

УДК 82:821

Түмөнбаева Ж.А.

*К.Тыныстанов ат. ҮМУ*

**ТИЛ ИЛИМИНДЕ МАТЕМАТИКАЛЫК ЛИНГВИСТИКАНЫН,  
ЛИНГВОСТАТИСТИКАЛЫК МЕТОДДУН ПАЙДА БОЛУШУ ЖАНА  
КАЛЫПТАНЫШЫ**

*Макалa тил илиминде жаны багыт болуп саналган математикалык лингвистика жана лингвостатистикалык методго арналып, анын мааниси, пайда болушу жана калыптанышы, түркологиядагы жана кыргыз тили илиминдеги изилдениш даражасы жөнүндө сөз болот.*

XX кылымда илимдердин ар түрдүү тармактарынын бирге аракеттешүү, өз ара байланышып өнүгүү тенденциясы башталды. Бара-бара өз алдынча илимдердин чеги кеңейди; гуманитардык, техникалык жана табийгый илимдердин байланышынын негизинде акыл ишмердигинин жаңы тармактары пайда боло баштады. Азыркы илимдин анык өзгөчөлүгү – структураны жана аны түзүүчү элементтерди үйрөнүүгө болгон умтулуу болуп эсептелет. Ошондуктан илимий теорияда жана практикада математикага көп маани берилип келет. Математика, бир жагынан, логика жана философия, экинчи жагынан, статистика менен чектешип, көп мезгилдер бою таза «гуманитардык» деп эсептелген сфераларга, анын ичинде тил илимине да улам барган сайын өтүп, алардын эвристикалык потенциалын кеңейтип келе жатат. Анткени математика илиминде көп бериле турган «канча» деген суроого жооп болуучу аргументтер башка илимдердеги «эмне» жана «кандай» деген суроолорду жараткан маселелерди дагы кеңири изилдөөгө жардам берет.

Өткөн кылымдын 50-жылдарынан баштап математика стилдердин түзүлүшүн (табийгый жана жасалма) сыпаттоо үчүн колдонула баштаган. Бирок ал дароо эле мындай практикалык колдонулушка ээ болгон эмес. Лингвистикадагы алгачкы математикалык методдор тил илиминин негизги түшүнүктөрүн тактоо үчүн колдонула баштады, бирок компьютердик техниканын өнүгүшү аны практикада колдонууга жол ачты. Биздин мезгилде кеңири белгилүү «математикалык лингвистика» деген термин менен так методдор (так методдор түшүнүгү математика менен байланышкан) колдонулган бардык лингвистикалык изилдөөлөр аталат. Айрым изилдөөчүлөр муну өзгөчө «лингвистика» катары жаңы термин деңгээлине көтөрбөстөн, изилдөөнүн жаңы методу колдонулган багыт катары гана кароо керек дешет.

Тилди математикалык жактан сыпаттап жазуу XX кылымдын башталышында белгилүү швейцариялык лингвист Фердинанд де Соссюрдун аны механизм катары эсептеген көз караштарында берилген. Тилди ар бир сөзү (системанын мүчөсү) өз алдынча эмес, башка мүчөлөр менен бирдикте каралган үч бөлүктөн (тил – langue, кеп – parole, кеп ишмердиги – langage) турган система катары караган теориясы анын концепциясынын негизин түзөт. Белгилүү даниялык лингвист Луи Ельмслев белгилегендей, Соссюр «тилге структуралык мамиле кылуу, башкача айтканда, бирдиктердин арасындагы катышты каттоо жолу менен тилди сыпаттоо керектигин айткан» [14, 156; 15, 187].

Математикалык лингвистиканын бир методу лингвостатистикалык метод болуп эсептелет. Анда, жогоруда айтылгандай эле, илим математикалык так методдорду колдонгондо гана толук максатына жетет деген илимий негиз жатат. Сандык анализдин теориялык жактан негизделиши жана алгоритмдерди түзүү, аны лингвистикада практика жүзүндө пайдалануу лингвостатистика деп аталган илимдин өзгөчө тармагынын предмети болуп саналат.

