

ҚАЗАҚ ТІЛІНДЕГІ ЭМОЦИОНАЛДЫ ЛЕКСИКАНЫ КОРПУС НЕГІЗІНДЕ АНЫҚТАУ

Нурғалиева А.Е

6B01701-«Қазақ тілі мен әдебиеті» ББ.,

2-курс студенті

nurgalievaa2006@gmail.com

Қ.Жұбанов атындағы Ақтөбе өңірлік университеті

Қазақстан, Ақтөбе

Аннотация. Мақалада қазақ тіліндегі эмоционалды лексиканы корпус негізінде анықтау мәселесі қарастырылады. Эмоционалды лексиканың теориялық негіздері, оның тіл біліміндегі орны және корпус лингвистикасының мүмкіндіктері жан-жақты талданады. Сонымен қатар қазақ тіліндегі эмоциялық бірліктерді автоматты түрде тану әдістері, көңіл күйді талдау (sentiment analysis) жүйелері және эмоциялық сөздіктерді құрастыру мәселелері сипатталады. Зерттеу барысында жиілік талдауы, контекстік талдау, коллокациялық әдіс және компьютерлік лингвистика тәсілдері қолданылды. Зерттеу нәтижелері қазақ тілінің агглютинативті табиғаты эмоционалды лексиканы автоматты түрде анықтауда белгілі бір қиындықтар туғызатынын көрсетті. Мақалада қазақ тіліндегі эмоционалды лексиканы цифрлық ортада зерттеудің болашағы мен NLP технологияларының маңызы айқындалады.

Тірек сөздер: эмоционалды лексика, корпус лингвистикасы, қазақ тілі, көңіл күйді талдау, эмоция, NLP, эмоциялық сөздік.

Annotation. The article examines the issue of identifying emotional vocabulary in the Kazakh language based on corpus analysis. The theoretical foundations of emotional vocabulary, its role in linguistics, and the possibilities of corpus linguistics are comprehensively analyzed. In addition, the study describes methods for the automatic recognition of emotional units in the Kazakh language, sentiment analysis systems, and the development of emotional dictionaries. The

research employs frequency analysis, contextual analysis, collocational methods, and computational linguistics approaches. The findings reveal that the agglutinative nature of the Kazakh language creates specific challenges in the automatic identification of emotional vocabulary. The article also highlights the prospects of studying emotional vocabulary in the digital environment and the importance of NLP technologies in the development of Kazakh language research.

Keywords: *emotional vocabulary, corpus linguistics, Kazakh language, sentiment analysis, emotion, NLP, emotional dictionary.*

Қазіргі тіл білімінде эмоцияны тіл арқылы зерттеу маңызды ғылыми бағыттардың біріне айналды. Тіл - адамның ішкі сезімін, көңіл күйін және бағалау қатынасын білдіретін негізгі құралдардың бірі. Осыған байланысты эмоционалды лексиканы зерттеу психолингвистика, когнитивті лингвистика, стилистика және компьютерлік лингвистика салаларында кеңінен қарастырылып келеді.

Қазақ тіл білімінде эмоционалды-экспрессивті лексика мәселесі дәстүрлі түрде қарастырылғанымен, корпус негізіндегі автоматтандырылған зерттеулер жеткіліксіз. Сондықтан қазақ тіліндегі эмоционалды лексиканы корпус материалдары негізінде анықтау қазіргі қазақ компьютерлік лингвистикасының маңызды міндеттерінің бірі болып табылады.

Эмоционалды лексика дегеніміз – адамның сезімін, эмоциялық күйін, психологиялық әсерін және субъективті бағасын білдіретін тілдік бірліктер жиынтығы. Мұндай сөздер сөйлеушінің тек ақпарат беру мақсатын ғана емес, оның ішкі жан дүниесін, көңіл-күйін, қуанышын, өкінішін, ренішін немесе таңданысын білдіреді. Эмоционалды сөздер арқылы адамның дүниені қабылдау ерекшелігі, бағалау сипаты және ұлттық дүниетанымы айқын көрінеді.

