

**МЕН/БЕН/ПЕН ЖАЛҒАУЛЫҚ ШЫЛАУЛАРЫНЫҢ СТИЛЬДЕР БОЙЫНША  
КВАНТИТАТИВТІ САРАЛАНУЫ**  
**МЕДЕТБЕКОВА Перизат, филол.ғ.к., доцент**  
[medet\\_73@mail.ru](mailto:medet_73@mail.ru)  
Алматы қ-сы, әл-Фараби ат.ҚазҰУ

«Тіл білімінде сандық, математикалық ойлауды жиі қолданып, оны мейлінше нақты ғылымдарға жақындату керек» [1], – деп ғалым И.А. Бодуэн де Куртенэнің айтқанындай, тіл ғылымында сандық зерттеулердің қажеттігін математикалық лингвистиканы жақтаушылармен бірге, аса көрнекті дәстүрлі тіл білімінің өкілдері мойындап, барлық грамматикалық категорияларды сандық тәсілмен талдау олардың бүкіл әдеби тіл сфераларындағы функционалдық салмағын да көрсетуге әлі де мол мүмкіндік беретіні анық.

Қазақ тілін лингвостатистикалық тұрғыдан зерттеу 1968 жылдары Қазақ ССР Ғылым академиясы Тіл білімі институтында «Лингвистикалық жұмыстарды автоматтандыру және тіл статистикасы» тобының құрылуымен байланысты әрі жүйелі, әрі нәтижелі бола бастады. Оған бірден-бір атсалысқан: А. Ахабаев, Қ. Бектаев, А. Белботаев, А. Зекенова, С. Мырзабеков, А. Жұбанов және тағы басқа тілші-ғалымдарды атап өтуге болады [2].

Ғалым К. Ахановтың сөзіне жүгінер болсақ, тіл білімі ғылымы математика ғылымымен, біріншіден, машиналық аударма мәселелері бойынша байланысса, екіншіден, тілді зерттеуде математикалық статистика әдістерінің қолданылуы жағынан тығыз байланысады. Математикалық статистика әдісі сөздіктерді жасауда оларға енетін сөздер тізбегінің аумағын белгілеуге және байланыс техникасында дыбыстардың қолданылу жиілігін айқындауға көмектеседі [3, 30 б.].

«Лингвистика мен математика салаларының аралығынан туындаған – «математикалық лингвистика» атты жаңа пән өмірге келді. Бұл пәннің ХХ ғ. 50-60 жылдары қалыптасуы бұрыннан да тіл білімінің барлық ішкі дамуының өзінен де көңіл аударудан, тілдік жүйе ретінде қарастырудан және оның микрожүйелерден тұруы күрделілігін жете түсінуден, лингвистикалық нысандар мен олардың атрибуттары аралығындағы қатынастарды танып-білуден байқалып, математикалық лингвистика пәнінің дамуына объективті жағдай туғызды деуге болады», – дейді профессор А. Жұбанов [4, 11 б.]

Сондай-ақ, жазба әдебиеттерге квантитативті-статистикалық зерттеу жүргізу көршелес орыс ғалымдарында бізден бұрын және қазірдің өзінде де анағұрлым жақсы дамыған. Ең бірінші орыс ғалым-тілшілері статистиканы тілдегі дыбыстық бірліктерді анықтау үшін қолданған. 1911 жылы В. Петров пен П. Александров француз және ағылшын тілдерінің дыбыстық бірліктерін меңгере отырып, сапасына қарай сандық талдау жасауға тырысты.

