

ПРОФЕССОР Қ. ЖҰБАНОВ ЕҢБЕКТЕРІНДЕГІ СТАТИСТИКАЛЫҚ ӘДІСТЕР

Садуакас Н.А.¹, *Кушкимбаева А.С.², Абдрахманова Ж.М.³, Асанов Ж.⁴
¹ф.ғ.к., доцент, Қ. Жұбанов атындағы Ақтөбе өңірлік университеті,

Ақтөбе, Қазақстан, e-mail: sadu_ngd@mail.ru,

*²PhD, Қ. Жұбанов атындағы Ақтөбе өңірлік университетінің доценті,

Ақтөбе, Қазақстан, e-mail: alitok@mail.ru,

³ф.ғ.к., доцент, Х. Досмухамедов атындағы Атырау университеті,

Атырау, Қазақстан, e-mail: azhm63@mail.ru,

⁴ф.ғ.д., профессор, С. Бәйішев атындағы Ақтөбе университеті, Ақтөбе,

Қазақстан, e-mail: Shalkar1900@mail.ru

Аңдатпа. Мақалада қолданбалы тіл білімінде тіл элементтерін есептеудің сандық мәселелері, тілді зерттеудегі статистикалық әдіс-тәсілдердің қолданылуы, қазақ тіл білімінде статистикалық лингвистиканың алғашқы қалыптасу кезеңіндегі зерттеушілердің еңбектері, қазіргі лингвостатистикалық зерттеулердің бағыты, т.б. туралы айтылған.

Ғылыми зерттеудің мақсаты, негізгі бағыттары мен идеялары: тіл білімінің барлық саласы бойынша өзінің ғылыми озық ойларын жазып, осы еңбектері арқылы кейінгі ұрпаққа талантымен танылған профессор Қ.Жұбанов еңбектерінде қазақ тілін зерттеуде статистикалық әдіс-тәсілдерді қолданудың алғашқы үлгілері көрініс тапқандығы нақты деректер негізінде баяндалған.

Статистикалық әдістерді қазақ тіліндегі дыбыстардың фонетика-фонологиялық мәнін таныту, фонемалар құрамын, әліпбиге алынатын әріп санын білу, сөз бен буын саны мен буын құрамындағы дыбыс санын есептеу, сөз мағынасының фонема саны мен оның сөз құрамындағы орнына қатысын, әр фонеманың сөз ішіндегі тіркесуі мен позициясының есебін анықтау үшін қолданғаны дәлелденген. Жұмыстың ғылыми-практикалық маңыздылығы да осында.

Зерттеу жұмысының негізгі нәтижелері мен талдауы, қорытындылары: қазақ тілі дыбыстарының тіркесім заңдылықтарын, буынға бөліну ерекшелігін, тасымалдану-тасымалданбау жағдайларын анықтау орфографияның басты қажеттіліктерінің бірі болып табылатынын ескерген ғалым Қ. Жұбанов жазу емлесін бір жүйеге түсіру мақсатындағы зерттеулерін статистикалық ізденістермен ұштастыра жүргізген.

Бір тілдегі сөйлемді екінші бір тілге аударуда кездесетін қиындықтардың бірі – сөздің мағынасын дәл беру, яғни, бір сөздің жеке тұрғандағы мағынасы мен сөйлем құрамында басқа сөздермен байланысқа түскендегі аударма мағынасының тура келе бермейтінін статистикалық әдістер арқылы көрсетілген деректерге тілдік талдау жасалған.

Бұл жұмыстың нәтижелері Сандық лингвистика, Қазақ тілі фонетикасы саласындағы сандық технологияға қатысты тақырыптарды өту кезінде пайдалануға болады.

Қ. Жұбановтың еңбектерінде қолданылған алғашқы статистикалық деректерді қазіргі сандық лингвистика пәндерін оқу барысында пайдаланып, тіл элементтерінің сандық мәселелерін, жиілік сөздіктердің түрлерін, кері әліпби-жиілік сөздік, сөзнұсқағыш әліпби-жиілік сөздік, мәтіндегі грамматикалық бірліктер санын есептеу, компьютер көмегімен сөздердің алгоритмін жасау, қазақ мәтіні бойынша түзілген жиілік сөздіктерге сандық есептеулер сипаттамасын беру, т.б. жұмыстарды практика жүзінде жасай білуге қол жеткізіледі.

Тірек сөздер: тіл білімі, статистикалық әдістер, дыбыс саны, әріп саны, буын, сөз саны, сөз тіркесі, сөйлем.

Негізгі ережелер

Қазіргі тіл білімінде статистикалық лингвистика бойынша зерттеулер болғанымен, осы саладағы алғашқы ізденістерді анықтау, тілдік бірліктерді сандық есептеулермен байланыстыра зерттеудің жолдарын көрсету, тілді зерттеуде математикалық есептеулердің артықшылықтарын таныту және профессор Қ. Жұбанов еңбектеріндегі статистикалық әдістерге талдау жасауға назар аударылды. Қазақ тілін зерттеуде Қ. Жұбанов еңбектерінде көрініс алған алғашқы статистикалық әдістерді нақты деректер арқылы көрсету, ғалым қолданысындағы статистикалық зерттеудің негізгі бағыттарын анықтау және тілдік категорияларды теориялық тұрғыдан кеңірек танытуда сандық өлшемдерді пайдалануға талдау жасау мақсат етілді.

Кіріспе

Тілді зерттеу математикалық есептеулермен де байланысты болып, тілдегі коммуникациялық ақпараттар түрлі есептеулер мен сандық деректерге де мүдделі болып табылады. Осы қажеттіліктер статистикалық лингвистиканы дүниеге әкелгенін білеміз. Статистикалық лингвистика тілдің грамматикалық категорияларын, тілдегі барлық деректерді сандық тұрғыда тексеруді, тіл элементтерін өлшеудің сандық мәселелерін қарастыратын қолданбалы тіл білімінің бір жаңа саласы болып табылады. Тілдегі түрлі жиілік сөздіктер, сөздерді компьютер арқылы алудың алгоритмін жасау, т.б. да жұмыстар сандық есептеулер арқылы жүзеге асырылады. Ауа райы болжамында, табиғат құбылыстарында, тарихтың сан-алуан кезденінде, салт-дәстүрді де көркем тіл құдыретті арқылы сипаттайды [1, 46 б.]. Сол сияқты тіл деректері есептеу, сандық мәселелермен де байланысты қарастырылады. Қазіргі тіл білімінде статистикалық лингвистика саласында тілдегі сөз тұлғаларының құрамын сандық есептеу әдісімен зерттеп, оның формалды жақтарын айырып, сөз бөлшектерін сегментті ерекшеліктеріне қарай бөлшектей қарауға, сөз құрамын автоматты түрде талдау ісіне ерекше назар аударылады. Қазір қазақ тіл білімінде статистикалық лингвистика бойынша зерттеулер қарқынды түрде дамып келеді. Осы бағыттағы зерттеулердің алғашқы бастаулары А. Байтұрсынұлы мен Қ. Жұбанов еңбектерінен көрініс табады. Қазіргі тіл білімінде статистикалық лингвистика, лингвистикадағы модельдеу бағыттарының кейбір алғашқы бастамалары профессор Қ. Жұбанов еңбектерінде сандық деректер мен сапалық өлшемдерді таныту мақсатында қолданыс тапқаны кейінгі зерттеушілер тарапынан жиі айтылып келеді. Қолданбалы лингвистиканың осы аталған салаларының басын қосатын корпуслық лингвистика, яғни, компьютерлік технологияларды қолдана отырып, тілдік мәліметтердің лингвистикалық корпусын құрастыру мен оны қолданудың жалпы принциптерін зерттеумен айналысатын сала бүгінде қарқынды дамып келеді [2].