Қазақ тіліндегі эмоционалды лексика өте бай әрі көркем болып келеді. Эмоцияны білдіретін сөздер халықтың тұрмыс-тіршілігімен, мәдениетімен, дүниетанымымен және ұлттық танымымен тығыз байланысты қалыптасқан. Мысалы, қазақ тілінде қуаныш, қайғы, қорқыныш, ашу, таңдану, сүйіспеншілік сияқты сезімдерді білдіретін көптеген сөздер мен тұрақты тіркестер бар.

Қуаныш эмоциясын білдіретін сөздерге *керемет, тамаша, гажап, шаттық, қуаныш, мәз, бақыт* сияқты лексемалар жатады. Бұл сөздер адамның жағымды эмоциялық күйін, риза көңілін білдіреді. Мысалы: «Ол бұл хабарды естіп қатты қуанды», «Керемет жаңалық екен» деген сөйлемдерде сөйлеушінің жағымды эмоциясы анық байқалады.

Ал қайғы мен мұңды білдіретін эмоционалды сөздерге *мұң, өкініш, қасірет, қайғы, налу, шер* сияқты лексемалар кіреді. Мұндай сөздер адамның ішкі күйзелісін, жан азабын немесе өкініш сезімін білдіреді. Мысалы: «Жүрегін мұң кернеді», «Ол болған жағдайға қатты өкінді» деген сөйлемдерде жағымсыз эмоциялық реңк айқын сезіледі.

Ашу-ыза эмоциясын білдіретін сөздер де қазақ тілінде кең таралған. Оларға *ыза, ашу, долылық, жек көру, кек, қаһар* сияқты сөздер жатады. Бұл сөздер адамның наразылығын, ренішін немесе агрессиялық көңіл күйін сипаттайды. Мысалы: «Оның сөзін естіп іші удай ашыды», «Қаһарына мінді» деген тіркестерде өкініш, ашу эмоциясы көрінеді.

Қорқыныш пен үрейді білдіретін сөздерге *үрей, қорқыныш, шошу, сескену, зәре, қобалжу* сияқты лексемалар жатады. Бұл сөздер қауіпті жағдайға немесе белгісіздікке байланысты туындайтын эмоциялық күйді сипаттайды. Мысалы: «Қараңғыдан қорықты», «Зәресі ұшты» деген тіркестерде қорқыныш эмоциясы берілген.

Эмоционалды лексика көбіне экспрессивтілік ұғымымен қатар қарастырылады. Дегенмен бұл екі ұғымның арасында айырмашылық бар. Эмоционалдылық адамның сезімін, ішкі эмоциялық күйін білдірумен байланысты болса, экспрессивтілік тілдің әсер ету, көркемдеу және бейнелеу қызметін күшейтеді. Яғни эмоционалдылық - мазмұндық категория, ал экспрессивтілік – стильдік және әсерлілік категориясы. Кейбір сөздерде осы екі қасиет қатар ұшырасады. Мысалы, *сұмдық, масқара, ғажап* сияқты сөздер әрі эмоциялық реңкке, әрі күшті экспрессивтік сипатқа ие.

Қазақ тіліндегі эмоционалды-экспрессивті лексика ұлттық тілдің көркемдік қуатын, бейнелілік мүмкіндігін және эмоцияны жеткізу ерекшелігін айқын көрсетеді. Эмоционалды-экспрессивті сөздер сөйлеушінің белгілі бір құбылысқа, оқиғаға немесе адамға деген субъективті көзқарасын білдіріп қана қоймай, тыңдаушыға әсер ету қызметін де атқарады. Сондықтан мұндай лексика көркем әдебиетте, ауызекі сөйлеу тілінде, публицистикалық стильде және фольклорлық шығармаларда кеңінен қолданылады.

Қазақ тіліндегі эмоционалды-экспрессивті лексиканың басты ерекшеліктерінің бірі - оның бейнелілік сипаты. Қазақ халқы эмоцияны көбіне тура емес, астарлы әрі образды түрде жеткізуге бейім. Осыған байланысты тілде фразеологиялық тіркестер мен тұрақты сөз орамдары мол кездеседі. Мысалы: *жүрегі жарылды, іші күйді, төбе шашы тік тұрды, көзіне жас алды, зәресі ұшты, қаны басына шапты*. Бұл тіркестер адамның эмоциялық күйін дәл әрі әсерлі жеткізеді. Мәселен, «*іші күйді*» тіркесі адамның қатты қиналуын немесе күйзелуін білдірсе, «*төбе шашы тік тұрды*» тіркесі қорқыныш пен шошынуды бейнелейді. Мұндай тіркестердің эмоциялық әсері жеке сөздерге қарағанда әлдеқайда күшті болады. Адамға тән жақсы көру, жек көру, жеру, жирену, күйіну, сүйсіну т.б. да сезім, көңіл-күйлеріне қатысты тілдік бірліктерге лингвистикалық талдау жасағанда оларға тек тілдік тұрғыдан ғана

