерін, компьютерлік бағдарламаларын т.б. зерттеудің маңызы зор.

Қазақ тілінің ұлттық корпусы тиімді құрастырылып, өз дәрежесінде қолданысқа енгізілуі үшін әлемдік тәжірибедегі метабелгіленімдер қоюдың ортақ ұстанымдары мен әдіс-тәсілдерін зерттеу қажет. Ал бұл жетістік қазіргі дамыған еліміздің ғылыми-зерттеушілік әлеуетін, ғылым мен білімді жан-жақты дамытатын бірден-бір күш екендігі сөзсіз.

### ӘДЕБИЕТТЕР ТІЗІМІ:

[1] Национальный корпус русского языка. Параметры текстов. URL: <http://ruscorpora.ru> (дата обращения: 10.05.2015).

*МРНТИ 16.21.65*

#### **Г.Б.Көбденова**

А.Байтұрсынұлы атындағы Тіл білімі институтының  
аға ғылыми қызметкері, филология ғылымдарының кандидаты  
Алматы қ., Қазақстан

#### **«ҚАЗАҚ ЕСКІ ЖАЗУ ЛЕКСИКАСЫНЫҢ СӨЗДІГІНІҢ» СӨЗДІК МАҚАЛАСЫНЫҢ ЛЕКСИКОГРАФИЯЛЫҚ ПАРАМЕТРІ**

**Аннотация:** Мақалада ескі қазақ кітаби лексикасы сөздігіне лексикографиялық талдау жасалынған. Тарихи бағдарлы сөздіктің құрылымдық-типологиялық, жанрлық жіктемесі берілген. Тарихи, ұлттық, тілдік саналардың сөздік жасаудағы маңызына тоқталады. Сөздіктің макроқұрылымы және микроқұрылымы қарастырылды. Лексикограф – ұжымдық тарихи білімді жеткізуші кәсіби маман екендігіне айтлады. «Сөздік – оқырман» коммуникациясының сөздік жасаудағы басты талаптардың бірі ретінде қарастырылу қажеттігі айтылды.

Тіл – көне құбылыс, тарихи категория. Қазақ тілінің тілдік мұрасы, сөз байлығы бізге жазба ескерткіштер, халықтың ауыз әдебиеті мен ауызекі сөйлеу тілі арқылы жетті. Осы ескі тіл бірліктерді, көне формаларды сөздікте мүмкіндігінше семантикасын дәл тауып көрсете білу сөздікті құрастырушылардың алға қойған зор міндеті.

**Тірек сөздер:** сөздік мақала, тарихи лексикография, тарихи сөздік, ескі кітаби лексика, қазақ ескі жазу лексикасының сөздігі, сөздіктің жіктемесі, сөздіктің лексикографиялық параметрі.