

А 2009
19667к

ҚАЗАҚ
ТЕКСІНІҢ
СТАТИСТИКАСЫ



СТАТИСТИКА
КАЗАХСКОГО
ТЕКСТА



АКАДЕМИЯ НАУК КАЗАХСКОЙ ССР
ИНСТИТУТ ЯЗЫКОЗНАНИЯ

СТАТИСТИКА КАЗАХСКОГО ТЕКСТА

выпуск II

*ТРУДЫ ГРУППЫ
АВТОМАТИЗАЦИИ ЛИНГВИСТИЧЕСКИХ ИССЛЕДОВАНИЙ
И СТАТИСТИКИ РЕЧИ*

АЛМА-АТА

«Гылым»

1990

ҚАЗАҚ ССР ҒЫЛЫМ АКАДЕМИЯСЫ
ТІЛ БІЛІМІ ИНСТИТУТЫ

ҚАЗАҚ ТЕКСІНІҢ СТАТИСТИКАСЫ

II шығарылуы

*«ЛИНГВИСТИКАЛЫҚ ЖҰМЫСТАРДЫ
АВТОМАТТАНДЫРУ ЖӘНЕ ТІЛ СТАТИСТИКАСЫ»
ТОБЫ ҚЫЗМЕТКЕРЛЕРІНІҢ ЕҢБЕКТЕРІ*

АЛМАТЫ

«Ғылым»

1990

Қазақ тексінің статистикасы. II шығарылуы: («Лингвистикалық жұмыстарды автоматтандыру және тіл статистикасы» тобы қызметкерлерінің еңбектері). — Алматы: Ғылым, 1990. — 208 б.

Оқырмандарға ұсынылып отырған бұл жинақ қазақ текстерін лингвостатистикалық тұрғыдан зерттеу және сол зерттеулерді автоматтандыру мәселелеріне арналған.

Мақалаларда түркі тілдерінің машиналық қорын жасау жайы қарастырылады, сөзформасын автоматты түрде морфологиялық талдаудың алгоритмдері келтіріледі. Сондай-ақ ЭВМ-нің көмегімен алынған жілік сөздіктер негізінде текстерге статистика-морфологиялық, лексика-семантикалық және фонетикалық талдау жасалады.

Кітап ғылыми қызметкерлерге, жоғары оқу орындарының оқытушыларына, сондай-ақ, филология факультеттерінің мамандары мен математикалық тіл білімі саласындағы студенттерге, аспиранттарға арналған. Жинақ материалдарын оқулықтар мен оқу құралдарын құрастыруда да пайдалануға болады.

Редакциялық коллегия:

А. К. Жұбанов (жауапты редактор), К. Б. Бектаев, Ә. Ж. Жүнісбеков

Жауапты шығарушы

А. Б. Белботаев

Статистика казахского текста. Выпуск II: (Труды группы автоматизации лингвистических исследований и статистики речи). — Алма-Ата: Ғылым, 1990. — 208 с.

Сборник посвящен проблемам автоматизации лингвистических исследований и статистического изучения казахского текста. В статьях рассматриваются вопросы создания машинного фонда казахского языка, приводятся алгоритмы автоматического морфологического анализа словоформ, фонетической транскрипции казахского текста с помощью ЭВМ, а также принципы семантико-синтаксической организации казахского текста на уровне абзацев. Дается статистико-морфологический, стилистический и фонетический анализ текста на основе приведенных в приложении сборника частотных словарей повестей, рассказов и пьес М. О. Ауэзова.

Сборник рассчитан на научных работников и преподавателей, студентов и аспирантов по структурной, прикладной и математической лингвистике, а также для филологических специальностей. Материалы сборника могут быть использованы при составлении учебников и учебных пособий.

Редакционная коллегия:

А. Х. Джубанов (ответ. редактор), К. Б. Бектаев, А. Д. Джунисбеков

Ответственный за выпуск

А. Б. Белботаев

С 4602010000—090
407(05)—90—131.90

ISBN 5—628—00551—7

©Институт
языкознания
АН Казахской ССР,
1990

АЛҒЫ СӨЗ

Тілдің ішкі заңдылықтарын ашуға арналған зерттеулерде қазіргі кезде ғалымдар сандық тәсілді мейлінше кеңінен қолдана бастады. Әртүрлі текстерді электронды есептегіш машиналар — ЭВМ-дер арқылы өңдеу — машиналық аударма, автоматты түрде индекстеу, аннотациялау, рефераттау, сөйлеу тілін анализдеу және синтездеу жұмыстары лингвистикалық объектілерді квантитативтік экспликация тәсілімен алгоритмдеу арқылы іске асырылады.

Статистикалық тәсілдер үнді-европа тіл білімінде өткен ғасырдың соңғы кезінен бастап қолданыла бастаса, түркі тілдерінде ол 60-жылдардың бас кезінен басталды десе де болады. Түркі текстерін статистикалық тәсілдермен зерттеу — көптеген оқымыстылардың пайымдауынша бұл тілдердің ішкі табиғатына, яғни ол тілдердің агглютинативтік құрылымына негізделіп отыр.

Тіл ғылымында сандық зерттеулердің қажеттігін математикалық лингвистиканы жақтаушылармен қатар аса көрнекті дәстүрлі тіл білімінің өкілдері де мойындап отыр. Олардың пікірінше, жазба әдебиет пен сөйлеу тіліндегі түрлі типті сөздердің әр ыңғайдағы жиілігін анықтайтын мұндай тәсіл тілдің құрылымдық-грамматикалық, кейде стильдік-семантикалық айырымдарын ажырата отыра осы стильдерге тән қасиеттерді анықтауға көмектеседі. Сондай-ақ, барлық грамматикалық категорияларды сандық тәсілмен талдау олардың бүкіл әдеби тіл сфераларындағы функциональдық салмағын да көрсетуге мол мүмкіндік беретіні сөзсіз.

Қазақ тілін лингвостатистикалық тұрғыдан зерттеу 1968 жылы Қазақ ССР Ғылым академиясы Тіл білімі институтында «Лингвистикалық жұмыстарды автоматтандыру және тіл статистикасы» тобының құрылуымен байланысты әрі жүйелі, әрі нәтижелі бола бастады. Қазақстандық статист-тілшілер 20 жылдан астам уақыттың ішінде өздерінің он шақты монографиялық еңбектері және

ғылыми-зерттеу жиіліктері мен 200-ден аса ғылыми және ғылыми-көпшілік мақалаларын шығарды. Сондай-ақ, ЭВМ-нің көмегімен түркі тілдерінің көптеген тексттерінен әр түрдегі жиілік сөздіктер алынып, көпшілікке ұсынылды.

1973 жылы аталған топтың күшімен республика лингвостатисттерінің он жылдық ғылыми жұмыстарын қорытындылайтын «Қазақ тексінің статистикасы — 1» (көлемі 30 баспа табақтан аса) атты ғылыми жинақ баспадан шықты. Ал қазіргі кезде топ қызметкерлері ЭВМ-нің көмегімен М. О. Әуезовтің 20 томдық толық шығармалар жинағы негізінде мақсатқа сай неше түрлі жиілік сөздіктері мен автоматтандырылған картотекалық қорды жасау жұмысын аяқтау кезеңінде.

Оқырмандарға ұсынылып отырған «Қазақ тексінің статистикасы-2» жинағы да өз ретінде жоғарыдағы осы аттас коллективтік еңбектің заңды жалғасы. Бұған енгізілген материалдар бес түрлі тақырыптарға топтастырылып қазақ және орыс тілдерінде берілді.

Жинақтың алғашқы бөлімі тіл біліміндегі жаңа көзқарасты білдіретін жалпы мәселелерге, яғни лингвостатистика пәнінің тарихына, оның ішінде лингвистикада математика әдісін қолданудың бастапқы идеяларына, түркі тілдеріндегі ықтималдық-статистикалық тәсілдің қолданылу тәжірибелеріне арналған.

Жинақтағы алғашқы материалдардың бірі — «Түркі тілдерінің машиналық қорын жасаудың алғы шарттары» деп аталынатын А. Қ. Жұбановтың информациялық мақаласы қазақ тілін зерттеуші оқырмандарға арнайы жазылған. Онда Москва қаласында алғашқы өткізілген орыс тілінің машиналық қорын құру мәселесіне арналған Бүкілодақтық конференцияда және Совет түркологтары Комитетінің XIV Пленумында әңгіме болған машиналық қор жасау мәселелерінің кей сәттері кеңірек сөз болады.

Жинақтың екінші бөлімі қазақ тілінің дыбыстық құрылымын статистикалық тұрғыда зерттеген мақалалардан құралған. Түрколог мамандардың зерттеу жұмысына кедергі келтіретін жайт — жалпыға ортақ түркі тілдерінің фонетикалық ерекшеліктерін дұрыс түсінуге бағыттайтын транскрипцияның жоқтығы — барша түркі тілдеріне, оның ішінде қазақ тіліне де тән жағдай. Осы мәселенің төңірегінде Ә. Жүнісбековтің мақаласы біраз жайлардан хабар береді. Сондай-ақ, бұл бөлімде қазақ тіліндегі сингармо дыбыстарға да статистикалық талдау жүргізілгенін айту керек (К. К. Алдабергенова, Ә. Ж. Жүнісбеков).

Қазақ тіл білімінің актуальды міндеттерінің бірі — дауысты, дауыссыз дыбыстардың қазақ тексіндегі энтропиясын анықтау. Осы тақырыпқа арналған Д. Ә. Молдалиеваның мақаласында Шеннон — Пиотровский тәсілімен жүргізілген эксперименттің нәтижесінде қазақ тексіндегі тіл молшылығы («избыточность») —

энтропия жайына, оның тілдегі құрылымдық-семантикалық және лексико-грамматикалық тұрғысына да әсер стетіндігі жайлы маңызды қорытындылар жасалады.

Жинақтың үшінші бөлімінде қазақ тілі лексикасы қарақалпақ, өзбек, башқұрт, орыс тілдерімен салғастырылған, статистикалық пайымдаулар берілген. Мұндай зерттеулердің негізгі мақсаты — дәстүрлі әдістермен ажыратылмайтын әр түркі тілдерінің тек өзіне ғана тән сыр-сипаттарын айқындауға әрекет жасау. Сөйтіп, квантитативті-салыстыру тәсілімен зерттеу арқылы түркі тілдерінің ішінде көп замандардан бері бірге жасасып келе жатқан қарақалпақ, қазақ, өзбек тілдерінің лексикалық жақындастығы айқындалды.

Келесі мақалалардың бірінде М. О. Әуезовтің драмалық шығармаларындағы зат есімдердің лексикалық тобына статистикалық талдау жасалса, енді бірінде 10 томдық «Қазақ тілінің түсіндірме сөздігіндегі» сөздердің көп мағыналылығының лексико-грамматикалық ынғайы сандық сипаттамалармен беріледі. Ал мұндай зерттеулер болашақта қазақ тіліндегі омоним сөздерді автоматты түрде ажыратуға арналған алгоритмін жасауға бастары анық.

Жинақ авторлары өмір талабына сай, семантикалық-синтаксистік тұрғыдан қазақ тексін абзац көлемінде зерттеудің қажеттілігін сөз етеді.

Сондай-ақ, бұл бөлімде қазақ тіліндегі сөз таптары мен оның морфологиялық құрылымына жүргізілген статистикалық зерттеулер топтастырылған.

Сонғы буын ЭВМ-дерінің үлкен мүмкіндіктерін түркологиялық зерттеулерді автоматтандыру үшін пайдалану — ғылыми творчествоға даяғыл жол ашып, сөйтіп бұл саладағы ғылымды жаңа сатыға көтеретіні даусыз.

Жинақтағы статистикалық материалдарды қазақ тілін дәстүрлі түрдегі болсын, не қолданбалы түрдегі болсын, сондай-ақ түркі тілдері мен оған жақын байланыстағы тілдерді салыстыра-салғастыра зерттеулерде де пайдалануға болады.

ПРЕДИСЛОВИЕ

В исследованиях, посвященных внутренним закономерностям языка и речи, в настоящее время все шире применяются количественные методы. Обработка текста на ЭВМ (машинный перевод, автоматическое индексирование, аннотирование и реферирование, автоматический анализ и синтез устной речи) осуществляется по алгоритмам, использующим методы количественной экспликации лингвистических объектов.

Статистический подход, нашедший применение в индоевропейском языкознании с конца прошлого века, в тюркских языках вошел в обиход примерно с начала 60-х гг. Это объясняется тем, что реальные основания для статистического описания тюркских текстов заложены в самой их природе, т. е. в агглютинативной структуре этих языков.

Как известно, необходимость количественных характеристик в науке о языке была осознана не только сторонниками «математической лингвистики», но и представителями традиционного языкознания. По мнению последних, точные методы в области изучения различных частотных данных многообразных типов слов, отражающих в потоке речи разные параметры как литературного, так и общенародного языка, помогают установить структурно-грамматические, а отчасти и семантические различия между стилями, а сам количественный анализ всякого рода грамматических категорий дает возможность уяснить их относительно функциональный вес в разных сферах литературного языка.

Лингвистическое изучение казахского языка стало более системным и результативным благодаря образованию в Институте языкознания АН КазССР (1968 г.) специальной группы по «Автоматизации лингвистических работ и статистике речи». За свою двадцатилетнюю историю казахстанские статоллингвисты выпустили

около десяти монографий и научно-исследовательских сборников, опубликовали более 200 научных и научно-популярных статей. Кроме того, с помощью ЭВМ были получены различные типы частотных словарей по многим характерным текстам тюркских языков.

В 1973 г. был выпущен сборник научных статей «Статистика казахского текста 1» объемом более 30 печатных листов. Он стал итогом работы за последнее десятилетие.

В настоящее время имеются все возможности для создания с помощью ЭВМ частотных словарей различного назначения и объема, частотных словарей и создания автоматизированного картотечного фонда на базе текстов 20-томного собрания сочинений М. О. Ауэзова. Наряду с этим осуществляется процесс поэтапного создания машинного фонда тюркских языков.

Предлагаемый читателям новый сборник «Статистика казахского текста» является продолжением первого коллективного труда.

Материалы сборника, представленные на русском и казахском языках, посвящены пяти темам.

В сборнике рассматривается история предмета, прослеживаются изначальные идеи математизации лингвистических исследований, обобщается опыт применения вероятностно-статистических методов в тюркских языках, изучаются проблемы создания машинного фонда тюркских языков. Одна из таких статей — «Предпосылки создания машинного фонда тюркских языков», где обобщаются некоторые моменты создания машин, проблемы, которые рассматривались на Всесоюзных конференциях по проблемам создания машинного фонда русского языка и на XIV пленуме комитета Советских тюркологов в г. Москве. Эта статья в основном адресована казахским традиционным лингвистам, всем, кто интересуется проблемами автоматизации лингвистических исследований.

В сборнике рассматриваются также фонетические проблемы статоллингвистики. Известно, что специалисты-тюркологи в своей исследовательской практике постоянно ощущают отсутствие общепринятой транскрипции, способствующей правильному восприятию фонетических закономерностей тюркских языков, в том числе и казахского.

Изучение энтропии гласных и согласных казахского текста является одной из актуальных задач казахского языкознания. Итоги проведенной одним из авторов эксперимента по методу Шеннона-Пиотровского позволяют сделать выводы об избыточности элементов речи, что в конечном счете сказывается на структуре семантической и лексико-грамматической организации текстов того или иного языка.

В этой связи представляет интерес также работа З. А. Сиразит-

динова, касающаяся статистических характеристик романа «Кап» башкирского писателя Д. Юлтыя.

В сборнике даются статистические выкладки по лексике казахского языка в сопоставлении последнего с каракалпакским, узбекским, башкирским и русским языками. Основной целью подобных исследований является осознание глубинных и специфических процессов сдвига лексики тюркских языков. С помощью количественно-сопоставительной методики предприняты попытки установить близость лексического состава трех тюркских языков — каракалпакского, казахского и узбекского, которые с давних времен находятся в тесном контакте.

В других статьях изучается статистика лексических пластов существительных в драматургических произведениях М. О. Ауэзова, на материале десятитомного толкового словаря казахского языка определяется мера многозначности слов, на этой базе иллюстрируется количественная картина речевой репрезентации лексико-грамматической омонимии в казахском языке. Предполагается, что подобное исследование будет полезным при составлении алгоритмов автоматического снятия омонимии на уровне лексики казахского языка.

Авторы сборника уделяют внимание направлениям в лингвистике, подчеркивают важность изучения казахского текста в семантико-синтаксическом плане на уровне абзаца.

Интерес ученых вызовут и статистико-лингвистические исследования морфологических структур различных частей речи казахского языка.

Использование возможностей современных ЭВМ для автоматизации тюркологических исследований — эта тема также нашла отражение в сборнике статистики казахского текста. Предлагаемые статистические материалы могут быть полезными для исследования казахского языка как в прикладном, так и в традиционном плане, а также при сравнительно-сопоставительном изучении самих тюркских и контактирующих с ними языков.

ТҮРКІ ТІЛДЕРІНІҢ
МАШИНАЛЫҚ ҚОРЫН ЖАСАУ
ЖӘНЕ СТАТИСТИКАЛЫҚ
ЗЕРТТЕУ МӘСЕЛЕЛЕРІ

ВОПРОСЫ СТАТИСТИЧЕСКОГО
ИССЛЕДОВАНИЯ И СОЗДАНИЯ
МАШИННОГО ФОНДА
ТЮРКСКИХ ЯЗЫКОВ

Е. Х. ДЖУБАНОВ

У ИСТОКОВ ИДЕИ МАТЕМАТИЗАЦИИ ЛИНГВИСТИЧЕСКИХ ИССЛЕДОВАНИЙ

Познавательная деятельность человека, направленная на объективную действительность, так же, как и ее результаты, осуществляется в форме логических категорий. К их числу относятся категории материи: качества, количества и меры, пространства и времени, сущности и явления, формы и содержания, причины и следствия, необходимости и случайности и т. д. Логические категории обнаруживают тесную взаимосвязь друг с другом, что является отражением взаимосвязи всех явлений объективной действительности. По выражению Ф. Энгельса, «количество и качество ...соответствуют друг другу взаимно и обоюдосторонне» [1].

Посредством установления качественных определенностей познание объединяет свои предметы в совокупности, внутри которых они как бы уравниваются, отождествляются друг с другом. При этом возможные определенности стираются, не фиксируясь познанием. Частичное разрешение этого противоречия возможно в процессе измельчения классов, охватывающих изучаемую область. Но этот процесс не бесконечен. Следовательно, механизм абстрагирования, лежащий в основе вычленения качеств, должен быть чем-то дополнен.

Познание разрешило эту проблему: между предметами, уже объединенными в классы, устанавливаются отношения, упорядочивающие предметы друг относительно друга. Результатом является установление нового типа определенностей — количественных.

Иначе говоря, категория качества, предшествуя категории количества, по началу своего формирования в процессе исторического развития мышления человека, вместе с тем выступает и как начальный этап познавательной деятельности человека, направленной на какой-либо предмет или явление объективной действительности, то есть, между категориями мышления существует определенная

зависимость. По образной ленинской формулировке, «сначала мелькают впечатления, затем выделяется нечто, — потом развиваются понятия качества и количества» [2].

Исторически первой познавательной системой, создавшей собственную логику, явился миф. Мифу свойственна жесткая логика бинарного членения мира; она задает наборы парных противопоставлений (небесное — земное, свое — чужое, священное — низменное и т. д.), которые и выделяют отдельные плоскости, рассекающие по различным направлениям создаваемую мифом модель мира и качественные определенности предметов. Мифологические качества, таким образом, выражаются гораздо четче «языковых».

В этом принципе бинарного членения кроется возможность создания в будущем специализированных моделирующих систем, сводящих до минимума характерную для естественного языка «поливариантность» в передаче значений.

Особую роль в процессах классификации играет естественный язык. Он является как бы всеобщим классификатором. Предметы, обозначаемые одним словом, попадают в один класс, уравниваются в определенном плане, хотя в других планах различия между ними могут быть и значительными. Язык тем не менее обнаруживает себя системой, хорошо приспособленной для вычленения в ней определенностей качественного типа. Языковые возможности, в частности, не всегда обслуживают все потребности практики.

Возможность фиксирования качественных определенностей как бы автоматически «вырастает» из структуры естественного языка. Что же касается упорядочивающих отношений, то здесь естественные языки оказываются довольно бедными как система: естественный язык создает наивную, донаучную модель мира. Именно поэтому познание постоянно увеличивает арсенал формальных языков, представляющих значительно большие возможности.

Еще Ф. Энгельс отмечал, что наличие количественных закономерностей может быть прослежено во всех областях науки. А так как количественные отношения действительного мира изучаются математикой, то из этого непосредственно следует применимость математики ко всем наукам. Углубление в познании количества означает вместе с тем более глубокое познание качества. В этой связи нам представляется несостоятельным утверждение, что математика отражает лишь количественную сторону объектов и игнорирует качественную.

В одной из своих историографических работ академик А. Н. Кононов отмечал, что величайшей заслугой Петра I перед русской наукой явилось учреждение Академии наук (1724—1725), в которой наряду с математическим и физическим классами был и класс гуманитарный [3]. Со дня основания Петербургской Академии

наук математические и филологические дисциплины изучались и развивались взаимосвязанно, обуславливая расцвет культуры вообще, были главными и равнозначными объектами отечественной науки. В стенах Академии был налажен регулярный выпуск научных журналов и монографий, переводов произведений классиков мировой литературы, научно-популярных изданий. Эта работа объединяла усилия представителей множества отраслей наук, которые сообща вырабатывали принципы перевода текстов, создавали научно-технические термины, отбирали из общенародной русской речи наиболее гибкие и точные обороты и выражения для передачи многих понятий. Очевиден вклад первого русского академика М. В. Ломоносова, который, как и академик В. К. Тредиаковский, успешно занимался теорией русского стихосложения, писал первый академический учебник грамматики, вырабатывал принципы формирования русского литературного языка. Как пишут исследователи, «в 1740-х годах он много переводил. Благодаря ему перевод перестал быть ремеслом и превратился в искусство... Выступал он в качестве редактора чужих переводов, в чем был признан непрекаемым авторитетом» [4].

В дальнейшем вторжение математической мысли в область языкознания принимает еще более целеустремленный характер. И это прежде всего связано с деятельностью знаменитой «Петербургской математической школы», основанной П. Л. Чебышевым (1821—1894). Его труды, как известно, составили целую эпоху в развитии мировой математической науки; признание его заслуг перед ученым миром выразилось в том, что он был избран членом академий почти всех стран мира и многих научных обществ. Эту же школу представляли академик М. В. Остроградский (1801—1861), автор выдающихся работ по вариационному исчислению, теории вероятностей, небесной механике, математической физике, гидромеханике и баллистике; академик В. Я. Буняковский (1804—1889), которому принадлежат ценные труды по интегральному исчислению, теории чисел и теории вероятностей, академик А. М. Ляпунов (1857—1918), изучивший устойчивость различных форм вращающейся тяжелой жидкости и др. [5].

Будучи убежденными материалистами в своих мировоззрениях, русские математики рассматривали элементы речи как частицу материального мира, поддающуюся количественному анализу и подсчету. Ученым было важно выявить действующие в речевом потоке статистические закономерности, которые регулируют соотношения языковых величин на том или ином уровне, подобный количественный подход к изучению лингвистических данных, по их мнению, давал возможность сделать не только отвлеченные математические выводы, но и получить наиболее конкретные сведения

о природе того или иного языка, что необходимо и в собственно языковедческих изысканиях. Более ста лет назад В. Я. Буняковский, в частности, дал оценку истории развития статистики, ее применения к проблемам языка и речи [6].

Если В. Я. Буняковский был первым, кто высказал мысль о возможности применения статистики в языкознании, то другой выдающийся математик А. А. Марков в 1913 г. впервые применил статистику к языковому материалу [7]. Он отсчитал 20 000 букв, которые заполняют первую главу, 16 строк второй главы «Евгения Онегина». Проведя, таким образом, 20 000 связанных испытаний, каждое из которых давало букву, обозначающую либо гласный, либо согласный звук, он исследовал закон вероятности: гласной следовать за гласной, гласной следовать за согласной, гласной следовать за двумя гласными, гласной следовать за согласной, которой предшествует гласная и т. д. А. А. Марков не ставил перед собой лингвистическую задачу, буквы и их сочетания были использованы им для выявления свойств знаменитых марковских цепей, т. е. для иллюстрирования связей испытаний в цепи.

Этой же теме была посвящена работа Н. А. Морозова «Лингвистические спектры» [8], опубликованная в 1915 г. Автор изучал средство для отличия плагиатов от истинных произведений того или другого известного автора. Под «лингвистическими спектрами» он понимал графики, отражающие особенности употребления служебных слов тем или иным писателем, наподобие световых спектров, по которым можно определить химический состав любого тела. Свою работу Н. А. Морозов считал «стилеметрическим этюдом».

Надо сказать, что в зародыше этот метод был известен еще в начале XIX в. в работах западно-европейских исследователей, изучавших «слоги» Платона и некоторых латинских писателей. Русские ученые тогда еще не занимались вопросами стилеметрии [9].

Интересно, что на эту работу Н. А. Морозова А. А. Марков написал рецензию, имевшую большое значение для становления статистической лингвистики как науки. В частности, рецензент указал на отсутствие оценки достоверности в исследованиях Н. А. Морозова, на недостаточный объем его выборки. «Только значительное расширение поля исследования (подсчет не 5 тыс., а сотен тысяч слов), — писал А. А. Марков, — может придать заключению некоторую степень основательности, если только границы итогов работы различных писателей окажутся резко отдельными, а не обнаружат другое — весьма вероятное — обстоятельство, что итоги всех писателей будут колебаться около среднего числа, подчиняясь общим законам языка» [10].

Несмотря на эти замечания, работа Н. А. Морозова представля-

ет значительное изыскание в области лингвостатистики в русском дореволюционном языкознании. В нем содержатся ценные наблюдения и выводы, не потерявшие значения сегодня.

Первые исследователи использовали статистику в основном для характеристики звукового строя. Так, еще в 1911 г. в «Ученых записках» Казанского университета появляются первые опыты В. Петрова и П. Александрова по вычислению процента участия отдельных звуков во французском и немецком языках. Оба исследователя подсчитали по 1 000 звуков из каждого языка соответственно. Наблюдения проводились только над художественными текстами. Однако выводы их малодостоверны, т. к. объем выборки недостаточен. Ценно то, что оба автора стремились сочетать количественный анализ фактов с их качественной характеристикой [11].

Почти в это же время В. А. Богородицкий в грамматике, изданной в Казани в 1913 г., описывает порядок употребительности звуков в русской речи, дает общее представление о количественных соотношениях в звуковом строе литературного языка [12].

В 20-е гг., когда статистический анализ языковых данных входит в обиход собственно лингвистических исследований, становятся известными работы А. М. Пешковского [13]. Он сумел зафиксировать большой объем разговорного текста, чтобы установить частоту появления отдельных звуков в устной речи. Особый интерес представляют его исследования по эвфонии (благозвучию) некоторых языков и текстов. А. М. Пешковский считает, что благозвучие языка зависит от величины соотношения музыкальных тонов и шумов. Музыкальные тоны — это совокупность гласных, сонорных, звонких шумных согласных, а шумы — это глухие шумные согласные. Путем статистического анализа звуков в рассказе И. С. Тургенева «Милостыня» (1500 звуков) и подсчета 10 тыс. звуков разговорной речи А. М. Пешковский составил весьма достоверную таблицу соотношения звуков и убедился, что в тургеневском тексте преобладают тоны над шумами. Он пришел к выводу, что отношение тонов к шумам почти одинаково в разных языках. Язык художественной литературы отличается от разговорного как по частоте появления отдельных звуков в потоке речи, так и по порядку следования звуков. Наибольшую величину благозвучия дают стихотворные тексты, наименьшую — научные и газетные. В языке художественной прозы благозвучия больше, чем в разговорной речи, но меньше, чем в стихотворных текстах.