Мысалы, Б.Н. Головиннің авторлығымен 1971 жылы шыққан «Тіл және статистика» атты монографиясын қазіргі орыс тіліндегі сөз таптарын статистикалық оқытудағы қажетті еңбектерінің бірі деп атауға болады. Монография толығымен статистикалық мәселелеріндегі лингвистикалық құбылыстарға арналған немесе мәтіндегі сөз таптарын меңгеру еңбектің

негізгі мақсаты болып саналады. Бұл жерде сөзжасам және әр түрлі сандық тәжірибенің синтаксистік категория – сөйлем түрлері, жеке сөйлем мүшелерінің қолданылуы, орыс тіліндегі буындардың белсенді қолданылу деректері жүргізіледі. Дегенмен, монография біз үшін практикалық басшылығы, методикасы, негізгі түсінігі, формулалар мен математикалық статистикамен байланысты, зерттеулердегі тілдік фактілер жағынан құнды еңбек болып саналады [5].

П.М. Алексеевтің пайымдауынша жиілік сөздік туралы ең бірінші оқу құралы жиырма жыл бұрын жарық көрген. Оның артынан көп ұзамай жиілік сөздікті құрастыру мәселелері және сөздіктің қолданылуы, қазіргі украин прозаларындағы лексикалық статистикасы бойынша монографиялар шыға бастады [6, 12 б.].

Тіл білімінде статистикалық әдістерді қолдану мәселелерін қозғай отырып, бірінші болып осы әдістің қолданылуын көрсеткен орыс математигі В.Я. Буняковскийді атап кетпеу мүмкін емес. И.А. Бодуэн де Крутенэ, М.Н. Петерсон, Е.Д. Поливанов, В.В. Виноградов және т.б. атақты орыс ғалымдарының тілдегі сандық элементтерге көңіл бөлуі кездейсоқ емес.

XIX ғасырда үндіуропа тіл білімінде кеңінен тараған статистикалық әдіс түркітануда, соның ішінде қазақ тілінде XX ғасырдың 60-жылдарынан бастап қана жүйелі және ғылыми тұрғыда қолданыла бастады. Бұдан бұрын да аталған әдіс бірыңғай қолданыс тауып, тілдің әр деңгейіндегі бірліктердің сандық және пайыздық арасалмағын анықтаумен шектеліп келген болатын.

Тілдегі қолданбалы мәселелерді шешуде жиілік сөздігінің атқарар қызметі орасан зор. Бұл жерде, әсіресе, бір тілден екінші тілге аударуда машиналық аударма қажет. Лингвистикалық және көптеген қолданбалы жұмыстарды жиілік сөздіксіз автоматтандыру мүмкін емес.

Тілді немесе сөзді зерттеуде сапалық және сандық әдістер қолданылады. Сапалы математикалық аппараттың көмегімен теориялық тіл білімінің бағыттары, анығырақ айтқанда математикалық лингвистика анықталған. Квантативті әдістер тіл білімінің кейбір бөлімдерін математикалық жүйелеуге келеді.

Жүйеге ғана тіл заңдылығы тән болғандықтан, статистикалық заңдылықтарды ашуға байланысты қолданылатын әдістер кезінде тілдің өзінше бір бөлек жүйе құратындығы туралы сөз ете бермейміз. Басқаша айтсақ, тіл – бір-біріне қатыссыз құбылыстар жиынтығы емес, керісінше, заңды түрде ұйымдасқан, көпөлшемді және тәуелсіз сипаттағы ақиқат.

Статистикалық деректерді өңдеуде компьютерлік технологияны пайдалану, тілдік мәселелердің шешімін тез арада табу, математикалық статистика әдістерін басқа ғылымдар тәсілдерімен (аппараттық теория, математикалық логика және т.б.) ұштастыру қазіргі кездегі отандық және шетелдік тіл білімінің даму бағытына сай келеді [4, 5 б.].

Статистикалық лингвистика, жиілік сөздік базасын қалыптастыру, мәтіндегі интерпретациялық сөз стилистикасы және тілдік заңдылық стилистикасының қорытындысын береді. Бөлек активті және жиіліктің

қолданылуы сөз таптарының бір бөлігі квантативті-статистикалық зерттеулерін қарастырады.