Материалдар мен әдістердің сипаттамасы

Осы аталған мақалада профессор Қ. Жұбанов еңбектеріндегі статистикалық әдістер жөнінде жан-жақты мәлімет беріп, лингвистикалық тұрғыда талдау жасау барысында зерттеудің сипаттамалы және тарихи әдістері негізінде қажетті деректер сұрыптап алынып, жинақталған деректер бойынша статистикалық әдістермен қатар, салыстыру, талдау, қорыту әдістері де қолданылды. Қазіргі тіл білімінде статистикалық әдістердің күрделі түрлері қолданылатын болса, алғашқы шыққан еңбектерде сандық өлшемдерді қолдану әдістері қарапайым мысалдармен көрсетілгеніне сипаттама берілді.

Статистикалық әдістер қазір тіл үйрету барысында да ескеріледі, сонымен қатар, тілді үйрену сол тілдің сөздік құрамындағы барлық лексиканы түгелдей білумен өлшенбейді. Бастапқыда үйренетін тілдің сөздік құрамындағы ең қажетті лексикалық минимумдарды сұрыптап, таңдап ала білу керек. Лексикалық минимумды топтастырудың өзіндік қағидаларының біріне статистикалық принциптер жатады десек, сандық есептеулер сөздердің қолданылу жиілігі мен таралу өрісін негізге алатын сандық мөлшерге қатысты болса, әдістемелік қағидалар оқыту бағдарламасы мен алынған тақырыптарға сүйенеді [3, 165 б.].

Талқылау және нәтижелер

Алғашқы болып ана тіліміздегі сөздерді қолдана отырып, қазақ тіл дыбыстарына ғылыми зерттеу жасаған Ахмет Байтұрсынұлы екен. Ғалымның еңбектерінде тілдің дыбыс жүйесі, дыбыстық құрамы, әртүрлі ерекшеліктеріне қарай олардың топтастырылуы, әріп, әліпби және жазу мәселелері жоғары деңгейде талданған. Зерттеуші 1920 жылы қазақ мұғалімдеріне арналған Баяншы деген еңбегінде ана тілімізде барлығы жиырма төрт дыбыс бар екенін айтып, “Оның бесеуі дауысты, он жетісі дауыссыз, екеуі жарты дауысты”, – дейді [4, 479 б.]. Бұл жерде дауыстылардың жуан түрін ғана көрсетіп, ал дауыссыздардың төл тілімізге қатыстыларын ғана алып, у мен и табиғаты жарты дауысты деген атаумен таныстырылған. Бес дауысты деп оның е-ден басқаларының тек жуан түрін көрсеткенін, ал жіңішкелерін дәйекшемен танытатынын есептегенде, осы бес жуан-жіңішкелі үні басым айтылатындар, дауысты дыбыстар, саны тоғыз болады. Сонда барлық дауысты-дауыссыздар санын қоса есептегенде, қазақ тіліне тән тоғыз дауысты, он тоғыз дауыссыз, барлығы жиырма сегіз төл дыбысымыз бар екенін көрсеткен. Бұдан А. Байтұрсынұлының тіл дыбыстарының құрамы мен фонемалық мәнін түсіндіруде санмен есептеп көрсетуге ерекше мән бергенін байқаймыз.

Ғалым қазақ тіліндегі осы 24 түрлі дыбыстың фонемалық сипаты бар екенін төмендегідей сандық есепке құрылған мысалдар арқылы көрсетіп, бұл дыбыстардың ішінде қ һәм ғ ылғи жуан айтылады, к, г һәм е ылғи жіңішке айтылады. Өзге 19 дыбыстардың һәр қайсысы бірде жуан, бірде жіңішке айтылады. Егерде бұл 19 дыбыстың жуан айтылуы үшін бір белгі, жіңішке айтылуы үшін бір белгі керек десек, онда 19 дыбысқа 38 белгі керек болар еді. Бұған дайым жуан айтылатын қ мен ғ-ны қосқанда, дайым жіңішке айтылатын

к, г, е- лерді қосқанда, бәріне 43 белгі керек, – дейді [4, 479 б.]. Қазақ тілінің сингармониялық табиғатына сай сөйлесім кезіндегі қасиеттерін жазылым үдерісінде дұрыс таңбалау жолын тауып, 43 түрлі дыбысты толық қамтып, оларды дәйекшіні қосқанда 25 таңбамен белгілеп, қазақы жазу техникасын ойлап тапқан Ахмет Байтұрсынұлының ғылыми құнды жаңалығын алғаш рет ғылыми ортада ғалым М. Жүсіпұлы өз зерттеуінде дәлелдеген еді. Қазақтың сингармониялық сөйлесімінде 24 әріптен және 1 дәйекшіден тұратын алфавиттегі таңбалар саны жеткілікті [5, 22 б.]. Осы деректер арқылы дыбыс пен фонеманың ортақ және айырым белгілерін тереңірек түсіндіру үшін А. Байтұрсынұлы статистикалық әдістерді қолданып, сандық есепке жүгінген.

Өзінің ғылыми зерттеулерінде профессор Қ. Жұбановта сандық есепке құрылған статистикалық әдістерді өз орнымен қолданған. Мәселен, Қ. Жұбанов қазақ тілінің фонетикасына арналған еңбегінде дыбыстың саны мен сапасы туралы түсіндіру барысында ақындардың өлең жолдарынан да мысалдар алып отырған. Мысалға, Ілияс Жансүгіровтың “Қойшы ойы” деген он екі жол өлеңінде төрт жүз сексен сегіз дыбыс бар екенін айтып, бірақ мұнда осыншама түрлі дыбыстың жоқ екенін, бұл жерде не бары жиырма сегіз ғана түрлі дыбыс бар екенін мысалға келтірген. “Өлеңде келтірген сөздердегі әртүрлі дыбыс саны жиырма сегіз болса, олардың “бірі – 10, бірі 50 рет қайталап барып 488 болған”,- дейді [6, 172 б.]. Осы мысалында мәтіннің көлемін, яғни, он екі жол екенін, ондағы дыбыстың барлық саны мен дыбыс түрлерін ажырата көрсету арқылы олардың мәтінде қайталанып қолдану жиілігі туралы да ой қозғаған. Ғалымның бұл талдауларынан қазіргі статистикалық лингвистиканың амал-тәсілдерін, дыбыстың, сөздің қолданылу жиілігін, сан мен сапаның айырмашылығын анық көреміз. Фонема аллофондары туралы теориялық түсініктерге жүгінетін болсақ, саны 28 түрлі болып тұрған дыбыс фонологиялық қызметіне байланысты мәтіндегі қолданысы барысында саны 488 болғанымен, олар не бары 28 түрлі фонема екенін түсіндірген. Бұл жерде тіл дыбыстарының фонемалық қасиеті мен фонологиялық қызметі деген түсінікті қалыптастыру мақсатында статистикалық әдістердің тиімділігі анық байқалады.