В конце 20 — начале 30-х гг. математические методы исследования языка проникли в Среднюю Азию и Казахстан, хотя здесь трудно обнаружить единство взглядов на проблему. Так, в период подготовки к латинизированному алфавиту профессор Н. Ф. Яковлев выводит свою «математическую формулу построения алфави-

та» [14]. Имя этого ученого трижды встречается в одномомнике избранных работ Е. Д. Поливанова, изданном в Москве в 1968 г. Е. Д. Поливанов упоминает имя последнего в числе компаративистов, занятых проблемой родственных отношений баскского языка, далее поясняет суть предложенного Н. Ф. Яковлевым термина «губно-дорсальные», характеризующего некоторые абхазские согласные. И, наконец, он отмечает: «Из молодых же советских лингвистов наибольшего внимания, по-моему, заслуживает Н. Ф. Яковлев с его теорией эволюционных типов грамматического строя, обуславливаемых различными стадиями в истории общественных форм» [15]. В книге Е. Д. Поливанова нет какого-либо отношения или оценки «математической формулы» Н. Ф. Яковлева, хотя сам Е. Д. Поливанов спустя два года опубликовал полемичную статью «И математика может быть полезной». В ней он предложил использовать математику, включая дифференциальные и интегральные исчисления, в анализах, связанных с экспериментальной фонетикой; методы демографической статистики — в диалектологических исследованиях; при определении относительной вероятности этимологий — как достоверных, так и гипотетических и, наконец, фантастических — приложить теорию вероятностей и т. д. [16].

Х. К. Жубанов в своем проекте казахского алфавита по этому поводу писал: «Что же касается существа дела, то премудрая «математическая формула построения алфавита» была лишь ученым плагиатом. Это перевоплощенная в математику казахская система сокращения букв, созданная для заметания следов, как и сознавался сам Яковлев» [17]. «Дело в том, — говорит он, — что каждый звук казахского языка, за исключением трех — *л*, *к*, *г*, — которые звучат всегда мягко, имеет по две разновидности звучания: твердое и мягкое. Это дает возможность обозначать их одной буквой и т. д.»

Как известно, Х. К. Жубанов не отвергал вообще количественного метода применительно к языку. Наоборот, ему были присущи математическая точность в определении лингвистических закономерностей, количественная характеристика фактов при оценке качества тех или иных явлений, подход к элементам речи как к объективно-реальным величинам действительности, познание сути синтаксической комбинаторики по аналогии из области чисел и другие соображения, общие как для филологии, так и для математической логики. Так, он пишет: «Элементы словесного выражения напоминают собой элементы графического изображения чисел. Последние имеют не только непосредственное значение, паличное у отдельно взятых цифр: 1, 2, 3 и т. д., но и относительное значение, которое они приобретают и не могут не приобретать, когда со-

считаются друг с другом, как в случаях: 123 или 321, или 231, или $2+1+3$ и т. д. ... Благодаря тому, что цифрам придается не только непосредственное, но и относительное значение, вытекающее из местоположения каждой по отношению к другим, мы в состоянии изобразить на письме бесконечные числа всего лишь десятью знаками. Числа бесконечны и мы не в состоянии были бы передать их графическим способом, если бы не прибегали к помощи «синтаксиса чисел» [18]. Далее следует аналогия относительно того, что если сочетание цифр дает графически выраженные их отношения, то сочетания слов обнаруживают собой грамматически выраженные отношения, т. е. синтаксис.

На деле статистические поиски предприняты Х. К. Жубановым еще тогда, когда он взялся построить казахскую орфографию на научной основе, подчинив количество графически изображаемых букв алфавита практической, утилитарной значимости тех или иных казахских фонем, частоте их употребления.

В наши дни, разумеется, ручная обработка статистического материала не может удовлетворить запросы современной науки. Применение в лингвистике новых, более точных методов анализа языковых явлений на базе структурного моделирования, статистического описания с использованием ЭВМ из года в год возрастает. Вслед за первыми работами Р. М. Фрумкиной и частотным словарем Э. А. Штейнфельдта стали известны коллективные исследования. Так, общесоюзная группа «Статистика речи» работает под руководством проф. Р. Г. Пиотровского. Она объединяет ленинградских, минских, казахстанских, молдавских, дагестанских ученых-языковедов, математиков, инженеров-программистов. Группа изучает кодовые (энтропийные) свойства языка; определяет статистические характеристики словаря, грамматики, фразеологии, осуществляет автоматизированный анализ текста [19].

Группа статоллингвистики Института языкознания АН КазССР является координирующим центром тюркоязычных научно-исследовательских учреждений страны и работает над созданием автоматизированного фонда казахского языка. В этой связи надо сказать о мыслительной категории количества. Она есть результат отражения количественной определенности бытия. Поэтому категория количества является универсальной, т. е. логической, необходимой ступенькой познания действительности. Количественная определенность в объективной действительности предстает, с одной стороны, как дискретное, прерывное количество, а с другой стороны, — недискретное, непрерывное количество. Первое из них определяется посредством счета, а второе — посредством измерения. Соответственно этому в мыслительной категории количества выделяются два основных момента — число и величина.

Категория количества разнообразными грамматическими и лексическими средствами получает свое выражение во всех современных языках. Во всех языках существуют лексические обозначения определенных количеств, т. е. чисел. При этом если в некоторых языках этот разряд слов ограничивается несколькими или даже одним названием чисел в пределах первого десятка, то в большинстве языков он получил значительное развитие, выделившись в самостоятельную часть речи — числительное.

Точно так же во всех языках существуют лексические обозначения неопределенных количеств (много, мало, все, некоторые, всякий, каждый и т. п.). В большинстве языков грамматическая категория числа обычно включает единственное и множественное число, а в отдельных языках наряду с ними имеются также двойственное и даже тройственное число. Кроме того, особыми словообразовательными суффиксами или самим лексическим значением существительного может выражаться и собирательное множество (тряпка — тряпье, или стадо, толпа и т. п.).

В языках, где не выделяется особая грамматическая категория числа, существуют не только лексические, но и особые грамматические средства выражения множественности. Так, например, для выражения идеи множественности в некоторых языках пользуются способом удвоения существительных. А в китайском языке существительное само по себе обозначает не отдельный индивидуальный предмет, а некоторую совокупность однородных предметов, типа рус. «класс», «домашняя птица», «красная рыба».

В системе глаголов многих языков значения однократности-многократности, мгновенности-длительности и им подобные выражаются или видовыми формами, или словообразовательными средствами, указывающими на способ действия. Как видно, категория количества пронизывает всю систему языка: и лексику, и грамматику. Анализ всех этих способов в их историческом развитии и в связи с развитием категории количества как категории мышления может составить предмет ряда специальных исследований.

Диалектико-материалистическая теория познания исходит из того, что все знания об окружающем нас мире люди черпают не из самих себя, а из внешнего мира, который существует независимо от человека и его мышления и является единственным источником наших ощущений, представлений, понятий, единственным источником знаний. Поэтому мышление само по себе не может быть ни источником знания, ни критерием истины. Исходя из этого, диалектический материализм рассматривает истину как достоверность наших знаний, как соответствие понятия объекту, как знание, правильно отражающее объективную действительность.

Понятие экономного описания фактов в структурной лингвистике

тике рассматривается как прямое следствие применения точных формализованных методов. «Мышление человека тогда экономно, — писал В. И. Ленин, — когда оно правильно отражает объективную истину» [20].

Идеалом многих представителей структурной лингвистики является превращение языкознания в точную науку, следствием чего стало проявление особого внимания к логико-математическим методам исследования, повышенный интерес к логическим и вообще формальным аспектам языковедческого знания и предпочтение дедуктивных методов исследования, применяемых в математике, индуктивным методам, широко используемым в традиционной лингвистике.

Отсюда вытекает, что искусство применения логико-математических методов состоит не в механическом их перемещении из области математики и логики в область языка, а в исследовании тех сторон языка, которые лучше всего допускают применение этих методов, в выяснении границ их применения.

Математический аппарат сам по себе обладает достаточной точностью и надежностью, но для того, чтобы эти качества эффективно проявились, нужно точно определить сферу его применения, найти такую адекватную область, где работа аппарата давала бы конкретные результаты.

ЛИТЕРАТУРА

1. Маркс К., Энгельс Ф. Соч. Т. 20. С. 385.
2. Ленин В. И. Полн. собр. соч. Т. 29. С. 109—114.
3. Кононов А. Н. История изучения тюркских языков в России. Л., 1972. С. 31.
4. Макеев В. Н. История создания «Российской грамматики» М. В. Ломоносова. М.; Л., 1961. С. 35.
5. 220 лет Академии наук СССР. М.; Л., 1945. С. 10.
6. Буняковский В. Я. О возможности введения определенных мер доверия к результатам некоторых наблюдений, преимущественно статистических // Современник. Т. 131. Отд. 2. Спб., 1847.
7. Марков А. А. Пример статистического исследования над текстом «Евгения Онегина» // Известия Академии наук. 1913. Сер. 6. № 3.
8. Морозов Н. А. Лингвистические спектры // Известия отделения русского языка и словесности Академии наук. 1915. Т. 20. Кн. 1—4.
9. Ермоленко Г. В. Лингвистическая статистика. Краткий очерк и библиографический указатель. Алма-Ата, 1970. С. 12.
10. Марков А. А. Об одном применении статистического метода // Известия Академии наук. 1916. Сер. 6. № 4.
11. Петров В. Звуковая характеристика французского языка по статистическим данным; Александров П. Звуковая характеристика немецкого языка по статистическим данным // Учен. зап. Казанского Университета. Казань, 1911.
12. Богородицкий В. А. Общий курс русской грамматики. Казань, 1913. С. 39.
13. Пешковский А. М. Десять тысяч звуков // Методика родного языка, лингвистика, поэтика. Л., 1925; *Он же*. Вопросы методики русского языка, лингвистика и стилистики. М.; Л., 1923. С. 78.

14. Яковлев Н. Ф. Математическая формула построения алфавита // Культура и письменность Востока. 1928. № 1.
15. Поливанов Е. Д. Статьи по общему языкознанию. М., 1968. С. 185. (Впервые статья была опубликована в сб. «За марксистское языкознание», М., 1931).
16. Там же. С. 288—290.
17. Жубанов Х. К. Исследования по казахскому языку. Алма-Ата, 1966. С. 240.
18. Там же. С. 53—54.
19. Бектаев К. Б., Кенесбаев С. К., Пиотровский Р. Г. Об инженерной лингвистике // Вопросы языкознания. 1973. № 2.
20. Ленин В. И. Полн. собр. соч. Т. 18. С. 176.

А. Б. БЕЛБОТАЕВ, А. Х. ДЖУБАНОВ,
З. А. СИРАЗИТДИНОВ

О ЛИНГВИСТИЧЕСКИХ ИССЛЕДОВАНИЯХ С ПРИМЕНЕНИЕМ ВЕРОЯТНОСТНО-СТАТИСТИЧЕСКИХ МЕТОДОВ В ТЮРКСКИХ ЯЗЫКАХ

Последние двадцать лет в советском языкознании активно используются математические методы. Применение их не только в лингвистике, но и в социологии, психологии, археологии и других гуманитарных науках свидетельствует о том, что эти науки вступают в новый качественный этап своего развития [1—3].

Среди разнообразных математических методов наиболее разработанным и часто применяемым методом в лингвистике является вероятностно-статистический метод. Этот метод позволяет выявить те закономерности, которые действуют в языке и обусловлены многими массовыми процессами языка. Выяснить полностью их состав не представляется возможным.

Надо сказать, что отечественные и зарубежные ученые достаточно глубоко исследовали вопрос об использовании вероятностно-статистического метода.

Он успешно применяется при исследовании тюркских языков, таких, как казахский, узбекский, азербайджанский, киргизский, каракалпакский. Впервые этот метод был взят на вооружение казахстанскими языковедами. Для этой цели при Институте языкознания Академии наук Казахской ССР была создана группа ста-

тистико-лингвистических исследований и автоматизации. Она является неотъемлемой частью общесоюзной группы «Статистика речи».

Сейчас в Казахстане две группы математической лингвистики: чымкентская и алма-атинская. Они занимаются автоматизацией лексикографических работ и статистическими исследованиями казахских текстов, а также проблемами машинного перевода. Итоги работы казахстанских статистико-лингвистов обобщены в сборнике «Статистика казахского текста», изданном в 1973 г. [4]. В нем нашли отражение важнейшие выводы и рекомендации казахстанской школы структурно-прикладной и математической лингвистики. Так, приведены результаты опытов по составлению алфавитно-частотного словаря газет, сказок, языка художественной литературы, графемно-фонемной частотности казахских текстов, исследования по энтропии казахского языка и измерения смысловой информации слова. Впервые затрагивается проблема автоматизации и моделирования не только казахских текстов, но и тюркских. Авторы предлагают машинные индексы для знаков орхон-енисейской письменности, фонем каракалпакского и ногайского языков, звуков узбекского.

Многие важные вопросы, касающиеся вероятностно-статистического изучения языка и автоматизации этого процесса с применением средств вычислительной техники разработал К. Б. Бектаев. Его монография «Математические методы в языкознании» является фундаментальным пособием для всех лингвистов, использующих этот метод в своих работах [5].

Проблемы статистико-информационной типологии были впервые рассмотрены К. Б. Бектаевым в его монографии «Статистико-информационная типология тюркского текста» [6]. Она посвящена последовательному применению лингвистических приемов, вероятностных и информационных моделей к исследованию тюркского текста. Исследуя текст с помощью формальных статистико-вероятностных методов, автор описывает типологические особенности тюркских языков, сравнивая их с индоевропейскими языками. В Институте языкознания АН КазССР уделяется внимание новому направлению в исследовании тюркских языков. Так, в Алма-Ате был проведен Всесоюзный семинар «Статистическое и информационное изучение тюркских языков» (1969 г.)

Можно выделить основные исследования по фонетике и фонологии языка, слога и слоговой структуры слова, лексико-морфологической структуры текста и словоформ, а также функциональных стилей, синтаксиса текста, информационного характера языка.

Широко использовался статистический метод в работах по изучению лексико-морфологической структуры современных литера-

турных текстов. Главным компонентом такого изучения являются частотные словари: алфавитно-частотные и обратнo-алфавитно-частотные словари основ и словоформ. Методика составления разработана достаточно. Проблема обсуждалась на многих всесоюзных конференциях, опубликованы подробные руководства по составлению частотных словарей и процессу автоматизации. Так, общие принципы индексации и кодирования современных тюркских текстов при автоматическом составлении с помощью ЭВМ были разработаны в Институте языкознания АН КазССР К. Б. Бектаевым и А. Х. Джубановым.

Первым частотным словарем в тюркском языкознании следует признать «Словарь языка Абая» [7]. Он представляет собой толковый словарь с указанием частоты употребления слов, который был составлен традиционным методом (вручную).

Сегодня имеются частотные словари, наиболее полно охватывающие язык, по казахскому и узбекскому языкам: частотные словари отдельных авторских произведений, языка публицистики, детской литературы, науки.

Если вести речь о словарях, то широко известный письменный памятник средневековья «Кодекс Куманикус» в свое время был переведен на различные иностранные языки (латинский, немецкий, французский), однако ни на один из языков народов СССР не переводился. Поэтому издание этого памятника на казахском языке с указанием частотных показателей стало заметным событием отечественной тюркологии [8].

Частотный словарь по роману М. О. Ауэзова «Абай жолы» («Путь Абая»), изданный в Алма-Ате в 1979 г. [9], явился началом для составления частотного словаря-словоуказателя языка произведений одного из основоположников казахской советской литературы М. О. Ауэзова по полному собранию его сочинений в 20-томых. В этом лексикографическом труде найдут отражение лексическое и фразеологическое богатство языка писателя, индивидуальные особенности стилистических приемов словоупотребления и использования художественных и образительных средств казахского языка. Он дает богатый материал как для изучения особенностей языка писателя, так и для исследования статистической структуры казахской лексики вообще.

Объем исследуемых текстов произведений М. О. Ауэзова составляет более 2,5 млн словоупотреблений. В дальнейшем планируется также получить на ЭВМ распределительный (по томам) алфавитно-частотный словарь словоформ с указанием соответствующего номера тома и страниц, а также частоты употребления слов как по отдельным томам, так и в целом. Все тексты и словари-словоуказатели будут храниться на магнитных лентах.

Составлены и находятся на стадии завершения частотные словари каракалпакского и киргизского языков, исследуется язык публицистики. В 1988 г. увидел свет частотный словарь киргизского эпоса «Манас» К. Дыканова в двух частях.

Составление частотных словарей отдельных произведений авторов, функциональных стилей, всего литературного языка очень актуально. На их основе можно разрешить такой сложный вопрос, как лексико-грамматическая норма языка. Эти словари могут быть использованы и при сравнительно-типологическом изучении тюркских языков, а также при создании машинного фонда тюркских языков. Начало такому исследованию уже положено.

Следует отметить, что составление частотных словарей важно не только в плане лексико-грамматического изучения языков, но и в методическом. Так, выяснение вида статистического распределения классов слов языка имеет большое значение при изучении «чисто» лингвистических явлений статистическим методом. Результаты анализа частотных словарей могут быть полезны при выяснении фундаментальных свойств и закономерностей конкретного языка, например параметров Ципфа для конкретного тюркского языка.

В плане нормативного изучения языка представляет интерес диссертационная работа А. Ахабаева, посвященная изучению лексико-морфологической структуры современной казахской публицистики [10]. В ней рассматривается, в частности, проблема статистического критерия нормативности одного из вариантов дублетных слов. На основе статистического анализа показано, что частотные характеристики могут выступать в качестве объективного критерия. Посредством статистико-морфологического анализа изучен вопрос о стилевой дифференциации вариантов грамматических форм казахского языка.

В целом же работа посвящена статистико-морфологическому анализу форм окончаний имен существительных и местоимений на основе сравнения газетных текстов за 1965—1966 гг. (150 тыс. словоупотреблений) и художественного текста (романа М. О. Ауэзова «Абай жолы», второй книги, 124 384 словоупотребления). Многие казахстанские ученые работают в этом направлении. А. Х. Джубанов сделал статистический анализ лексико-морфологических форм прилагательных по тексту романа М. Ауэзова «Абай жолы». Результаты анализа позволили выявить своеобразие языка автора в использовании класса слов — прилагательных казахского языка и их морфологических средств [11].

Если предыдущие авторы рассматривали статистико-морфологическую структуру имен казахского языка, то С. Мырзабековым исследована структура глагола современного казахского языка:

основ и аффиксов, образующих производные формы глагола [12].

В диссертационной работе К. Молдабекова рассматривается проблема количественных особенностей лексики и морфологии языка казахской детской литературы и учебников начальных классов [13].

Статистические исследования казахских текстов проводятся и для выявления стилевых характеристик языка. Интерес представляют исследования А. Р. Зекиной, которая на материале частотных словарей изучает язык пьес М. О. Ауэзова и других драматургов казахской литературы [14, 15].

А. Б. Белботаев на материале частотных словарей статистически изучил структурные формы окончаний казахского языка [16], а также проблему лингвистических терминов [17], исследовал лексику иноязычных слов, вошедших в произведения М. О. Ауэзова через русский язык [18, 19].

Проведены исследования количественной характеристики местоимений каракалпакского языка [20, 21]. А. Бабанаров составил частотный словарь языка современных турецких газет, исследовал лексико-статистическую структуру газетных текстов [22]. Ж. Жетешников рассмотрел статистику словоизменительных аффиксов имен существительных киргизской публицистики, провел сравнительный анализ с данными казахского языка [23]. Статистический анализ использован С. Мухамедовым при изучении казахского языка узбекских газетных текстов в его диссертационной работе [24].

М. Хусанов статистический анализ использовал для оценки результатов, достигнутых реформой в области турецкой лексики [25]. С. Мамедовым изучены количественные характеристики словообразовательных элементов азербайджанского языка [26].

Т. М. Гарипов использовал статистический анализ для сравнительного лексико-семантического изучения тюркских языков. Результаты позволили выяснить, насколько такие языки, как башкирский, казахский, татарский, чувашский, отличаются друг от друга как в плане количественном, так и в качественном [27].

Статистический метод с успехом применяется и при сравнительном изучении грамматических форм разных языков. Такие исследования проведены по казахскому и русскому, казахскому и английскому языкам [28, 29]. Так, А. Ергалауовым этот метод использован при изучении форм страдательного залога в английском и казахском языках [30].

Вероятностно-статистический метод в области фонетики применяется как при инструментально-экспериментальном подходе, так и при традиционном слуховом, при изучении как фонологической системы, так и графической [31, 32, 33].

Частотные характеристики графем, фонем, фонемосочетаний получены по казахскому, узбекскому, каракалпакскому языкам. В этой связи интерес представляют частотные словари сочетаний букв казахского языка, фонем каракалпакского языка, а также частотный словарь знаковосочетаний Орхоно-Енисейского памятника [34].

В статье А. Х. Джубанова «К вопросу о графемной статистике казахского текста» сделана попытка статистического изучения казахского письменного текста на графемном уровне. Материал исследования может служить основой и примером для дальнейших работ, посвященных анализу морфемной структуры языка [35].

С. А. Ризаевым рассмотрен вопрос о частотности двухфонемных сочетаний в современном узбекском литературном языке [36]. По узбекскому и казахскому языкам изучены факторы, влияющие на относительную частотность фонем. Выяснено, что употребление фонем и их количественных отношений в разных жанрах различно. Сопоставлялись частотности фонем публицистического, научного и художественного по узбекскому и тексты сказок с научным по казахскому языкам. Статистический анализ позволил выделить группы фонем, определяющих специфику жанров.

С. С. Кенесбаевой применен статистический метод при определении наличия или отсутствия протетического звука в словах арабского происхождения, начинающихся со звуков, не характерных для исконно казахских слов [37]. А. И. Рабиновичем использована статистика для подтверждения теории о фонетической интерференции при контактах казахского и немецкого языков в условиях аудиторного билингвизма [38].

Проведены статистические анализы фонем турецкого и кумандинского языков [39, 40].

А. Джунисбековым и Д. Маматовым проведены статистические анализы спектров фонем [41, 42]. Первым — гласных казахского языка, вторым — сложных согласных узбекского языка. По казахскому и узбекскому языкам проведены статистические исследования фонетической структуры языка: выявлены частотности индвидуальных фонем, позиционных вариантов. Начато такое исследование и по каракалпакскому языку [43].

Изучены комбинаторные свойства фонем азербайджанского языка [44].

По казахскому языку Д. А. Байтанаевой проведены статистико-информационные измерения всех фонем. Выяснено, что информация различна в разных жанрах языка [45].

С. С. Татубаевым исследованы статистические характеристики благозвучия певческой речи в сравнении с разговорной речью.

Выявлены количественные характеристики, определяющие благозвучие казахской певческой речи [46].

Вероятностно-статистический подход к изучению структуры слога и слоговой структуры слова применен в таких языках, как казахский, татарский, узбекский. К. Б. Бектаевым составлен частотный словарь слогов казахского языка на материале стилей художественных, научно-технических, публицистических текстов и двухтомного «Толкового словаря казахского языка», изданных в Алма-Ате в 1959, 1961 гг. В словаре указаны частота слога, по отношению его позиции в слове. Отдельно приводится словарь односложных слов. Д. А. Байтанаевой изучены статистико-информационные свойства слогов казахского письменного текста на основе статистической обработки опытно-экспериментальных данных [47, 48].

Т. И. Ибрагимовым выявлены состав и структурные типы слогов современного татарского литературного языка, определены комбинаторные свойства типов слогов, частоты употребления их в слове. Исследованы закономерности слоговой организации слова, в частности, выяснено, что предыдущий слог татарского языка несет 47% информации о последующем слове, в значительной мере определяя его структурный и фонетический состав [49].

С. А. Ризаев исследовал статистическую структуру современного узбекского языка. Им выявлены типы слогов, изучены позиционные характеристики слога в слове, выяснены соотношения открытых и закрытых слогов, установлена каноническая форма слова современного узбекского подъязыка детской литературы [50].

Статистический анализ используется и при исследовании некоторых вопросов синтаксиса языка.

С. И. Ибрагимов на основе такого анализа текстов научно-технического характера выявил синтаксисеско-функциональные классы существительных киргизского языка, определил виды синтаксических связей существительных с другими частями речи. Статистический анализ применен и при исследовании порядка слов в предложениях в казахском языке. Это позволило выявить особенности структуры простого предложения таких функциональных стилей, как литературно-художественный и научный. М. Р. Джумабаевой рассмотрены особенности порядка слов в простых предложениях английского и казахского языков [51, 52, 53].

В 1987 г. издана монография А. Х. Джубанова «Квантитативная структура казахского текста (опыт лингвистического анализа на ЭВМ)». В ней были рассмотрены опыт автоматического построения частотных списков лингвистических единиц тюркоязычных текстов, а также алгоритмы получения словаря-словоуказателя. Особо выделены автором вероятностно-статистическая модель распределе-

ния классов слов казахского языка и другие вопросы лингвистики [54].

В монографии Т. Садыкова «Проблемы моделирования тюркской морфологии (аспект порождения киргизской именной словоформы)» на основе системного подхода анализируется морфологическая структура киргизского языка, а результаты анализа используются в целях построения модели порождения реальных именных словоформ по заданной лексико-грамматической информации [55].

Приведенный нами обзор опубликованных работ по тюркским языкам, в которых применяется вероятностно-статистический метод, показывает, что круг исследований, в которых метод используется, довольно широк, охватывает все разделы языка. Это объясняется тем, что метод является универсальным, не зависящим от объекта исследования.

Однако не все анализируемые работы являются равноценными. Если по казахскому, киргизскому, азербайджанскому и узбекскому языкам проведены фундаментальные исследования на уровне диссертационных работ, то по другим языкам эти исследования только начаты, и о ходе работ можно судить только по кратким сообщениям и тезисам. Также неравномерно исследованы этим методом и сами области языка. Так, фундаментальные работы по выяснению статистической структуры в плане синтаксиса по какому-либо тюркскому языку еще отсутствуют. Это является отражением того, что статистико-вероятностный метод начал использоваться в исследованиях тюркских языков сравнительно недавно. Тем не менее он успел определить себя как весьма полезный метод, и постепенно привлекает внимание ученых. Объясняется это прежде всего тем, что автоматизация трудоемких процессов в тюркологических исследованиях все еще желает лучшего. Проблема создания машинного фонда тюркских языков является жизненно важным делом и требует серьезного отношения со стороны тюркологов.

ЛИТЕРАТУРА

1. *Количественные методы в социологии*. М., 1966.
2. *Количественные методы в гуманитарных науках*. Минск, 1979.
3. *Математические методы в историко-экономических и историко-культурных исследованиях*. М., 1977.
4. *Статистика казахского текста*. Алма-Ата, 1973.
5. *Бектаев К. Б., Пиотровский Р. Г.* Математические методы в языкознании. Ч. 1: Теория вероятности и моделирование нормы языка. Алма-Ата, 1973. С. 282; *Бектаев К. Б., Пиотровский Р. Г.* Математические методы в языкознании. Ч. 2: Математическая статистика и моделирование текста. Алма-Ата, 1974. С. 336.
6. *Бектаев К. Б.* Статистико-информационная типология тюркского текста. Алма-Ата, 1978. С. 183.