емес, философиялық, физиологиялық, әлеуметтік, психологиялық, педагогикалық тұрғыдан да келу керек [2;147-б.].

Эмоционалды-экспрессивті лексиканың тағы бір ерекшелігі - бағалауыштық сипатында. Сөйлеуші белгілі бір затқа, құбылысқа немесе адамға оң не теріс баға беру арқылы өзінің эмоциялық қатынасын білдіреді. Мысалы, *тамаша, керемет, гажан* сияқты сөздер жағымды бағалауды білдірсе, *сорақы, сұмдық, масқара* сөздері жағымсыз бағалау мәнін береді. Қазақ тілінде эмоцияны білдіруде интонация, қайталау, күшейткіш сөздер және одағайлар да маңызды қызмет атқарады. Мысалы: *на, шіркін!әттеген-ай! ойпырмай! масқара-ай!* деген одағайлар сөйлеушінің таңдану, өкініш, реніш немесе сүйсіну сезімдерін білдіреді. Сонымен қатар эмоционалды-экспрессивті лексика ұлттық мәдениетпен тығыз байланысты. Қазақ халқының салт-дәстүрі, дүниетанымы және тұрмыс-тіршілігі тілдегі эмоциялық бірліктердің қалыптасуына ықпал еткен. Мысалы, қазақ тілінде жүрек сөзі эмоциямен жиі байланыстырылады: *жүрегі елжіреді, жүрегі ауырды, жүрегі тас төбесіне шықты* деген тіркестер адамның ішкі сезімін жүрек ұғымымен байланысты қабылдау ерекшелігін көрсетеді.

Ш.Нұрғожина сөйлеу тіліндегі эмоционалды-экспрессивті лексиканы алты түрге жіктейді. Олар: жалпы қолданыстағы лексика, қарапайым сөздер, тұрмыстық лексика, әдепсіз, балағат сөздер, диалогтер, варваризмдер. Зерттеуші бұлардың алты түріне де жеке-жеке тоқталып, әрқайсысына тілдік фактілер ұсынады және «халықтың ауызекі сөйлеу тіліндегі қарапайым сөздердің экспрессиясы күшті болып келеді, сондықтан олар ешбір өзгеріссіз сол қалпында кейіпкерлер тілінде беріледі» деп, оған мысалдар келтіреді: Ана сорлы таздың жамандығы ғой. Әшейінде адамсып, шәңгірлеп сөйлер еді, енді қайда қалды екен? (Б.Майлин. Раушан-коммунист) немесе «Жалап-жұқтау»

деген сөз бе екен, - дейді Қайрақбай кейіп, - «сіміріп жұтқанын неге айтпайсындар? (С.Мұқанов. Мөлдір махаббат) [1;37-б.].

Қазіргі компьютерлік лингвистикада эмоционалды-экспрессивті бірліктерді автоматты түрде анықтау мәселесі ерекше өзекті. Әсіресе, қазақ тіліндегі эмоционалды тіркестердің көпмағыналылығы мен бейнелілігі автоматты талдау жүйелеріне қиындық туғызады. Осыған байланысты эмоционалды-экспрессивті лексиканы корпус негізінде зерттеу, олардың қолданыс жиілігін, контекстік ерекшелігін және эмоциялық реңкін анықтау қазақ тіл біліміндегі маңызды ғылыми бағыттардың бірі болып саналады.