Осындай есеп-қисаптан кейін, Қ. Жұбанов мынандай тұжырым жасаған: Қазақ тілінде де сапасы әртүрлі болатын дыбыстардың арнаулы саны бар [6, 172 б.]. Бұл деректер қазақ тіл біліміндегі статистикалық әдістің профессор Қ. Жұбанов еңбектерінде мақсатты түрде ерекше білгірлікпен қолданылғанын айна-қатесіз дәлелдейді. Бұл жөнінде тіл білімі жүйесін математикалық есептеулермен байланысты зерттеудің алғашқы маманы, профессор Қ.Б. Бектаев естелік мақаласында: Бүгінде лингвистикалық статистика деп аталып жүрген тіл білімінің жаңа саласының алғашқы нышандары мен белгілері. Қазақ тіл білімінің бұл саласының бастауында да проф. Құдайберген Жұбанов тұрды. Осы бағытта қазақ тіл біліміндегі көптеген мәселелердің Әліппесін ғалым өзі бастап, өзі жазып кеткен болатын. Осылайша, біз сөз етіп отырған тілді зерттеудегі сандық мәліметтер мен деректерді пайдалану әдісі де аты аталып, түсі түстеліп, арнайы түрде қарастырылмағанмен, ол жөнінде жанама болса да сөз етіліп, белгілі бір игі мақсаттың жолын дәл мегзеп берген еді,-деп

жазған болатын

[7, 109-110 бб.]. Осы айтылғандай, әліпби құрастыру үшін алдымен тілдегі фонемалар құрамын анықтап, одан кейін жазу барысындағы емле ережелерін дұрыс көрсету үшін де дыбыстардың жазылымдағы фонологиясын нақты сандық деректер арқылы талдап алуға мән берген. Фонетика мен фонологиялық танымның арасындағы айырым белгісін түсіндіру үшін де сандық есептеулерге жүгінеді. Дыбыс пен мағына арасындағы байланыс пен олардың ерекшеліктерін танытудағы Қ. Жұбановтың өз сөзін келтірсек, ол сөздегі дыбыстардың біреуі көп не аз, не саны өзгеше болып тұрса, немесе дыбыстар саны мен түрі бірдей болып келсе де, “дыбыстарының тек тізілу тәртібі басқа болса да, сөздер бірінен-бірі басқа болады”, -дейді [6, 171 б.]. Зерттеуші фонема атауы мен оның ерекшелігін таныту үшін бұл ойын одан әрі тереңдете түсіп, дыбыстың сапасы мен санының арасындағы өзгешеліктерді, сапасы бірдей болса да, саны әлденеше бола беретінін де сандық мысалдармен баяндайды. Осы сөздерін зат-сат, ат-сат, сат-тас деген сөздердегі дыбыстардың саны мен орналасуындағы ерекшеліктерін, баба, бөле деген сөздерде төрт-төрттен дыбыс болса да, бірінде сапасы екі түрлі, екіншісінде сапасы төрт түрлі дыбыс бар екенін айтады. Дыбыстардың өзіндік ерекшелігі және оның орналасуы, яғни, тіркесу қабілеттері де арнайы сандық есептеулермен де ұштасып жататынын байқауға болады. Қазіргі кезде қолданбалы және теориялық тіл білімінде көрші дыбыстардың бір-біріне ықпалын анықтауда міндетті түрде статистика-комбинаторлық әдістер арқылы талдау жасалады [8]. Мұндай әдістерді Қ. Жұбанов та өз зерттеулерінде тиімді жолдардың бірі ретінде қолдана отырып, тілдегі сөзден келесі сөздің өзгешелігі оның ішкі құрамын құрап тұрған барлық дыбыстың бірдей келуі, не сапасы басқаша болуы міндет емес, сондағы тек бір ғана дыбыстың өзге сапада болғаны, сондай-ақ, нысанаға алған сөз бен екіншісі құрамындағы дыбыс саны, дыбыстардың көрші тұрып, тіркесу тәртібі басқаша болуы жеткілікті, олардың екі бөлек фонема болуын танытатынын айта келіп, “яғни, тіл дыбыстарының табиғатын тани білуде фонология заңдылықтарына сүйенген” [9, 50 б.]. Ғалымның осы секілді дыбыс табиғатына байланысты айтқан пікірлерінде статистикалық әдістерді ретті жерінде орынымен қолдана біліп, дыбыс заңдылықтарын мақсатты түрде сандық есептер арқылы баяндау жолын көрсеткен деп білеміз. Қ. Жұбанов ана тіліміз үшін орфография мәселесіне ерекше қарап, осы жазу заңдылықтарын таныту үшін сандық деректер арқылы терең түсіндіруге болатынын байқатып, келгенін, жазу емлесін ғылыми тұрғыда қалыптастыру, дамыту мақсатында қолданған. Әліпбидегі дыбыс санын есептеуге, жазуда өз тілімізге тән дыбыстар санын дәл анықтауға, қазақ тіліндегі барлық фонеманы өз орнында қолдану үшін олардың санын да есептеп отыруға, іс жүзінде әр дыбыстың қажеттігіне, қолданыс жиілігін белгілеуге мән берген екен. Қ. Жұбанов өзі редакторлық еткен мемлекеттік терминологиялық комитетінің №4 Бюллетенінде емле туралы жарияланған мақаласында бұрынғы әріптермен жазғанда 1532 әріп-таңбадан тұратын мәтінді мысалға алып, оны әліпбидің жаңа жобасы бойынша жазғанда тиімділігі қысаң ы, і 126 жерде, ә, ө 33 жерде кем жазуға болатынын есептеп шығарған [10, 30 б.]. Бұл да

статистикалық әдістің тиімді тұстарын көрсетеді. Ғалым өзінің айтпақ ойын осындай цифрлық деректермен байланыстыра түсіндіруі арқылы тілдік талдауды статистикалық деңгейге жеткізуге ұмтылған.

Тіл теориясы, қай тіл болмасын, оның дыбыс жүйесіндегі фонемалар санын нақты анықтауды қажет етеді. Фонемалардың саны мен сапасын анықтау – әліпби құрастыру жұмысында маңызды мәселе, сондықтан сандық есептеу әдісі – барлық тілде де қолданылатын зерттеу әдістерінің бірі болып табылады. Мысалы, орыс тілінде 5 дауысты (и-э-а-о-у) бар, ал ы фонема емес. Орыс тілінің сөйленісінде и-ы дыбыстары бірдей фонетикалық қоршауда кездеспейді, и тек жіңішке дауыссыздардан кейін ғана келсе, ал ы, керісінше, тек жуан дауыссыздардан кейін ғана келеді. Басқаша айтқанда, бұлар бір фонеманың варианты болатындықтан и болатын жерде ы кездеспейді, ал ы болатын жерде и болмайды. Сондықтан орыс тіліндегі и фонемасының жуан дауыссыздан кейін келетін варианты ы болып табылады [11, 19 б.]. Тілдегі фонема санын оның аллофондарынан ажырата тану керек десек, олардың сандық құрамын да біліп, статистика жүргізудің мәні зор деп есептейміз. Әр тілдің фонемалық құрамы да әртүрлі болады. Мысалы, ағылшын тілінде 26 әріп, оның ішінде 6 дауысты, қалғаны дауыссыздың таңбасы болса, барлық фонема саны – 44, оның ішінде 20 дауысты, 24 дауыссыз бар [12, 11 б.]. Ал осы ағылшын тіліндегі 20 дауыстының 12-сі – жалаң дауыстылар, 8-і – дифтонгтар болып табылады [12, 21 б.].