7. *Абай тілі сөздігі*. Алматы, 1968.
8. *Құрышжанов А. К., Жұбанов А. Х., Белботаев А. Б.* Куманско-казахский частотный словарь. Алма-Ата, 1978. С. 276.
9. *Бектаев К. Б., Жұбанов А. Қ., Мырзабеков С., Белботаев А. Б.* М. Әуезовтің «Абай жолы» романының жиілік сөздігі (Частотный словарь по роману М. Ауэзова «Путь Абая»). Алматы, 1979. 334-б.
10. *Ахбаев А.* Статистический анализ лексико-морфологической структуры языка казахской публицистики. Алма-Ата, 1971. С. 24.
11. *Жұбанов А. Х.* Статистическое исследование казахского текста с применением ЭВМ (на материале романа М. О. Ауэзова «Путь Абая»). Алма-Ата, 1973. С. 37.
12. *Мырзабеков С.* Статистико-лингвистический анализ структуры глагола современного казахского языка (по роману М. О. Ауэзова «Путь Абая»). Алма-Ата, 1973. С. 32.
13. *Молдабеков К.* Лингвистические исследования казахских текстов для младших школьников (на материале текстов учебников начальных классов). Алма-Ата, 1985.
14. *Зекенова А. М. О.* Әуезов пьесаларындағы есім сөздердің морфологиялық құрылысы жайында. // Известия АН КазССР. Сер. филолог. 1976. С. 56—62.
15. *Зекенова А. О* лингвистическом исследовании драматургии М. О. Ауэзова // Материалы семинара «Статистическая оптимизация преподавания языков и инженерная лингвистика». Чимкент, 1980. С. 219—220.
16. *Белботаев А. Б.* Казак тіліндегі жалғаулар және олардың құрылымдық формалары туралы // Материалы республиканской научно-теоретической конференции молодых ученых по общественным наукам. 24—25 февраля 1978 г. Алма-Ата, 1978. С. 181—183.
17. *Белботаев А. Б., Алдабергенова Қ. К.* Казак тілі біліміндегі статистикалық терминдер мен сөз тіркестері туралы // Материалы республиканской научной конференции молодых ученых и специалистов по общественным наукам, посвященной 110-й годовщине со дня рождения В. И. Ленина (17—18 апреля 1980 г.). Алма-Ата, 1981. С. 178—179.
18. *Алдабергенова Қ. К., Белботаев А.* Статистическое исследование морфологической структуры иноязычных слов, вошедших в произведения М. Ауэзова через русский язык // Материалы семинара «Статистическая оптимизация преподавания языков и инженерная лингвистика». Чимкент, 1980. С. 149—150.
19. *Алдабергенова Қ. К., Белботаев А. Б.* Количественно-морфологический анализ слов из русского языка, вошедших в рассказы М. О. Ауэзова // Вопросы тюркского языкознания. Алма-Ата, 1985. С. 164—166.
20. *Бектаев К. Б., Насыров Д. С., Аимбетов М. К.* Количественная характеристика личных и лично-притяжательных местоимений каракалпакского языка // Вестник Каракалпакского филиала АН УзССР. 1982. № 1. С. 67—71.
21. *Насыров Д. С., Бектаев К. Б., Аимбетов М. К.* Количественная характеристика указательных и указательно-притяжательных местоимений каракалпакского языка, употребляемых в публицистике // Вестник Каракалпакского филиала АН УзССР. 1981. № 4. С. 79—82.
22. *Бабанаров А.* Частотный словарь и автоматический словарь для машинного перевода турецких газетных текстов // Инженерная лингвистика и оптимизация преподавания иностранных языков. Л., 1980.
23. *Жетешиков Ж.* Статистический анализ словоизменительных аффиксов имен существительных киргизского языка (на материале газетных текстов) // Материалы семинара «Статистическая оптимизация преподавания языков и инженерная лингвистика». Чимкент, 1980. С. 210—211.
24. *Мухамедов С.* Статистический анализ лексико-морфологической структуры узбекских газетных текстов. Ташкент, 1980. С. 25.
25. *Хусаинов С.* Из опыта статистического анализа частотности употребления

исологизмов в современном турецком литературном языке // Советская тюркология. 1981. № 4. С. 74—79.

26. *Мамедов Э. С.* Количественные характеристики словообразовательных элементов // Научные труды Азербайджанского пединститута рус. яз. и лит.-ры. Сер. яз. и лит. Баку, 1979. № 1. С. 105—110.

27. *Гарипов Т. М.* Лексикостатистические отношения тюркских языков Урало-Поволжья. // Языковые контакты в Башкирии. Уфа, 1972. С. 218—267

28. *Ишанов К. И., Садиков А. В., Мырзабеков С.* Частотно-сопоставительная характеристика надежных форм русского и казахского языков // Тезисы докладов и сообщений всесоюзного семинара «Статистическое и информационное изучение тюркских языков». Алма-Ата, 1969. С. 56—58.

29. *Ахметов Е. Г.* Статистическая характеристика родительного падежа имени существительного в английском и казахском языках // Тезисы докладов и сообщений всесоюзного семинара «Статистическое и информационное изучение тюркских языков». Алма-Ата, 1969. С. 51—54.

30. *Ергалауов А.* Статистическая характеристика форм страдательного залога в английском и казахском языках // Тезисы докладов и сообщений всесоюзного семинара «Статистическое и информационное изучение тюркских языков». Алма-Ата. 1969. С. 62—63.

31. *Джубанов А. Х.* К вопросу о графемной статистике казахского текста // Вопросы казахской фонетики и фонологии. Алма-Ата, 1979. С. 79—86.

32. *Джубанов А. Х., Джунисбеков А.* Преобразование с помощью ЭВМ казахских текстов в фонетическую запись // Материалы семинара «Статистическая оптимизация преподавания языков и инженерная лингвистика». Чимкент, 1980. С. 317—319.

33. *Исенгельдина А. А.* Факторы, определяющие относительную частотность фонем // Статистика казахского языка. Алма-Ата, 1973. С. 659—664.

34. Статистика казахского текста, Алма-Ата, 1973.

35. *Джубанов А. Х.* К вопросу о графемной статистике казахского текста // Вопросы казахской фонетики и фонологии. Алма-Ата, 1979. С. 79—86.

36. *Ризаев С. А.* К вопросу о частотности двухфонемных сочетаний в современном узбекском литературном языке // Опыт экспериментального и структурного изучения языка. Ташкент, 1982. С. 45—92.

37. *Кенесбаева С. С.* К вопросу исследования протетического звука в словах арабского происхождения статистическим методом // Тезисы докладов и сообщений всесоюзного семинара «Статистические и информационные изучения тюркских языков». Алма-Ата, 1969. С. 43—44.

38. *Рабинович А. И.* О статистическом исследовании казахско-немецкой фонетической интерференции // Там же. С. 65—66.

39. *Чикадзе В. Г.* Статистический анализ фонем в односложных словах турецкого языка // Сообщения АН СССР. Тбилиси. Т. 96. № 1. С. 201—209.

40. *Селютина И. Я.* Относительная частотность согласных в языке кумандинов // Экспериментально-фонетические исследования сибирских языков. Новосибирск, 1981. С. 76—92.

41. *Джунисбеков А.* Гласные казахского языка. Алма-Ата, 1972. С. 94.

42. *Маматов Д.* Количественный анализ сложных согласных // Опыт экспериментального и структурного изучения языка. Ташкент, 1982. С. 108—111.

43. *Насыров Д. Х., Аимбетов А.* О статистике фонем каракалпакского языка // Материалы семинара «Статистическая оптимизация преподавания языков и инженерная лингвистика». Чимкент, 1980. С. 273—274.

44. *Вейсалов Ф. Е., Керимов Я. К., Сафаров М. Г.* Комбинаторика фонем современного азербайджанского языка // Вопросы азербайджанского языкознания. Баку, 1981.

45. Байтанаева Д. А. Информационные характеристики казахского текста. Алма-Ата, 1985. С. 18.

46. Татубаев С. С. Статистические подходы к исследованию певческой фонетики казахского языка // Тезисы докладов и сообщений всесоюзного семинара «Статистическое и информационное изучение тюркских языков». Алма-Ата, 1969. С. 38—40.

47. Бектаев К. Б. Алфавитно-частотный словарь слогов казахского языка // Статистика казахского текста. Алма-Ата, 1973. С. 566—612.

48. Байтанаева Д. А. Энтропии распределения слогов казахского письменного текста // Тезисы докладов и сообщений всесоюзного семинара «Статистическое и информационное изучение тюркских языков». Алма-Ата, 1969. С. 80—82.

49. Ибрагимов Т. И. Некоторые статистические данные о слогах татарского языка: вероятностные методы и кибернетика // Учен. зап. Казанского ун-та. Вып. 4. Т. 25. Кн. 6. 1966. С. 74—78; *Его же*. Изучение слогов и структуры их сочетаний в татарском литературном языке. Казань, 1970.

50. Ризаев С. А. Из опыта исследования статистической структуры слога в узбекском языке // Общественные науки в Узбекистане. Ташкент, 1972. № 1; *Его же*. Статоллингвистический анализ структуры слова в современном узбекском языке // Материалы семинара «Статистическая оптимизация преподавания языков и инженерная лингвистика». Алма-Ата, 1980. С. 180—182.

51. Ибрагимов С. И. Некоторые статистические характеристики имен существительных киргизского языка // Тезисы докладов и сообщений всесоюзного семинара «Статистическое и информационное изучение тюркских языков». Алма-Ата, 1969. С. 34—38.

52. Ахабаев А., Белботаев А., Молдабеков К., Осижигитова Н. Статистика порядка слов в предложении // Там же.

53. Джумабаева М. Р. О статистической характеристике простого предложения в английском и казахском языках // Там же. С. 62—63.

54. Джубанов А. Х. Квантитативная структура казахского текста (опыт лингвистического анализа на ЭВМ). Алма-Ата, 1987. С. 147.

55. Садыков Т. Проблемы моделирования тюркской морфологии (аспект порождения киргизской именной словоформы). Фрунзе, 1987. С. 121.

А. К. ЖУБАНОВ

ТҮРКІ ТІЛДЕРІНІҢ МАШИНАЛЫҚ ҚОРЫН ЖАСАУДЫҢ АЛҒЫШАРТТАРЫ

Электронды техниканың дүниеге келуі қазіргі кезде инженерлік-экономикалық міндеттерді шешумен бірге тіл білімі саласында да өз көмегін тигізуде. Мысалы, электронды есептеу машиналары арқылы бір тілден екінші тілге аударуға, ғылыми-техникалық текстердің рефераттары мен аннотацияларын автоматты жолмен жасауға, қажетті информацияларды жедел тауып алуға кең жол ашылууда. Электронды есептеу машиналар (бұдан әрі — ЭВМ) бұл күндері практикалық тіл білімі саласындағы әсіресе еңбекті көп ті-

лейтін операцияларды ішінара автоматтандыруға, текст түріндегі материалдарды әралуан белгілеріне қарай сұрыптауға жиі пайдаланылып отыр.

Электрониканың өркендеуі, дәл методтардың қолданыс табуы тіл білімінің сан алуан салаларын, соның ішінде, лингвистика ғылымын да жеделдете дамытуға ықпалын тигізуде. Үнді-европа тіл білімі мамандары қолданған математикалық және кибернетикалық методтардың әсері түркітану бағытындағы зерттеулерге де тимей қойған жоқ. Алайда бұл құбылыс 60 жылдарға дейін айтарлықтай еленіп, ескерілген жоқ.

Біз түркі тілдерін ЭВМ-ді қолдана отырып санамалы (количественный) метод жолымен зерттеудің белгілі бір теориялық мәні барлығын атап айтуымыз қажет. Өйткені үнді-европаның копармалы (флективті) аналитикалық тілдері бойынша жинақталған ұлан-ғайыр статистикалық материалдарды жалғамалы (агглютинативті) түркі тілдерінің соңғы статистикалық деректерімен салыстырудың мүмкіншілігі артады. Зерттеулеріміздің нәтижелеріне қарағанда, түркі тілдерінің табиғаты сөйленіс ағымында жекелеген сөз тұлғаларының түбегейлі фонетикалық және морфологиялық өзгерістерге ұшырауын қажет етпейтіндігі, олар негізінен өздерінің бітім-бейнелерін сол қалпында сақтайтындығы айқындалып отыр. Дыбыстардың үнді-европа тілдеріндегі толассыз тоғысулары, әлсіреп редуциялануы тағы басқа қасиеттер түркі тілдеріне тән емес. Ана тіліміздің осы тәрізді ерекшеліктеріне сүйеніп, оның алуан сырлы құбылыстарын математикалық жолмен зерттеуге дәл және даусыз хаттауға түркі тіл білімінің бірсыпыра процестерін автоматтандырудың тигізер пайдасы мол.

Лингвистикалық зерттеулерді автоматтандыру кезінде «адам—машина—адам» жүйесі бойынша жұмыс істеледі. Сондықтан адам мен ЭВМ-нің қызметіндегі кейбір ұқсастықтар мен айырмашылықтар мыналар: машина (ЭВМ) шапшаң әрекет жасайды және барлық амалдарын дәлме-дәл орындайды, оның жадындағы информациялар ұзақ мерзім бойында сақталады, әрі қажетті сәтінде қолма-қол іздеп табуға мүмкіншілік береді. Ал адамның ми қызметінің ерекшелігі — эвристикалық немесе тапқыштық қабілетінде (мидың қызметінің басқа қабілетін айтпағанның өзінде). Сонымен лингвистикалық зерттеулерді автоматтандыру жұмысын ЭВМ мен адамның ойлау қабілетіндегі ерекше сипаттарды ескере отырып жүзеге асыруға тиіспіз. Шет елдердегі, сондай-ақ Совет Одағындағы ғылыми топтардың қазіргі жұмыс тәжірибелерінің негізінде мынадай тоқтамға келуге болады: ЭВМ-нің көмегімен тек қолданбалы тіл білімі ғана емес, тіпті басқа да тіл зерттеу саласындағы өзіндік ішкі теориялық проблемаларды шешуде де қомақты нәтижелерге қол жететіні даусыз. Яғни электронды машиналардың кө-

мегімен лексиканы, грамматиканы, фонетиканы, қазіргі және көне жазба текстердің басқа да сипаттарын хатқа түсіруде оның берері мол. Лингвистикалық процестерді автоматтандыру жұмысы ғалымдардың тұжырымы бойынша екі деңгейде болмақ: а) текстерді комбинаторлы-статистикалық әдіспен суреттеп шығу, ә) текстерді семантикалық шешімдер қабылдау тұрғысынан қарастыру.

Лингвистикалық зерттеулерде ЭВМ-ді қолдану көбіне әріптердің, фонемалардың, сөздердің, сөз тұлғаларының, сөз тіркестерінің т. б. жиілік тізбелерін (сөздіктерін) жасаудан басталады. Әр дәуірдегі түрлі авторлардың шығармаларының текстеріне сәйкес шығарылған жиілік көрсеткіштерінің арқасында қазіргі кезде жалпы лексикологияның мәселелері жалпыхалықтық лексиканың құрылымы оның әр алуан қабаттарының арақатынасы, қазіргі замандағы тілдік норма, жазушы тілінің функционалды стилі мен ерекшелігі, сөз қорының тарихи дамуы т. б. аса маңызды проблемалар принциптік тұрғыдан жаңаша пайымдалып отыр.

Тек ЭВМ-ді қолданудың арқасында ғана кері алфавитті жиілік сөздігі тәрізді құралдарды шығару мүмкін болды. Мұндай сөздік әрбір сөздің грамматикалық сипатын түсіндіруге аса оңтайлы болғандықтан морфология саласындағы түрлі зерттеу жұмыстарының міндеттерін анағұрлым жеңілдетеді. Мәселен, оны белгілі бір тілдегі түрлендіруші және сөз жасаушы тәсілдердің толық сипатын айқындауға санамалы қатынастарды білуге, сондай-ақ фонетика мен фонология өрісіндегі толып жатқан мәселелерді шешуге пайдалануға мейлінше болады. Өйткені, кері алфавитті сөздікте жазылуын реттестіретін сөздердің жиіны бесенеден белгілі болып көрініп тұрады. Қазірде ЭВМ арқылы әралуан жиілік сөздіктерді жасаудың тәжірибелері бірсыпыра жинақталды. Мұндай сөздіктер түрліше мақсатпен және әрқелкі көлемде үнді-европа тілдері бойынша да, сондай-ақ түркі тілдері бойынша да көптеп жасалуда.

ЭВМ-ді лингвистикалық немесе жалпы филологиялық мақсаттардағы зерттеулерге пайдаланумен бірге сол ғылыми ізденістердің нәтижелерін іске асыру мақсатында да қолдануға болады. Мұның өзі ғылыми-техникалық процестерді жеделдету туралы КПСС XXVII съезі алға қойған міндеттерге сай ғылыми-зерттеу істерін және қазіргі өндірістік процестерді кеңінен автоматтандыру туралы талаптарға барынша үндесіп жатады. Осылайша электронды машиналарды пайдалана отырып тілдерді зерттеу және ғылыми қорытындыларды табиғи тілдердің информациясын автоматты жолмен танып-білу мүмкіндігіне саналы түрде көз жеткізудің арқасында ғалымдар орыс тілінің машиналық қорын жасау қажеттігі жөніндегі идеяларды жүзеге асыра бастады. Осы мақсатпен 1983 жылы шақырылған бірінші және 1987 жылы шақырылған екінші Бүкілодақтық конференцияларда орыс тілінің машиналық

қорын жасау проблемалары күн тәртібіне қойылып, онда машинаға түсірілген тілдің құрамы мен құрылымы қандай болуы тиіс екендігі, бірінші кезекте мұндай тіл қандай міндеттер атқаратыны, машина қоры қамтитын тіл көздерінің ауқымы қандай болмақтығы (минимум және максимум мөлшерде), машиналық тіл қорын дәйекті түрде тұтыну үшін қандай ұйымдық және техникалық шаралар атқарылуы керектігі талқыланды.

Осы программаны жүзеге асыруға, «қағазсыз информацияның» технологиялық тәсілдерін игеруге, сөйтіп өз еңбегінің өнімділігін арттыруға академиялық ғылыми мекемелер мен жоғары оқу орындарындағы тіл мамандарының мейлінше ынталы екендігін конференция айқын танытты.

Егер алғашқы конференцияда орыс тілінің машиналық қоры қандай болуы қажет деген ділгерлігі күшті проблемалар талқыланған болса, екінші конференцияда осы салада орындалған жұмыстардың нәтижелері сөз болды және жаңа саладағы жұмыс барысында орыс тілінің және басқа ұлттар тілдерінің машиналық қорларын жасаумен байланысты күн тәртібіне тосыннан қойылған соны проблемалар сөз болды. Атап айтқанда СССР халықтары тілдерінің машиналық қорларын жасау мәселесі өз алдына жеке тақырып ретінде қаралды. Ерекше атап кетерлік бір жайт мынау болды: конференцияда орыс, украин, грузин, эстон тілдері жөніндегі баяндамалармен қатар түркі тілдерінің, оның ішінде қазақ тілінің машиналық қорын жасау мәселесі жеке тыңдалды. Сөйтіп, қазақ текстерін ЭВМ-ге енгізу және автоматты жолмен зерттеу бағытындағы біздің көпжылдық тәжірибелеріміз бен қол жеткен нәтижелеріміз ескеріліп, екінші конференцияның шешімдерінде түркі тілдерінің машиналық қорын жасайтын координациялық орталық Қазақ ССР ҒА Тіл білімі институтының құрамынан ашылсын деген дербес қаулы қабылданды.

Біз осы мақаламызда тілдің машиналық қорының мән-жағдайы жайлы өз ойларын ортаға салған ірі ғалымдардың жоғарыда сөз болған ғылыми форумдағы кейбір тұжырымдарына тоқтала кетпекпіз.

Машиналық қор жайындағы мәселені ең алғаш рет белгілі совет ғалымы, информатика саласының маманы А. П. Ершов 1978 жылы «Іс қағаздары прозасының феномены: диалогтар жүйесін жасау методологиясына» атты мақаласында сөз етті¹. Бұл мәселенің ойдағыдай шешілуіне ғалым үлкен мән беріп, егер орыс тілінің машиналық қорын жасау мақсаты ойдағыдай орындалса, модельдеу мен алгоритмдерді құрастыру саласындағы қол жеткен табыстарымыздың бәрі академиялық жаттығулар қалпында қалып

¹ Машинный фонд русского языка: идеи и суждения. М., 1986. С. 7—12.

қоймай, іс жүзіне асқан болар еді, мұндай өте-мөте қажетті проблеманың шешілуі ғылым, мәдениет және қолданбалы техникалық тәсілдер тұрғысынан мейлінше құнды болмақ деген болатын. Оның пікірінше, маман лингвистер таралынан орыс тілінің машиналық қорын жасау ісі өндірістік лингвистикалық жүйелерді іске қосудан озық болуы тиіс, өйткені тек осы жағдайда ғана толып жатқан қажетсіз қайталаулардан құтылып, тілдің жанды өзегін озбырлар мен дүмшелердің қиянатынан аман сақтап қалуға болады.

«Орыс тілін сүйетін қоғамның (формальды қоғам) танымынша, дейді ол, орыс тіл білімі дегеніміз әзірге ұйықтап жатқан Гулливер тәрізді, сондықтан оның оянатын және өзі туралы бар дауыспен жар салатын кезі енді жетті. Ал, ЭВМ бұл ретте таптырмайтын қоңыраулы сағат секілді», — дейді ғалым. А. П. Ершов болашақ жасалатын орыс тілінің машиналық қорып тіл жөніндегі білімдердің аса қуатты концентраты болады деп сенеді. Мұндай формальды жүйе табиғи тілдің организммен сәйкес әрі тең көлемді болуы қажет, сөйте тұра ол (модель) анатомиялық тұрғыдан мүшеленген, ашық, айқын бақылауға, зерттеуге қолайлы әрі өзгермелі болуы тиіс. Алайда, қазіргі жансыз материяның формаларын модельдеуге қатысты білімдеріміздің аясын осы күнгі тіл білімінің статусымен және ауқымымен салыстырып қарайтын болсақ, біз аталған салада жүріп өтпек жолымыздың бастауында тұрғанымыз айқын сезіледі.

Ю. Н. Қарауловтың «Лингвистикалық зерттеулердің методологиясы және орыс тілінің машиналық қоры» атты баяндамасы² орыс тілінің машиналық қоры дегеннің ғылыми түсінігі — көп тармақты әрі терең автоматтандырылған жүйе дегенді аңғартады. Бұл жүйе информациялық барлау міндеттерін де, сондай-ақ, тіл білімі саласындағы, кенірек алғанда, жалпы филологиялық бағыттағы зерттеушілік міндеттерді де атқара алады. Орыс тілінің машиналық қорын жасаудың методологиялық негіздері ретінде екі тезис басшылыққа алынған: бірінші, — тіл жөніндегі деректердің қандай болмасын лексикографиялық пішінде ұсынылады. Егер кейбір тіл жүйелерінің тарамдары немесе тұтас бір деңгейлері (мысалы, синтаксис) лексикография тұрғысынан сипаттауға әлі икемделіп жетпеген болса да, принциптік жағынан оны жүзеге асыру толық мүмкін деп саналады. Екінші, — лексикографиялық жағынан жүйеленген тілдік мәліметтердің барлығы да машинаға лайықталып, алгоритм тіліне аударуға келеді.

Сонымен, тілдің машиналық қоры деген ұғымның түпкі төркіні лексикографиялық, яғни сөздік түріндегі жұмыстардың идеясына келіп тіреледі. Осы түсінікке сәйкес және басқа да ойларды қамти

² Бұл да сонда. 13—25-б.

келе Ю. Н. Қараулов машиналық қордағы тілдік информацияларды шоғырландыратын бірегей түйін немесе түп қазық жеке сөз деп есептейді. Өйткені, сөз тілдің ең табиғи бөлшегі болып табылады. Ол морфемалармен, сөз орамдарымен немесе сөйлемдермен салыстырғанда да ішкі-сыртқы қызметтері жағынан яғни морфемаға, буынға, әріпке, мағынаға талданатындығы, сондай-ақ, сөздерді өзгертетіндігі, жаңа сөз тудыратындығы, синтаксистік, грамматикалық, ассоциялық және концептуалды қатарлар жасайтындығы жағынан ерекшеленеді. Осыған орай, информация сақтағыштық қасиеттің негізінде мынадай бір кепілдеме болады: ол сөз — белгілі бір жүйедегі сөз қорын танытады, оның белгілі мөлшерде өзіндік сегменттері болады, олар грамматикалық, стилистикалық, фонетикалық, контекстік т. б. параметрлердің, демек, сол сөз жайындағы барлық мәліметтердің бүкіл жүйесін сәулелендіреді.

Мынадай екі жақты құбылысты қарастыра кетуге болады: тілдің машиналық қоры жасалып болғаннан кейін ол тек архив тәрізді қызмет ете ме, әлде машиналық тіл қоры тіл жөніндегі жаңа білімдер қосатын, жаңа мәліметтер тауып беретін зерттеу лабораториясы болмақ па? Бұл ретте айтарымыз: екі жақты құбылыстың соңғысы алғашқы қисынның барлық қасиеттерін иемденетіндігі себепті берер мүмкіншіліктері де молырақ. Демек, тіл қоры бізге белгілі тілдік заңдылықтар жөніндегі түсініктерді бойына жинақтаған тілдің синхронды және диахронды процестерін моделдей алатын әрекет үстінде болуы қажет. Сонда ғана машиналық қор түрлі деңгейдегі тіл бөлшектерін, жүйелерді негізге алып, қызмет ететін болады. Басқаша айтқанда тілдің машиналық қоры дегеніміз — барлық мәліметтерді жинақтаушы банк, бірақ ол «ақылды», мамандандырылған білім қоры болып табылады. Оны кітап ретінде пайдалануға болады, тиісті ұсыныстарын зерттеуші белгілі бір мақсаттарда іске асыра алады. Ғалымдардың айтуынша, машина қоры дегеніміз — ғылым адамының өз зерттеу объектісіне жаңаша тұрғыдан көз салудың мүмкіндігі болып табылады. Мұндай қор неғұрлым қомақтырақ болса, солғұрлым тіл құрылысының сырын тереңірек ашуға болады, сөйтіп зерттелетін объект жөніндегі түсініктердің шеңбері кеңиді, адамның білім өрісіндегі «ақтандақтардың» бедер-бейнесі айқындала береді.

Тілдің машиналық қоры адамдардың зерттеушілік қалып-кабілетін әлденеше есе арттырады, творчестволық қуат көздері ашылады, сөйтіп бұл жана мүмкіндіктер тілдің жүйелілік қасиеттерін жетілдіруге, яғни тіл жүйесін мұқият тануға жұмсалатын болады. Машиналық қор жасаудың программасы комплексті сипатта болуы тиіс, демек, ондай программа ең қажетті деген теориялық және экспериментті, яғни практикалық мәліметтерді бірлікте қамтуы қажет.

Машиналық қор жасайтын программа сонымен бірге белгілі бір нысананы көздеуі тиіс, өйткені, программаны жүзеге асырудың нәтижесінде жаңаша бір зерттеу объектісі, немесе материалдық жүйе, бұрын болмаған жаңа аппараттар жасалып, тіл жүйелерінің барлық қасиеттері мен көріністері жан-жақты зерттелетін болады.

В. М. Андрищенко «Орыс тілінің машиналық қорын жасаудың концепсиясы мен архитектурасы» атты баяндамасында³ тілдің машиналық қоры деген ұғымның ең жалпы түсінігі деп төмендегідей ұғындырады. Ол — тілді зерттеп, тексеретін автоматтандырылған комплексті жүйе болмақ. Оған түрлі лингвистикалық мәліметтер, сол мәліметтердің текстермен, картотекалармен, сөздіктермен, грамматикалық т. б. лингвистикалық білім көздерімен деректендірілген білімдер жинақталмақ. Осы білім көздерінің бәрі де программаға сәйкес қолданыс тауып жаңа тілдік объектілерді, яғни сөздіктерді, грамматикаларды, тіл дамуының түрліше процестерін құрастыруға қызмет етеді. Машина қоры жинақталған білім көздерінің бәрі де қорландырудың жаңа единицалары ретінде ғылыми мүмкіндіктерді молайта бермекші.