XX ғасырдың соңғы ширегінен бастап тіл білімінде жаңа технологияларға негізделген зерттеу бағыттары қарқынды дами бастады. Солардың бірі - корпус лингвистикасы. Бұл сала тілдік материалдарды электронды мәтіндер базасы арқылы зерттеуге бағытталған және қазіргі заманғы лингвистиканың маңызды ғылыми салаларының біріне айналды. Корпус лингвистикасы тілдік құбылыстарды үлкен көлемдегі нақты мәтіндік деректер негізінде талдауға мүмкіндік береді. Дәстүрлі тіл білімінде зерттеуші көбіне жекелеген мысалдарға сүйеніп қорытынды жасаса, корпус лингвистикасында миллиондаған сөзқолданыстар арнайы бағдарламалар арқылы өңделіп, статистикалық, семантикалық және прагматикалық талдаулар жүргізіледі. Осы ерекшелігі бұл саланы объективті әрі дәл ғылыми зерттеу құралы ретінде танытады.

Корпус дегеніміз – белгілі бір ғылыми талаптарға сәйкес жинақталған, құрылымданған және электронды форматта сақталған мәтіндер жиынтығы. Корпусқа көркем әдебиет шығармалары, ғылыми еңбектер, публицистикалық мақалалар, ресми құжаттар, медиа мәтіндер, ауызекі сөйлеу үлгілері және интернет материалдары енгізіледі. Мұндай мәтіндер арнайы таңбаланып, морфологиялық, синтаксистік және семантикалық белгілермен

толықтырылады. Соның нәтижесінде зерттеуші тілдік бірліктердің нақты қолданысын автоматты түрде талдай алады. Корпус лингвистикасының негізгі мақсаты – тілдің табиғи қолданысын көрсету және тілдік заңдылықтарды нақты деректер негізінде анықтау. Бұл тәсіл тілдің динамикалық сипатын, сөздердің қолданыс жиілігін, контекстік ортасын және мағыналық өзгерістерін зерттеуге мүмкіндік береді. Сонымен қатар корпус материалдары арқылы белгілі бір кезеңдегі тілдік үрдістерді, жаңа сөздердің пайда болуын және стильдік ерекшеліктерді бақылауға болады [5;127-б.].

Корпус лингвистикасының эмоционалды лексиканы зерттеудегі маңызы ерекше. Эмоционалды сөздер мен тіркестердің қолданылу ерекшеліктерін, олардың мағыналық реңктерін және эмоциялық қызметін үлкен мәтіндік база арқылы ғана жан-жақты анықтауға болады. Мысалы, «қуаныш», «қайғы», «ыза», «қорқыныш» сияқты эмоциялық сөздердің қандай стильде жиі қолданылатынын, қандай сөздермен тіркесетінін және қандай контексте эмоциялық реңкке ие болатынын корпус материалдары нақты көрсетеді. Корпустық зерттеудің маңызды артықшылықтарының бірі - объективтілігі. Сонымен қатар корпус лингвистикасы эмоциялық бірліктердің уақыт өте келе өзгеруін, жаңа эмоциялық реңктердің қалыптасуын және тілдегі бағалауыштық қатынастың трансформациясын бақылауға жағдай жасайды. Ш.И. Нұрғожинаның «Эмоционально-экспрессивная лексика казахского разговорного языка» атты кандидаттық диссертациясында қазақ сөйлеу тілі лексикасындағы эмоционалды-экспрессивті сөздердің жасалу жолдары, айырымдық белгілері мен функционалдық ерекшеліктері қарастырылады. Зерттеуші қазақ ауызекі сөйлеу тіліндегі эмоционалды және экспрессивті сөздер тобын анықтау арқылы эмоционалды-экспрессивтіліктің өзіндік ерекшелігін зерттеуді мақсат етіп қояды [2;144-б.].

Эмоционалды лексиканы корпус негізінде зерттеу бірнеше маңызды ғылыми міндеттерді шешеді. Біріншіден, эмоциялық сөздердің қолданыс жиілігі анықталады. Мысалы, «қуаныш», «өкініш», «ыза», «қорқыныш» сөздерінің мәтіндерде қаншалықты жиі кездесетіні есептеледі. Екіншіден, эмоциялық бірліктердің контекстік ерекшеліктері зерттеледі. Кейбір сөздердің эмоциялық реңкі контекске байланысты өзгертіндіктен, олардың нақты мағынасы сөйлем ішінде ғана анықталады. Үшіншіден, эмоциялық сөздердің коллокациялық байланыстары зерттеледі, яғни қандай сөздермен жиі тіркесетіні талданады. Бұл эмоциялық өрістерді анықтауға мүмкіндік береді.