Лингвостатистикалык методдун маңызы тилдик көрүнүштөрдүн сапаттык жаңыланууларын пайда кылган сандык өзгөрүүлөрдү аныктоодо турат. Лингвостилистиканын алкагында тилди изилдөөдө математикалык методду колдонуу

кептеги тигил же бул тилдик элементтердин жыштыгы белглиүү бир статистикалык мыйзамдарга (мыйзам ченемдүүлүктөргө) баш иет. Бул статистикалык маалыматтардын негизинде тилдик бирдиктердин жана тексттин түзүлүшүнүн кызматынын белгилүү бир мыйзам ченемдүүлүктөрүн түшүндүрүүгө жол берет.

Лингвостатистикалык метод азыркы лексикологияда жана стилистикада тилдин да, кептин да көрүнүштөрүн үйрөнүүчү колдонулат. Мисалы, лингвостатистикалык методдун жардамы менен лингвисттер кептин стилдик жана автордук ар түрдүүлүгүндө сөздүк составдын сандык мүнөздөмөсүн иликтешет. Лингвостатистикалык иликтөөнүн натыйжасында жыштык сөздүктөр пайда болду.

Тилди үйрөнүүдөгү статистикалык мыйзам ченемдүүлүктөр менен генетиканын негиздөөчүсү Георг Мендель алек боло баштаган. 1968-жылдарда гана филологдор анын өмүрүнүн акырында математикалык методдорду тилдик кубулуштарды аныктоодо колдонуп иштегенин билишти. Мендель аббат болуп иштеген Брно шаарындагы ыйык Томашо монастырында аягы “mann”, “bauer”, “mayer” менен бүткөн фамилиялардан турган саптар жана кандайдыр бир бөлчөк, эсептөөлөр жүргүзүлгөн баракчалар табылган. Фамилиялык аттардын келип чыгышынын формалдуу закондорун табууга умтулуп жатып, Мендель татаал эсептөөлөрдү жүргүзөт, аларда немец тилиндеги үндүү жана үнсүздөрдү, карап жаткан сөздөрдүн жалпы санын, фамилиялардын санын жана башкаларды эсептеп чыккан.

Ал эми Россияда структуралык лингвистика Батыштагыдай эле XIX-XX-кылымдар аралыгында өнүгө баштаган. Тилди система катары түшүнүү Ф. де Соссюр менен бир мезгилде эле Казань университетинин профессорлору Ф.Фортунатов менен И.А.Бодуэн де Куртэнэнин эмгектеринде өнүктүрүлө баштаган. Куртэнэ көп мезгилге чейин Ф. де Соссюр менен кат алышып, мунун натыйжасында Женева жана Казань лингвистикалык мектептери кызматташ болгон. Эгерде Соссюр тил илиминдеги «так» методдордун идеологу катары эсептелсе, Бодуэн де Куртэнэ аны практикалык жактан колдонууга негиз салган. Ал биринчи болуп статистикалык методдорду жана функционалдык көз карандылыкты колдонуучу так илим катары лингвистиканы филологиядан (рухий маданиятты тил жана кеп аркылуу үйрөнүүчү гуманитардык илимдердин жалпылыгынан) бөлүп караган [9, 102].

Бодуэн де Куртэнэ өзүнүн фонетикалык альтернация (алмашуу) жана альтернанттар жөнүндөгү окуусунда математикалык формулаларды фонетикалык альтернацияларды классификациялоодо, салыштырууда абсолюттуу тактыкка жетишүүчү колдонгон [4, 280-345].

Лингвистикага математикалык методдорду жайылтууну «сыноочу» объект катары фонология алынган. Тыбыштар тилдик системанын «атому» катары сыпаттоонун формалдуу, так методдору үчүн ыңгайлуу материалы болгон.

Бара-бара так методдор фонетикада гана эмес, синтаксисте да колдонула баштады. Тилдин структуралуулугу менен лингвисттер жана математиктер алектене баштаган. 1950-60-жылдарда математика менен лингвистиканын өз ара аракеттенүүсүнүн машиналык которууларды иштеп чыгуу менен байланышкан жаңы этабы башталган. 1947-жылы А.Бутт жана Д.Бриттен тарабынан ЭЭМдин жардамы менен сөзмө-сөз (пословно) которуунун коду ойлонуп табылган, бир жыл кийин Р.Риченс машиналык которууда сөздөрдү негиз жана мүчөлөргө бөлүүнүн эрежесин сунуш кылды. Ал кездеги эсептөөчү машиналар азыркы машиналардан кескин айырмаланып, өтө кымбат жана чоң болушкан. Ошондой эле аларда иш алып баруу үчүн оператор, программист, инженерлердин бир топ штаты талап кылынган. 1954-жылы Массачусетс техникалык университетинде «Джоржтаун эксперименти» - ИБМ-701 машинасында орус тилинен англис тилине которууну көрсөтүштү. Эксперименттин ийгиликтүү болгондугу Д.Ю.Панов тарабынан «Математикада» жарыяланган (Перевод с одного языка на другой