Тілдің машиналық қорын жасау үшін оны мынадай құрамдармен қорландыру керек: 1) машина қорының негізгі сөзтізбегі; 2) иллюстрациялық текстер қоры; 3) терминология қоры; 4) академиялық сөздік — грамматика қоры; 5) тіл процессорларының қоры; 6) лингво-статистикалық база; 7) лексикографиялық база; 8) лингвистикалық алгоритмдер мен программалар қоры.

Машиналық қордың осы айтылған жобасының ерекшелігі деп мынаны айтуға болады: бұл қор лингвистикалық мәліметтердің сала-салаға жүйеленген банк ретінде қызмет атқарады, яғни оның компоненттері дербес күйінде әр салада сақталады да, оларды тұтыну қажеттігіне сәйкес керек кезінде қайта жишакталады.

Бұл күндері жоғарыдағы ғылыми негіздерге сүйеніп СССР-дағы түрлі ұлттар тілінің машиналық қорын жасау да жоспарланып отыр. Бұл істің ұзын ырғасы он бес жылға созылмақ. Осы мерзім ішінде тіл зерттеуші барлық институттар мен ірі-ірі жоғары оқу орындарында лингвистикалық жұмыстарды автоматтандыру мен тіл мамандарына информациялық-анықтамалық қызмет көрсетудің қалыпты комплекстік жүйелері іске қосылуы тиіс. Тіл жөніндегі деректерді жинаудың, сақтаудың, талдаудың және салыстыра зерттеудің жаңа әдістеріне көшу, сондай-ақ лингвистикалық дерек көздерінің автоматтандырылған сөздіктер мен грамматикалар тәрізді жаңа түрлерін игеру ісі өмірге жанасымды және тиімді болу үшін бұл салада жалпыфилологиялық дәстүрлермен мәдениетке,

³ Бұл да сонда. 26—44-б.

тілді терең танып білуге, практикалық жаңа міндеттер мен тәжірибелерді мұқият ескеруге тиіспіз. Бұл ретте табиғи тілдің деректерін зерттеу жүйесіне негізделетін жаңа информациялық технологияны жетілдірудің бірден-бір шарты — адам мен машинаның қарым-қатысына келіп тірелетін интеллектуальды орта деп білеміз.

А. С. Гердтың «Орыс текстерінің типтері және орыс тілінің машиналық қорын құру» атты баяндамасында⁴ тілдің машиналық қоры дегенді ол өзіне тән кішігірім жүйелер мен микросистемаларды қамтитын автоматтандырылған, аса үлкен макросистема деп түсіндіреді. Мұндай жүйелердің бірі сөйлеу фактілерін тіркеп отыруды нысанаға алатын тіл қорлары, атап айтқанда, текстер, сөздер, сөз формалары, сөз тіркестері мен сөйлемдер. Мұндай қорлар кішігірім жүйелердің тіркемелі типіне жатады.

Бұл жүйелердің максаттары мен міндеттері қандай болмақ? Мұндай жүйелерде тұтас текстер сақталады және тұтынушының сұранысы бойынша сол текстердің үзінділері, нақты адресі және сөз формаларының статистикасы, сөз нұсқағыштардың (словоуказатели) толық тізімдері немесе солардың фрагменттері, сөздерге телулі шаталар, жиілікті және кері алфавитті сөздіктер беріледі. Осы арқылы сөздің әралуан тұлғаларының ең түпкі бірегей нұсқасы айқындалады. Болашақ машиналық қордың лексикасының сипаты түпнұсқа текстердің типіне байланысты болып келеді.

Бұл айтылғандардың кейбір мысалы ретінде яғни тілдің жүйесін танытатын сөйленіс жағдаяты деп мыналарды айтуға болады: көркем әдебиеттің тілі, газеттің тілі, ғылым мен техниканың тілі, ұйымдастыру-нұсқау документтерінің тілі, заң орындары мен сот мекемелерінің тілі, мектеп т. б. түрлі оқу орындарының тілі, жиналыстардың тілі, күнделікті әңгімелесудің тілі, транспорт, магазин, аптека, ателье, поликлиникалардағы сөйлеу тілі, жеке адамдардың күнделіктері мен қойын жазбаларының тілі т. б.

Әрине сөйленіс типтерінің бұл түрлері біркелкі емес, көпшілігінде жазбаша, ал енді біреулері ауызша жүзеге асады. Алайда жазба текстердің қорына даусыз енетін сөз үлгілері деп көркем әдебиеттің, ғылымның, техниканың, заңның, қаулы-қарардың, іс қағаздарының тілін айтамыз.

Қазіргі ғылыми-техникалық және іс қағаздары тексінің негізгі өзегі — терминдер, әйтсе де, қай тілдің болсын машиналық қорына ғылыми-техникалық текстерді түгел енгізу керек пе, жоқ па деген мәселе әлі де ойлануды қажет етеді. Жалпы сөйлеу тілінің аясындағы сөздердей емес, терминдердің тілдік қоршауы, яғни контексті өзгеше бір тілдік жүйе, басқаша айтқанда терминдер өрісі болып табылады. Осы аяда өмір сүретін термин үшін барлық ғылыми

⁴ Бұл да сонда. 67—75-б.

шығармалардың сөз қолданыс машықтары етене, ортақ болып келеді. Сырт қарағанда белгілі бір ғылым саласындағы терминдер өзінің аумағы жағынан бүкіл жасалатын қордан асып кететін тәрізді. Олай болғанда басқа ғылымдар саласындағы терминдерге машинаның жадынан үлес тимей қалуы да мүмкін. Сондықтан бұл мәселені терминологтардың алқасына салып, келелі кеңес жолымен ойланып-толғанып шешу керек. Себебі ғылым мен техника тілінің лексикасы дегеніміз тек термин сөздер ғана емес, бұл ұғымға ең алдымен бүкіл терминдік тіркесімдер, символдар, теңгерме элементтер немесе идентификаторлар да кіреді. Бұлардың саны күн санап өсе бермек, өйткені бұл қызметті атқаратындар кәсіби етістіктер, шартты символика т. б. болады да олардың әрқайсысы белгілі бір ғылыми ұғымдар мен концепцияларды танытады. Машиналық қорға текстер енгізуде қиынырақ тиетін мәселелер диспуттар, сот процестері, жиналыстар, күнделікті әңгімелесулер, реплика тастау, айтыс, телефон арқылы сөйлесу, радиомен хабарласу кезіндегі ауызекі сөйлесу тілінің үлгілерін жазып алу болып табылады.

Текстердің әр дәуірдегі белгілерін ажыратып, тануда және оларды сол дәуірлер бойынша жазып алу аса маңызы жұмыстардың бірі. Текстердің хронологиялық жігін айырып тану мақсатында диалектілерді машиналық қорға қосудың ойланатын жақтары жоқ емес. Диалектілер синхронды түрде өмір сүреді. Бірақ өздерінің мәні жағынан олар әрі диахронды болып келеді. Диалект сөздер кез-келген ұлт тілінен әсіресе әдеби тілден ескі, әріден келе жатқан тіл құбылыстары. Ал, қазіргі тіліміздегі диалектілердің лексикасы біртектес емес, әркелкі. Олар бір жағынан ертедегі сөздер мен қолданыстардың ұлтқа, жазуға дейінгі үлгілері болып келсе, екінші жағынан басқа тілдерден енген сөздерді тіл ішінде кейіннен пайда болған, бірақ окшау, томаға-тұйық қолданыста болып, таралу әрісі шектеліп қалған сөздер болуы да ықтимал. Халық тілі сол халықтың жанды қазынасы болғандықтан соның бүкіл болмысын барлық ерекшеліктерімен қоса қамту машиналық қордың міндеті.

Диалектілер секілді әрі синхрониялық әрі диахрониялық сипатта күн кешетін тіл — фольклордың тілі және жер-судың тілі немесе топонимика. Осы аталған тіл үлгілерін жинастырып, машиналық қор жасағанымызда, оның бұл аталғандардан басқа да салалары мен арналары қамтылары сөзсіз. Қысқасы машиналық қор жасау мақсатымен байланысты бүгінгі таңда күн тәртібіне қойылатын мәселелер мыналар деп білеміз: тіл маманына ең алдымен қажет болатын мәліметтердің сипаты. Керек болған белгілі бір сөзді қандай тіл көздерінен, текстерден, тіл айырымдарынан іздестіру қажет және сол сөз тиісті тексте қамтылған ба, жоқ па, графикалық

және фонетикалық пішіндерін қоса есептегенде сөздің мүмкін болған барлық түрлері мен тұлғалары айқындалып болды ма, белгілі бір сөзге қатысты контекстердің жалпы аумағы, сипаты анықталды ма, белгілі бір сөз осы семантикалық және грамматикалық сипатта өткен дәуірлердегі, көне ғасырдағы текстерде қолданылған ба, егер қолданылса, қай замандардан бергі сөздіктерден көрініс тапқан.

1988 жылдың май айында Москвада Совет тюркологтары комитетінің XIV Пленумы болып өтті. Пленум мәжілісінің күн тәртібінде: «Түркі тілдерінің машиналық қорын жасау» проблемалары талқыланды. Онда негізгі баяндамашылар: Р. Г. Пиотровский, А. М. Щербак, В. Г. Гузев өз ой-пікірлерін ортаға салып, түркі тілдерінің машиналық қорын (*бұдан әрі — ТТМҚ*) жасаудың қажеттілігімен қатар қиыншылықтарын да атап көрсетті⁵.

ТТМҚ-ын жасаудың болашақта екі жақты қажеттілігі (ұтымдылығы) болатындығы ескерілді. Біріншіден, салыстыру-салғастыру (сравнительно-сопоставительный) тұрғысынан қарастырғанда лексикографиялық, грамматикалық, стилистика-статистикалық, тарихи-этимологиялық мәліметтерді жинау және реттеу (тез іздеп табу); екіншіден, тұтынушыларға (тілші-зерттеушіге) бір тілдің тұрақты единицаларының «туу» (порождение) ережелерінің көмегі арқасында қажетті тілдік элементтерді әралуан текстен автоматты түрде бөліп алу мүмкіншілігіне ие болу.

Түркі тілдеріне ортақ машиналық қордың маңызы сөз бола келе, сол қорды жасаушылар, яғни тілшілер мен инженер-программистер теориялық жағынан маңызы зор екі түрлі әдістің біреуін: «жоғарыдан төменге» (нисходящий) немесе «төменнен жоғарыға» (восходящий) қарай зерттеудің жолын күн ілгері тандап алу қажеттігі сөз болды.

Қомақты жүйелерді құрастыру кезінде (ТТМҚ-ын құрастыру да соған жатады) «төменнен жоғарыға» қарай тексеру принципінің елеулі қиыншылықтарға әкеліп соғатыны нақты ескерілді. Себебі бұл принцип бойынша автоматты жүйе құрастырғанда бөлек-бөлек жағдайларға (есептерге) негізделген қарапайым модульдер жасалып, олар бірте-бірте бірігіп, іріленіп, күрделі блоктарға айналып, нәтижесінде кішігірім жүйелердің орнына қомақты біртұтас жүйе құру мақсаты көзделеді. Алайда мұндай бөлекше модульдер өз алдына жеке тұрып жұмыс істегенмен, біртұтас біріккен жүйеге (демек бүкіл түркі тілдерінің машиналық қорын жасау жүйесіне) айналғанда алға қойған мақсатты шешетіні, шешпейтіні бимәлім. Көп жағдайда мақсат орындала да бермей-

⁵ Гузев В. Г., Пиотровский Р. Г., Щербак А. М. О создании машинного фонда тюркских языков // Советская тюркология. 1988. № 2. С. 98—101.

ді. Басқаша айтсақ, жеке тұрып лексика-семантикалық, этимологиялық, морфологиялық, тіпті, синтаксистік есептерді шешетін программалардың жиынтығын және ЭВМ жадына жазылған неше түрлі лексикографиялық материалдардың тұтастығын тілдің машиналық қоры осы деп есептеуге болмайды.

Осы айтылғандарды қорыта келе, авторлар түркі тілдерінің машиналық қорын жасау деген ұғымды екі түрлі концепция (стратегия) тұрғысынан қарау қажеттігін ұсынды. Бірінші концепция түркі тілдерінің ғылыми машиналық қорын жасау жайында болса, екінші концепция бойынша түркі тілдерінің көп аспектілі машиналық қорын жасау міндеті көзделеді.

Мұндай концепцияларға бөлу мына көзқарастан туады. Біздегі ғылыми-зерттеу орталықтарының нақты жағдаяттарын ескеретін болсақ, тіл мамандарының өзіне бекітіліп берілген ЭВМ орталықтарының жоқтығы, қажетті маман инженер-программистер мен тілшілердің жетіспеуі т. б. салдарынан таяудағы 5—7 жыл шамасында түркі тілдерінің көпәспектілі машиналық қорын жасау мүмкіншілігі екіталай. Сондықтан да жұмысты түркі тілдерінің шағын көлемді ғылыми машиналық қорын жасаудан бастау жөн болады.

Осы тұрғыдан қарағанда, ең алғаш түркі тіліндегі сөздердің грамматикалық формаларының жасалу (порождения) процестерін зерттеумен шектелген жөн. Толығырақ айтқанда, ғылыми машиналық қор түркі тілдеріндегі бір буынды сөздердің әртүрлі құрылымдық-фонетикалық үлгілерінің толық жиынтығын қамтуы тиіс және ол дыбыстардың тіркесуі мен қисындасуы жағынан ата түркілік күйден шет шықпағаны жөн. Бұл жиынтық әрі тұрақты, әрі өзгермейтін болуы шарт. Бір буынды сөздердің алғашқы қалпын анықтау фонетиканы салыстырмалы анализ әдісімен зерттеуді қажет ететіні түсінікті. Түркологтердің компаративтік зерттеулері негізінде алған нәтижелерінің өзі ата түркілік күйдегі сөз тұлғасының біршама тиянақты моделін құруға мүмкіндік береді. Ал, модельді жетілдіре түсу жолдарын табу үшін жана зерттеулердің қажет екендігі түсінікті.

Екінші концепция бойынша түркі тілдерінің көпәспектілі машиналық қорын жасау қолданбалы ілімдердің мұқтажына сай жүзеге асырылуы тиіс. Мысалы, оны баспа, радио, телевизия, театр қызметкерлерінің, мұғалімдердің түркі тілдерін пайдалану қажеттігіне орайластыруға әбден болады. Мұндай мұқтаждықты өтеу үшін түркі тілдерінің көпәспектілі машиналық қоры әрі жалпы-түркілік тілдің жиынын, әрі жекелеген нақты бір тілдердің жиынын модельдеу арқылы жасалады.

Төменде осы айтылған машиналық қордың түрліше салалары (блоктары) жалпылама сипатталмақ.

1. Ең алғаш құрастыруды қажет ететін нәрсе — лексикалық блок. Ол орфографиялық, орфоэпиялық, түсіндірме, терминологиялық қолданыс жиілігі және аударма (орысша-түркіше, түркіше-орысша) типіндегі сөздіктерді қамтитын көпшілік қолды лексикографиялық жүйеден тұруы тиіс. Бұл ретте жалпытүркі лексиконын жасау (құрастыру) ісінің өте-мөте қиындыққа түсетінін ескеру қажеттігі сөзсіз.

2. Грамматикалық блок — лексикалық және грамматикалық морфемалардың тізбелерінің жиынтығынан, сондай-ақ синтаксистік схема-модельдерді қамтитын тілдік материалдардан тұрады. Әрбір жеке тіл бойынша морфемалармен синтаксистік схемалардың жиынтығы қосыла келе жалпытүркілік тілдің машиналандырылған грамматикалық тезаурусын жасауға мүмкіндік береді. Оның түркі тілдерінің ғылыми-машиналық қоры ретінде де кәдеге жарайтындығы ескерілмек. Бұған типті тезаурустың түп-тірегі ретінде аффикстер мен грамматикалық көрсеткіштердің жиынтығы алынғаны жөн.

3. Түркі тілдерінің машиналық қорындағы ең бір шағын жүйе — фонологиялық және морфонологиялық блоктар.

Фонологиялық блок әр тілдің фонемалық құрамын (состав), әр фонемаға тән бүкіл дифференциалдық (айырым) белгілердің түсіндірмелерін қорландыратын жиынтықтан тұруы қажет.

Морфонологиялық блок төмендегідей информациялық мәліметтерді қорландырып сақтайды:

1) тілдік единицалардың фонемдік құрылысының типтері: негіздер, форманттар, сөздердің жекелеген кластарына (мысалы, бейнелеуіш сөздерге) тән фонемдік құрылыстың ерекшеліктері жөніндегі деректер;

2) белгілі бір тіл үшін міндетті сингармонизм заңдары, сондай-ақ дауыссыз дыбыстардың ассимиляциясы, катандауы мен ұяндауы жайындағы мәліметтер;

3) түбір-негіздер мен сөз-формаларға түсетін екпін туралы мәліметтер т. б.

Әралуан түркі тілдеріне қатысты морфонологиялық суреттемелерді салыстыру нәтижесінде жалпы түркілік морфонологиялық белгілердің терең сырларын ашуға, сондай-ақ жекелеген тілдердің өзіндік қасиеттерін тануға қажетті мүмкіндіктер туады.

Міне осы аталған мәселелердің бәрін жан-жақты қамти отыра толық та тұжырымды жауап алуымыз үшін электронды есептегіш техниканы пайдалана отырып, тілдің машиналық қорын жасау — лингвистикалық ғылым саласындағы қайта құрудың теориялық әрі практикалық міндеттерді жаңаша ойлап, жаңаша шешудің бүгінгі таңдағы ең тиімді жолы болып табылады.

ФОНЕТИКА БОЙЫНША
ЗЕРТТЕУЛЕР

ИССЛЕДОВАНИЯ
ПО ФОНЕТИКЕ

ОРФОГРАФИЯ МЕН ТРАНСКРИПЦИЯ ЖАЙЫНДА БІРЕР СӨЗ

Тюрколог-мамандар өздерінің ғылыми жұмыстарын жазғанда орфографияның кейбір кемшіліктері мен оңтайлы транскрипцияның (дыбыс жазуының) жоқтығынан қатты қиналады. Көп жағдайда олар өз ана тілінің алфавитін фонетикалық жазудың негізі ретінде пайдаланады. Алайда, мұның өзі ғылыми еңбекті өзге тіл мамандарының еркін пайдалануына қиыншылық туғызады және түркі тілдеріне тән көптеген фонетикалық құбылыстардың ескерусіз қалуына әкеп соғады.

Қазақ жазуы негізінен орыс алфавитіне құрылғаны және қазақ тілі дыбыс жүйесінің ерекшеліктерін ескерген толықтырулар бары белгілі. Қазақ мектептерінде орыс тілі негізгі пәндердің бірі ретінде жүргізілетін болғандықтан орыс графикасын қабылдаудың практикалық мәні бар.

Орыс алфавиті ұлт тілдеріндегі жазуды бір жүйеге түсіруге, екі тілде бірдей сауат ашуға, қоғамдық-саяси және ғылыми-техникалық терминдерді игеруге көп септігін тигізді. Дегенмен алфавит пен орфографияның өзара үйлесімін табу үшін әлі де болса фонологиялық зерттеулердің қажет екендігі сонғы кезде біліне бастады. Қазақтың төл сөздерінің жазылуына И, У, Ю, Я әріптерінің кірістірілуі де үстірт болған. Өйткені мұндай тың «жаңашылдықтың» нәтижесінде, мысалы, бір ғана «И» әрпі 8 түрлі оқылатын болып шыға келді: 1) Ы мен Й (жинақ — жыйнақ), 2) І мен Й (сирек—сійрек), 3) Ұ мен Й (құнжиды—құнжүйдү), 4) Ү мен Й (дүнне—дүнүйе) дыбыстарының тіркесі, сондай-ақ, 5) Ы (сия—сыйа), 6) І (ию—ійүү), 7) Ұ (сұрқия—сұрқұйа), 8) Ү (өңкию—өңкүйүү).

Осындай жайттардың салдарынан буын жігін ашу мен тасымалдау жолдары шым-шытырық күйге түсті. Орфографиялық буын жігі мен орфоэпиялық буын жігі үйлеспейтінді шығарды. Мысалы,

ти-ын	тый-ын	ты-йын
су-у	сұу-ұу	сү-уұу
ти-ю	тій-үү	ті-йүү
у-іл	үү-үл	ү-уұл

Егер И, У, Ю, Я әріптерінің оқылуы әзірге қиындық туғызбай жүрген болса, ол, тек, қазақ сөзінің құрамындағы буындар мен дыбыстардың басын құрастырып тұрған сингармонизм заңының арқасы. Мысалы, Тәңір *рия* (ірійә) ғамалды қабыл көрмесе (ҚӘ, 11.07.88). *Суси-суси* (сүусұй—сұусұй) алыстап, көңіліміз суысқан (ҚӘ, 14.06.85). *Ти* (тій) десем тимейді.., (жұмбақ). *Көк ми* (мый), *су ми* (сұу мый), *ту* (түү) қоян, бас *ию* (ійүү), көңіл *суу* (сұууу), *ии* (іій) бастады (тұрақты тіркестер).

Осы келтірілген мысалдардағы асты сызылған сөздердің оқылуын (айтылуын) дұрыс нұсқайтын орфографиялық ереже (лингвистикалық белгі) жоқ, оны тек жаттап алу керек.

Шынтуайтқа келгенде, қазақ (жалпы түркі) жазуы үшін сингармонизмнің нұсқасын дауыссыз дыбыстың әрпі көрсете ме (көне түркі жазуындағыдай) немесе дауысты дыбыстың әрпі көрсете ме (осы күнгі жазуымыздағыдай), жоқ, әлде, арнайы үстеме (диакритика) белгі көрсете ме (кейбір транскрипциядағыдай) бәрі бір. Егер сөздің сингармонизмдік нұсқасын: дауыссыз әріп көрсетіп тұрса, онда көне түркі жазуындағыдай дауыссыз әріптердің саны көбейеді, сөйтіп, дауысты әріптерде сингармонизм белгісі қалмайды да олардың саны кемиді; дауысты әріп көрсетіп тұрса, онда қазіргі жазуымыздағыдай дауысты әріптердің саны көп болады, сөйтіп дауыссыз әріптерде сингармонизм белгісі қалмайды да, олардың саны кемиді. Көне түркі жазуы мен қазіргі жазу жүйесінің арасындағы негізгі айырмашылық осы. Оның үстіне дауыссыздарды тірек еткен жазу буын жазуға айналады да текст таңбалардың (әріптердің) үнемді тізбегінен құралады, ал дауыстыларды тірек еткен жазу дыбыс жазуына айналады да таңбалардың тізбегі созылып кетеді.

Жалпы сингармонизм белгісінің дауысты әріптерге телінгені «дауыстылардың үндесуі» (гармония гласных) деген жаңсақ пікірге әкеп соқты. Дауысты дыбыстар емес, буындар үндесетіндігін зерттеуші ғалымдар мен практик ұстаздар естен шығармағаны жөн. Түркі тілдерінің, оның ішінде қазақ тілінің «буыншыл» (Қ. Жұбанов) тіл екенін кездейсоқ құбылыс деп қарамай, түркі тілдерінің іргелі белгілерінің бірі деп есептеу қажет. Осы жерде жол-жөнекей болса да, қазіргі жазуымыздағы дауысты әріптің графикалық белгі ретінде үш бірдей қызмет атқаратындығын айта

кетейік: 1) буын құраушы элементті көрсетеді; 2) дауыстының ашық, қысаң не дифтонг екенін көрсетеді; 3) буынның сингармонизм сазын көрсетеді (жуан, жіңішке, еріндік, езулік).

Сонымен үнді-европа (оның ішінде орыс тілінде) тілдерінде сөзді оқудың (айтудың) орфографиялық ережелерін сөздік екпін реттеп отырса, түркі (оның ішінде қазақ тілінде) тілдерінде сөздердің оқылу (айтылу) нұсқаларын сингармонизм реттейді. Сингармонизмді түркі тілдеріне тән іргелі белгілердің бірі деп отырғанымыз осыдан. Үнді-европа тілдерінде дыбыстардың басын біріктіріп, сөз етіп ұйыстыратын сөздік екпін, ал түркі тілдерінде — сингармонизм.

Жазудың сөздің айтылуына ықпалы зор екендігін ескере келсек, жазу жүйесін ұйымдастыруға жауапкершілікпен қарау керек. Епдеше жазу ережелері сол тілдің фонологиялық заңдылықтарына телінгенде ғана жүйелі жазу шығады. Фонологиялық заңдылықтар ескерілмеген жағдайда немесе қазақ жазуын «И», «У» әріптерімен «жетілдіру» сипатты шешімдердің әсерінен тетелес ұрпақтың-ақ сөз сазында күрт өзгерістер болып жатады. Соның нәтижесінде үлкеннің сөзі (дыбысталуы) кішіге, кішінің сөзі үлкенге түрпідей естіліп жатады.

Қазіргі кезде түркі тіл білімінде ортақ оңтайлы транскрипция жоқ, сондықтан да түркі тілдеріндегі әртүрлі дыбыстар бір белгімен беріліп жатады.

Қазақ жазуының айтылумен сәйкес келмейтін сәттерін атап өтейік.