Қ. Жұбанов өз кезіндегі 33 әріптен тұратын латын графикасына негізделген қазақ әліпбиін жетілдіру барысында 25 әріп пен 1 белгіден тұратын өзі құрастырған жаңа әліпбиін ұсынған екен [6, 495 б.]. Міне, осы фонемалардың сапалық құрамымен қатар сандық есебі де тілдің теориялық ерекшелігін таныта түсуге септігі бар. Сондықтан сан мен сапаны қатар ескеруі арқылы зерттеуші Қ. Жұбанов статистикалық тұрғыдағы әдістердің қажеттілігін, оны қолданудың дұрыстығын біліп, көрегендікпен танып, цифрлық әдістердің болашағын да терең сезінген деп есептейміз.

Фонетист Қ. Жұбанов өз тіліміздің ерекшеліктерін, яғни фонологиялық табиғатын терең түсініп, дұрыс түсіндіре білген ғалым еді. Сөйлеу дыбыстарының сөз бен буын ішіндегі орналасуын қарастырған, кезінде ұ-у, ү-у, ы-й, і-й қосынды дыбыстар екендігін көрсеткен, осы қосар тіркестерді қолданыста жалаң түрде бір ғана у және и әріптерімен таңбасымен беру арқылы жазу сауаттылығын арттыруға шақырған ұсынысын да кейіннен қабыл алғанымызды, жазуда осы көзқарас басшылыққа алынып келгенін де жоққа шығара алмаймыз. Бұл ұлттық жазудың алғашқы балаң кезеңі үшін тиімді болған еді. Сауатты жазу үшін сол кезеңдегі алғашқы латын таңбасын иеленген әліпбиімізде дауыстылар таңбасы 13 әріппен (17 таңба) арқылы белгіленген, ғалым енді оны (ұу, үу, ый, ій әрбіреуі 2 әріптен, яғни 8 таңбадан тұратын тіркесімдерді) жазу барысында бұрынғыдан да қысқартып, тіліміздің дұрыс таңбалануына кедергі келтірмей, дыбыс пен әріпті бірдейлестіріп таңбалауға болатындығын орысша мақаласында “7 букв, 7 знаков” деп нақты статистикалық есеппен көрсеткен [6, 494 б.].

Статистикалық мәселелер, Қ. Жұбанов еңбектерінде ескерілгендей, қазіргі әліпби құрастыру барысында да көңіл аударуды қажет ететін маңызды жұмыстардың қатарында болып қала береді. Мысалы, қазақ әліпбиін бұрынғы жазуын латын таңбасына өткізу мәселесі қозғалғанда әліпбидің бірнеше нұсқасы ұсынылғанын білеміз. Сол әліпбилердің алғашқы нұсқасында әріп – дыбыс бірде-бір деген ұстанымға сай келмей, әсіресе, қазақ тіліндегі төл дауысты дыбыстардың қосарланған әріптермен (диграфпен) берілген болатын. Батыс елдерінің, яғни роман, герман, славян тілді елдердің тілдері флективті тілдерге жатады. Олардың тілінде сөздің құрамы да, сыртқы тұрқы да қысқа келіп, түрленіп айтылуға бейім келеді. Бұл тілдердегі сөздер сөйлеу барысында қанша өзгеріске түсіп, дыбыс құрамы өзгерсе де, құбылса да, сыртқы көрінісі қысқа күйін сақтайды. Солай болғандықтан мұндай құрылымды тілдерге дыбыс пен таңба саны бір-біріне тең келуі міндетті болмай, оларда бір дыбыс және олар екі таңбамен де берілетіндіктен, жазуда диграф ұстанымы кедергі болмайды. Осы топтағы тілдер жазуда оп-оңай қосар әріпті пайдалана береді. Ал өзіндік табиғаты, құрылымы жалғамалы болып келетіндіктен, қазақ және басқа түркі тілдерінде сөздің түбірінен кейін қосымшалар көп жалғанады, сондықтан қосымшалар арқылы сөздің сырты да созыла түсіп, мағынасы да өзгереді. Егер жазу кезінде жалғыз дыбысты қос әріппен, диграфпен, алатын болсақ, сөздің сыртқы құрылымы да өсіп, әріп саны көбейе түседі. Бұндай өзгерістер ана тіліміздің дұрыс жазылуына кедергі келтіріп, жазуға ауыртпалық түсіруі мүмкін деп есептейміз. Сол себепті әліпби құрастыруда дыбыс - әріп байланысының статистикалық жағына да назар аударылып отыр. Осы орайда, ғалым А. Шәріпбай өз зерттеуінде қазақ әріптері мен әріп тіркестерінің жиілігін қазақ жазуын латын графикасына көшірудегі емле ережелеріне қажетті статистикалық мәліметтер ретінде негізге алған [13].

Зерттеудің құндылығы. Қазіргі таңда қазақ тіл білімінде статистикалық лингвистика талаптары бойынша тілдегі сөз тұлғаларының құрамы мен оның формалды жақтарын есептеу негізінде айырып зерттеу жүйелі түрде іске асырылып келеді. Қазақ тілі құрылымы жағынан агглютинативті тілдер қатарына жататындықтан оның сөз байлығын лексикалық және морфологиялық талдау барысында әр бөлшегін сегменттік ерекшеліктеріне қарай жіктеп отырамыз. Сондықтан қазақ сөздерін буын жігіне қарай сегменттеу, қазақ тіліндегі буын жүйесінің фонетикалық және фонологиялық ерекшеліктерін білу маңызды болып табылады. Сөз құрамындағы буын санын есептеп, олардың аражігін ажырата білу, сандық құрамы мен әр бөлшектің грамматикалық мағынасы мен тілдік қызметін тануымыз білім беру ісінде, оқулықтар жазуда да маңызды болып табылады. Бұл жұмыстың нәтижелері Сандық лингвистика, Қазақ тілі фонетикасы саласындағы сандық технологияға қатысты тақырыптарды өту кезінде пайдалануға болады. Буын тақырыбын қарастырған кезде де Қ. Жұбанов статистикалық әдістерді өз зерттеулерінде жүйелі түрде пайдаланғанына назар аударып көрейік. Ғалымның 1934 жылы жазған “Буын жігін қалай табуға болады?” деген еңбегінде қолданған әдістерін статистикалық теорияның алғашқы базасы деп

білуге болады. Буын жүйесін тіліміздегі сөздердің қолданысын, олардың көркемдігін тани түсу үшін де маңызы ерекше деп есептейміз. Ғалым буынды жете зерттеуге алуының себебі де оның қызметі қазақ тілінің ерекшелігін танытатындығын көрсетуге негізделген еді. Буынның қызметі мен жалпы құрылымын бірлікте қарай отырып, буынды білмесе, тілді де терең игере алмайтынымызды түсіндірген. Буын арқылы жазу емлені де жақсы білуге болады, жалпы морфологиялық заңдылықтар да буынмен байланысы бар екенін, сол сияқты сөз екпін, интонация секілді сөйлем жүйесіне қатысты тіл құбылыстарын да жете білуде буынның маңызын түсіну керектігін жазып кеткен. Ажарлап сөйлеу, әсерлеп оқу сияқты сабақтарды да дұрыс буын жайына қанық болғанда терең түсінеміз деген ой білдірген. Өлең құрылысын да буынмен өлшеп, сол арқылы мәнерлі жеткізуге болатынын айтқан. Қазақ сөздерінің де табиғаты осыған қатысты болатынын айта келе, “буын жүйесін жете білу керек”, – деген еді [6, 349 б.].