Жалпы жиілік сөздіктерінің басқа сөздіктерден айырмашылығы, кез келген жиілік сөздік үшін талдауға алынған мәтіндегі сөздердің қолданылу жиілігінің көрсеткіші анықталып, құрастырылуы жағынан көбіне «алфавиттік-жиілік» болып беріледі.

Лингвостатистикалық талдау лингвистикалық бірліктердің (фонема, әріп, буын, морфема, сөз, т.б.) сөйлеуде (ауызша не жазбаша) жұмсалып жиілігіне негізделсе, ал біздің негізгі зерттеу объектіміз – жазбаша сөйлеу тілі, оның ішінде көркем стиль (Жазушы Ә. Нұршайықовтың шығармалар жинағы, Р.Б. Нұртазина мен А.Д. Сейсенованың соавторлығымен жарық көрген балалар әдебиеті және т.б. көркем мәтіндер), ғылыми стиль (Лингвистика, әдебиеттану, құқықтану, журналистика, т.б. ғылымдары бойынша қорғалған кандидаттық диссертациялар) және БАҚ тілі («Дала мен Қала» газетіндегі жарияланымдар бойынша), яғни электронды мәтін, оның ішінде автордың жалғаулық шылауларды қолдануына сандық тәжірибе, квантативті-статистикалық талдау жүргізу.

Қазақ тілінің грамматикалық сипаттарына қарай, шылаулар ішінара үш жікке бөлінеді – септеуліктер, жалғаулықтар және демеуліктер. Соның ішінде біздің жоғарыда көрсетілген көркем стиль, ғылыми стиль, БАҚ тілінің ішіндегі *мен/бен/пен, және* жалғаулық шылауларының сандық жиілігін анықтау біздің ғылыми жұмысымыздың негізгі мәселелері болып табылады.

Сандық талдау жүргізілген көркем стиль – 1879 беттен, 460 000 сөзден, ғылыми стиль – 1593 беттен, 460 000 сөзден, ал публицистикалық стильде – 1723 беттен, 460 000 сөзден тұратынын анық болғандықтан, жалғаулық шылаулардың сандық индексін төмендегі кестеден көруге болады.

№ 1 кесте – *Мен/бен/пен, және* жалғаулық шылауларының түрлі стильдерде қолданылуының сандық индексі

стиль	<i>мен/бен/пен,</i>	<i>және</i>
публицистикалық	141,81	199,74
ғылыми стиль	231,86	449,66
көркем әдеби тіл	305,86	774,42
барлық индекс саны	679,53	1423,82

Үш мәтіндегі *мен/бен/пен* жалғаулық шылаулары анализденген мәтінде әрбір 679,53 сөз – бұл жалғаулық шылауларының фонетикалық варианттылығы.

Публицистикалық стиль бойынша *мен/бен/пен* жалғаулық шылауы – 141,81 сөз жалғаса, оның әрбір 177,06 сөзін – *мен* шылауы; 4 600 сөзін – *бен* шылауы; ал 808,44 сөзін – *пен* шылауы;

Ғылыми стильде 231,86 сөзді – *мен/бен/пен* жалғаулық шылауларының жалғануы арқылы құралса; ал әрбір 292,99 сөзді – *мен* шылауы, 6969,70 сөзді – *бен* шылауы, 1321,84 сөзді – *пен* жалғаулық шылауы;

Көркем әдеби 305,86 сөзді – *мен/бен/пен* жалғаулық шылауының жалғануы арқылы, оның ішінде әрбір 414,05 сөзді – *мен* шылауы, 7666,67 сөзді – *бен*, 3756,2 сөзді – *пен* жалғаулық шылауы жалғануы арқылы құралады.