Буын тіркесіне байланысты:

тіркес	жазылуы	айтылуы
1. и+я	ақияды	ақыйады
2. и+ю	ақию	ақыйуу
3. и+ю	тию	тійуу
4. и+ю	мөлию	мөлүйуу
5. и+я	қонжияды	қонжүйады
6. и+ю	көнтию	көнтүйуу
7. и+дауысты дыбыс	нық	ыйық
8. дауыссыз+и	ақи	ақый
9. дауыссыз+и+дауыссыз	айниды	айныйды
10. дауыссыз+и+дауысты	абиыр	абыйыр
11. и+дауыссыз	изең	ійзең
12. дауыссыз+и+дауысты	шиет	нийет
13. и	и	ій
14. дауыссыз+и	әлди	әлдій

15. дауысты+и	ау	ау
16. еріндік+и	монти	монтұй
17. еріндік+и	дүнки	дүнкүй
18. еріндік+у	күлу	күлүу
19. еріндік+у	оку	оқуу
20. у+дауысты	уыс	ууұс
21. у+дауыссыз	улы	ұулу
22. у	у	ұу
23. у+дауыссыз	уледі	ууледу
24. у+дауысты	уәде	ууәде
25. еріндік+е	бөле	бөле
26. еріндік+ы	бұдыр	бұдұр
27. еріндік+і	көңіл	көңүл
28. ю+ы	оюы	ойуу
29. жуан	жария	жарийа
30. жінішке	дүрия	дүрүйә
31. жуан+ю	аю	айуу
32. жінішке+ю	ую	үйуу
33. и+ю	акию	акыйуу
34. и+ю	тик	тійуу
35. р+жуан	рас	ырас
36. р+жінішке	рет	ірет
37. р+еріндік	ру	ұруу
38. л+жуан	лақ	ылақ

Екі морфеманың жігінде

39. н+м	ағайынмен	ағайыммен
40. дауысты+к+дауысты	аки	акый
41. н+г	сенген	сеңген
42. н+к	айлаңкес	айлаңкес
43. н+ғ	ағайынға	ағайынға
44. з+ш	ағызшы	ағышшы
45—46 з+с	ағызса	ағысса
47. с+ш	оқысшы	оқұшшұ
48. н+б	сынбау	сымбау
49. фразаның соңында	барады	барат
50. фразаның соңында	келеді	келег
51. фразаның соңында	оқиды	оқұйт
52. фразаның соңында	үреді	үрет
53. ш+с	пішсе	пішше
54. ж+ш	монтажшы	монташшы
55. д+т	наградтау	награттау

56. дауысты + п + дауысты	алып ұшу	алыбұшуу
57. з + ж	аз жыл	ажжыл
58. п + б	айып беру	айыпперуу
59. сонант + к + дауысты	ай қабақ	айғабақ
60. дауысты + к + дауысты	азғана күн	азғана гүн
61. сонант + к + дауысты	аз кем	аз гем
62. н + б	алтын балық	алтымбалық
63. дауысты + дауысты	ағалы-інілі	ағалінілі
64. н + п	тон пішер	томпішер
65. р + і	бөрі басар	бөрбасар
66. ш + ж	көш жер	көшшер
67. к + сонант	көк ми	көгмый
68. к + ұян	өкпек жел	өкпөг жел
69. ұян + к + дауысты	аз қалды	азғалды
70. к + ұян	сәбилік дәуір	сәбйліг дәуүр
71. дауысты + к + ұян	ақ жауын	ағ жауұн
72. дауысты + к + дауысты	ақ айыл	ағайыл
73. дауысты + к + дауысты	бара қой	бара ғой
74. дауысты + к + дауысты	бөбек үні	бөбег үнү
75. с + ш	қыс шағы	қышшағы
76. з + ш	сызшы	сышшы
77. сонант + к + дауысты	бақай құрт	бақай ғұрт
78. сонант + к + дауысты	адал көңіл	адал гөңүл
79. с + ж	алыс-жұлыс	алышшұлұс
80. н + к	аман келу	амаң гелуу
81. н + к	азан-қазан	азаң-ғазан
82. дауысты + к + ұян	ақ жар	ағжар
83. дауысты + к + ұян	көкжал	көгжал
84. сонант + к + дауысты	әрқалай	әрғалай
85. сонант + к + дауысты	әркім	әргім

Бұл тізімнен қазақ тілінің жазу ережелері мен оқылу ережелерінің арасындағы алшақ жайлардың шамадан тыс мол екенін көреміз. Әрине, тіл дыбыстарын жазуда түп-түгел дәл белгілей беру мүмкін емес. Кейбір әріптердің орфографиялық ережелерге сәйкес әртүрлі оқыла беретіні қай тілдің жазуын алсақ та кездесе береді. Алайда әртүрлі оқылу неғұрлым кем болса, орфография соғұрлым дұрыс, құнды болып есептеледі. Ендеше бір әріп бір-ақ түрлі оқылуға тиіс. Олай болмаған жағдайда орфографиялық ескертпелердің (исключения) мүмкіндігінше кем болғаны жөн.

Қазіргі қазақ орфографиясында орфографиялық шалағайлықтардың басы баршылық. Мұның өзі мектепте еміле ережелерін үй-

ретуге, ғылымда тілдің дыбыс жүйесін зерттеуге көп нұқсан келтіреді. Сондықтан да бір дыбысқа бір әріп сай келетін фонетикалық жазуды (транскрипцияны) кеңінен пайдаланып отыру керек. Сонда ғана оқушы қазақ сөзінің айтылу (оқылу) үлгісін дұрыс игереді, ал акустикалық, статистикалық, этимологиялық т. б. зерттеулер жаңсақ текстен аулақ болады. Сөйтіп қазақ тілі дыбыс жүйесінің төл заңдылықтарының басын ашуға болады. Ең бастысы қазақ сөзінің дыбысталу үлгісіне нұқсан келмейді, ұрпақтан ұрпаққа қыры сынбай ауысып отырады. Өйткені шалағай орфография бірте-бірте шалағай сөйлеуге әкеп соғады. Ендеше қазақ жазуын ғылыми жолмен ретке келтіру — бүгінгі күннің көкейкесті мәселесі.

Л. ГРЖЕБИЧЕК

ЭСТЕТИЧЕСКИЕ ФУНКЦИИ ГАРМОНИИ ГЛАСНЫХ В ПОЭЗИИ АБАЯ КУНАНБАЕВА

Приступая к исследованию в фонетическом аспекте поэзии классика казахской литературы Абая Кунанбаева, мы рассматриваем проблему эстетической функции гармонии гласных в стихотворном тексте. Полагаем, что данный вопрос является вполне актуальным, так как он касается одной из наиболее значительных фонетических закономерностей казахского языка. В частности, имеется в виду аспект гармонии гласных, основанный на том, что все слова тюркского текста делятся на палатальные — П и веллярные — В, которые в потоке речи выступают в последовательности. Бесспорно, что здесь наличествуют заимствованные и сложные слова, в которых гласные обеих групп встречаются в одной и той же словесной единице. В таких случаях составные слова мы относим к П или В-группе, исходя из законов казахской грамматики**.

* Печатается по тексту «Archiv Orientalni». № 32, АН ЧССР. Прага, 1964. С. 100—103 (пер. с англ. А. Е. Жубановой).

** Это упрощение не может умалить наши результаты. Когда к такому слову присоединяется аффикс, происходит гармония гласных; в случае отсутствия аффикса можно говорить о потенциальной гармонии гласных. Однако количество таких образований относительно мало.

В случае появления инвариантных аффиксов, гласная в основе является решающей. Исследуя сочетающиеся звенья П и В слов, мы не берем в расчет последовательность финального слова в предыдущей строке и начального слова в последующей строке; иначе, говоря по-другому, наши исследования включают в себя только последовательность внутри одной строки, т. е. стиха.

Мы находим, что количественный метод — это единственно возможный подход при решении этой проблемы, хотя наши заключения не являются результатом статистических подсчетов, так как мы не задаемся такой целью, а ограничиваемся только установлением главной тенденции.

Любой казахский поэтический текст может быть переписан в символах П и В. Мы вводим новый символ нуль — О, который может быть определен как граница между двумя соседними строками. Внутри одной строки могут быть реализованы следующие комбинации символов: ПП, ПВ, ВВ, ВП. Символ О появляется только в комбинации с первым и последним словом строки, в результате чего мы получаем дополнительные четыре комбинации символов: ОП, ОВ, ПО, ВО. Очевидно, комбинация ОО не может быть реализована. Частота этих восьми последовательностей была подсчитана по 30-ти сочинениям, выбранным наугад из сборника стихотворений Абая, а полученные данные были зафиксированы в следующей форме:

↓	П	В	О
П	546	793	457
В	806	1111	699
О	444	712	0

Вероятности, соответствующие этим данным, были подсчитаны со следующими итогами:

↓	П	В	О
П	0,30	0,30	0,40
В	0,45	0,43	0,60
О	0,25	0,27	0,00

Средняя энтропия этой системы — 1,42 бит [3]***.

Данный результат мы сравнили с поэтическими фрагментами словаря Махмуда Кашгари (в 400 строк) и с турецким мани [3]

*** Из «Дивану-лугат-ит тюрк тержюмеси» взяты первые 100 четверостиший в том порядке, в каком они даны в тексте, из «Осман-торок» взят каждый четвертый мани.

также в 400 строк. Энтропия их была подсчитана тем же способом. Из турецких мани нами получены следующие вероятности:

↓	П	В	О
П	0,33	0,32	0,54
В	0,35	0,34	0,46
О	0,32	0,34	0,00

Здесь средняя энтропия — 1,44 бит. По стихотворным текстам словаря Махмуда Кашгари мы получили вероятные данные:

↓	П	В	О
П	0,26	0,28	0,39
В	0,41	0,41	0,61
О	0,33	0,31	0,00

Средняя энтропия здесь равна 1,42 бит.

Из этого небольшого сопоставления видно, что во всех трех случаях результаты оценки средней энтропии приблизительно равнозначны. В турецких мани мы наблюдали постоянную тенденцию к повышенной вариативности в цепочке П и В. Исследуя более тщательно форму матрицы, мы заметили, что благодаря введению символа О, первые и последние слова строк приобрели отличительную позицию. Мы могли предположить, что это именно символ О сильно влияет на результаты энтропий.

Наше предположение подтвердили расчеты средней энтропии в каждом стихотворении выборки (см.: табл. 1). Самая низкая оценка — 1,07, высшая — 1,44. Вместе с тем не удалось найти взаимосвязь между этими энтропиями и содержанием соответствующих стихов или их зависимость от таких количественных оценок, как длина стихотворения, выражающаяся в числе слов, или число слогов внутри одной строки данного стихотворения и т. д. После упорядочения этих поэтических текстов по высоте средней энтропии стало очевидным то, что стихотворения с наиболее яркой и экспрессивной поэтической формой имеют самую низкую энтропию. Например, стихотворение «Моз болады болысын» [5; 76], имея самую низкую энтропию в нашей выборке — 1,07, предстает в виде схемы рифмы: абаб/сбсб, дбдб и т. д., и насыщено аллитерацией. Стихотворение «Тугызган ата-ана жок» (П, 8) с энтропией — 1,08 имеет схему рифмовки: ааба/ааса/аада; 7 из 12 строк этого стихотворного текста аллитерируются в начале стиха. Экспрессивную форму мы нашли также в стихотворении «Жайнаган туын жыгылмай» (1, 160) с энтропией 1,12, где 14 и 16 стихов

целостного произведения рифмуются одинаково, а все строки имеют однообразную пинципальную аллитерацию.

Мы не можем по любому случаю брать среднюю энтропию для измерения «формального совершенства» стихотворения или поэзии. Возможно, что звуковое качество стихотворения основывается не столько на последовательности элементов П и В, как мы это виде-

Таблица 1. Величины H_1 [энтропия, составленная на основе системы, включающей символ О] и H_2 [энтропия, подсчитанная без учета этого символа]

Стихотворение	H_1	H_2	Стихотворение	H_1	H_2	Стихотворение	H_1	H_2
I 76	1,07	0,77	II 33	1,34	0,99	I 101	1,40	0,96
II 8	1,08	0,74	I 62	1,35	0,91	I 223	1,41	0,98
I 160	1,12	0,87	I 161	1,36	0,94	II 199	1,40	0,98
II 47	1,20	0,90	I 216	1,36	0,87	I 38	1,41	0,99
I 181	1,27	0,85	II 45	1,37	0,83	I 193	1,41	0,95
I 123	1,28	0,93	I 143	1,38	0,87	II 94	1,41	0,98
I 163	1,32	0,85	II 143	1,39	0,89	II 114	1,41	0,97
I 15	1,33	1,00	I 137	1,39	0,92	I 47	1,43	0,93
I 33	1,34	0,99	I 189	1,39	1,00	I 119	1,43	0,99
I 64	1,34	0,91	II 90	1,39	0,99	II 145	1,14	1,00

ли на примерах, сколько на разнице между ними, то есть на нарушении этой последовательности. Наиболее высокая средняя энтропия в нашей выборке стихотворений принадлежит известному стихотворению Абая «Домбырага кол сокпа» (II, 145), хотя мы не можем сказать, что его форма менее экспрессивна, чем другие стихи поэта. Однако различие должно состоять как в технике рифмовки, так и в количестве аллитераций, особенно в тех случаях, когда аллитерацию оформляют не только начальные согласные, но и последующие гласные.

Отсюда следует, что гармония гласных как эстетический фактор действует только в сочетании с рифмой и аллитерацией. Мы доказали данный вывод подсчетом средних энтропий стихотворений той самой выборки из произведений Абая, но на этот раз без введения символа О. Последовательность финального слова в предыдущей, и начального в следующей строке нами не рассматривались вообще. Оценочные данные средних энтропий стоят в этом случае в ряду от 0,74 до 1,00 битов (см. вторую строку в выводах).

Таким образом, нами получены две различные оценки для каждого стихотворения: H_1 — представляет оценочные данные, подсчитанные на основе системы, включающей символ O ; и H_2 — представляющий оценочные данные без этого символа, коэффициент корреляции этих двух оценок очень низкий, почти равный нулю. Тем самым правильность наших выводов доказана. Следует подчеркнуть, что дальнейшее исследование фонетических свойств абаевских произведений требует общего, суммарного анализа всех его черт.

ЛИТЕРАТУРА

1. См.: *Рос Ашби В.* Кибернетика. Прага, 1961. С. 217—218.
2. *Бесия Аталая Б.* Дивану-лугат-ит тюрк тержюмеси. Анкара, 1940. Т. 1. 2; *Кунша И.* Осман-торок неколкошени гуйтемень. Будапешт, 1889. С. 181—252.
3. Абай Кунанбаев шығармаларының толық жинағы. Алматы, 1954. Т. 1. С. 76.
4. Там же. Т. 1. С. 8.
5. Там же. Т. 1. С. 160.
6. Там же. С. 145.

А. Е. ЖУБАНОВА

СТАТОЛИНГВИСТИЧЕСКОЕ ИЗУЧЕНИЕ КАЗАХСКОГО ПОЭТИЧЕСКОГО ЯЗЫКА

Изучение поэтического текста как эстетического целого, можно сказать, не было предпринято в полном смысле слова в традиционной казахской лингвистике. Если и были отдельные работы о языке того или иного писателя, или же фольклорного образца, то они, главным образом, носили фрагментарный характер, то есть из общего текстового массива брались те или иные моменты, удобные для изолированного показа, с тем чтобы выявить языковую особенность, присущую определенному автору или жанру художественного произведения. Таким образом описывались метафорические словоупотребления И. Джансугурова, фразеологические обороты М. Ауэзова, лексико-стилистические особенности И. Алтынсарина, Б. Майлина, Г. Мусрепова, Г. Мустафина и т. д. Однако авторы подобных лингво-стилистических работ нарушали концептуальность замысла художника, для которого все элементы речи

важны и стилистические оправданы только в их тесной взаимосвязи и органическом переплетении.

Исходя из этого нам представляется теоретически наиболее полновесным среди других лингво-стилистических разработок исследование Р. Г. Сыздыковой, посвященное языку Абая Кунанбаева [1]. В этой фундаментальной работе рассматривается творчество великого казахского поэта, с разных сторон бытования образного поэтического речетворчества. Можно сказать, что исследование Р. Сыздыковой представляет явление в процессе становления казахской лингвистической поэтики.

Свою лепту внесли и другие казахские ученые. Так, в первом сборнике статистики казахского текста [2] были опубликованы статьи А. Х. Джубанова, Е. Х. Жубанова («Статистика лексико-морфологических форм прилагательных в романе М. Ауэзова «Абай жолы» («Путь Абая»)), К. Б. Бектаева («Алфавитно-частотный словарь языка казахских сказок»)), К. Б. Бектаева, А. Б. Белботаева, К. Молдабекова («Алфавитно-частотный словарь языка повести Г. Мусрепова «Кездеспей кеткен бир бейпе»)).

К числу исследований поэтического текста можно отнести и статью Е. Агманова, А. Х. Джубанова «Распределения частот появления сочетаний знаков в Орхоно-Енисейской письменности», хотя язык древнетюркских памятников письменности не является собственно казахским, а представляет собой искусственный тип общетюркского литературного языка раннего средневековья.

Все исследования в основном посвящены статистической характеристике текста типа алфавитно-частотного списка, частотного распределения различных параметров и т. д. Исключение составляет первая статья, где подвергаются анализу формы прилагательных на -кай в поэтике романа «Абай» М. О. Ауэзова, и оценивается лингвистическая оправданность данного образования в массиве словоформ текста.

Однако следует отметить, что казахский поэтический текст как таковой в стато-лингвистическом плане был объектом изучения чешского ученого Л. Гржебичека. Свою научную деятельность он начал с изучения эстетической структуры творчества Абая [3]. Так, в 1963 г. в Институте востоковедения в Праге Л. Гржебичек защитил диссертацию на тему: «О некоторых количественных свойствах стиля Абая». В ней были рассмотрены некоторые фонетические, морфологические и лексические особенности языка поэта. В статье «Эстетические функции гармонии гласных в поэзии Абая Кунанбаева» [4], являющейся фрагментом диссертации, рассматривается одна из наиболее значительных фонетических закономерностей казахского языка, в частности гармония гласных.

Отметим, что о роли гармонии гласных в природе тюркских

языков вообще существуют и другие мнения. Согласно последним, предпочтение отдается не только созвучиям вокализмов, но и сингармонизации звуков в целом. Видимо, поэтому Л. Гржебичек допускает, что звуковое качество стихотворения основывается не столько на последовательности элементов П и В, как видно из его статьи, сколько на разнице между ними, т. е. на нарушении этой последовательности.

Л. Гржебичек считает, что выразительность поэтического текста не всегда зависит от величины средней энтропии, так как стихотворения с высокой средней энтропией также обладают большой долей экспрессивности, зависящей еще и от техники рифмовки, и количества аллитераций. Отсюда следует, что гармония гласных как естественный фактор действует только в ассоциации с рифмой и аллитерацией.

Другая статья автора посвящена опыту применения количественного метода при изучении движения в языке на материале русских заимствований в казахском языке [5]. В числе использованных им текстов семь статей из газеты «Казак эдебиети», исследования М. Ауэзова об Абае, перевод статьи Н. К. Гудзия о Л. Н. Толстом, а также выдержки из художественных произведений Г. Мустафина, С. Муканова, М. Ауэзова, Т. Алимкулова и других. Поскольку статья была опубликована, отметим, что в ней автор продемонстрировал количественный метод на материале русских заимствований в казахском языке и показал зависимость частоты заимствований на функциональном стиле текста. Результаты анализа посредством статистического подхода можно считать проверкой интуитивных результатов (по Л. Гржебичеку).

В статье «О некоторых количественных свойствах лексики Абая» [6] Л. Гржебичек дает некоторые характеристики количественного распределения слов в поэзии Абая и сравнивает их с характеристиками его переводов стихов русских классиков. В статье также приводятся таблицы величины избранных стихотворений Абая и его переводов с указанием автора оригинала.

Таким образом, чешский ученый Л. Гржебичек одним из первых обратился к казахским текстам и сделал статистический анализ главным образом данных художественной речи. Его работа как научный опыт заслуживает внимания.

В советской статоллингвистике перспективной является новая программа «Скальд» по моделированию поэтического творчества для ЭВМ А. М. Кондратова и А. В. Зубова [7]. На наш взгляд, она представляет интерес и для аппаратной обработки текстов тюркоязычного, а именно казахского эпического сказания. Программа предполагает отбор слов и порождение текстов по ритмическому, эвфоническому и семантическому параметрам. В основу модели

исследователи положили закономерности поэтического творчества скандинавских скальдов, поэзия которых признается «самой трудной в мире» как по образной схеме, так и по формальным требованиям, предъявляемым к звуковой организации стиха. Автоматическая модель исследований предполагает компьютерное рождение самой основной и характерной особенности поэзии скальдов — так называемых «кеннингов», в которых метафора уступает место условной схеме, т. е. названию любого объекта того же класса, что и описываемые целые, а определения — названию любого конкретного предмета из сферы целого.

Программа «Скальд» распадается на несколько этапов. Им предшествует разметка словаря или создание банка данных, на основе которого генерируются стихотворения. Авторами создан и введен в ЭВМ в вычислительном центре Минского пединститута иностранных языков словарь всех односложных слов русского языка — существительных, прилагательных, числительных и т. д. — с индексацией их грамматической принадлежности. С помощью этого словаря, состоящего примерно из 6 000 слов, могут быть написаны все возможные стихотворения — брахиколоны, каждая строка которых образована одним односложным словом.

При удачной переориентации программы на закономерности тюркского стихосложения и внедрения в машинную память формульных характеристик казахской эпической поэзии, возможно, откроются широкие возможности и перед казахстанской группой «Статистика речи».

ЛИТЕРАТУРА

1. Сыздықова Р. «Абай шығармаларының тілі» (лексика мен грамматикасы). «Гылым». Алма-Ата, 1968. С. 335; *Ее же*. «Абай өлеңдерінің синтаксистік куралысы». «Гылым». Алма-Ата, 1970. С. 176.
2. Статистика казахского текста. Алма-Ата, 1973. С. 731.
3. Хусаинов К. Ш., Таужанова К. Из истории изучения казахского языка за рубежом // Развитие казахского советского языкознания. Алма-Ата, 1980. С. 160.
4. *Там же*.
5. Гржебичек Л. Опыт применения количественного метода при изучении движения в языке на материале русских заимствований в казахском языке // Исследования по тюркологии. Алма-Ата, 1969. С. 113—118.
6. *Он же*. О некоторых количественных свойствах лексики Абая // Казахская литература в оценке зарубежной критики. Алма-Ата, 1971. С. 40.
7. Кондратов А. М., Зубов А. В. Программа «Скальд» — опыт моделирования поэтического творчества для ЭВМ // Кибернетика. 1984. № 5. С. 86—88.

ЭНТРОПИЯ ГЛАСНЫХ И СОГЛАСНЫХ БУКВ КАЗАХСКОГО ТЕКСТА

Изучение информационно-статистического строения текста является одной из актуальных проблем современной лингвистики. Она продиктована необходимостью применения объективной, количественной меры при анализе грамматического строя и семантико-синтаксической организации текста, а также при сравнительно-сопоставительном изучении языков.

Изучение энтропии гласных и согласных казахского текста нами проводилось на основе статистической обработки лингво-психологического эксперимента, организованного по методу Шеннона-Пиотровского, сущность которого заключалась в последовательном угадывании букв связного текста [1—3].

Исследования проводились на основе текстов произведений А. Куанбаева, М. Ауэзова, С. Муканова, И. Джансугурова и др., а также текстов из газет, журналов, брошюр, магнитофонных записей спонтанной устной речи [4, 5]. Результаты обработки по формулам:

$$\bar{H}_n = \bar{I}_n = (1 - q_0) \log_2 (1 - q_0) - \sum_{k=1}^N q_k \log_2 k; \quad (1)$$

$$\underline{H}_n = \underline{I}_n = \sum_{k=2}^N k(q_k - q_{k+1}) \log_2 k, \quad (2)$$

где \bar{H}_n (\bar{I}_n) — верхняя, а H_n (I_n) — нижняя оценки интервала, в котором находится истинная информация (синтаксическая информация) парадигмы, стоящей в n -й позиции текста, q_k — вероятность угадать букву, стоящую в n -й позиции текста с k -й попытки ($k = 0, 1, \dots, n-1$), q_0 — вероятность достоверных продолжений текста в n -й позиции [6].

В работе осуществлен опыт по угадыванию 8 717 букв (3 775 гласных и 4 942 согласных), образующих 1 454 текстовые словоформы. Мы планировали выяснить, отражается ли качественно-количественная специфика позиционного употребления казахских звуков в энтропийных характеристиках, уплотщающих эти звуки, буквах; оценить величины энтропий, которые дают начальные, се-

редные и конечные буквы словоформы; осуществить классификационную работу гласных и согласных букв, исходя из тех лингвистических принципов, которые обычно применяются в фонетико-фонологических исследованиях по казахскому и другим тюркским языкам.

Из 15 букв казахского алфавита, которые обозначают гласность (А, Э, Е, Ё, И, О, Ө, У, Ү, Ү, Ы, І, Э, Ю, Я), угадыванию подверглись все, кроме «Ё», которая в обследованных текстах не встретилась.

Как известно, для казахского и вообще тюркского вокализма характерно использование по крайней мере трех пар различительных признаков [7—9]. Чтобы выяснить, как отражается эта фонетико-фонологическая парадигматика в энтропии графемного материала, буквы были сгруппированы тремя способами (по степени раствора полости рта; по месту образования обозначаемых ими звуков; по участию губ).

Из 25 согласных букв (Б, В, Г, Ғ, Д, Ж, З, Й, К, Қ, Л, М, Н, Ң, П, Р, С, Т, Ф, Х, Һ, Ц, Ч, Ш, Щ) угадыванию подверглись все, кроме Ч, Һ, которые в текстах не встретились.

Парадигматическая разбивка согласных осуществлена по традиционной схеме, построенной на трех фонетических параметрах [10—12]: по участию голоса; по способу образования, по месту образования.

Анализ полученных данных показывает, что малые величины энтропии дают те лингвистические парадигмы, которые характеризуются небольшим разнообразием элементов с значительной их статистической предсказуемостью, т. е. те парадигмы, которые содержат наименьшее число вариантов угадывания, показывают сравнительно малую величину энтропии. Те парадигмы, которые обладают большим разнообразием вариантов угадывания, дают высокий уровень энтропии.

Как показывают обобщенные данные, для всех разновидностей казахской речи характерно резкое убывание величины энтропии по мере движения от начальной позиции, через срединные к конечной позиции слова (табл. 1, 2).

Эта тенденция характерна и для отдельных букв, особенно для А, Е, которые регулярно встречаются с достаточно высокой частотой во всех трех позициях слова, а также и у тех букв, которые встречаются только в начальных и срединных положениях (Э, О, Ө, Ү, Ү, И) или только в срединных и конечных позициях (Ы, І). Соотношение энтропийных характеристик начальной, срединной и конечной буквенных позиций в казахском слове характеризуется следующими пропорциями:

Для языка в целом:	3,03	:	0,62	:	0,29
публицистики:	3,08	:	0,72	:	0,31
беллетристики:	2,80	:	1,04	:	0,34
устно-разг. речи	3,03	:	1,02	:	0,46
поэзии	2,79	:	1,09	:	0,23

Темп уменьшения энтропии по мере движения к концу слова у гласных значительно выше, чем у согласных, что свидетельствует о том, что на выбор гласных наряду с обычными контекстными

Таблица 1.

Сводные данные о распределении энтропии гласных по позициям

Позиция в слове	Абс. част. F	Отн. част. f	Язык в целом*	
			\bar{H}	\bar{H}
Нач.	450	0,109	2,4632	3,5890
Сер.	2816	0,784	0,6025	1,6379
Кон.	519	0,137	0,1235	0,5637
Всего:	3775	1,000	0,7949	1,5425

* $H_{cp} 1,1687$

Таблица 2.

Сводные данные о распределении энтропии согласных по позициям

Позиция в слове	Абс. част. F	Отн. част. f	Язык в целом*	
			\bar{H}	\bar{H}
Нач.	1078	0,218	2,1287	3,4539
Сер.	2998	0,697	0,9315	1,7312
Кон.	866	0,175	0,3854	0,8486
Всего:	4942	1,000	1,4795	2,1088

* $H_{cp} 1,8042$

ограничениями накладываются дополнительные по сравнению с согласными воздействия. Очевидно, эти ограничения связаны с законом небной гармонии гласных [13].

Согласные также повсеместно дают падение величины энтропии по мере движения к концу слова. Об этом можно судить по следующим соотношениям усредненных энтропийных характеристик для начальной, средней и конечной позиций:

Для языка в целом:	2,99	:	1,33	:	0,62
публицистики:	2,99	:	1,14	:	0,46
беллетристики:	3,16	:	1,49	:	0,65
устно-разг. речи	3,10	:	1,48	:	0,64
поэзии	2,62	:	1,52	:	0,75

Эти соотношения прослеживаются особенно отчетливо на энтропийных характеристиках букв К, К, М, Т, передающих исконные тюркские фонемы *iki*, *imi*, *iti*. Энтропийные характеристики остальных букв выражены менее отчетливо ввиду того, что одни из них, например, Б, Д, Ж, Ш, употребляются только в начальных и средних позициях, а другие (З, И, Л, Н) встречаются почти исключительно в середине и конце слова.

Уменьшение энтропии согласных по мере движения к концу слова объясняется нарастанием ограничений на выбор буквы не только со стороны окружающих букву графем, но и со стороны значимых элементов словоформы — лексических и грамматических морфем, а отчасти и предшествующих данной словоформе лексических единиц. Однако, как уж говорилось, темп убывания энтропии у согласных букв значительно отстает от сокращения энтропии гласных под влиянием сингармонизма.