Қазіргі таңда корпус лингвистикасы компьютерлік лингвистикамен, жасанды интеллект технологияларымен және табиғи тілді автоматты өңдеу бағыттарымен тығыз байланыста дамып келеді. Әсіресе sentiment analysis, emotion recognition, opinion mining сияқты салаларда корпус деректері негізгі ресурс ретінде пайдаланылады. Бұл бағыттар әлеуметтік желілердегі пікірлерді талдау, медиамәтіндердің эмоционалды реңкін анықтау және автоматты диалогтық жүйелер құру сияқты практикалық міндеттерді шешуде кеңінен қолданылады. Қазақ тілінің ұлттық корпусын қалыптастыру, электронды сөздіктер жасау, автоматты морфологиялық талдау жүйелерін дамыту және эмоциялық бірліктерді белгілеу жұмыстары қазақ тілінің цифрлық кеңістіктегі дамуына негіз болуда. Сондықтан корпус лингвистикасы қазақ тіліндегі эмоционалды лексиканы зерттеудің ғылыми-теориялық және практикалық базасын қалыптастыратын маңызды сала болып саналады. Корпустық деректер теориялық лингвистикаға қандай септігін тигізеді деген сұраққа келсек, әрине, олар тілшілердің лексика мен грамматика жайындағы пайымдауын да алмастыра алмайды, бірақ олар мамандарға бай репрезентативтік эмпирикалық материал береді. Түптеп келгенде, корпусар тілдік зерттеулер барысында пайдалануға болатын үш типті деректер бере

алады: эмпирикалық қолдау, жиілігі бойынша ақпарат, экстралингвистикалық ақпарат (метаақпарат) [4;85-б.].

Ұлттық корпус қазақ тілінің қазіргі қолданысын, функционалдық стильдерін және тілдік заңдылықтарын зерттеуге арналған маңызды ғылыми ресурс болып табылады. Қазақ тілінің ұлттық корпусына әртүрлі жанрлар мен стильдердегі мәтіндер енгізіледі. Олардың қатарына: көркем әдебиет мәтіндері, ғылыми стильдегі еңбектер, публицистикалық мақалалар, ресми құжаттар, медиа материалдар, интернет мәтіндері, ауызекі сөйлеу үлгілері жатады. Осындай әртүрлі материалдар қазақ тілінің эмоциялық мүмкіндіктерін жан-жақты зерттеуге жағдай жасайды.

Көркем әдебиет мәтіндері эмоционалды лексиканың ең бай көздерінің бірі болып табылады. Қазақ жазушылары мен ақындары адамның ішкі сезімін бейнелі әрі әсерлі жеткізу үшін эмоциялық-экспрессивті сөздерді, теңеулерді, эпитеттерді және тұрақты тіркестерді кеңінен қолданады. Мысалы, Мұхтар Әуезов, Ғабит Мүсірепов, Ілияс Есенберлин шығармаларында адамның қуаныш, қайғы, өкініш, қорқыныш және ашу сезімдері терең суреттеледі. Сондықтан көркем әдебиет мәтіндері эмоцияның ұлттық-мәдени ерекшеліктерін анықтауда маңызды дереккөз болып саналады.

Ақын, жазушыларымыздың сөз құдіретін қастерлей білуі сезгіш, аңғарғыш болуы – заңды құбылыс. Сондықтан да олар ойды жеткізуші сөздердің әрсіз, сезімсіз болмай, жалын атып, ұшқын шашып тұруын, аса уытты, әсерлі болуын, сезім мен ойдың бірдей дөп тиіп жатуын қалайды. Осы үшін де экспрессивті-эмоционалды сөздердің атқаратын қызметі жоғарғы дәрежеде тұр деп айтуға болады. Эмоционалдық пен экспрессивтілік теориялық және практикалық тұрғыдан әлі терең, жан-жақты саралануға тиісті категория болып саналады. Дегенмен де, оның өрісі кең, қолданысы бай

құбылыс екені анық көрінеді. Көркем шығарма тіліндегі эмоционалды-экспрессивті лексика әлі де болса зерттеуді қажет етеді [3;54-б.].