Көріп отырғанымыздай, №2 кесте қорытындысында *мен/бен/пен* жалғаулық шылаулары публицистикалық стильде ең көп, яғни – 141,81 сөздерді құраса, ғылыми стильде – 231,86 сөзді, ал ең азы көркем әдеби тілінде – 305,86 сөз құрайды. Ал *және* жалғаулық шылауы ең жоғары индексті көркем стильде, яғни – 813,96 сөзді, БАҚ тілінде – 605,48 сөзді, ал ғылыми стильде – 93,21 сөзді құрайды екен. *Мен/бен/пен*, *және* жалғаулық шылаулары көркем стильге, БАҚ тіліне қарағанда ғылыми стильде аз сөз құрайды екен.

Жалпы сөз құрауы жағынан, яғни индекс қорытындысы бойынша публицистикалық стильде *мен/бен/пен* жалғаулық шылауы (141,81) *және* жалғаулық шылауынан (199,74) аз, ғылыми стильде де *мен/бен/пен* жалғаулық шылауы (231,86) *және* жалғаулық шылауынан (449,66) аз, ал көркем әдеби стилінде, *мен/бен/пен* жалғаулық шылауы (305,86), *және* жалғаулық шылауынан (774,42) сөздерді аз жалғайды екен, осының салдарынан *мен/бен/пен* жалғаулық шылауы *және* жалғаулық шылауынан сөз санын көп құрайды.

Сандық талдауға алынған үш мәтін (көркем стиль, ғылыми стиль, БАҚ тілі) бойынша барлық сөздегі анализденген жалғаулық шылаулардың жалпы салмағын жоғарыдағы №1 кестеден көруге болады.

*Мен/бен/пен* жалғаулық шылауларының көркем стиль, ғылыми стиль, БАҚ тілі мәтіндеріндегі таза салмағы: 1, 49% , ал *және* жалғаулық шылауы – 0,87% құрайды.

Талдауға алынған көркем әдебиет мәтінінде:

*Бұл кітапты қолыңызға алғанда ой мен қиялдан туған балдай тәтті тамаша роман оқиды екенмін деп, дәметіп қалмаңыз. Бала күнінде бірін-бірі жақсы көрген қыз бен жігіт артынан гашығынан көз жазып қалады. Сол себептен де екеуіміз тамақ пен киімнен басқа бар ақшамызды кітап жиюға жұмсайтынбыз. Міне кеше гана Қызылорда облысының Арал қаласы, Сейфуллин көшесі, №7 үйде тұратын Кеңес Өмірзақұлы деген көрмеген оқырманымнан және бір хат келді (Ә.Н.), деп жалғаулық шылаулары беріледі.*

Ал, ғылыми стильде:

*Фразеологизмдерді бұл әдіспен зерттеу нәтижесінде жеке тілдерге тән ерекшеліктер мен тіларалық фразеологиялық сәйкестіктерді жан-жақты зерттеуге мүмкіндік туады. Ұлттық әдебиетіміз бен рухани мұраларымызға да осы басылым арқылы қол жеткізілді. Оның үстіне айтыстың тууы, өмір сүруі, өнер табу, шеберленуі, өткірлік пен кесірлік табу тәжірибесі ұзақ гасырларға ұласатын болғандықтан, бұл саладағы*

асыл қазына да өте молшылық. Жалпы меңгеру байланысу формасының өзіндік ерекшелігі – бағыныңқы сыңардың барыс, табыс, жатыс, шығыс, көмектес септіктерінде тұрып, басыңқы сыңары етістік, баяндауыш, есім сөздерден болған баяндауыш және құрама баяндауыштармен сөз тіркесін құрауы (Ғ/с).

БАҚ тілінде:

Жоғары оқу орындарының өкілдері мен студенттік құрылыс отрядтары командирлерінің жиыны облыстық ішкі саясат департаментінде өтті. Біріншіден, бүгінгі күні жастарымыз өз тарландарымыз бен майталмандарымыздың елге, өнерге сіңірген еңбегі мен туындыларын жете білмейді. Сондай-ақ, мемлекеттік органдар мен үкіметтік емес ұйымдар арасындағы әріптестік қатынастардың тұрақты әрі тиімді жүйесін құру жөніндегі заңнамалық негіз жасау, шетелде тұратын отандастарға мәдени ағарту және тілдік қолдау көрсету, мәдениет, мұрағат ісі ақпарат және қоғамдық келісім саласындағы нормативтік құқықтық базаны жетілдіру, халықтың рухани-білім беру және зияткерлік мәдени деңгейін көтеру, ұлттық әдебиет пен жазба әдебиеттің көптеген ғасырлар бойғы бай тәжірибесін қорыту, әлемдік ғылыми ой, мәдениет пен әдебиеттің үздік жетістіктері негізінде мемлекеттік тілдегі әдебиеттің толыққанды қорын құру, тарих, археология мен сәулет ескерткіштерінің, тарихи орталықтардың сақталуын қамтамасыз ету мақсаттары көзделіп отыр. Шешімдердің мүмкін көп, жазуға болады және латын алфавитімен, бар болу қазақ, қырғыз болады, башқұрттық кириллицасымен (кириллица транслитерациямен диакритикалық белгілермен), ал пайдалануға болады және транскрипциямен (казновица) бұлар әдістер таласпайды аралық өзімен, қарапайым әрбір өз мақсаттарым олардан шешеді (Д/Қ), деп берілген сөйлемдерді мысалға келтіре кетуге болады.

Жалғаулық шылаулар өзара бір-бірімен ыңғайласып құралатын, тең дәрежедегі сөздер мен сөз тіркестерін жалғастыра отырып, солардың ыңғайласқан қатынасын білдіретін бұл тобына жоғарыда атап өткен, яғни біз зерттегелі отырған мен/бен/пен, және жалғаулық шылауларын жатқызамыз.

Талданған көркем стиль, ғылыми стиль және БАҚ тілі, көріп отырғанымыздай, көлемі жағынан анағұрлым айтарлықтай өзгеріс болмағанымен, сөз саны жағынан әр түрлі. Әсіресе БАҚ тіліндегі сөз саны айрықша көрініп тұр, бұл мәтіндегі стильдік ерекшеліктерге байланысты болуы мүмкін. Ал жалғаулық шылаулардың қолданылуы мүлде өзгеше. Бұдан біз, тағы да әр мәтін авторының өзіндік стиліне байланысты дейміз.

Мысалы, М. Әуезовтің «Абай жолы» роман-эпопеясының жиілік сөздігі бойынша мен жалғаулық шылауы – 2261 рет, пен жалғаулық шылауы – 560 рет, бен жалғаулық шылауы бір де бір қолданылмаған, ал және жалғаулық шылауы – 963 рет қолданылысқа түскенін аңғаруға болады [92].

Қорыта отырып айтарымыз, математикалық статистика кез келген лингвистің қолындағы көмекші құралы ретінде қолданылуы тиіс деген тұжырымға келеміз.

### **Пайдаланылған әдебиеттер:**

- 1 Бодуэн де Куртенэ И.А. Языкознание, или лингвистика XIX века // Бодуэн де Куртенэ И.А. Избранные труды по общему языкознанию. – М., 1963. – Т.2. – С. 168.
- 2 Ахабаев А., Бектаев Қ., Белботаев Лингвистикалық жқмысты автоматтандыру және тіл статистикасы// Қазақ ССР Ғылым академиясы Тіл білімі институты. – 1968 ж.
- 3 Аханов К. Тіл біліміне кіріспе. – Алматы: Санат, 2000. – 30 б.
- 4 Жұбанов А.Қ. Қолданбалы лингвистика: қазақ тілінің статистикасы: Оқу құралы. – Алматы: Қазақ университеті, 2004. – 209 б.; 22-б.
- 5 Головин Б.Н. Язык и статистика. – М.: Просвещение, 1971. – С. 18.
- 6 Алексеев П.М. Частотный словарь. – СПб., 2001. – 12 б.
- 7 Бектаев К.Б. Статистико-информационная типология тюркского текста. – М.: Наука, 1978. – 714 с.