Таблица 3.
Сводные данные о распределении энтропии гласных, взятых по группам

Энтропия	Группы букв	Абс. част. F	Отн. част. f	Язык в целом		
				\bar{H}	\bar{H}	$H_{ср}$
По способу образования	Широкие (А, Э, Е, О, Ө)	2204	0,584	0,8632	1,6562	1,2587
	Узкие (Ы, І, И, Ң, Ү, У)	1571	0,416	0,6633	1,3912	1,0275
По месту образования	Твердые (А, О, Ы, Ү, У, И)	2174	0,576	0,6479	1,4382	1,0430
	Мягкие (Э, Ө, І, Ү, Е, И, У)	1631	0,424	0,8824	1,6536	1,2680
По участию губ	Негубные (А, Э, Е, Ы, І, И)	3232	0,856	0,5993	1,2372	0,9182
	Губные (О, Ө, Ү, У, У)	543	0,155	0,9373	2,8961	2,4167
	Гласные буквы	3775		0,7949	1,5425	1,1687

Проведенное количественное обследование употребительности графем обнаружило в современных казахских текстах ряд обще- и регионально-тюркских фонетико-фонологических закономерностей. Это неупотребительность губных гласных в абсолютном исходе слова, неупотребительность сонатов И, Л, Н, Р в начале слова, отражающая древнетюркскую тенденцию не использовать сонанты в абсолютном начале слова; общетюркская слабая употребляемость губно-губных Ф, В и их практическое отсутствие в абсолютном начале и конце слова; общетюркская неупотребительность заднеязычного (фарингального) [14—16]. Обращает на себя внимание

малая употребительность узких гласных, особенно И, Ы в абсолютном начале слова.

Ввиду малой употребительности названных букв в указанных позициях слова или вообще в казахских текстах их энтропийные характеристики не рассматривались.

Таблица 4.
Сводные данные о распределении энтропии согласных, взятых по группам

Энтропия	Группа букв	Абс. част.	Отн. част.	Язык в целом		
				\bar{H}	\bar{H}	$H_{ср}$
По участию голоса	Сонанты	2054	0,416	1,6953	1,6241	1,6097
	Звонкие	129)	0,261	1,4518	2,2251	1,8884
	Глухие	1598	0,323	2,3163	2,4958	2,4061
По способу образования	Смычные	1910	0,383	1,3540	2,2296	1,7918
	Щелевые	3032	0,614	1,1357	2,0016	1,5685
По месту образования	Губно-губные	708	0,153	1,5152	2,3034	1,9108
	Губно-зубные	16	0,003	1,0933	1,7983	1,4452
	Зубные	1376	0,278	0,3142	2,2014	1,7578
	Согласные, произносящ. конч. языка	991	0,201	0,6989	1,3793	1,0392
	Переднеязычные	918	0,186	1,3255	2,2418	1,7833
	Среднеязычные	346	0,070	1,2107	2,3500	1,7964
	Заднеязычные	587	0,119	1,3278	2,2287	1,7782
Согласные, взятые в целом	4942	1,4995	1,4995	2,1085	1,8342	

Энтропийные характеристики букв и буквенных классов, имеющих достаточно высокие частоты, показывают такие явления, как крайне низкая энтропия негубных гласных ($H_{ср} = 0,92$ дв. ед.), противопоставленная высокой энтропии губных гласных букв ($H_{ср} = 2,42$ дв. ед.) (табл. 3); это противопоставление прослеживается по всем позициям (за исключением конечной, в которой губные практически не употребляются) и стилям; более высокая по сравнению с энтропией сонантов ($H_{ср} = 1,61$ дв. ед.) и звонких ($H_{ср} = 1,89$ дв. ед.) энтропия глухих согласных ($H_{ср} = 2,41$ дв. ед.) (см: табл. 4); это явление отмечается относительно большинства буквенных позиций и стилей.

Если учесть, что и в том и другом случае речь идет об отражении противопоставления маркированных фонем и немаркированных (*губные гласные*: негубные гласные; *глухие согласные*: неглухие согласные), то оба эти явления могут быть интерпретированы в

общефонологическом плане в том смысле, что маркированность фонемы или группы фонем не только связана с их более низкой частотностью, как это предполагал Н. С. Трубецкой [17], но и обнаруживается в более высокой неопределенности выбора воплощающих их букв в тексте (энтропии), а отсюда в более высокой по сравнению с немаркированной фонемой или фонемным классом информативностью.

Вместе с тем обращают на себя внимание и чисто тюркологические аспекты этих явлений. Дело в том, что противопоставление неогубленных гласных огубленным является, очевидно, одной из древнейших корреляций в фонологической системе тюркских языков [18], причем лабиализация неогубленных обычно оказывалась активным флективно-фонологическим процессом, выполнявшим какую-то важную структурную функцию. Вполне возможно, что высокий энтропийно-информационный вес казахских губных гласных является отражением этих внутривидовых особенностей тюркского вокализма.

ЛИТЕРАТУРА

1. *Пиотровский Р. Г.* Информационные изменения языка. Л., 1968. С. 15—56.
2. *Он же.* Текст, машина, человек. Л., 1975. С. 367.
3. *Шеннон К. Е.* Предсказание и энтропия английского печатного текста // Работы по теории информации и кибернетики. М., 1963. С. 669—686.
4. *Багирова Х. З., Байтанаева Д. А., Бектаев К. Б.* Энтропия и избыточность двух тюркских языков (предварительные данные) // Энтропия языка и статистика речи. Минск, 1966. С. 87—89.
5. *Байтанаева Д. А., Бектаев К. Б.* Энтропия казахского текста. Алма-Ата, 1973. С. 664—695.
6. *Пиотровский Р. Г.* Информационные измерения языка. Л., 1968. С. 15—56.
7. *Богородицкий В. А.* Введение в татарское языкознание в связи с другими языками. Казань, 1934. С. 35.
8. Современный казахский язык: Фонетика и морфология. Алма-Ата, 1962. С. 18—39.
9. *Шербак А. М.* Сравнительная фонетика тюркских языков. Л., 1970. С. 27—42.
10. *Кенесбаев С. К., Мусабаев Г.* Современный казахский язык // Қазыргі қазақ тілі (Лексика, фонетика). Алма-Ата, 1975. С. 256—258.
11. *Мельников Г. П.* Некоторые способы описывания и анализа гармонии гласных в современных тюркских языках // Вопросы языкознания. 1963 № 6. С. 31—53.
12. *Шербак А. М.* Сравнительная фонетика тюркских языков. Л., 1970. С. 78—100.
13. *Джунисбеков А.* Гласные казахского языка: экспериментально-фонетическое исследование. Алма-Ата, 1972. С. 84—89.
14. *Там же.* С. 86.
15. *Кенесбаев С. К., Мусабаев Г.* Современный казахский язык // Қазыргі қазақ тілі. (Лексика, фонетика). Алма-Ата, 1975. С. 237—238.
16. *Шербак А. М.* Сравнительная фонетика тюркских языков. Л., 1970. С. 77, 79, 81, 106.

17. Трубецкой Н. С. Основы фонетики. М., 1960. С. 83.
18. Щербак А. М. Сравнительная фонетика тюркских языков. Л., 1970. С. 39—42.

З. М. БАЗАРБАЕВА

К ВОПРОСУ ОБ ИНТОНАЦИИ И ЕЕ СВЯЗИ С СЕМАНТИКОЙ

(на материале казахского языка)

В научной литературе много внимания уделяется интонации, выполнены работы по описанию общих и частных аспектов этой темы, по взаимосвязи компонентов интонации. Имеются исследования сопоставительного характера, а целостная теория интонации еще не создана, хотя в этом направлении ведутся изыскания отечественных и зарубежных лингвистов. В настоящее время идет процесс осмысления и обобщения полученных результатов, и проблема интонации остается актуальной, представляя научный интерес как в теоретическом отношении — в сфере изучения интонуологии, так и в практическом — для преподавания неродного языка.

У разных лингвистов свой подход к изучению интонации: одни стараются выявить интонационные структуры различных коммуникативных типов предложения или дать описание некоторых грамматических структур [1, 2], другие устанавливают различные виды интонации [3, 4]. В последнее время известные интонуологи [5, 6] страны выявляют наличие интонационных единиц, функционирующих в речи, и приходят к выводу, что их количество ограничено.

Много споров возникает, когда речь идет о функциях интонации, но никто не возражает против ее многофункциональности.

В рамках данной статьи, когда речь идет о связи интонации с семантикой, привлекают внимание функции, выделяемые Т. М. Николаевой [7] функции членения, связи и передачи смысловых отношений между сегментами. Две последние функции проявляются в предложениях с вводными единицами, где с помощью интонации передается характер связи, обусловленный смысловыми отношениями, между вводными единицами и включающими предложениями.

И. Г. Торсуева [8, 9], говоря о перспективах развития интонологической науки, указывает, что одной из нерешенных проблем является проблема соотношения интонации и семантики. На тесное соотношение интонационного контура с семантическим планом высказывания указывают также Н. А. Баскаков [3] и В. Б. Касевич [5].

В этой работе ставилась задача выявить интонации вводных единиц в зависимости от их семантического содержания, обусловленного контекстом, и определить смысловые отношения между вводными единицами и включающими предложениями на материале казахских текстов.

При подборе экспериментального материала за основу была взята семантическая классификация вводных единиц, по которой все вводные единицы распределены по 9 семантическим группам [10]. Среди них группы:

1. Вводных элементов, выражающих оттенок уверенности: эрине, элбетте, сөз жок, дау жок, рас, шың, дұрысында и т. д. Напр.: **Өз қолына, элбетте**, тура табыс етемін. Бұлар біздің адамдар, **сөз жок**,

2. Вводных единиц, выражающих оттенок предположения: мүмкін, тәрізі, шамасы, сірә, бәлки, кім біледі, құдай біледі. Напр.: **Бәлкім**, есімін бұрмалаған шығар. Бастың ауруы, *мүмкін*, шәйдікі шығар.

3. Вводных единиц, выражающих эмоциональное отношение говорящего к фактам изображаемой действительности: бақытымызға қарай, бір ғажабы, амал нешік, амал не, ең қызығы и др. Напр.: **Әттең, не керек**, патша құламай қалды. Топырақ тартып әкетсе, **амал не**.

4. Вводные единицы могут указывать на последовательность сообщений, на степень их значимости: біріншіден, екіншіден, ақырында, ең әуелі, ең алдымен, оның үстіне и т. п. Напр.: Жасырары жок, **біріншіден**, бұл біздің үлкен мәселеміз. Бұған, **екіншіден**, көп уақыт керек.

5. Вводные единицы, уточняющие, конкретизирующие и подытоживающие высказанную мысль: айтпақшы, айтқандай, сонымен, сөйтіп, сондықтан, мысалы, жалпы и т. д. Напр.: Абайдың айналасы, **көбінше**, әйелдерге толы. Сол жылдың жазы екен-ау, **айтпақшы**.

6. Эта группа включает вводные единицы, выражающие значение противопоставления содержания данного предложения или его части ранее сказанному: әйтпесе, керісінше, қайта, немесе. Напр.: Өкпелей де алған жок. **Қайта**, ол өкпелегендей болды. Әскерден де ол осы кемтарлығынан босанған көрінеді. **Әйтпесе**, соғыста жүре бсереді екен.

7. Вводные единицы, указывающие на источник сообщения: меніңше, сіздіңше, оның ойынша, әдетте, біреудің сөзіне қарағанда и др. Напр.: Мұрат, *әдетте*, көңілді бала болатын. Солар дұрыс айта ма деймін, **менің ойымша**.

8. Вводные единицы, указывающие на приемы и способы оформления мыслей: қысқасы, сайып келгенде, жалпы алғанда, бір сөзбен айтқанда, ұзын сөздің қысқасы и др. Напр.: **Сайып келгенде**, жұмыстың түйіні сізде. Мен оныңызға, **расымды айтсам**, жауап бере алмас едім.

9. Вводные единицы, выражающие призывы к собеседнику, с целью привлечь его внимание к сообщению: ғафу етіңіз, білесің бе, көрдіңіз бе, қарай гөр, несін айтасың и др. Напр.: **Көрдің бе**, әңгімені оларға ұнаған. Ал, **сіз білесіз бе**, мен сорлыға сондай қыян.

Эксперимент, о котором мы ведем речь, проводился в лаборатории экспериментальной фонетики Института языкознания АН КазССР. Экспериментальный анализ заключал следующие этапы: подбор из художественных произведений (М. Ауэзова, Г. Мусрепова, Г. Мустафина, Х. Есенжанова, М. Иманжанова, Л. Абешева) ситуаций, включающих простые повествовательные предложения с вводными единицами в контексте; семантический анализ отобранного материала; подбор дикторов и запись материала; аудиторский, электроакустический анализ и статистическую обработку данных.

Анализ вводных единиц различных семантических групп в пре-, интер- и постпозиции был произведен по пяти главным акустическим параметрам: 1) ширине частотного диапазона (в полтонах), 2) тональному уровню (в герцах), 3) диапазону интенсивности (в децибелах), 4) среднему уровню интенсивности (в децибелах), 5) среднеслоговой длительности (в миллисекундах).

По частотному диапазону в пре-, интер-и постпозиции выявлено, что вводные единицы почти всех семантических групп противопоставляются включающим предложениям с более узкими частотными диапазонами в силу меньшего коммуникативного веса вводных единиц в структуре всего предложения, только в препозиции, в третьей семантической группе, в произнесении обоих дикторов по сравнению с усредненными данными всех семантических групп обнаружено примерное равенство частотных диапазонов в связи с тем что третья семантическая группа выражает эмоциональное отношение говорящего к фактам изображаемой действительности (табл. 1).

Сравнение вводных единиц всех семантических групп с их включающими предложениями по среднечастотному уровню показало, что в препозиции вводные единицы воспринимаются как имеющие значительную информационную нагрузку в структуре

предложения, что проявляется в более высоких среднечастотных уровнях вводных единиц. Особенно явно тенденция повышения голоса вводных единиц по сравнению с включающими предложениями наблюдалась в третьей семантической группе, выражающей эмоции.

Таблица 1. Частотный диапазон вводных единиц (в. е.) и включающих предложений (в. п.) в препозиции (в полутонах)

Семант. группа	1 диктор			2 диктор		
	в. е.	отно- шения	в. п.	в. е.	отно- шения	в. п.
3	12,3	≈	11,7	9,3	>	9,6
Усредненные данные	8,04	<	12,04	7,2	<	10,2

В интерпозиции сопоставление вводных единиц всех семантических групп с их включающими предложениями по этому же параметру выявило тенденцию ступенчатого понижения среднечастотных уровней от первой части включающего предложения к вводной единице, от вводной единицы ко второй части включающего предложения. А в первой и пятой семантических группах, выражающих оттенок уверенности и уточняющих, конкретизирующих высказанную мысль, выявлен более высокий среднечастотный уровень по сравнению с обеими частями включающего предложения (табл. 2, 3).

Таблица 2. Средний уровень частоты основного тона вводной единицы и включающего предложения в препозиции, Гц

Семант. группа	1 диктор			2 диктор		
	в. е.	отно- шения	в. п.	в. е.	отно- шения	в. п.
3	129,4	>	117,5	128,5	>	117,1
Усредненные данные	128,8	>	122,2	127,5	>	123,8

В постпозиции сопоставление анализируемых сегментов по среднему уровню частоты основного тона выявило смысловую сопоставленность вводных единиц и включающих предложений, которая проявляется в более высоких среднечастотных уровнях включающих предложений, организующих содержательный план всего предложения. А вводные единицы всех семантических групп значительно

Таблица 3. Средний уровень частоты основного тона вводной единицы и включающего предложения в интерпозиции, Гц

Семантич. группа	1 диктор					2 диктор				
	1-я часть в. п.	отношения	в. е.	отношения	2-я часть в. п.	1-я часть в. п.	отношения	в. е.	отношения	2-я часть в. п.
1	123,3	<	126,6	>	100,8	107,8	<	116,7	>	108,4
5	119,4	<	148,3	>	116,4	109,4	<	135,0	>	132,2
Усредн. данные	125,5	<	117,6	>	108,5	121,7	<	118,8	>	112,9

уступают им по объему информации, играя второстепенную роль в предложении.

На рис. 1, 2 обозначено ступенчатое понижение среднечастотных уровней от первой части включающего предложения (1 ч. в. п.) к вводной единице (в. е.) и от вводной единицы ко второй части включающего предложения (2-я часть в. п.). Далее показан более

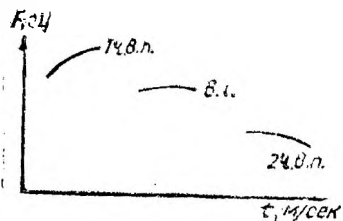


Рис. 1. F — частота основного тона, Гц.

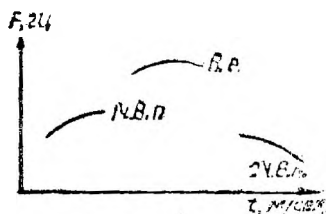


Рис. 2. t — время звучания, мс.

высокий среднечастотный уровень вводной единицы по сравнению со среднечастотным уровнем обеих частей включающего предложения.

По ширине диапазона интенсивности выявлено, что в интерпозиции вводные единицы всех семантических групп контрастируют с включающими предложениями. При этом включающие предложения, несущие фактическую информацию и являющиеся стержнем всего высказывания, характеризовались более широким диапазоном интенсивности в отличие от вводных единиц, предваряющих основную информацию, имеющих более узкие диапазоны интенсивности.

В интерпозиции вводные единицы почти всех семантических групп противопоставляются обеим частям включающего предложе-

ния более монотонным произнесением. Благодаря этому включающие предложения воспринимаются как имеющие больший смысловый вес по сравнению с вводными единицами. Только вводные единицы девятой семантической группы произносились дикторами в более широком диапазоне интенсивности, чтобы привлечь внимание собеседника к сообщению (табл. 4).

Таблица 4. Диапазон интенсивности вводных единиц и включающих предложений в интерпозиции, дециб.

Семан- гич. группа	1 диктор					2 диктор				
	1-я часть в. п.	отно- шения	в. с.	отно- шения	2-я часть в. п.	1-я часть в. п.	отно- шения	в. с.	отно- шения	2-я часть в. п.
9 Усредн. данные	2,4	<	10,6	<	15,7	3,7	<	5,7	<	11,6
	5,1	<	4,6	<	15,9	4,3	>	3,1	<	9,5

В постпозиции контраст анализируемых сегментов проявляется в различной ширине диапазона интенсивности. Во всех семантических группах включающие предложения выступают как основной информативный центр всего высказывания. Их смысловая и интонационная выделенность проявляется в расширенных диапазонах интенсивности по сравнению с вводными единицами, которые произносятся монотонно. Анализ вводных единиц всех семантических групп по среднему уровню интенсивности в препозиции показал, что они характеризуются более энергичным произнесением, проявляющимся в более высоком среднем уровне интенсивности по сравнению с включающими предложениями. Это свидетельствует о значительном смысловом весе вводных единиц в предложении.

В интерпозиции вводные единицы и обе части включающего предложения имеют ступенчатый характер понижения среднего уровня интенсивности. А в первой семантической группе, выражающей оттенок уверенности, в произнесении обоих дикторов вводные единицы интонационно выделяются, в результате чего контраст между вводной единицей и первой частью включающего предложения стирается (табл. 5).

В постпозиции второстепенный характер вводных единиц и их малая информативность по сравнению с включающими предложениями проявляется в их более низких уровнях интенсивности, причем разница между анализируемыми сегментами значительна.

Сравнение вводных единиц разных семантических групп по среднеслоговой длительности показало, что в препозиции они ха-

рактируются более медленным, чем включающие предложения, темпом произнесения, что свидетельствует о значительном коммуникативном весе вводных единиц в этой позиции.

В интерпозиции вводные единицы произносятся быстрее, чем первая, и медленнее, чем вторая часть включающего предложения за исключением пятой семантической группы, уточняющей и конк-

Таблица 5. Средний уровень интенсивности вводных единиц и включающих предложений в интерпозиции, дециб.

Семантич. группа	1 диктор					2 диктор				
	1-я часть в. п.	отношения	в. с.	отношения	2-я часть в. п.	1-я часть в. п.	отношения	в. с.	отношения	2-я часть в. п.
1 Усредн. данные	20,3	≈	20,7	>	16,5	24,3	≈	24,8	>	20,2
	23,0	<	20,8	>	16,3	24,0	>	22,0	>	18,9

ретизирующей высказанную мысль. Вводные единицы этой группы произносятся в более медленном темпе, чем обе части включающего предложения, что объясняется, видимо, их большей семантической наполненностью. В постпозиции смысловая соотносительность анализируемых сегментов почти во всех группах проявляется в более быстром по сравнению с включающими предложениями темпе произнесения. Смысловая и интонационная выделенность

Таблица 6. Среднеслоговая длительность вводных единиц и включающих предложений в интерпозиции, м/сек

Семантич. группа	1 диктор					2 диктор				
	1-я часть в. п.	отношения	в. с.	отношения	2-я часть в. п.	1-я часть в. п.	отношения	в. с.	отношения	2-я часть в. п.
5 Усредн. данные	177,1	<	213,3	>	137,6	169,0	<	202,2	>	142,2
	220,7	>	170,7	>	163,6	216,0	>	175,8	<	133,9

вводных единиц отмечена в первой семантической группе, выражающей оттенок уверенности. Это проявляется в их замедленном темпе произнесения по отношению к включающим предложениям (табл. 6, 7).

Как показало наше исследование, интонация находится в тесной взаимосвязи с семантикой. В результате эксперимента удалось

выявить, что акустические параметры интонации могут «говорить» о большем или меньшем смысловом весе тех или других семантических групп вводных единиц. Причем, как правило, акустические параметры интонации вводных единиц (тональные уровни, уровень интенсивности, среднеслоговая длительность) наиболее ярко выступают в препозиции, т. к. в начальной позиции все внимание кон-

Таблица 7. Среднеслоговая длительность вводных единиц и включающих предложений в постпозиции, м/сек

Семантич. группа	1 диктор			2 диктор		
	в. п.	отношения	в. с.	в. с.	отношения	в. п.
1	138,2	<	190,0	153,4	<	200
Усредн. данные	161,9	>	150,3	172,2	>	161,4

центрируется на вводных единицах, и они в полной мере реализуют свое содержание. В интер- и постпозиции интонационные характеристики вводных единиц нивелируются. Особенно это заметно в постпозиции, т. к. в конечной позиции отношение говорящего к высказываемой мысли выражается уже после того, как формируется основная мысль.

Итак, на примере вводных единиц и включающих предложений были выявлены характер связи, обусловленный смысловыми отношениями между ними в пре-, интер- и постпозиции, а также интонационная специфика исследуемых единиц в зависимости от их семантического содержания и позиции по отношению к включающему предложению.

Интонация, как выяснилось, наряду с лексическими и грамматическими средствами, играет большую роль в передаче тех или иных смысловых оттенков, являясь формой выражения содержания, и прямо пропорциональна семантической наполненности. И, как показали данные просодических параметров речи, чем больше смысловая нагрузка, тем ярче интонация.

ЛИТЕРАТУРА

1. Артемов В. А. К вопросу об интонации русского языка // Учен. зап. КГПИИЯ. 1953. Т. 6.
2. Базарбаева З. М. Семантико-грамматическая характеристика вводных единиц во французском, казахском и русском языках // Известия АН КазССР. Сер. филол. 1980. № 3.

3. *Баскаков Н. А.* Об интонации и фразовом ударении в тюркских языках // Советская тюркология. 1984. № 1.
4. *Брызгунова Е. А.* Практическая фонетика и интонация русского языка. М., 1963.
5. *Касевич В. Б.* Фонологические проблемы общего и восточного языкознания. М., 1983.
6. *Михеева В. Т.* Интонация повествовательного предложения с обособленными определениями в современном французском языке: Автореф. дис. ... канд. филол. наук. М., 1974.
7. *Николаева Т. М.* Фразовая интонация славянских языков. М., 1977.
8. *Торсуева И. Г.* Теория интонации. М., 1974.
9. *Торсуева И. Г.* Современная проблематика интонационных исследований // Вопросы языкознания. 1984. № 1.
10. *Туркбенбаев Н. У.* Интонация простых вопросительных предложений в казахском языке: Автореф. дис. ... канд. филол. наук. М., 1966.

А. Д. ДЖУНИСБЕКОВ, С. АЛДАБЕРГЕНОВА

СТАТИСТИЧЕСКИЙ АНАЛИЗ СИНГАРМОЗВУКОВ В КАЗАХСКОМ ЯЗЫКЕ

Сингармонизм — это не просто фонетический факт тюркских языков, как представляют и описывают его большинство тюркологов-фонетистов, а фонологическая суть тюркского слова. Несмотря на кажущуюся простоту сингармонизма, порожденную представлением о нем как о гармонии гласных, уровень его фонологической и фонетической сложности, безусловно, таков, что и словесное ударение в акцентных и тонема в слоговых языках. Это свидетельствует о том, что если в этих языках сегментное описание на уровне звуков строго обусловлено словесным ударением и тоном, то в тюркских языках — сингармонизмом. Следовательно, сегментация речи на отдельные звуки и их фонологическое и фонетическое описание без опоры на сингармоническую организацию слова не отражает реальной языковой действительности, а также не определяет и не описывает реализацию фонологических единиц. Так, например, определение и описание акустической характеристики или же артикуляции согласного должно охватить как минимум четыре его реализации, тогда как в современных тюркских грамматиках дается единственная обобщенная его характеристика, которая не тождественна ни одной из форм его реализации.

Таким образом, исследование звуковой системы казахского языка должно начинаться с определения просодики слова, ибо только она определяет порядок и линейную величину членения речи, основные типы сочетания звуков (гласных и согласных), артикуляционное и акустическое описание этих сочетаний с точки зрения надсегментных единиц. Так показывают лингвистическая практика, исследования индоевропейских и восточных языков.

Любой анализ и описание звукового строя казахского языка должны отталкиваться от сингармонизма, через слово и слог к звуку. По крайней мере, такая направленность анализа должна соблюдаться в исследовательских работах. Такой порядок анализа, во-первых, позволит охватить все варианты реализации фонетических единиц, во-вторых, избавит от той неразберихи в тюркских грамматиках (фонетиках), где густо смешаны ассимилятивные (позиционно-комбинаторные) и сингармонические модификации звуков.

Под сингармонизмом мы понимаем однородную тембральную организацию всех звуков, как гласных, так и согласных, составляющих весь фонетический облик слова. Это касается и односложного, и многосложного слов, где однородный тембр объединяет уже слоги. Варианты палатального и лабиального сингармонизма ведут себя как равноправные единицы языка, наделены одинаковой функцией — словоразличительной. Вопреки традиционному мнению, принципиальной функциональной разницы между ними нет. Она скорее всего артикуляторная. Следовательно, «губная гармония» не самостоятельное явление, а лишь разновидность сингармонизма.

Вопрос о роли согласных в системе сингармонизма привлекал не всех исследователей. Этим подчеркивались несущественность фонетической характеристики в сингармонической просодике тюркского слова и их зависимость от характера вокального сегмента слога, о чем свидетельствует и термин «гармония гласных», широко распространенный прежде, но не менее распространенный, к сожалению, и ныне.

До сих пор считается, что каждый тюркский согласный имеет одно (единственное) артикуляционное и акустическое описание. Тогда как в сингармонизме один согласный — это система, состоящая из четырех сингармонизмов, которые кроме общих черт, присущих этому согласному, имеют взаимоисключающие артикуляционно-акустические характеристики. Количество значимых типов согласных в тюркских языках несравненно меньше, чем количество их сингармонических вариантов (оттенков). Так, например, в тюркских языках система сингармонизмов каждого согласного обра-

зуется путем противопоставления признаков по твердости, мягкости, огубленности и неогубленности.