Қ. Жұбанов қазақ сөздеріндегі буынның алты түрін көрсетіп, оларға теориялық анықтама беріп, мысалдар арқылы сипаттама жасап, осы буындардың өзара жігін айырудың өзіндік формалды әдісін көрсеткен. Ғалым қазақ тіліндегі буындардың алты белгісін анықтап, қазіргі ғылыми ортадағы компьютерлік лингвистика тілімен айтқанда, буындардың жігін айырудың алгоритмін жасап берген еді. Қ. Жұбанов қазақ тілінің төл сөздерінде бір буынның құрамында дыбыстардың саны төрттен аспайтынын да атап көрсеткен. Сөздің соңынан бастап, кері қарай санағанда ұшырасатын әрбір дауысты дыбыстың алдында тұрған дауыссыз дыбыс қосылып барып, буын жігіне ажыратылуы тиіс деп есептеген. Ғалым дауыссыз дыбысты буынның жігін танытатын, буын жігін нұсқайтын сегмент ретінде танытқан. Сонымен бірге Қ. Жұбановтың қазақ сөзінің сегменті, яғни, сөздің буын жігін ажыратуды сөздің соңында тұрған буыннан бастауды ұсынғанын да атап айту керек. Осы деректерге сүйене отырып, ғалым зерттеулерінде буын құрамын, ондағы дыбыстар санын сапасына қарай ажырата отырып, буын құрылымын алты түрге бөліп, статистикалық әдіспен талдап түсіндірген. Оның осы алты түрлі буында санамалап тұрып белгілеу әдісі қазіргі қолданбалы тіл білімінің жаңа саласының статистикалық талаптарына сай келеді. Қазақ тіліндегі буын құрылымдарының (Д-ты; Д-сыз+Д-ты; Д-ты+Д-сыз; Д-сыз+Д-ты+Д-сыз; Д-сыз+Д-ты+Д-сыз+Д-сыз) алты вариантын шартты белгілермен таңбалау алгоритмін жасаған деуге болады [6, 87 б.]. Ғалым Қ. Жұбановтың Буын жігін қалай табуға болады? мақаласының ғылыми әрі практикалық маңыздылығы – тек сол өз кезеңі үшін ғана емес, қазіргі кезде де, ертеңгі күн үшін де қажет екендігі. Қазақ тіліндегі буын жігін табу, буынға бөлу, буын құрылымына фонетикалық сипаттама беруде профессор Қ. Жұбановтың осы мәселеге қатысты ғылыми талдаулары бүгінгі күні де өзекті болып тұрса, оның сан мен есепке құрылған анықтамалары қазіргі статистикалық талдаулармен ұштасып жатқанын да ескеруіміз керек. Қазіргі таңда санды лингвистикада буынның цифрлық мәнін жете зерделеу қажет болып отыр. Буын теориясы – фонетикадағы ең күрделі мәселелердің бірі болып табылады, ал оның корпустық лингвистикаға қатысты жағы одан да маңызды деп білеміз. Қ.

Жұбановтың статистикалық есепке құрылған теориялық талдаулары буын модельдерінің түрлерін анықтап, буындарды таңдау алгоритмдерін жасауға негіз болды. Бұл буынның бөліну алгоритмдері тек қазақ тілінің ғана емес, сонымен бірге туыстық түркі тілдерінің де барлық сөздерді автоматты құру жүйелерінің негізін құрайтын болып табылады [14, 60 б.]. Өз ана тілімізді зерттеу тарихында цифрлық деректерге көңіл аударған тұңғыш маман Қ. Жұбановтың қалдырған мұрасы туралы зерттеушілер айтқан пікірлерге қарағанда, ол тілді есептеу, санау сияқты амалдар арқылы зерттеуге жүгінгенін, есептеу мен тіл заңдылықтарын бір мақсатқа пайдалану арқылы есеп-санаққа ерекше мән берген. Оның осы мәселеге қатысты еңбектері “ретті жерінде математиканың тәсілдеріне сүйенгендігіне” дәлел бола алады [15, 23 б.].

Тіл дыбыстарының акустикалық табиғатын, өзіндік ерекшеліктерін бір-бірінен ажырата көрсету үшін ғалым оларды бір-бірінен арасын ажырату үшін кесіп тастау қиын екенін айтып, оны есеп сандарын мысалға алып түсіндіреді. Ол сандарды ретімен санайтынымызды, дыбыстарды айтуда да олардың осы секілді тәртібі барын айтып, санды ретімен “тәртібін сақтамақ болғанда 3,1,4,2,5 деп жазбайтынымыз сияқты”,-деп түсіндірген [6, 64 б.]. Одан әрі дыбыстың жеке сапасы мен қызметі, оның сөз ішіндегі тіркесі, орнына қарай сөз мағынасын да өзгертуге себі тиетінін, “семсер деген сөздегі дыбыстар саны – 6, сапасы – 4”,-деп сандық есептеулер арқылы талдаған [6, 65 б.]. Сандық есептеулер статистикалық зерттеулердің маңызды бөлігі десек, осы әдістің алғаш тіл саласында қолданылуына ғалым Қ. Жұбановтың да үлесі бар деп есептеуге болады.

Статистикалық әдістер қазір тіл үйрету барысында да ескеріледі, сонымен қатар, тілді үйрену сол тілдің сөздік құрамындағы барлық лексиканы түгелдей білумен өлшенбейді. Үйренетін тілдің сөздік құрамындағы алғашқы кезекте ең қажетті лексикалық минимумдарды сұрыптап, таңдап ала білу керек. Лексикалық минимумды топтастырудың өзіндік қағидаларының біріне статистикалық принциптер жатады десек, осы сандық есептеулер сөздердің қолданылу жиілігі мен таралу өрісін негізге алатын сандық мөлшерге қатысты болса, әдістемелік қағидалар оқыту бағдарламасы мен алынған тақырыптарға сүйенеді [3, 165 б.]. Профессор Қ. Жұбанов тілді терең меңгеру үшін оның тарихына да терең үңілуге шақырады, бұған көп дәлелдер келтіре келіп, “арабтың мән деген сөзі, орыстың меня дегені, қазақтың мен, өңге түркілердегі бән, біз, бәрі дегендерді айырып тастау қиын” – дейді [6, 86 б.]. Сонымен бірге сөздерді таптастыру мәселесінде де жауаптылықпен қарау керектігін, әр сөз табының өзіндік ерекшелігін ескермей, оларды бірдей қарауға болмайтынын айтып, бұған “кісі” мен “үшін” сөзінің біреуі түбір, біреуі шылау болса, оларды қатар қойып талдау, сөздің бір табы демеу-деу – қарыс пен құлашты қосып санағандай болып шығатынын айтады. Егер “2 құлаш пен 3 қарыстың жиыны 5-еу дегенмен бір есеп”, “бір сом мен 50 тиынды қосып 51 шығарғанмен есеп бола ма?” дейді [6, 130 б.]. Осы мысалдарды санмен көрсету арқылы теориялық мәселелерді түсіндірудің тиімді жолы ретінде сандық есептеулерге жүгінгенін байқаймыз.