при помощи машины: отчет о первом успешном испытании», 1954, № 10).

Машиналык которуу ишине Д.Ю.Панов И.К.Бельскаяны тарткан, кийинчерээк ал СССР ИАда математика жана так эсептөө техникасы институтун жетектеген. 1955-жылдын аягында БЭСМ машинасында англис тилинен орус тилине которуу тажрыйбасы жүргүзүлөт. Анын программасын Н.П.Трифонов жана Л.Н.Королев түзүп чыгышкан. Алардын кандидаттык иштери машиналык которуу үчүн сөздүктөрдү түзүү ишине арналган эле. Бирок машиналык которуу системасынын биринчи муунунун кемчиликтери бар эле. Алардын бардыгы удаалаштык которуунун алгоритмине негизделген, «сөз артынан сөз», «фраза артынан фраза» - сөздөрдүн жана сүйлөмдөрдүн арасындагы маанилик байланыштар эске алынган эмес. Синонимдердин, антонимдердин жана өтмө маанилердин маанисин компьютерге киргизүү оор болчу. Келечектүү багыт катары машиналык системаны иштеп чыгуу керек болгон.

Мезгилдин өтүшү менен түз которуу Т-системасы (англис тилинде «transfer» - кайра өзгөртүлгөн) менен алмашылган. Анда которуу синтаксистик структуранын деңгээлинде жүргүзүлгөн.

Кириш сүйлөмдүн маанилик түшүнүгүн алууга негизделген машиналык которуунун системасын түзүү иши азыркыга чейин эң алгылыктуу деп эсептелет. Мындай системалар И-системалар («интерлингва» деген сөздөн) деп аталат. Бирок аларды жаратуу маселеси өткөн кылымдын 50-жылдарынын аягында, 60-жылдардын башында коюлганы менен бүгүнкү күнгө чейин толук чечиле элек. Машиналык которуунун теоретикалык негиздерин жаратууга болгон муктаждык математикалык лингвистиканын калыптанышына алып келген. Мында А.А.Ляпунов, О.С.Кулагина, В.А.Успенский сыяктуу математиктер В.Ю.Розенцвейг, П.С.Кузнецов, Р.М.Фрумкина, А.А.Реформатский, И.А.Мельчук, В.В.Иванов сыяктуу лингвисттер негизги ролду ойношкон.

1960-жылдын 6-майында СССР ИА Президиумунун «Тилди изилдөнүн структуралык жана математикалык методдорун өнүктүрүү жөнүндөгү» Токтому менен Тил илими институтунда жана Орус тили институтунда ага ылайыктуу бөлүмдөр ачылган. 1960-жылдан баштап өлкөнүн башкы гуманитардык ЖОЖдорунда – МГУнун филология факультетинде, Ленинград, Новосибирск университеттеринде текстти автоматтык иштеп чыгуу тармагы боюнча кадрлар даярдала баштаган.

70-жылдарда советтик өлкөдө ХХ кылымдагы батыш лингвисттеринин эмгектери которулуп, үйрөнүлө баштайт. И.А.Мельчук, Н.Хомскийдин чыгармаларын, Н.А.Слюсарёвүнүн «Теория де Соссюра в свете современной лингвистики» деген эмгегинде соссюрдук постулаттарды 70-жылдардагы лингвистиканын актуалдуу маселелери менен байланыштырат. Лингвистиканы андан ары математикалаштыруу тенденциялары уланат. Белгилүү жогорку окуу жайларда «Математикалык (теориялык, колдонмо) лингвистика» адистиктери боюнча кадрларды даярдоо жүрөт. Ошол эле мезгилде батышта эсептөө техниканын өнүгүшүндө кескин секирик жасалат.