Что касается гласных, то признание сингармонизма как пресодической единицы в тюркских языках заставляет пересмотреть некоторые моменты системы гласных. В частности, роль гласных в тюркских языках может быть только слогаорганизирующая, а различные гласные являются вариантами (оттенками) реализации сингармотембра. Следовательно, количество значимых типов гласных в казахском языке меньше, чем количество их сингармонических вариантов. Девять гласных казахского языка образуют три сингармотипа: гласный верхнего подъема (реализуется в виде четырех сингармозвуков), гласный нижнего подъема (реализуется в виде двух сингармозвуков) и дифтонг (реализуется в виде трех сингармозвуков).

Статистический анализ сингармозвуков был проведен на тексте казахских сказок. Текст сказок был заранее полностью затранскрибирован. Как известно, в казахском языке принят русский алфавит с некоторыми дополнениями, что имеет большое практическое значение, так как во всех казахских школах изучается русский язык, и иметь единую графику наиболее целесообразно с точки зрения обучения. Казахская письменность в основном составлена так, что каждый звук обозначается отдельной буквой. Однако есть и исключения, определяющиеся только правилами орфографии казахского языка, которые не всегда дают правильное представление о количестве звуков в составе слога, слова или морфем. Например, разные оттенки некоторых согласных передаются отдельными буквами или же, наоборот, сочетания двух, трех самостоятельных звуков — одной буквой, поэтому необходимость как фонологической, так и фонетической транскрипции очевидна. Поэтому текст сказок был заранее полностью затранскрибирован.

В рассматриваемых нами текстах сказок были выявлены следующие статистические данные:

твердых неогубленных согласных — 10665
мягких неогубленных согласных — 7891
твердых огубленных согласных — 3490
мягких огубленных согласных — 3403

Если взять весь текст затранскрибированных сказок за 100%, а это 66220 звуков, то в процентном соотношении выявляется следующая картина:

твердых неогубленных согласных — 20,4%
мягких неогубленных согласных — 15,7%
твердых огубленных согласных — 9,2%
мягких огубленных согласных — 8,4%

То есть 5 367% всего текста являются согласными. Основные тембральные характеристики согласных распределены следующим образом:

палатальный тембр согласных — 36,1%
лабиальный тембр согласных — 17,6%
твердый тембр согласных — 29,6%
мягкий тембр согласных — 24,1%

При анализе гласных звуков были получены следующие данные: всего гласных в тексте — 46,3%, из них:

твердых негубных гласных — 7 770
мягких негубных гласных — 4 410
твердых губных гласных — 1 310
мягких губных гласных — 2 100

Из общего текста сказок для гласных звуков выявляется следующая картина:

твердых негубных гласных — 24,2%
мягких негубных гласных — 13,4%
твердых губных гласных — 6,9%
мягких губных гласных — 2,8%

Основные тембральные характеристики гласных распределены следующим образом:

палатальный тембр гласных — 36,6%
лабиальный тембр гласных — 9,7%
твердый тембр гласных — 31,1%
мягкий тембр гласных — 16,2%

Таким образом, оказалось, что тембральные компоненты сингармонизма в тексте распределены неравномерно: твердых сингармозвуков больше, чем мягких, палатальных сингармозвуков больше, чем лабиальных. Полученные количественные характеристики распределения сингармосегментов будут полезны для составления оптимальной системы машинной графики и машинной обработки текста при составлении машинного фонда казахского языка, а также для решения вопросов теории словесной просодии.

ЛЕКСИКА БОЙЫНША
ЗЕРТТЕУЛЕР
ИССЛЕДОВАНИЯ
ПО ЛЕКСИКЕ

ЛИНГВОСТАТИЧЕСКИЕ ИССЛЕДОВАНИЯ ЛЕКСИЧЕСКОЙ БЛИЗОСТИ КАРАКАЛПАКСКОГО, КАЗАХСКОГО И УЗБЕКСКОГО ПОДЪЯЗЫКОВ ПУБЛИЦИСТИКИ

В работах многих ученых, на которые ссылаемся, рассмотрены общие количественные характеристики тюркской лексики [1—14]. Применяя статистико-информационную методику, мы провели сравнение последних с аналогичными параметрами по тюркским и индоевропейским языкам. Они показали, что по общим лексико-статистическим и информационным показателям тюркские языки заметно отличаются от индоевропейских. Однако использованные методы не дали возможности провести детальное сравнение количественных сходств и различий между рассматриваемыми тюркскими языками. Эти методы не обладают достаточной разрешающей способностью для того, чтобы вскрыть особенности языков.

Мы должны применять такой исследовательский аппарат, который позволил бы нам фиксировать сходство и различия в означаемых и означающих лексических единицах (словах) в системах сопоставляемых языков, также сделать наблюдения за схождением и расхождением речевой употребительности слов в этих языках. Воспользуемся количественно-сопоставительной методикой исследования лексической близости каракалпакского, казахского и узбекского подязыков публицистики на уровне языковой и речевой системы.

1. **На уровне языковой системы.** В течение многих веков и особенно в настоящее время каракалпакский язык интенсивно контактирует с казахским и узбекским. С первым каракалпакский язык объединяют древние исторические связи — оба языка принадлежат к кыпчакско-ногайской подгруппе кыпчакской группы западнохуннских языков [15]. С узбекским языком, принадлежащим к карлукской группе западнохуннских языков, каракалпакский язык связывает более далекое родство. Однако оба находятся в последние десятилетия в самом тесном вторичном кон-

такте в связи с вхождением Каракалпакской АССР в состав Узбекистана.

Задача состоит в том, чтобы численно оценить степень количественно-лексической близости этих языков (точнее, их публицистических подязыков) и определить место современной газетной лексики каракалпакского языка по отношению к лексике двух соседствующих с ним языков. Это позволит рассмотреть важную для тюркских общетеоретическую проблему о соотношении первичного родства и вторичных контактов, основанных на культурно-хозяйственных связях между родственными языками. Необходимо также оценить лексическую близость каракалпакского языка к казахскому и узбекскому языкам на уровне текста.

Решая первую задачу, мы подвергнем количественному анализу первые 200 слов частотного словаря (они покрывают примерно 45,16% взятого наугад каракалпакского публицистического текста) с эквивалентами в частотных словарях казахского [16] и узбекского [17]. Все эти словари получены путем переработки частотных списков словоформ, полученных этими авторами, а также нами.

При преобразовании частотных словарей словоформ в частотный словарь слов и при их объединении в таблицу 1 применялись следующие правила:

1. Все именные формы, относящиеся к определенной основе, объединялись в одно множество, представителем которого выступала исходная (словарная) форма имени — форма единственного числа, именительного падежа:

ккалп. жыл (жыл, жылдын, жылдын, жылга, жылда... и т. д.).

2. В качестве глагольной словарной формы использовалась основа, совпадающая с формой второго лица, единственного числа настоящего времени:

ккалп. ал — алдым, алдын алды, ..., алыу ... и т. д.

Далее нами в таблице указаны 200 каракалпакских слов, расположенных в порядке убывания частот и возрастания рангов (столбцы 1 и 4).

В столбцах 2 и 3 приведены казахские и узбекские эквиваленты каракалпакских слов, а в столбцах 5 и 6 даны ранги (порядковые номера) этих эквивалентов. Если казахский или узбекский эквивалент попадает в зону высоких рангов, находящуюся выше порядкового номера = 300, то сам ранг не указывается, а проводится помета 300. Если ранговая или частотная характеристика узбекского эквивалента в частотном словаре Мухамедова отсутствует, то на пересечении соответствующей строки и столбца 6 ставится прочерк (—).

Прямое наблюдение указанных трех ранговых списков пока-

зывает различные формальные, семантические, а также количественные соотношения между каракалпакским словом, с одной стороны, и казахским и узбекским эквивалентами, — с другой. Формальные и семантические сопоставления отражают языковые соотношения лексических парадигм (словарей) каракалпакского, казахского и узбекского языков. Сопоставление же рангов и стоящих за ними частот служит средством обнаружения сходства и различия в речевом функционировании лексических единиц в трех рассматриваемых нами языках.

Различия в речевом функционировании контрольных слов мы рассмотрим подробно ниже.

Если говорить о семантических и формальных соотношениях, то здесь можно выделить следующие типовые соотношения между каракалпакским словом и его эквивалентами.

1. Каракалпакское слово имеет совпадающие с ним по значению и форме казахские и узбекские эквиваленты. Такая ситуация возникает либо в силу общих исторических условий развития общетюркских основ — ср. ккалп. *бер*—(i-17), каз. *бер*—(i-20), узб. *бер*—(i-20), либо в результате заимствования новых слов из других языков при условии их одинаковой адаптации фонетико-фонологическим системам всех трех языков — ср. ккалп. *совет* (i-8), каз. *совет* (i-8), узб. *совет* (i-21) «советский». Обозначим эту ситуацию кодом 1/1.

2. Каракалпакское слово может иметь близкие к нему по значению и форме эквиваленты в двух других языках, причем фонетические различия в этих случаях закономерны с точки зрения тех процессов, которые определили формирование фонетики и фонологии каждого из трех языков — ср. ккалп. *исле*—(i-25), каз. *исте*—(i-101), узб. *шила*—(i-41) «работать». Для обозначения этой ситуации используем код 2/2.

3. Каракалпакское слово и его казахский и узбекский эквиваленты восходят к разным этимологическим источникам, либо к разным тюркским основам — ср. ккалп. *шөлкем* (i-80), каз. *үйым* (i-43), узб. *ташкилот* (i-59) «организация», либо к тюркской и нетюркским основам — ср. ккалп. каз. *қала* (i-14), узб. *шахар* (i-149). Кодом этой ситуации будет последовательность 3/3.

Если рассматривать пары языков (каракалпакский — казахский, каракалпакский — узбекский), то соотношение каждой двойки слов из частотного словаря будет описываться одной из трех приведенных выше ситуаций. Если же рассматривать тройки слов (каракалпакское — казахское — узбекское), то здесь возможны еще комбинированные ситуации, возникающие в связи с тем, что каракалпакское слово, будучи идентичным казахскому слову, может не совпасть или только частично совпасть с узбекским словом, и

Таблица 1. Первые двести наиболее частных слов каракалпакских публицистических текстов и их казахские и узбекские эквиваленты (скобками и звездочкой помечены слова, которые отсутствуют в зоне $i=1-300$)

Каракалпакский	Казахский	Узбекский	Ранг			Русский перевод
			Ккалп.	Каз.	Узб.	
1	2	3	4	5	6	7
Ҳәм	Жәле	Ҳам	1	5	9	И
Бол—	Бол—	Бул—	2	1	4	Быть
Ол	Ол	У	3	2	6	Он
Менен	Мен	Билан	4	4	3	С
Бул	Бул	Бу	5	3	2	Это, этот, эта
Жыл	Жыл	Йил	6	17	10	Год
Ет—	Ет—	Эт—	7	30	73	Десять
Совет	Совет	Совет	8	24	21	Советский
Ал	Ал	Ол	9	10	13	Братъ, взять
Ушын	Ушын	Учун	10	23	11	Последлог для,
Мийнет	Еңбек	Мехнат	11	41	42	ради Труд
Атыр—	Отыр—	(Утир)*	12	6	$i>300$	Вспомогательный глагол, выражаю- щий длительность действия
Өз	Өз	Уз	13	9	15	Сам, свой
Жумыс	Жұмыс	Иш	14	19	8	Работа
Халық	Халық	Халк	15	18	14	Народ
Комитет	Комитет	Комитет	16	39	16	Комитет
Бер	Бер	Бер	17	20	20	Дай
Биз	Біз	Биз	18	27	67	Мы
Де—	Де—	Де—	19	8	13	Скажи, говори
Бир—4	(бир)*	Бир	20	$i>300$	17	Некоторый
Партия	Партия	Партия	21	13	16	Партия
Ҳәр	Эр	Ҳар	22	213	28	Каждый, любой
Республика	Республика	Республика	23	26	71	Республика
Жер	Жер	Ер	24	29	65	Земля
Исле—	Исте	Ишла—	25	101	41	Работа
Бойынша	Бойынша	буйнча	26	148	207	По (последлог)
Күн	Күн	Кун	27	32	55	День
Ис	Ис	Иш	28	22	8	Дело
Орайлык	Орталық	Марказий	29	59	18	Центральный
Пахта	(Мақта)*	Пахта	30	$i>300$	168	Хлопок
Барлык	Барлық	Ҳамма	31	48	29	Все, весь
Пенен	Пен	Билан	32	110	3	И
Адам	Адам	(Одам)*	33	31	$i>300$	Человек
Хожалық	Шаруашылық	Хужалик	34	25	39	Хозяйство
КПСС	КПСС	КПСС	35	111	—	КПСС

1	2	3	4	5	6	7
Және	Және	Яна	36	5	134	И, еще
Жана	Жана	Янги	37	34	31	Новый
Үлкен	Үлкен	Катта	38	112	36	Большой
Кел	Кел	Кел	39	12	22	Приходи
Бир-А	Бір	Бир	40	16	17	Один
Район	Аудан	Район	41	81	135	Район
Совхоз	Совхоз	Совхоз	42	68	198	Совхоз
Каракал- пакстан	(Каракал- пакстан)*	(Қорақол пакстон)*	43	—	—	Каракалпакия
Көрсет	Көрсет	Курсат—	44	42	117	Показывать
Өзбек- стан	(Өзбек- стан)*	(Ўзбе- кстон)*	45	—	—	Узбекистан
План	Жоспар	План	46	62	94	План
Бес	(Бес—)*	Беш	47	i>300	116	Пять
Мың	Мың	Минг	48	80	90	Тысяча
Комсо- мол	Комсо- мол	Комсо- мол	49	225	215	Комсомол
Шық	Шық	Чық	50	40	54	Выходи
Сөз	Сөз	Сүз	51	85	300	Слово
Орынла	Орында	Бажар	52	98	287	Выполняй
Биринши	Бірінші	Биринчи	53	133	58	Первый
Жокары	Жоғары	Олш	54	141	106	Верховный
СССР	СССР	СССР	55	203	—	СССР
Уактында	(Уақтын- да)*	(Вақти- да)*	56	i>300	i>300	Вовремя
Усы	Осы	Шу	57	14	7	Это, эта
Де	Де	Де	58	15	—	Также; тоже
Басқа	Басқа	Бошқа	59	160	40	Другой, нной
Жылдық	(Жылдық)	Йиллик	60	i>300	97	Годовой
Жас	Жас	(Ёш)*	61	46	i>300	Молодой
Коллек- тив	Коллек- тив	Коллек- тив	62	197	107	Коллектив
Тонна	Тонна	Тонна	63	288	271	Тонна
Социали- стик	Социали- стик	Социали- стик	64	107	31	Социалистический
Ел	Ел	Эл	65	21	i>300	Родина
Бригада	(Брига- да)*	Бригада	66	i>300	i>300	Бригада
Сол	Сол	Шу	67	37	7	Тот
Мийет- кеш	Еңбекші	Мехнат- каш	68	150	95	Трудолюбивый
Отыр	Отыр	(Утир)*	69	6	i>300	Сиди
Еди	—	Эди	70	—	60	Неполный глагол со зна- чением совер- шенного дей- ствия

1	2	3	4	5	6	7
Секретарь Ал-ш	(Секретарь)* Ал	Секретарь Лекин	71 72	i > 300 52	34 232	Секретарь Союз а
Ет-	Ет-	Ут- v	73	33	33	Проходить
Жолдас Бар-	Жолдас Бар-	Уртоқ Бор-	74 75	180 77	124 181	Товарищ Идти
Автономиялы Қарар	(Автономиялы)* (Қарар)*	(Автоном)* Қарор	76 77	— i > 300	i > 300 80	Автономия Постановление
Қурылыс Ағза	Қурылыс Мүше	Курилиш Льзо	78 79	89 233	114 63	Строительство Член
Шөлкем Мектеп	Ұйым Мектеп	Ташкилот Мактаб	80 81	43 209	59 300	Организация Школа
Жарыс Жоқарғы	Жарыс Жоғары	Мусобақа Олий	82 83	155 141	191 106	Соревнование Верховный
Тур- v	Тур	Тур- v	84	168	47	Стоять, находиться
Оқыушы v	Оқушы	(Уқувчи)*	85	i > 300	i > 300	Ученик
Ауыл	Ауыл	Қишлоқ v	85	95	100	Село
Қол Арттыр	Қол Арттыр	Қул Ошир	87 88	87 118	193 68	Рука Увеличивая
Гектар Ерис-	Гектар Жетіс-	(Гектар)* Эриш	89 90	219 69	i > 300 261	Гектар Достичь
Жаксы Атындағы	Жаксы Атындағы	Яхши (Номдаги)*	91 92	124 283	185 i > 300	Хорошо Имени
Сиясий Бала	Саяси Бала	Сиёсий Бола	93 94	182 231	46 179	Политический Мальчик
Процент	Процент	Процент v	95	184	102	Процент
Жол Табыс	Жол Табыс	Йул Муваф- факият	95 97	36 174	29 204	Дорога Достижение
Тур-	Тур-	Тур-	98	168	12	Стоять, находиться
Мәселе Хызмет	Мәселе Қызмет	Масала Хизмат	99 100	65 164	50 240	Задача Служба
Турмыс Ең	(Турмыс)* Ең	(Турмуш)* Энг	101 102	i > 300 84	i > 300 104	Жизнь Самый
Мәмлекетлик	Мемлекетлік	Давлат	103	215	23	Государственный
Съезд Ҳәзирги	Съезд Қазіргі	Съезд Ҳозирги	104 105	83 208	101 128	Съезд Настоящий
Бар-	Бар-	Бар- v	106	35	218	Есть, имеется
Көп	Қөп	Қул v	107	53	52	Много

1	2	3	4	5	6	7
Керек	Керек	Керак	108	102	99	Нужно, надо
Окью	Оку	(Укиш)*	109	236	i > 300	Учеба
Есап	Есеп	Ҳисоб	110	122	270	Учет
Колхоз	Колхоз	Колхоз	111	161	169	Колхоз
Уақыт	Уақыт	Вакт	112	90	113	Время
Белгиле-	(Белгиле-)	(Белгила-)*	113	i > 300	i > 300	Отмечать
Председатель	(Председатель)*	Раис	114	i > 300	66	Председатель
Өнім	Өсу	(Унум)*	115	227	i > 300	Рост, произрастание
Зурәт	Өнім	Ҳосил	116	72	286	Урожай
Пайдалан	Пайдалан	Фойдалан	117	131	277	Используй
Қатар	(Қатар)*	(Қатор)*	118	i > 300	i > 300	Ряд
Тил	(Тил)*	Тил	119	i > 300	i > 300	Язык
Күш	Күш	Куч	120	74	51	Сила
Мен	Мен	Мен	121	210	237	Я
Асыр	Асыр	Ошир	122	128	216	Увеличивай
Басла	Баста	(Бошла)*	123	104	i > 300	Начинай
Хызметкер	(Хызметкер)*	(Хизматчи)*	124	i > 300	i > 300	Служащий
Рауажландыр	Дамыт	Ривожлантир	125	272	91	Развивай
Еки	Екі	Икки	126	61	25	Два
Жағдай	Жағдай	(Ҳола)*	127	51	i > 300	Положение
Жаса	Жаса	Яша	128	28	236	Жить
Рабочий	Жұмысшы	Ишчи	129	167	115	Рабочий
Тапсыр	(Тапсыр)*	(Топшир)*	130	i > 300	i > 300	Сдавай
Министр	Министр	Министр	131	181	26	Министр
Қатнас-	Қатыс-	Қатнаш-	132	151	260	Участвовать
Коммунистлик	Коммунистлик	Коммунистлик	133	93	62	Коммунистический
Пленум	Пленум	(Пленум)*	134	198	i > 300	Пленум
Тийкар	Негіз	(Асос)*	135	223	i > 300	Основа
Кең	Кең	Кенг	136	214	194	Широкий
Бирге-ш	Бирге	Бирга	137	200	200	Вместе
Бас	Бас	Бош	138	60	32	Главный, генеральный
Жәрдем	(Жәрдем)*	Ёрдам	139	i > 300	92	Помощь
Президиум	(Президиум)*	Президиум	140	i > 300	201	Президиум
Улы	Ұлы	Буюк	141	258	285	Великий
Ара	Ара	(Ора)*	142	163	i > 300	Промежуток
Нокис	Нөкіс	Нукус	143	—	—	Нукуа
Илим	Ғылым	Фап	144	297	252	Наука

1	2	3	4	5	6	7
Таярла	(Дайында)*	(Тайёрла)*	145	1 > 300	1 > 300	Приготовлявай
Оныншы	(Оныншы)*	(Упинчи)*	146	1 > 300	—	Десятый
Атыз	(Дала)*	(Дала)*	147	1 > 300	1 > 300	Поле
Билим	(Білім)*	(Билим)*	148	1 > 300	1 > 300	Знание
Кала	Қала	Шахар	149	109	149	Город
СССР	(СССР)*	(СССР)*	150	1 > 300	1 > 300	СССР
Машина	Машина	Машина	151	189	205	Машина
Тарау	Сала	Соха	152	216	85	Отрасль
Кет	Кет	Кет	153	86	148	Уходи
Емес	Емес	(Эмас)*	154	1 > 300	—	Именное отрицание
Алтын	(Алтын)*	(Олтин)*	155	1 > 300	1 > 300	Золото
Ишиндеги	(Ишиндеги)*	(Ичидаги)*	156	1 > 300	1 > 300	Внутри
Кал	Қал	Қола	157	50	56	Останья
Тағы	Тағы	Яна	158	138	134	Еще
Центнер	(Центнер)*	(Центнер)*	159	1 > 300	1 > 300	Центнер
Қабыл	Қабыл	Қабул	160	235	108	Прием
Жаз	Жаз	(Ез)*	161	261	1 > 300	Пиши
Кой	Кой	Куй	162	94	129	Ставь
Аслам	Астам	(Ортнк)*	162	293	1 > 300	Свыше
Мәдени	(Мәдени)*	(Маданий)*	164	1 > 300	1 > 300	Культурный
Бөлім	(Бөлім)*	(Булим)*	165	1 > 300	1 > 300	Отдел
Мал	Мал	(Мол)*	166	146	1 > 300	Скот
АССР	(АССР)*	(АССР)*	167	—	—	АССР
Толық	Толық	(Тулик)*	168	205	1 > 300	Полный
Ҳазир	Қазір	Ҳозир	169	92	226	Сейчас
Салы	(Салы)*	(Шоли)*	170	1 > 300	1 > 300	Неочищенный рис
Бурын	(Бұрын)*	(Илгари)*	171	1 > 300	1 > 300	Раньше
Табыслы	(Табысты)*	Музаффиякятли	172	1 > 300	294	Успешно
Шөлкем-лестир	Ұйымдастыр	Ташкил	173	278	119	Организуй
Жүр	Жүр	(Юр)*	174	58	1 > 300	Ходи
Пүткіл	Бүкіл	Жахон	175	75	162	Всех, все
Аралық	(Қашықтық)*	(Оралик)*	176	1 > 300	1 > 300	Промежуток
Орынла	Орында	Баяр	177	98	287	Выполняй
Жоқ	Жоқ	Йуқ	178	55	267	Нет, не имеется

1	2	3	4	5	6	7
Талап	Талап	Талаб	179	132	126	Требование
Бирақ	Бірақ	Лекин	180	103	180	А, но, однако
Экономикалық	Экономикалық	Иқдисодий	181	108	74	Экономический
Қыз	(Қыз)*	(Қиз)*	182	i>300	i>300	Девочка
Ғана	Ғана	(Гина)*	183	57	i>300	Только, лишь
Беккемле	(Беккемде)*	Мустахамла	184	i>300	79	Укрепил
Мәмлекет	Мемлекет	Мамлакат	185	130	19	Государство
Кейин	Кейін	Кейин	186	105	292	После того со значением после
Сапаат	Өнеркәсіп	Саноат	187	115	86	Промышленность
Жәмийетлик	Қоғамдық	(Ижтимоий)*	188	265	i>300	Общественный
Басшылық	Басшылық	рахбарлик	189	277	70	Руководство
Быйылғы	(Быйылғы)*	(Бу йилги)*	190	i>300	—	Этого года
Былай	Былай	—	191	136	—	Так, этак
Миллион	Миллион	Миллион	192	191	150	Миллион
Сапа	(Сапа)*	Сифат	193	i>300	299	Качество
Табыл	Табыл	(Топил)*	194	241	i>300	Находись
Ақ	(Ақ)*	(Оқ)*	195	i>300	i>300	Белый
Газета	Газет	Газет	196	176	153	Газета
Усындай	Осындай	Шундай	197	149	136	Такой
Жууап	(Жауап)	(Жавоб)	198	i>300	i>300	Ответ
Актив	Актив*	(Актив)*	199	i>300	i>300	(Актив)*
Екинши	Екінші	(Иккинчи)*	200	154	—	Второй

наоборот. Таких ситуаций шесть. Обозначим каждую кодом из двух цифр, первая из которых указывает на соотношение каракалпакского и казахского слова, а вторая отмечает соотношение каракалпакского и узбекского слова, и получаем следующий ряд:

— 1/2 ккалп. жыл (i-6), каз. жыл (i-17), узб. йил (i-10) «год»;
 — 1/3 ккалп. жумыс (i-14), каз. жұмыс (i-19), узб. иш (i-8) «работа»;

— 2/1 ккалп. газета (i-196), каз. газет (i-176), узб. газета (i-153) «газета»;

— 2/3 ккалп. жоқары (i-54), каз. жоғары (i-141), узб. олий (i-106) «верховный»;

— 3/1 ккалп. район (i-41), каз. аудан (i-81), узб. район (i-135) «район»;

— 3/2 ккалп. хэм (i-1), каз. және (i-5), узб. хам (i-9) «и».

Располагая средствами, оценивающими степень языковой близости наиболее частых каракалпакских слов с их казахскими и узбекскими эквивалентами, попытаемся количественно определить общую парадигматическую близость каракалпакского языка в его публицистическом варианте к казахскому и узбекскому подъязыкам публицистики.

В таблице 2 представлено количественное распределение зафиксированных в табл. 1 троек слов по приведенным выше девяти ситуациям. Приписав соотношению «1» два балла, соотношению «2» — один балл, а соотношению «3» — нуль баллов, суммируем общее количество баллов по всем ситуациям относительно соотношения каракалпакский — казахский (столбец 3) и каракалпакский — узбекский (столбец 4), получим количественную оценку парадигматической близости указанных пар языков. Как показывают результаты этих подсчетов, близость каракалпакской лексической языковой системы к казахской системе примерно в полтора раза превышает относительно наиболее частых слов близость каракалпакской лексики к лексике узбекского языка. Такое соотношение полностью отвечает генетическим отношениям рассматриваемых языков. Как уже говорилось, каракалпакский и казахский языки находятся в очень близком родстве, входя в одну и ту же кыпчакско-ногайскую подгруппу. В то же время узбекский язык, входящий в карлукскую группу, находится по отношению к каракалпакскому языку в более далеком историческом родстве.

2. На уровне речи. Если обратиться к изучению соотношений в употребительности каракалпакских слов и их казахских или узбекских эквивалентов в тексте, то здесь снова выделяется три типовые ситуации:

а) каракалпакское слово и его эквивалент имеют примерно одинаковую употребительность, которая оценивается разностью

Таблица 2. Количественные оценки лексической близости каракалпакского, казахского и узбекского подъязыков публицистики на уровне языковой системы

Код ситуаций	Количество ситуаций	Количество баллов	
		соотношение каракалпакской и казахск. лексических систем	соотношение каракалпакской и узбекской лексических систем
1/1	23	46	46
1/2	40	80	40
1/3	14	28	0
2/2	11	11	11
2/1	3	3	6
2/3	11	11	0
3/3	7	0	0
3/1	2	0	2
3/2	7	0	14
Всего	118	179	119

рангов этих слов, не превышающую 10 — ср. ккалп. *бул* (i-5), каз. *бұл* (i-3), узб. *бу* (i-2) «это»;

б) каракалпакское слово и его эквивалент имеют среднее расхождение в употребительности, оценивающееся разностью рангов каракалпакского слова и его эквивалента, находящейся в пределах $10 \div 50$, ср. ккалп. *және* (i-36), каз. *және* (i-5) «и, еще»;

в) каракалпакское слово и его эквивалент имеют большое расхождение в употребительности, которое оценивается разностью рангов, превышающей 51 — ср. ккалп. *және* (i-36), узб. *яна* (i-134) «и, еще».