Қ. Жұбанов сөйлемдегі сөздердің орын тәртібін көрсету үшін де санмен есептеу арқылы түсіндіруге жүгінген. Ол сөйлем мүшелерін санмен белгілеп, олардың байланысуын сандардың орын ауыстыру арқылы танытады. Ғалым сөйлем мүшелерінің орын-орны туралы да нақты айтып кеткен. Сөйлемде баяндауыш соңында келсе, ал оның бастауышы өзімен қиысатын баяндауышынан бұрын тұратынын айтады. Сол сияқты анықтауыш, толықтауыш пен пысықтауыш та өздерінің байланысатын сөздерінен бұрын келеді, “анықтауыш өзінің анықтаған сөзінен, толықтауыш өзінің толықтаған сөзінен бұрын келеді”, - деген анықтама береді [6, 151 б.]. Қ. Жұбанов статистикалық тұрғыда төмендегі сөйлемді санмен белгілеп көрсетеді: Бұл (1) – мал(2) күткен(3), малға(4) ықылас(5) берген(6) бейілдің(7) көрмесі(8). Осы сөйлемдегі сөздердің байланысы төмендегіше көрсетілген: Бұл(1) – көрмесі(8); малға(4) ықыласты(5); бейілдің(7) көрмесі(8); берген(6) бейілдің(7); күткен(3) бейілдің(7); Малды (2) күткен(3); малға(4) берген(6); ықылас(5) берген(6); Осы сөз тіркестерінің санмен берілуі сөз тіркесі мен сөздердің байланысу ерекшелігі туралы терең түсінік береді. Ғалым Қазақ тіліндегі сөздердің тіркесу түрлері туралы деген мақаласында да 15 сөзден құралған сөйлемді алып, оның ішінде қосымша жалғанған 3 сөз, қалған 12 сөз түбір не туынды түбір күйінде тұрғанын айтып, төмендегі сөйлемді әр сөзін санмен белгілеп көрсетеді: “Сен(1) көрген(2) қымбат(3) алпауыттан(4) да(5) Қызыл(6) Құрт(7) Ертай(8) ұстаның(9) Темірбайдан(10) алған(11) торы (12) бесті(13) аты(14) жақсы(15)” [5, 450 б.]. Осы сөйлемді орыс тіліне аударған кездегі сөздердің орын тәртібінің өзгерісін де санмен көрсетіп, қазақ тіліндегі сөздің мағынасы мен тіркесімі басқа тілде өзгеше болатынын, сондықтан оны басқаша құрастыруды қажет ететінін бір-бірінен өзгеше 5 түрлі сөйлем етіп құрастырып көрсеткен [6, 455 б.]. Қ. Жұбанов өзінің статистикалық ізденістерінде осы сөйлемдердегі сөздердің басқа тілде орындары өзгеріп кететіндігін сөздеріне қарап жатпай-ақ, сандардың тіркесу реті араласып кеткенінен-ақ байқауға болатынын дәлелдеген. Сонымен қатар аударма жұмысында да сөздің баламасын таба білуде статистикалық нақтылық қажеттігін байқауға болады. Жоғарыдағы сөйлемді орыс тілінің ыңғайына қарай аударғандағы сандардың тізілу ретінің өзгерісі төмендегіше болған: “Чему(4) дорогой(3) помещику(4) виденный(2) тобою(1), лучше(15) и(5) гнедой(12) пятилетний(13) мерин(14), полученный(11) кузнецом(9) Ертаем(8) кызыл-куртовцем (6) от (7) Темирбая(10)” [6, 451 б.]. Аударма барысында бір сөздің бірнеше баламасын көрсетіп, құрт сөзін орысша курт, творг, червь сөздерінің қайсысына сәйкес келетінін, ал қымбат сөзіне бұл жерде орыс тіліне аударылғандағы дорог, дорогой, дороже сөздерінің қайсысы дәл келетінін талдап түсіндірген. Ғалым мақаласы соңында сөздердің орын тәртібі мен сөз тіркесінің маңыздылығын, осы мәселені арнайы зерттеу керектігін айтып, осы ойын әр сөзді санмен белгілеп, статистикалық әдістер арқылы дәлелдей білген.

Құдайберген Жұбанов өзінің “О построении речи на казахском языке” еңбегінде сандық ұғымдарға жүгініп, мысалға 16 сөз қолданыстан құралған сөйлемді келтіріп, олардың формалды белгілерін көрсетіп, әрбір сөзге тікелей қатысты параллель тіркестердің орын тәртібіне назар аударып, логикалық

екпінге құрылғандығын танытатын сызбаны ұсынған екен [6, 458 б.]. Алынған сызбалар статистикалық лингвистикада сөйлемнің синтаксистік модель үлгісіне айқын дәлелі бола алады. Ал аударма жұмысының синтаксистік және семантикалық көрінісінің үлгісі ретінде “Этюды к переводу трех стилей” деген еңбегін тануға болады, ғалым осы еңбегінде ғылыми, газет-журнал стиліндегі бір-бірінен өзгеше үш түрлі алынған мәтінді бір-бірімен салыстыра-салғастыра талдап, әр мәтіннің өзіндік айырмашылықтарына қарай сөйлем ішіндегі сөздердің өзара бір-бірімен қатысын 5 түрлі өрнекті сызбалармен көрсетіп, орысша берілген сөйлемді тікелей қазақшаға тәржімалауға керекті өзіндік семантика-синтаксистік модель-үлгі жасаған еді [6, 504 б.]. Ғалым еңбектеріндегі синтаксистікке қатысты атаулар “сөйлем мүшелері”, “жетек сөз”, “жетекші сөз”, “сөйлем мүшелерінің орын- тәртібі” туралы шағын тақырыптарда, сөздердің тіркесуі, тіркесу тәртібі, аударма туралы орысша жазылған еңбектерінде де бой көрсеткен 12 түрлі сурет бейнесімен алынған сызбалар үлгісі, қолданбалы лингвистика атаулары мен айтқанда, қазақ сөйлемдерінің синтаксистік модельдерінің үлгілері ретінде тануымызға негіз болады.

Профессор Қ. Жұбанов статистикалық талдауларды әдеби шығармалардың мәтіндерін зерттеу ісінде де пайдаланғанын байқаймыз. Ол Қожа Ахмет Яссауидың туған жылын анықтау мақсатында да сандық есептеулер жасайды. Хожа Ахметтің өмір сүрген кезін есептеп, мұсылмандардың ай санау календарын қазіргі күн санау календарына айналдыра есептей келе, Хожа Ахметтің қайтыс болған жылын, яғни 1166 жылы қайтыс болғанын айтады. Ахметтің 1041 жылы туған адам екенін де есептеп шығады. “Демек, ол Науайдан аттай 400 жыл бұрын туған”, -деп жазған [6, 291 б.]. Қожа Ахмет Яссауидың қалдырған жыр толғауының мәтініне талдау жасау барысында: Диуани хикмәт ішіне 4400 жолдық жыр-толғау сидырған кітап,-деп жыр жолдарының да статистикалық есебін шығарғанын көреміз [6, 310 б.]. Демек, барлық саладағы сияқты, тіл білімінде де статистикалық әдістердің қажеттігін Қ. Жұбанов еңбектерін оқыған сайын көз жеткізуге болады.