Ошентип, ХХ кылымда так илимдер менен гуманитардык илимдердин жакындашуу процесси жүрөт. Математиканын лингвистика менен бирге аракеттешүүсү практикалык колдонулушка ээ боло берет.

Лингвостатистиканын башатын орус тил илими менен байланыштырган изилдөөчүлөр да бар. Алар муну М.В.Ломоносовдун, П.А.Чебышевдин, М.В.Остроградскийдин, В.Я.Буняковскийдин математика менен тил илимин маанилүү ролго койгон көз караштары менен далилдешет [10].

Эгерде В.Я.Буняковский биринчи жолу статистиканы тил илимине колдонуу мүмкүндүгү жөнүндө оюн билдирсе, башка бир атактуу математик А.А.Марков 1913-жылы биринчи жолу статистиканы тилдик материалга колдонгон [10, 8-9].

1960-жылдары лингвистикадагы статистика боюнча Р.М.Фрумкинанын бир катар эмгектери белгилүү. Тагыраак айтканда, Пушкиндин лексикасынын статистикалык структурасына арналган изилдөөсү [16], ошондой эле статистикалык методдордун тил

илиминде колдонулушуна арналган көлөмдүү макалалары жарыяланган [17].

Б.Е.Головин өзүнүн “Тил жана статистика” аттуу эмгегинде [5] статистикалык методдордун тилди жана кепти үйрөнүүдө колдонулушунун реалдуу негизи болорун, тилге да сандык белгилердин жана сандык мүнөздөмөлөрдүн объективдүүлүкү таандык экендигин: «Опыты языковедческой и литературоведческой наук, накопленные к настоящему времени, знания о языке и его стилях позволяет утверждать, что одним из реальных оснований применения статистики в изучении языка и речи нужно признать объективную присущность языку количественных признаков, количественных характеристик» деп түшүндүрөт да [5, 11], элементтери тигил же бул ыктымалдык-статистикалык закондорго баш ийип өнүккөн структура катары каралышы мүмкүн экендигин белгилеп өтөт [5, 16-17]. Андан ары окумуштуу статистиканын негизги түшүнүктөрү болгон жыштык, орточо жыштыкты эсептөөнүн жолдору, үлүштөрдү салыштыруу ж.б. жөнүндө түшүнүк берет.

Лингвостатистикага арналган эмгектердин дагы бири И.А.Носенко [10] таандык. Автор лингвостатистиканын актуалдуулугу кийинки мезгилде өсүп жаткандыгын белгилеп кетет да, өзүнүн эмгегинин максаты - лингвостатистика багытындагы жаш изилдөөчүлөрдү статистикалык изилдөөлөрдүн ыкмалары менен жана математикалык аппаратты туура колдонуу менен тааныштыруу деп эсептейт [10, 8], ошондой эле лингвостатистикадагы негизги түшүнүктөр: кокустук чоңдук (случайная величина), үзгүлтүксүз кокустук чоңдук (непрерывная, случайная величина), окуянын салыштырма жыштыгы (сравнительная частота события), үлүш, таралыш (распространенность), статистикалык ыктымалдуулук (статистическая вероятность), орточо арифметикалык чоңдук (средняя арифметическая величина), четтеп кетүү (отклонение), статистикалык гипотеза (статистическая гипотеза) деген ж.б. түшүнүктөрдү чечмелеп берген [9, 17]. Өз сөзүндө ал айрымдар жаңылыш эсептегендей, лингвостатистика математикалык лингвистика эмес экендигин, лингвостатистика статистиканын математикалык аппаратын колдонгон салттуу лингвистиканын гана бир тармагы экендигин эскерте кеткен.

1988-жылы М.В.Араповдун “Квантитативдик лингвистика” аттуу монографиясы жарык көрөт. Монографияда тилге квантитативдик (сандык) мамиленин негиздери баяндалат жана тилди сапаттык жана математикалык статистиканын методдору менен изилдөөнүн жыйынтыктарын байланыштыруу, тилдик бирдиктердин информациялык баалуулугу менен сөздөрдүн сандык мүнөздөмөлөрүн (колдонулуу жыштыгы, көп маанилүүлүк даражасы) аныктоо аракеттери көрүлгөн [2].