Публицистикалық және медиа мәтіндер қоғамдық пікір мен бағалауыштық қатынасты зерттеуде ерекше рөл атқарады. Бұқаралық ақпарат құралдары мен интернет журналистикасында эмоциялық реңк көбіне қоғамдық пікір қалыптастыру мақсатында қолданылады. Әсіресе қазіргі әлеуметтік желілердегі мәтіндер эмоциялық лексиканың белсенді қолданылатын ортасына айналды. Әлеуметтік желі қолданушылары өз сезімдерін ашық әрі еркін білдіретіндіктен, мұндай мәтіндер қазіргі қазақ тіліндегі эмоционалды қолданыстың маңызды көрсеткіштерінің бірі болып есептеледі.

Ауызекі сөйлеу корпусы эмоционалды лексиканы зерттеуде ерекше маңызға ие. Ауызекі сөйлеуде эмоция көбіне интонация, одағай, қайталау және бейнелі тіркестер арқылы беріледі. Мысалы, «ойпырмай», «әттеген-ай», «па, шіркін» сияқты одағайлар сөйлеушінің таңдану, өкініш немесе сүйсіну сезімдерін білдіреді. Ауызекі сөйлеу тілінде эмоциялық реңк жазба мәтіндерге қарағанда анағұрлым айқын көрінеді.

Қазақ тілінің ұлттық корпусы эмоционалды лексиканы автоматты түрде анықтауға және белгілеуге мүмкіндік береді. Корпус негізінде эмоциялық сөздердің жиілігі, жағымды және жағымсыз бағалау, эмоцияның мәтіндегі таралуы, эмоциялық реңктің стильдік ерекшеліктері, эмоциялық тіркестердің қолданысы анықталады. Ұлттық корпус қазақ тілін цифрлық ортада дамытуға, тілдік ресурстарды жүйелеуге және қазақ тілінің халықаралық ғылыми кеңістіктегі орнын нығайтуға мүмкіндік береді. Сондықтан қазақ тілінің ұлттық корпусы тек лингвистикалық зерттеулер үшін ғана емес, сонымен қатар ақпараттық технологиялар, жасанды интеллект және білім беру салалары үшін де стратегиялық маңызды ресурс болып табылады [6;75-б.].

Қазақ тіліндегі эмоционалды лексиканы корпус негізінде анықтау - қазіргі қазақ тіл біліміндегі маңызды әрі перспективалы бағыттардың бірі. Корпустық зерттеу әдістері эмоциялық бірліктерді объективті түрде сипаттауға мүмкіндік береді. Сонымен қатар sentiment analysis және NLP технологиялары қазақ тілінің цифрлық дамуына ықпал етеді. Қазақ тілінің агглютинативті табиғаты мен фразеологиялық байлығы эмоционалды лексиканы зерттеуді күрделендіргенімен, бұл ерекшеліктер қазақ тілінің эмоциялық жүйесін тереңірек тануға жол ашады.

Пайдаланылған әдебиеттер

1. Нұрғожина Ш.И. Эмоционально-экспрессивная лексика казахского разговорного языка. Дисс. на соискание уч. степ. канд. филол. наук. – Алматы, 1989. – С. 145
2. Иманалиева Ғ.Қ. Эмоционалдылық пен экспрессивтілік категориялары туралы // *ҚазҰУ хабаршысы. Филология сериясы.* – Алматы, 2010.
3. Қазанбаева А.З. Қазақ тіліндегі эмоционалды-экспрессивті лексиканың көркем шығармада қолданылуы // *Успехи современного естествознания.* – 2013.
4. Пірманова К.Қ., Жаңабекова А.Ә., Барменқұлова А. Ұлттық корпустарға негізделген лингвистикалық зерттеулер жүргізу (қазақ, орыс, ағылшын тілі материалдары негізінде) Al-Farabi Kazakh National University Филология сериясы. №3 (187).2022
5. Жұбанов А.Қ. Корпустық лингвистика / А.Қ. Жұбанов, А.Ә. Жаңабекова. – Алматы: Қазақ тілі, 2017. – 336 б.
6. Қоңыратбаева Ж.М. Қазақ тілінің академиялық корпусын әзірлеу: мақсаты, міндеті, маңызы (гуманитарлық ғылымдар бағытында) // *Қарағанды университетінің хабаршысы. Филология сериясы.* – 2025, 30, 3(119).