Қорытынды

Қорыта айтқанда, лингвосандық есептеу әдіс-тәсілдері тілдік деректерге ықтималды немесе дедуктивті болжамдар жасауға мүмкіндік береді. Қазіргі таңда профессор Қ. Жұбанов еңбектерінде алғашқы бастаулары көрініс берген қазақ тіл біліміндегі статистикалық лингвистиканың мүмкіндіктері арқылы тіл элементтерін өлшеудің сандық мәселелерін, жиілік сөздіктердің түрлерін, кері әліпби-жиілік сөздік, сөзнұсқағыш әліпби-жиілік сөздік, мәтін мен оның жиілік сөздігі бірліктерімен арақатынасы, сөздерді компьютер арқылы алудың біріккен және іріленген алгоритмін жасау, қазақ мәтіні бойынша түзілген жиілік сөздіктерге сандық есептеулер қолдану, мәтін және оның жиілік сөздігі ішіндегі сөздің ақпараттық сипаттамасын беру, т.б. ұғымдарға қатысты зерттеулерге қол жеткізілді. Сондықтан бүгінде, күрделі технологиялар мен коммуникация заманында, әлемдік ғылым мен техника

қарқын алып, онымен қатар тіл білімі де жаңа теориялық жаңалықтармен толыға түскен кезеңде, сандық технология, статистика, автоматтандыру бағыты кең етек алып, дами түсті. Қазақ тілтанымы саласы да статистикалық ізденістерге кең жол ашып, қазақ тілінің осы бағыттағы заманауи әдіс-тәсілдермен зерттелуіне ықпал етті. Соңғы кезеңде Қ. Бектаев, А.Қ. Жұбанов, Ә. Ахабаев, Қ. Молдабек, т.б. еңбектерінде цифрлар арқылы деректер келтіріліп, математика мен тіл деректерін ұштастыра алып, статистикалық есептерді келтіріп, тіл заңдылықтарды ашуға мүмкіндік беретін әдіс-тәсілдер белсенді пайдаланылды. Қазіргі зерттеу деректеріне қарағанда, қазақ тіл білімінде статистикалық лингвистика саласының бастауында Құдайберген Жұбановтың алғашқы ізденістерінің ізі сайрап жатқаны байқалады. Қазақ тілтануында алғаш рет Қ. Жұбановтың еңбектерінде көрініс тапқан қолданбалы тіл білімінің статистикалық лингвистикасы бүгінгі таңда әр алуан сипатта жедел түрде дамып келеді.

ӘДЕБИЕТ

- [1] Тер-Минасова С.Г. Тіл және мәдениетаралық коммуникация. – Алматы: Ұлттық аударма бюросы қоғамдық қоры, 2018. – 320 б.
- [2] Жұбанов А.Қ., Жаңабекова А. Корпустық лингвистика. – Алматы: «Қазақ тілі» баспасы, 2017. – 336 б.
- [3] Оразбаева Ф. Тілдік қатынас. Оқу құралы. – Қарағанды, 2014. –184 б.
- [4] Байтұрсынұлы А. Қазақ тіл білімінің мәселелері. –Алматы:«Абзал-Ай» баспасы, 2013. – 640 б.
- [5] Жүсіпұлы М. Ахмет Байтұрсынов және қазіргі қазақ тілі фонологиясы. – Алматы: Ғылым, 1998. – 216 б.
- [6] Құдайберген Жұбанов. Шығармаларының толық жинағы. – Ақтөбе: «Жұбанов университеті», 2019. – 584 б.
- [7] Бектаев Қ., Белботаев А., Смайылов Т. Қ.Б. Қазақ тіл біліміндегі лингвистикалық статистиканың алғашқы көріністері // Құдайберген Жұбанов және қазақ совет тіл білімі. – Алматы: «Ғылым», 1990. – 108-116-бб.
- [8] Андреев Н.Д. Статистико-комбинаторные методы в теоретическом и прикладном языкознании. – М: Наука, 2017. – 403 с.
- [9] Садуақас Н.А. Профессор Қ.Жұбановтың фонетика-фонологиялық зерттеулері. Монография. – Алматы: Эверо, 2018. –216 б.
- [10] Бюллетень гостерминкома КазССР, №4. –Алма-Ата, 1935.
- [11] Камчатнов А.М., Николина Н.А. Введение в языкознание: учеб.пособие филологических факультетов университета и педагогический вузов. 2-е изд. – М.Флинта: Наука, 2000. – 232 с.
- [12] Баданбекқызы З. Ағылшын тілі фонетикасы. – Алматы: Эпиграф, 2016. – 264 б.
- [13] Шәрипбай А. Проблемы перевода казахской письменности на латинский алфавит. – Астана, 2017. – 138 с.
- [14] Сиразитдинов З.Ә. Қазақстанда қолданбалы лингвистиканы қалыптастыруда Жұбановтардың үлесі туралы //«Ұлы дала тұлғалары: Жұбановтар тағылымы және жазу реформасы» / ҚР БжҒМ А.Байтұрсынұлы атындағы Тіл білімі институты, Абай атындағы Қазақ ҰПУ, Еуразия ғылыми-зерттеу институты ұйымдастырған профессор Құдайберген Жұбановтың туғанына 120 жыл және оның ұлы ф.ғ.д., профессор Есет Жұбановқа 90 жыл толуына арналған халықаралық ғылыми-теориялық конференция материалдары (19.XII.2019). – Алматы: «Қазақ кітабы» баспасы, 2019. – 59-63 бб.
- [15] Жұбанов А.Қ. Қолданбалы лингвистика: Қазақ тілінің статистикасы. Оқу құралы. – Алматы: «Қазақ университеті», 2007. – 210 б.

REFERENCES

- [1] Ter-Minasova S.G. Til zhane madenietaralyk kommunikaciya (Language and intercultural communication). Almaty: Ul'tytk audarma buyrosy kogamydyk kory, 2018. 320 b. [In kaz.]
- [2] Zhubanov A.K., Zhanabekova A. Korpustyk lingvistika (Corpus Linguistics). Almaty: "Kazakh tili" baspasy, 2017. 336 b. [In kaz.]
- [3] Orazbaeva Ph. Tildik Katynas (Linguistic relationship). Oku kuraly. Karagandy, 2014. 184 b. [In kaz.]
- [4].Baitursynuly A. Kazakh til biliminin maseleleri (Problems of Kazakh language education). Almaty: "Abzal-Ai", 2013. 640 b. [In kaz.]
- [5] Zhussipuly M. Akhmet Baitursynov zhane kazirgi Kazakh tili fonologiyasy (Ahmet Baitursynov and the phonology of the modern Kazakh language). Almaty: Gylym, 1998. 216 b. [In kaz.]
- [6] Kudaibergen Zhubanov. Shygarmalarynyn tolyk zhinagy (Complete collection of works). Aktobe: "Zhubanov University", 2019. 584 b. [In kaz.]
- [7] Bektayev K., Beltobaev A., Smayilov T.K. Kazakh til bilimindegi lingvistikalyk statistikanyn algashky korinisteri//Kudaibergen Zhubanov zhane Kazakh sovet til bilimi (The first manifestations of linguistic statistics in Kazakh language education // Kudaibergen Zhubanov and Kazakh Soviet language education). Almaty: "Gylym", 1990. 108-116 b. [In kaz.]
- [8] Andreev N.D. Statistiko-kombinatornye metody v teoreticheskom i prikladnom yazykoznanii (Statistical combinatorial methods in theoretical and applied linguistics). M: Nauka, 2017. 403 s. [In Rus.]
- [9] Saduakas N.A. Professor K.Zhubanovtyn fonetika-fonologiyalyk zertteuleri (Phonetics and phonological studies of Professor K. Zhubanov). Monografiya. Almaty: Evro, 2018. 216 b. [In kaz.]
- [10] Buylleten gosterminkoma KazSSR (Phonetics and phonological studies of Professor K. Zhubanov). №4. Alma-Ata, 1935. [In kaz.]
- [11] Kamchatnov A.M., Nikolina N.A. Vvedenie v yazykoznanie: uchebnoe posobie philologicheskikh fakultetov universiteta i pedagogicheskii vuzov (Introduction to language learning: ucheb.posobie philological faculties of the university and pedagogical universities). 2-e izd. M.Phlinta: Nauka, 2000. 232 s. [In Rus.]
- [12] Badanbekkyzy Z. Agylyshyn tili fonetikasy (English phonetics). Almaty: Epigraph, 2016. 264 b. [In kaz.]
- [13] Sharipbai A. Problemy perevoda kazakhskoi pismennosti na latinskii alphavit (The problem of translation of Kazakh writing into the Latin alphabet). Astana, 2017. 138 s. [In Rus.]
- [14] Sirazitdinov Z.A. Kazakhstanda koldanbaly lingvistikany kalyptastyruda Zhubanovtardyn ulesi turaly// "Uly dala tulgalary: Zhubanovtar tagylymy zhane zhazu rephormasy" / KR BzhGM A.Baitursynuly atyndagy Til Bilimi instituty, Abai atyndagy Kazakh UPU, Eurasia gylymi-zertteu instituty uymdastyrgan professor Kudaibergen Zhubanovtyn tuganyna 120 zhyl zhane onyn uly ph.g.d., professor Eset Zhubanovka 90 zhyl toluyna arnalgan khalykaralyk gylymi-teoriyalyk konpherenciya materialdary (About the contribution of the Zhubanovs in the formation of applied linguistics in Kazakhstan // "Great steppe personalities: the teaching of the Zhubanovs and writing reform") (19.XII.2019). Almaty: «Kazakh kitabi», 2019. 59-63 b. [In kaz.]
- [15] Zhubanov A.K. Koldanbaly lingvistika: Kazakh tilinostatistikasy (Applied linguistics: Statistics of the Kazakh language). Oku kuraly. Almaty: "Kazakh universiteti", 2007. 210 b. [In kaz.]