Эмгектин кыйла бөлүгүн автордун “Илимий-техникалык маалымат”, “Семиотика жана информатика”, “Информациялык теориянын жана практиканын маселелери”, “Системалык изилдөөлөр”, “Квантитаивдик лингвистика” аттуу жыйнактарда жарык көргөн макалалары жана автор тарабынан 1977-1982-жылдарда МГУнун филология факультетинин Структуралык жана колдонмо лингвистика бөлүмүндө “Лингвистикадагы математикалык методдор” аттуу курсунда окулган материалдар түзөт [2, 4].

Ал эми түркологияда математикалык ыкма аркылуу тилдик изилдөөлөрдү жүргүзүү иши XX кылымдын 20-30-жылдарында Орто Азия жана Казакстанга тарай баштаган. Профессор Р.Г.Пиотровский баштаган “Статистика речи” деген топтун салымы, К.Б.Бектаевдин тил илиминдеги математикалык методдор жөнүндөгү эмгеги баштаганказак лингвостатисттеринин 20 жылдын ичиндеги 10дон ашык монография менен илимий жыйнактары, 200дөн ашуун илимий жана илимий популярдуу макалалары бул жаңы багыттын андан ары өнүгүшүнөбөлгө болду.

А.Х.Жубановдун кийинчерээк чыккан “Казак текстинин квантитативдик структурасы” деген эмгегинде [7] М.Ауэзовдун “Абай жолу” романынын статистикалык структурасы изилдөөгө алынат да, зат атооч, этиш, тактооч, сын атооч сөздөрдүн класстарга бөлүштүрүүнүн ыктымалдуулук-статистикалык модели каралат. Эмгектин

редактору академик С.К.Кенесбаев туура белгилегендей, автор казак тил илиминде биринчи болуп М.Ауэзовдун романынын тилинин жана стилинин бардык нюанстарын кеңири анализге алат. А.Х.Жубановдун эмгегинин мааниси статистикалык изилдөөнүн логикалык удаалаштык методун бул жазуучунун чыгармасына гана эмес, башка чыгармаларды изилдөөгө ылайыкталып иштелип чыгышында. Романдагы 456 966 сөз колдонулушу кылдаттык менен анализденип, жыйынтыгында, лингвистикалык изилдөөдө ЭМДи колдонуу боюнча булак катары кызмат кылуучу оригиналдуу монография болгон. Эске сала кетсек, ЭМ аркылуу жыштык сөздүктөрүн түзүүнүн принциптерин А.Х.Жубанов жана К.Б.Бектаевдер түзүп чыгып, ушул багыттагы алгачкы зор эмгек сөздүн түшүндүрмөсү менен колдонуш жыштыгы кошо берилген “Абай тилинин сөздүгү” (Абай тили. - Алматы, 1968.) болуп эсептелет.

Анализденген тексттердеги сөз түркүмдөрүнүн сандык катышы боюнча А.Б.Белботаев менен А.Р.Зекенова да “Казак тексттериндеги сөз түркүмдөрүнүн статистикасы” деген макаласында кызыктуу жыйынтыктарды келтиришет. Алардын иликтөөсү менен казак тексттеринин негизги үлүшүн зат атооч жана этиш түзөт: суммардык жыштыгы болжол менен 80%ды түзөт. Ар түрдүү жанрлардагы сөз түркүмдөрүнүн колдонулушу белгилүү бир туруктуулук менен мүнөздөлөт. Мисалы, роман, повесть, очерк, публицистика, илимий тексттердин негизги бөлүгүн зат атоочтор түзөт (26-41%), ал эми пьеса жана аңгемелерде алар этиштен кийинки орунда. Илимий жана публицистикалык стилдеги тексттерде сөз түркүмдөрүнүн колдонулуш өзгөчөлүгү көркөм стилдегидей экендиги да көңүлгө аларлык жагдайдан, башкача айтканда, аларда зат атооч, этиш, сын атооч, байламталар, ат атоочтор, тактооч, сан атооч, модалдык сөз, тууранды жана сырдык сөздөр тартибинде азайып отурат. Бирок илимий стилде сырдык сөздөр менен тууранды сөздөр колдонулбагандыгын бул жанрдагы тексттерге ылайык келген типологиялык индекс катары кароого болот деген жыйынтыкка келет [3, 177]. Ушундай эле бутумго А.Б.Белботаев өзүнүн “Казак текстинин сөз түркүмдөрүнүн лингвостатистикалык мүнөзү” деген темадагы кандидаттык диссертациясында да келет (А.Б.Белботаев, 1992).

Бүгүнкү күндө кыргыз тил илиминде да математикалык лингвистика, лингвостатистика багытында изилдөөлөр жүргүзүлүп, диссертациялар корголуп жатат.

1975-жылы С.Ибрагимов тарабынан “Структурно-вероятностный анализ синтаксических отношений киргизского языка и их функционирования в речи” деген темада алгач ирет кандидаттык диссертация (1975) корголгон. Иштин негизинде окуу-методикалык мүнөздөгү окуу куралы [8] даярдалып, Кыргыз Республикасынын Билим берүү жана илим министрлигинин сунушу менен 1994-жылы жарык көргөн. Эмгекте структуралык-ыктымалдуулук методунун зарылдыгы жана анын милдеттерин белгилейт Структуралык-ыктымалдуулук маселесине арналган буга чейинки изилдөөчүлөр Н.А.Андреев, Б.Н.Головиндин эмгектерине токтолуу менен, кыргыз тилин структуралык-ыктымалдуулук методу түрк тилдерине биринчи колдонулуп жаткандыгын баса белгилейт [8, 3]. Эмгек кыргыз тилин структуралык-ыктымалдуулук методдун жардамы менен синтаксистик планда изилдөөгө арналган Мында статистикалык анализдин негизинде кыргыз тилиндеги зат атоочторду синтаксистик-функционалдык жактан классификациялап, зат атоочтун башка сөз түркүмдөрү менен болгон синтаксистик байланышы аныкталган.

Иштин алдына койгон максаттары болуп автор өзү белгилегендей, структуралык-ыктымалдуулук методду кыргыз тилинин синтаксисине пайдалануу, кыргыз тилинин функционалдык класстары термининде синтаксистик мамилелерди изилдөө, кыргыз илимий-техникалык текстинин статистикалык жана лингвистикалык законченемдүүлүктөрүн ачып берүү, сөз түркүмдөрүнүн жана сүйлөм мүчөлөрүнүн жалпы ыктымалдуулук жана ыктымалдуулук дифференциалдык белгилерин аныктоо,

алынган жыйынтыктарды лингвистикалык жана башка максаттарда колдонуу мүмкүнчүлүктөрүн көрсөтүү болуп саналат [7, 6].

1979-ж “Манас” эпосунун жыштык сөздүгүн түзүү принциптери боюнча Т.Садыковдун “К исследованию квантитативно-лингвистической структуры кыргызского текста на ЭВМ” аттуу макаласы жарык көргөн (1979).

Кыргыз тилиндеги гезит тексттеринде колдонулган сөз өзгөртүүчү мүчөлөрдүн Ж.Жетешиков түзгөн статистикалык анализи ЭВМ ЕС-1045 эсептөөчү машинасында иштеп чыгуу менен алынган жыштык сөздүктүн негизинде түзүлгөн. Тандап алуулардын (выборки) көлөмү 100 миңден ашык сөз колдонулуш болгон [12]. Бул эмгек кыргыз тилиндеги гезит тексттерине жүргүзүлгөн биринчи статистикалык илик болуп калды.

К.Дыйканов тарабынан жөнөкөй саноо жолу менен “Манас” эпосунун академик Б.М.Юнусалиевдин редакциясына 1958-жылы басылып чыккан 1-томунун негизинде түзүлгөн жыштык сөздүгү [6] жарыяланган.