СТАТИСТИЧЕСКИЕ МЕТОДЫ В ТРУДАХ ПРОФЕССОРА К. ЖУБАНОВА

Садуакас Н.А.¹, *Кушкимбаева А.С.², Абрахманова Ж.М.³, Асанов Ж.⁴

¹к. ф.н., доцент, Актюбинский региональный университет имени К. Жубанова, Актобе, Казахстан, e-mail: sadu_ngd@mail.ru,

^{*2}PhD., доцент Актюбинского регионального университета имени К. Жубанова, e-mail: alitik@mail.ru,

³к.ф.н., доцент, Атырауский университет имени Х. Досмухаммедова, Атырау, Казахстан, e-mail: azhm63@mail.ru,

⁴д.ф.н., профессор, Актюбинский университет имени С. Байшева, Актобе, Казахстан, e-mail: shalkar1900@mail.ru

Аннотация. В статье речь идет о количественных проблемах исчисления элементов языка в прикладном языкознании, применении статистических методов, подходов в изучении языка, трудах исследователей на начальном этапе становления статистической лингвистики в казахском языкознании, направлении современных лингвостатистических исследований.

Цель, основные направления, идеи научного исследования: на основе фактических данных отражены первые примеры применения статистических методов, приемов при изучении казахского языка в работах профессора К. Жубанова, признанного своими передовыми научными идеями во всех областях языкознания.

Доказано, что статистические методы использовались для выражения фонетико-фонологического значения звуков в казахском языке, знания состава фонем, количества букв, заимствованных в алфавите, подсчета количества слов, слогов, определения значения слова для определения количества фонемы, ее места в слове, определения комбинации, положения каждой фонемы в слове. В этом научная и практическая значимость работы.

Основные результаты и анализ, выводы исследовательской работы: учитывая, что определение закономерностей сочетания звуков казахского языка, специфики деления на слоги, условий транспортабельности-нетранспортабельности является одной из главных потребностей орфографии, ученый К. Жубанов провел исследования орфографии в сочетании со статистическими изысканиями.

Одной из трудностей, с которыми мы сталкиваемся при переводе предложения с одного языка на другой, является точная передача значения слова; языковой анализ данных, представленных статистическими методами, показывает, что словарное значение слова, значение отдельного слова и значение слова в определенном контексте, контактирующего с другими словами в составе предложения, не совпадают.

Результаты данной работы могут быть использованы при изучении тем, связанных с цифровыми технологиями в области цифровой лингвистики, фонетики казахского языка.

Практическая значимость итогов работы заключается в том, что первые статистические данные, представленные в трудах К. Жубанова при изучении современных дисциплин цифровой лингвистики, можно использовать при решении следующих задач: численные задачи элементов языка, типы частотных словарей, обратный алфавит-частотный словарь, дублирующий алфавит-частотный словарь, вычисление количества грамматических единиц в тексте, разработка алгоритма слов с помощью компьютера, предоставление описания численных расчетов по частотным словарям, сформированным по казахскому тексту. В целом достигается умение выполнять работы на практике.

Ключевые слова: лингвистика, статистические методы, количество звуков, количество букв, слоги, количество слов, словосочетание, предложение

THE STATISTICAL METHODS IN THE WORKS OF professor K. ZHUBANOV

Saduakas N.A.¹, *Kushkimbayeva A.S.², Abdrakhmanova Zh.M.³, Assanov Zh.⁴

¹Candidate of Philological Sciences, associate Professor

K. Zhubanov Aktobe Regional University,
Aktobe, Kazakhstan, e-mail: sadu_ngd@mail.ru,

*² PhD, associate professor, K. Zhubanov Aktobe Regional University, Aktobe,
Kazakhstan, e-mail: alitok@mail.ru,

³Candidate of philological sciences, associate Professor
Kh. Dosmukhamedov Atyrau University, Atyrau, Kazakhstan,
e-mail: azhm63@mail.ru,

⁴Doctor of philology, professor, S. Baishev Aktobe University,
Aktobe, Kazakhstan, e-mail: shalkar1900@mail.ru

Abstract. Article describes calculating linguistic elements in applied linguistics quantitative problems, statistical methods use, language-learning approaches, achievements of first researchers in statistical linguistics in Kazakh linguistics formation, contemporary linguistic statistics research directions.

Scientific research's aims, main directions, ideas: First use of statistical methods, techniques based on real data in Kazakh linguistic research is reflected in Professor K. Zhubanov's work, known for his high scientific ideas in all areas of linguistics.

Scientific and practical significance of the work. Statistical methods have proven to represent the phonetic meaning of Kazakh sounds, knowledge of phoneme composition, the number of borrowed letters of the alphabet, the number of words, syllables, word meanings, phonemes number, phonemes position in a word, phoneme combinations determination.

Main results, analysis, conclusions of study. Scientist K. Zhubanov conducted study of spelling in conjunction with statistical research, main needs of spelling are combination pattern of Kazakh sounds determination, determination of their division into syllables, movement and transfer conditions determination.

Research methodology. Difficulty in translating a text from one language to another is to convey exact words' meanin. Linguistic analysis of the data presented by the statistical method shows that the meaning of a word is in a different context and the translated meaning is not in direct contact with other words in the sentence.

Value of research (research's contribution to relevant knowledge field). Digital technology can be used to study problems related to Kazakh phonetics in digital linguistics field.

Research results' practical significance. In modern digital linguistics research, using initial statistical data used in K. Zhubanov's research, numerical problems of linguistic elements, types frequency dictionaries' types, reverse alphabetic frequency dictionaries, duplicate alphabetic frequency dictionaries, textual grammatical units number calculation, computer algorithm development, numerical calculations description, frequency dictionaries provision built on words in Kazakh texts basis, ability to work in practice.

Keywords: linguistics, statistical methods, sound number, letter number, syllable, word number, phrase, sentence

Статья поступила 08.10.2022