Ал эми буга чейин колго алына элек кыргыз акын, жазуучуларынын колдонгон сөздөрүнүн тизмеси, сөздүгүнтүзүү ишине И.Сагынбеков проф. Т.К.Ахматовдун жетекчилигинде киришип, Ч.Айтматовдун кыргыз тилинде жарык көргөн чыгармаларынын жалпы көлөмү 199428 сөз колдонуштан турган жыштык сөздүгүн электрондук эсептегич машина ЕС 10-46 да 1988-ж алган. Алынган жыштык сөздүктүн негизинде «Кыргызча-орусча минимум сөздүк» (И.Сагынбеков, 1992) жана “Ч.Айтматовдун чыгармаларынын алфавиттик жыштык сөздүгү” (И.Сагынбеков, 1994) өз-өзүнчө китеп китеп кылып чыгарган. Кийинчерээк ал Ч.Айтматов кыргыз тилинде жазган “Асма көпүрө”, “Ак жаан”, “Түнкү сугат” аңгемелери жана “Биринчи мугалим”, “Бетме-бет”, “Жамийла”, “Саманчынын жолу” повесттери боюнча жыштык сөздүктөрдү түздү. Сөздүктү түзүүдө Ч.Айтматовдун эң акыркы жолу жарык көргөн сегиз томдугунан сегиз чыгарма бөлүнүп алынып, андагы омонимдерди, кошмок сөздөрдү белгилөө иштери жүргүзүлгөн жана информатик адистер менен бирдикте тиешелүү статистикалык маалыматтар алынган

И.Сагынбековдун “Ч.Айтматов тилинин сөздүгү” жазуучунун нукура кыргыз тилинде жазылган чыгармаларынан тургандыгы, сөздүктө чыгарманын бүтүндүгүзске алынгандыгы, ар бир чыгармага түзүлгөн сөздүктөр атайын программанын негизинде бириктирилип, бириккен жыштык сөздүк алынгандыгы, сөздүктөгү сөз колдонулуштар ар бир чыгарманын чегинен чыкпай, семантикасын жоготпой, кайсы сөз түркүмүнө тиешелүү болсо, ошо боюнча индекстелгендиги жана кош, кошмок сөздөрдү, омонимдерди, татаал сөздөрдү индекстөө учурунда сөз колдонулуштардын семантикасы сакталгандыгы менен кыргыз тил илиминде, кыргыз лексикографиясында жаңылык болуп эсептелет.

Филология илимдеринин доктору, профессор Ж.Мамытовдун жалпы илимий жетекчилиги астында Ысык-Көл мамлекеттик университетинин тилчилер жана информатиктер тобу ЭЭМди (электрондук эсептөөчү машина) колдонуу менен улуу манасчылар Сагынбай Орозбаковдун 4 томдук “Манасынын” жана Саякбай Каралаевдин 5 томдук (“Манас”, “Семетей”, “Сейтек”) трилогиясынын тексти боюнча түзүшкөн, болжолу 165 басма табак көлөмүндөгү “Манас” эпосунун лексикасынын жыштык сөздүгү” даярдалышы кыргыз тил илиминде статистикалык методдун өнүгүшүндө зор мааниге ээ болуп, мунун негизинде бир канча изилдөөлөргө жол ачты.

Бул сөздүктүн негизинде да кандидаттык диссертациялар корголду. Атап айтканда, Н.Ж.Мазекованын “Манас” эпосунда колдонулган сөз формаларына лингвостатистикалык илик (зат атооч)” (2000), Б.Д. Шаршембаевдин “Манас” эпосунун жаңы типтеги сөздүктөрүн түзүү жана чечмелөө технологиясы” (2004) аттуу эмгектери жазылды.

Н.Ж.Мазекованын “Манас” эпосунда колдонулган сөз формаларына лингвостатистикалык илик (зат атооч)” аттуу кандидаттык диссертациялык эмгегинде

эпостун жалпы текстинин жана Сагынбай менен Саякбайдын варианттарынын негизинде түзүлгөн түрдүү жыштык сөздүктөрдү териштирип талдоо аркылуу зат атооч түркүмүндөгү сөздөрдүн морфологиялык түзүлүшү, сөз формаларын түзүү мүмкүнчүлүктөрү, мындай сөз формалары менен аларды уюштуруучу категорияларынын тексттеги колдонуш өзгөчөлүктөрү лингвостатистиканын так жана объективдүү ыкмаларынын жардамы менен аныкталган. Иштеги “Манас” эпосунун текстинде колдонулган зат атоочтун грамматикалык формаларынын ар тараптан иликтениши кыргыз тили илими менен түркологиядагы алгачкы саамалык болгон.

Ал эми Б.Д.Шаршембаевдин жогоруда аталган диссертациялык иши компьютер техникасы менен маалыматтар технологиясынын, инженердик лингвистика менен лингвостатистиканын жетишкендиктерин, кыргыз жана башка түрк тилдери боюнча топтолгон сөздүк түзүү тажрыйбасын пайдалануу аркылуу “Манас” эпосунун бириктирилген текстинин жана Сагынбай менен Саякбайдын варианттарынын негизинде ар канда типтеги сөздүктөрдү түзүү жана чечмелөө технологиясына сыпаттама берилген жана тексттин квантитативдик типологиялык көрсөткүчтөрү тастыкталган. Тагыраак айтканда, сөз формаларынын ишенимдүү аралыгы, сөз формаларынын орточо кайталанышы, сөз формаларанынын орточо кайталанышынын өсүшү, синтетизм коэффициенти, жыштык сөздүктүн зоналарынын текстти камтуу өзгөчөлүктөрү сөз формаларынын информациялык наркы жыштык сөздүктүн негизинде аныкталган, сөз көрсөткүчтөрүнүн сөздүн түзүү технологиясы да каралган. Иштин мазмунунда, ошондой эле, текстти иликтөөнүн лингво-семиотикалык негиздери каралып, тилдик белгинин компоненттерине, кадимки жана жасалма тилдердин өзгөчөлүктөрү да мүнөздөлгөн [18].

Жогоруда сөз кылган эмгектер тилдин бардык денгээлдерин изилдөөдө математикалык лингвистиканын, лингвостатистикалык методдун колдонулуу мүмкүнчүлүгүнүн кеңири экендигин жана жаңы багыттагы изилдөөлөрдө универсалдуу экенин көрдүк.

#### Адабияттар:

1. Андреев Н.Д. Статистико-комбинаторные методы в теоретическом и прикладном языковедении. – М., 1967.
2. Арапов М.В. Квантитативная лингвистика. - М., 1988.
3. Белботаев А.Б., Зекенова А.Р. О статистике частей речи в казахском тексте. //Тюркология-88. V Бүткүл союздук түркология конференциясынын докладдарынын жана билдирүүлөрүнүн тезистерин. –Фрунзе: Илим, 1988.
4. Бодуэн де Куртэнэ И.А. Избранные труды по общему языкознанию. – М.: Издательство Академии Наук СССР, 1963.
5. Головин Б.Н. Язык и статистика. - М.: Просвещение, 1970.
6. Дыйканов К. “Манас” эпосунун алфавиттүү сөздүгү: 1-бөлүк. - Фрунзе, 1988., “Манас” эпосунун жиги алфавиттүү лугаты: 2-бөлүк. - Фрунзе, 1988.
7. Жубанов А.Х. Казак текстинин квантитативдик структурасы. – Алматы, 1987.
8. Ибрагимов С. Структурно-вероятностный анализ синтаксических отношений киргизского языка и их функционирование в речи. - Бишкек, 1994.
9. Лингвистические наследие Бодуэна де Куртэнэ на исходе XX столетия: Тезисы докладов международной научно-практической конференции 15-18 марта 2000 года. – Красноярск, 2000.
10. Мазекова Н.Ж. “Манас” эпосунда колдонулган сөз формаларына лингвостатистикалык илик (зат атооч). Канд.диссертация. – Бишкек, 2000.
11. Носенко И.А. Начала статистики для лингвистов. - М., 1981
12. Сагынбеков И., Садыков Т., Жетешиков Ж. О статистическом анализе лексики

и морфологии киргизского текста. //Тюркология-88. V Бүткүл союздук түркология конференциясынын докладдарынын жана билдирүүлөрүнүн тезистери. –Фрунзе: Илим, 1988.

13. Сагынбеков И. Ч.Айтматов тилинин сөздүгү. – Бишкек, 1994.

14. Слюсарёв Н.А. Теория Ф. де Соссюра в свете современной лингвистики. – М.: Наука, 1975.

15. Соссюр Ф.де. Заметки по общей лингвистике. – М.: Прогресс, 2000.

16. Фрумкина Р.М. Статистическая структура лексики Пушкина. //Вопросы языкознания, 1960, № 2.

17. Фрумкина Р.М. Применение статистических методов в языкознании. //Вопросы языкознания. 1960, № 4.

18. Шаршембаев Б.Д. “Манас” эпосунун жаңы типтеги сөздүктөрүн түзүү жана чечмелөө технологиясы. - Бишкек, 2004.