Для оценки соотношения этих ситуаций в целом относительно выделенного нами массива наиболее употребительных слов мы применили метод попарного сравнения рангов (порядковых номеров) в частотном словаре каракалпакской и казахской, а также каракалпакской и узбекской лексики. Результаты этого сравнения оцениваются с помощью коэффициента корреляции Спирмена, который реализуется выражением

$$\rho = 1 - \frac{\sum d_i^2}{n(n^2 - 1)}$$

Здесь ρ — коэффициент ранговой корреляции по Спирмену, d_i — разность рангов каракалпакской и соответствующей ей казахской (или узбекской) лексической единицы, n — количество учтенных в эксперименте единиц в обоих сравниваемых частотных словниках. При этом в обоих словниках происходит компрессия рангов таким образом, чтобы в каждом из словников последовательность рангов лексических единиц представляла бы последовательность натуральных целых чисел от 1 по n .

Метод ранговой корреляции применяется нами в двух лингвистических аспектах:

во-первых, для выявления содержательной близости каракалпакского публицистического текста и аналогичных текстов казахского и узбекского языков;

во-вторых, для определения материальной близости этих текстов.

В первом случае сравниваются ранги эквивалентных по значению каракалпакских и казахских, а также каракалпакских и узбекских лексических единиц независимо от того, выражены ли они в сопоставляемых языках одинаковыми словами или словами, восходящими к общему тюркскому корню, или они передаются формами совершенно различного происхождения.

Во втором случае учитываются только те слова, которые имеют совпадающие по форме эквиваленты (ср. ккалп. *совет*, каз., узб.

совет «советский») или эквиваленты, восходящие к одному тюркскому корню (ср. ккалп. *ушын*, каз. *үшін*, узб. *учун* «ради, для»).

Отметим, что и в том и другом случае учитываются только те слова, которые имеют в сопоставляемых языках точно определенные рамки в интервалах $i = 1 \div 300$.

Численные значения коэффициентов ранговой корреляции относительно пар каракалпакский — казахский и каракалпакский — узбекский языки показаны в табл. 3. Для сравнения представлены меры семантического и формального схождения наиболее частой лексики казахского и узбекского языков.

Таблица 3. Значения коэффициента ранговой корреляции ρ наиболее частых словоформ в публицистических текстах каракалпакского, казахского и узбекского языков

Семнотические уровни (планы)	Значения ρ для пар языков		
	Каракалпакский — казахский	Каракалпакский — узбекский	Казахский — узбекский
План содержания (семантика)	0,62	0,57	0,49
План выражения (форма)	0,62	0,52	0,45

Сопоставление значений величин ρ , характеризующих речевые схождения каракалпакской лексики с казахской и узбекской лексикой, с одной стороны, с языковыми соотношениями лексических систем указанных языков, с другой, обнаруживает несколько неожиданные результаты. Если расхождение лексической системы каракалпакского языка примерно в полтора раза ниже, чем расхождение каракалпакской и узбекской лексической систем, то на уровне речи коэффициенты ранговой корреляции, характеризующие близость семантического и формального распределения лексики в указанных парах языков, дают незначительные расхождения, особенно на семантическом уровне. Близость каракалпакской и казахской лексики на семантическом уровне текста только на 8% выше схождения каракалпакской и узбекской лексики. На формальном же уровне каракалпакская лексика ближе к казахской на 16%, чем к узбекской.

Наблюдение над рангами лексических единиц и их употребительностью в публицистических текстах рассматриваемых языков показывает (табл. 3), что наибольшее схождение дают:

1) служебные слова древнетюркского происхождения — ккалп. *менен* ($i=4$), каз. *мен* ($i=4$), узб. *билан* ($i=3$) «и», ккалп.

ушын ($i=10$), каз. *үшін* ($i=23$), узб. *учун* ($i=11$) послелог «для, ради»;

2) местоимения — ккалп. *ол* ($i=3$), каз. *Ол* ($i=2$), узб. *у* ($i=6$) «он», ккалп. *бул* ($i=5$), каз. *бұл* ($i=3$), узб. *бу* ($i=2$) «этот»;

3) вспомогательно-знаменательные глаголы — ккалп. *бол* ($i=2$), каз. *бол* ($i=1$), узб. *бул* ($i=4$) «быть»;

4) существительные, обозначающие основные реалии труда и жизни народа — ккалп. *жыл* ($i=6$), каз. *жыл* ($i=17$), узб. *йил* ($i=10$) «год», ккалп. *мийнет* ($i=11$), каз. *еңбек* ($i=41$), узб. *меҳнат* ($i=42$) «труд», ккалп. *халық* ($i=15$), каз. *халық* ($i=18$), узб. *халқ* ($i=14$) «народ»;

5) интернациональные общественно-политические термины — ккалп. *партия* ($i=21$), каз. *партия* ($i=13$), узб. *партия* ($i=16$), ккалп. *процент* ($i=95$), каз. *процент* ($i=184$), узб. *процент* ($i=102$), ккалп. *съезд* ($i=104$), каз. *съезд* ($i=83$), узб. *съезд* ($i=$

Значительные расхождения обнаруживаются в обозначениях, специфических для той или иной республики реалиях. Ср., например, калп. *Қарақалпақстан* ($i=43$) «Каракалпакия» при отсутствии этого слова в казахском и узбекском ЧС, ккалп. *автономиялы* ($i=76$), узб. *автоном* ($i>300$) «автономная (республика)», в казахском ЧС отсутствует, ккалп. *пахта* ($i=30$), каз. *мақта* ($i>300$), узб. *пахта* ($i=168$) «хлопок».

Итак, несмотря на то, что лексика каракалпакского языка генетически близка к лексике казахского языка (оба языка входят в одну подгруппу) и достаточно далека от лексики далекородственного узбекского языка, на уровне речи не только каракалпакская и казахская, но и каракалпакская и узбекская лексика обнаруживают значительное сходжение. Эта близость выше на семантическом уровне, чем на формальном.

Чем же объяснить, что генетически далекие каракалпакский и узбекский языки показывают на уровне речи примерно такое же высокое сходжение употребительной лексики, что и близкородственные каракалпакский и казахский языки? Причину следует искать в интенсивных народнохозяйственных и культурных контактах каракалпакского и узбекского народов. Чем интенсивнее и длительнее становятся эти контакты, тем сильнее становится общность и употребительность тех понятий (точнее, значений лексических единиц), которые они отражают. Этот процесс отчетливее всего прослеживается в публицистических текстах, отражающих народнохозяйственную и культурную жизнь республики. Динамические речевые лексические процессы не сразу воздействуют на лексическую систему языка, которая благодаря своей устойчивости и консервативности все еще отражает древние генетические соотно-

шения каракалпакского языка с его соседями — близкородственным казахским и более далеким по родству узбекским языками.

ЛИТЕРАТУРА

1. *Ахабаев А.* Статистический анализ лексико-морфологической структуры языка казахской публицистики (на материале газетных текстов за 1965—1966 гг.): Автореф. дис. ... канд. филол. наук. Алма-Ата, 1971. С. 33.
2. *Он же.* Алфавитно-частотный словарь казахской газетной лексики. Статистика казахского текста. Вып. 3. Алма-Ата, 1973. С. 344—464.
3. *Мырзабеков С.* Статистико-лингвистический анализ структуры глагола современного казахского языка (по роману М. О. Ауэзова «Абай жолы»): Автореф. дис. ... канд. филол. наук. Алма-Ата, 1973. С. 32.
4. *Джубанов А. Х.* Квантитативная структура казахского текста (опыт лингвистического анализа на ЭВМ). Алма-Ата, 1987. С. 147.
5. *Бектаев К. Б.* Статистико-информационная типология тюркского текста. Алма-Ата, 1978. С. 183.
6. *Пиотровский Р. Г., Бектаев К. Б., Пиотровский А. А.* Математическая лингвистика: учебное пособие для пединститутов. М., 1977. С. 383.
7. *Бабанаров А.* Разработка принципов построения словарного обеспечения турецко-русского машинного перевода: Автореф. дис. ... канд. филол. наук. Л., 1971. С. 17.
8. *Мухамедов С. А.* Алфавитно-частотный словарь узбекского языка. Ташкент, 1982. С. 110.
9. *Мухамедов С. А., Пиотровский Р. Г.* Инженерная лингвистика и опыт системно-статистического исследования узбекских текстов. Ташкент, 1986. С. 160.
10. *Садыков Т.* Проблемы моделирования тюркской морфологии. Фрунзе, 1987. С. 120.
11. *Айымбетов М.* Опыт лингво-статистического анализа лексики и морфологии каракалпакского публицистического текста: Дис. ... канд. филол. наук. Ташкент, 1986.
12. *Молдабеков К.* Лингво-статистические исследования казахских текстов для младших школьников (на материале текстов учебников начальных классов и литературы для детей): Автореф. дис. ... канд. филол. наук. Алма-Ата, 1985. С. 18.
13. Статистика казахского текста. Вып. 3. Алма-Ата, 1973. С. 732.
14. *Бектарев К. Б., Джубанов А. Х., Мырзабеков С. М., Белбогитов А. Б.* Частотный словарь по роману М. О. Ауэзова «Путь Абая» (М. О. Әуезовтің «Абай жолы» романының жиілік сөздігі). Алма-Ата, 1979. С. 336.
15. *Баскаков Н. А.* Алтайская семья языков и ее изучение. М., 1981. С. 135.
16. *Ахабаев А.* Алфавитно-частотный словарь казахской газетной лексики // Статистика казахского текста. Вып. 3. Алма-Ата, 1973. С. 334—464.
17. *Мухамедов С. А.* Алфавитно-частотный словарь узбекского языка. Ташкент, 1982. С. 110.

ЛЕКСИКОГРАФИЧЕСКОЕ ОФОРМЛЕНИЕ СУФФИКСАЛЬНЫХ СУЩЕСТВИТЕЛЬНЫХ В РУССКО-КАЗАХСКОМ СЛОВАРЕ

В данной работе анализируется словообразование суффиксальных существительных, мотивированных существительными со значением «носитель предметного признака» в сравнительном анализе продуктивности казахских аффиксов при их передаче. Сущность словообразовательной модификации заключается в добавлении к основному значению мотивирующего слова дополнительного смыслового значения.

Суффиксальные существительные, мотивированные существительными со значением «носитель предметного признака», исходя из «Русской грамматики», образованы с помощью 45 суффиксов (без небольших единичных групп). Данные существительные переданы на казахский язык с помощью 39 казахских аффиксов, а всего на материале «Русско-казахского словаря» выявлено свыше трех тысяч слов с данным признаком. При их анализе выделены наиболее и наименее продуктивные суффиксы. Эти данные отражены в выводах:

Кол-во	Продукт. суфф.	Кол-во	Непродукт. суфф.
517	-ник	2	-ад(а)
252	-щик	2	-над(а)
204	-нст	2	-ол
209	-нзм	3	-нич(ий)
197	-и/у/	1	-уш
306	-ство	1	-ш(а)

Существительные с наиболее продуктивным суффиксом -ник называют лицо, животное, растение, вместилище, неодушевленные предметы. Наибольшее количество существительных данного типа со значением «лица», например: фокусник (фокус), ключник (ключ).

Существительные с суффиксом -щик называют лицо, характеризующееся отношением к предмету, явлению, названному мотивирующим словом, например: табунщик (табун), трамвайщик (трамвай) и т. д.

Существительные с суффиксами -ник, -щик переданы на казахский язык наиболее продуктивным казахским аффиксом -шы/-ші, который в сочетании с именными основами образует следующие значения: наименование лица по роду его деятельности, по профес-

сии, по специальности, например: сапожник — етікші (от етік «сапог»). Соотношение передачи казахским средством — собственным эквивалентом с казахским аффиксом -шы/-ші русских суффиксальных существительных со значением «носитель предметного признака»:

Кол-во	Продуктивные русские суф.	Кол-во	Непродуктивные русские суффиксы
111	-ник	2	-ач
153	-щик	1	-ак/-як
45	-ист	1	-ант
18	-атор	1	-иц(а)
17	-ец	1	-изм
12	-ер	1	-ик(а)
11	-арь	1	-ушк(а)
6	-ик	1	-ш(а)

Как известно из выводов, наиболее продуктивным способом передачи существительных с суффиксами -ник, -щик на казахский язык является собственный эквивалент с казахским аффиксом -шы/-ші, например: охотник (охота) — аншы, взломщик (взлом) — каракшы.

Существительные с суффиксами -ник, -щик переданы на казахский язык при помощи таких казахских аффиксов, из которых одни являются продуктивными, а другие непродуктивными (табл. 1).

Таблица 1. Соотношение продуктивных и непродуктивных казахских аффиксов для передачи существительных с суффиксами -ник, -щик

Продуктивные казахские аффиксы	Русские суффиксы (кол-во)		Непродуктивные казахские аффиксы	Русские суффиксы (кол-во)	
	-ник	-щик		-ник	-щик
-шы/ші	111	153	-ша/-ше	4	2
-лық	14	—	-шық/-шақ	2	2
-қар	17	3	-папа/-хана	2	—
-лас	18	1	-қы/-ғы	2	1
-ки/-ғи	21	1	-ыш	2	—
-қыш/-ғыш/-қаш	10	2	-ша, -шак, -тай,	1	1
			-ке, -ш, -жан		
-у/-ық, -у/-ік	15	—	-сымақ	2	1

Как видно из выводов, наиболее продуктивным способом передачи существительных с суффиксами -ник, -щик со значением «носитель предметного признака» является собственный эквивалент с казахским аффиксом -шы/-ші, а наименее — аффиксы отглагольного об-

разования, сложные аффиксы, аффиксы экспрессивной оценки. Для перевода существительных с суффиксами -ник, -щик с вышеуказанным значением не использованы казахские аффиксы -лак/-дақ/-так, -ек, -к, -мақ, -ма, -ым, -қа, -ын, -да, -дай, -шылдық, -ластық, -герлік, -сыздық, -ғыштық.

Всего для передачи суффиксальных существительных, мотивированных существительных со значением «носитель предметного признака», использованы казахские аффиксы, количественное соотношение частоты употребления которых выглядит следующим образом:

Кол-во	Наибольшее	Кол-во	Наименьшее
381	-шы/-ші	1	-лак/-дақ/-так
178	-лық/-дық/-тық	1	-ек, -к
68	-ау(-лау, -дау, -тау)	2	-мақ/-бақ/-пақ
30	-лас/-дас/-тас	1	-ын,(-ін, -н)

Продуктивный тип существительных с суффиксом -ство передан на казахский язык в основном собственным эквивалентом с казахским аффиксом -лық (-лік, -дық/-дік, -тық/-тік), например: фразёрство (фразёр) — мылжындық, франтовство (франт) — кербездік, сәнқойлық. Наиболее продуктивным способом передачи существительных данного типа является собственный эквивалент с казахским аффиксом -лық:

Кол-во	Продуктив. рус. суфф.	Кол-во	Непродукт. рус. суфф.
118	-ство	1	-ич
21	-ец	4	-ин(а)
14	-ник	4	-ин/-у/
8	-изм	1	-ак
7	-ан/-ин	1	-ар

Наряду с вышеуказанными продуктивными существительными выделяется группа также продуктивных существительных с суффиксами -ист, -изм, -и/-у/, которые отличаются способом передачи их на казахский язык. Они переведены, в отличие от существительных с суффиксами -ник, -щик, -ство, путём прямого заимствования, и степень передачи их собственными эквивалентами с казахскими аффиксами составляет небольшой процент (табл. 2).

Непродуктивные суффиксы, такие как -ат, -ит, -ин, -ад(а), -иад, -ак, -смен, -мейстер, -оз, -ол, -орий, -ов(ой), -нич(ий), -ей/-ачей, -анк, -ух, -уш, -ушк(а), -ш(а), переданы на казахский язык в основном путём прямого заимствования, например: почтмейстер, аэрарий, бензол, городничий, полисмен, абордаж и т. д. Единичны примеры использования казахских аффиксов для перевода существи-

тельных с вышеуказанными суффиксами, например: гнилушка (гниль) — шіріген, сокольный (сокол) — құсбегі. Употребление казахских аффиксов при переводе существительных с непродуктивными суффиксами незначительно:

Казахские аффиксы	Русские суффиксы	Кол-во
-қы/-ғы, -лық, -и/-у,	-ат, -нт, -аж, -ов(ой),	1
-ық	-ух, -ушк(а)	
-кан, -ыр, -ап	-анк	1

Для русского, как и для казахского языка, аффиксальный способ словообразования (при помощи словообразующих аффиксов или суффиксов) является наиболее продуктивным. Продуктивные словообразующие элементы, основы и аффиксы являются весьма устойчивыми в языке. Вступая в соединение с разными основами, словообразующие аффиксы (суффиксы) образуют слова с общим значением данной основы, внося при оформлении слова к употреблению и свое дополнительное значение: лица, предмета, действия, состояния, абстрактного понятия и т. п. Для имени существительного, помимо других способов словообразования, наиболее продуктивным, характерным является в русском и казахском языках морфологический, аффиксальный. Именно его спецификой определяется структура имени, существительного как части речи.

Таблица 2. Степень частотности употребления казахских аффиксов для перевода существительных с указанными суффиксами

Казахские аффиксы	Русские суффиксы (кол-во)		
	-ист	-изм	-и/-и/
-шы/-ші	45	1	—
-лық	—	8	4
-лас	4	1	—
-қар	1	1	—
-қи/-ғи	1	—	—
-қыш	3	—	—

М. О. ӘУЕЗОВТИҢ «ИНДИЯ ОЧЕРКТЕРІНДЕГІ» ЖАЛҚЫ ЕСІМДЕР СТАТИСТИКАСЫ

Бұл күнде советтік тіл білімінің қарқындап дамуымен байланысты қазақ тіл біліміндегі ономастика да ғылым ретінде қалыптасып, оның зерттеу аясы кеңейіп отыр [1—6]. Сондықтан да оның әр түрлі салалары (топонимика, антропонимика, этнонимика, астронимика т. б.) тіл білімінің кең көлемдегі зерттеу объектілеріне айналууда. Солардың бірі — көркем әдебиеттердегі жалқы есімдердің қолданылу аясы жайлы.

Әрине, Абай Құнанбаев, Жамбыл Жабаев, С. Сейфуллин, Б. Майлин, І. Жансүгіров, М. Әуезов, С. Мұқанов, Ғ. Мүсірепов, Ғ. Мұстафин, Қ. Аманжолов, Т. Жароков, А. Тоқмағамбетов, Х. Есенжанов, І. Есенберлин т. б. көрнекті ақын-жазушыларымыздың шығармаларындағы жалқы есімдерді мұқият жинап, оларға тарихи-лингвистикалық, әдеби-стилистикалық [7, 65] және лингвостатистикалық тұрғыдан зерттеудің мәні айрықша зор екендігі белгілі.

Осы ретте біз зерттеу объектісі етіп М. О. Әуезовтің «Индия очерктерін» [8] алдық.

Бұл кітаптың текісі негізінде кибернетикалық машина — ЭВМ-нің көмегімен алынған «Индия очерктерінің» алфавиттік жиілік сөздігі бізге очерктегі қолданылған жалқы есімдерге лингвистикалық талдау жасауымызға мүмкіндік беріп отыр.

Осы күнге дейін ақын-жазушылар шығармаларынан болсын, немесе жалпы қазақ тіл білімінде болсын жалқы есімдер жайы лингвостатистика тұрғысынан арнайы сөз етілген жоқ болатын. Тек кейбір мақала-зерттеулерде ғана қаралынып отырған тектегі жалқы есімдердің саны туралы қысқаша хабарлар кездесетін [7, 66-б.; 9, 159-б.; 10, 12-б.]. Сондықтан да бұл жұмысты лингвостатистика саласында орындалып отырған алғашқы эксперименттік бақылау-барлау жұмысы деп айтса да болатындай. Рас, бұл салада мұндай мазмұндағы зерттеулер басқа тілдердің материалдарынан бізден бұрын да сөз болып жүргендігін айтқанымыз жөп [11—13].

Жалқы есімдер мәселесін лингвостатистикалық тұрғыдан қарау үшін М. О. Әуезов шығармаларының ішінен тек «Индия очерктерін» ғана таңдап алуымыз да жай кездейсоқтық емес. Себебі, жазушы бұл шығармасын жазған кезде Үнді—Совет достығы қоғамының Қазақстандық бөлімінің председателі еді. Осы себепті де ол ежелгі шығыс елдерінің бірінен саналатын Индияда (Үнді-

станда) екі рет (1955, 1960) болып қайтты. Ұлттық көне мәдениеті, дәстүрі, әдет-ғұрпы өзгеше ғажайып бұл елде және жол жөнекей басқа да елдерде көрген-білген жайлар ұлы жазушыға қатты әсер етті. Сөйтіп, сондағы алғашқы сапарының нәтижесі ретінде «Индия очерктері» жарық көрді.

Очерктерде ұлы жазушының Индия сапарында көрген-білген, аңғарып түйген жайлары бүкпесіз баян етілді. Әрине, бұл баяндаудың барысында шығармада казакша есім атаулармен қатар бөтен жердегі көптеген кісі аттары, елді мекен, жер-су атаулары, жекелей мекеме, қоғам не ғимарат-жайлар, құрылыстар атаулары т. б. жалқы есімдер ауызға алынып, очерктің өн бойында қолданыс табады. Сол кезде ол жалқы есімдер әртүрлі сөзформаларда, әр түрлі жиіліктерде қолданылғаны анық. Бірақ, мұндағы (очерктердегі) қолданылған жалқы есімдердің саны қанша, олар қандай жиіліктерде қолданылған, қандай топтардан тұрады, тектегі қолданылу формалары қалай, олар тіліміздің қандай грамматикалық қосымшаларымен жұмсалынады деген сияқты жайлар нақты сөз етуді керек ететіні анық. Мақаланың баяндар мақсаты да — сол.

М. О. Әуезовтің «Индия очерктері» 38004 сөзқолданыстан тұрады екен. Егер, ондағы бүкіл сөзқолданыстағы сөздерді сөз таптарына бөліп қарағанда, оның 40,3 проценті зат есімдер дейтін болсақ тектегі басқа сөз таптарының сандық үлестері мынадай: етістік — 22,8%, сын есім — 12,7%, есімдік — 8,7%, шылау — 6,0%, үстеу — 4,3%, сан есім — 2,6%, модаль сөздер — 2,3%, еліктеуіш сөздер — 0,042%, одағай — 0,010%, жалқы есімдердің бұдан алатын үлесі — 4,14% (1573 сөз қолданыс). Яғни, тектегі кейбір сөз таптарына (сан есім, модаль, еліктеуіш, одағай) қарағанда тек жалқы есімдердің өзі «Индия очерктерінде» едәуір.

Жоғарыда атап айтқанымыздай, ЭВМ-нің көмегімен алынған «Индия очерктерінің» алфавитті жиілік сөздігінен ондағы жалқы есімдер бөлек теріліп алынып, оның негізінде:

1. Очерктердегі жалқы есімдердің жиілігі көрсетіліп қысқаша түсіндірме сөздік жасалынды (қараңыз, сөздік жұмыстың соңында беріледі).

2. Осы сөздіктің негізінде: а) мұндағы жалқы есімдердің топтары ажыратылды; ә) лексикалық формаларына статоморфологиялық талдау жасалынды.

«Индия очерктеріндегі» жалқы есімдердің түсіндірме сөздігі

Сөздікте 327 атау (жалқы есімдер) бар. Ол атаулар тектешінде әр түрлі формаларда — 528 сөзформада 1573 рет қайталанып қолданылған.

Егер, сөздіктегі жалқы есімдерді өздеріне сай белгілі бір топтарға бөлсек, олардың саны мен кездесу жиілігі былайша көрінген болар еді (1-кестеге қараныз).

1-кесте «Индия очерктеріндегі» жалқы есімдер

Жалқы есімдер топтары	Атаудың саны	Кездесу жиілігі
Кісі аттары	92	256
Құдай, пайғамбар, әулие аттары	8	93
Мемлекет аттары	27	436
Қаланың аттары	53	426
Қыстақ аттары	2	14
Өзен, плотина, канал, дария аттары	10	23
Теңіз, көл, мұхит аттары	7	15
Тау, шың, биік аттары	11	18
Шөл аттары	2	4
Қойнау, мүйіс, уәлаят, арал, шат, аймақ, тоғай, аудан аттары	25	124
Ағаш, бақ, парк аттары	9	12
Мектеп, мешіт, университет	13	22
Ескерткіш аттары	11	23
Мұнара аттары	4	9
Көрме аты	1	3
Қонақ үй аттары	4	10
Аурухана аты	1	1
Аэродром аты	1	1
Компания, қоғам аттары	9	19
Қақпа аты	1	1
Сыйлық аты	1	1
Заман (дәуір) аты	1	1
Шығарма аттары	11	22
Храм, шіркеу, собор, сарай	23	39
Барлығы	327	1573

Мұнда не жок, бәрі бар: кісі аттары да, қала-қыстақ, мемлекет аттары да, өзен-су, тау-шөл аттары да, орман-тоғай, жергілікті жер атаулары да, оқу орындары, көрме, қонақ үй, қоғам аттары да, тарихи ескерткіштер, мұнаралар, храм, шіркеу, собор, сарайлар атаулары да, тіпті аурухана, аэродром, қақпа және әр саладағы шығармалар аттары да бар. Бұлардың ішінде ең көбі де, ең жиі қолданылатындары да — мемлекет, қала, кісі аттары. Бұл үшеуінің сөздіктегі саны — 172 де, қолданылу (кездесу) жиілігі — 1118. Яғни, текстегі бүкіл жалқы есімдердің 71 проценті.

Бұл кестеден тағы бір байқайтынымыз, мұндағы жалқы есім атауларының әр түрлілігі, сан тақырыптылығы. Мұнда, біздіңше,

сөз жоқ ұлы қаламгердің асқан ойшылдығы, жан-жақты білімділігі, шебер әңгімешілігі және жалқы есімдерді қолданылудағы қас талғампаздығының жатқаны даусыз.

«Индия очерктеріндегі» жалқы есімдердің қолданылуы

Тексте жалқы есімдер түбір сөзформада да, қосымшалар (жалғау-жұрнақтар) қабылдаған формада да қолданылған. Олардың сандық көрінісі мынадай (2-кестеге қараңыз).

2-кесте. «Индия очерктеріндегі» жалқы есімдердің түбір сөз және қосымшалармен жұмсалынуының статистикасы

Жалқы есімдер	Сөзформа		Сөзқолданыс	
	саны	%	жиілігі	%
Түбір сөз күйінде	278	52,65	893	55,77
Қосымшалармен:	250	47,35	680	43,23
Жалғаулармен	221	41,86	600	38,14
Жұрнақтармен	29	5,49	80	5,09
Барлығы	528	100,00	1573	100,00

Мұнда (очерктерде) жалқы есімдер 278 түбір формада 893 рет қайталанып қолданылған. Яғни, текстегі бүкіл жалқы есімдердің 56,77%. Ал, қосымшалармен қолданылған жалқы есімдер 250 формада 680 рет қайталанып жұмсалған (43,23%). Яғни, жартысына жуығы.

Жалқы есімдердің текст ішінде қолданылғанда әр түрлі грамматикалық формаларды қабылдау қасиеті қай кезде де басқа да текстерге тән болуы мүмкін. Өйткені, біздің қолымыздағы тілдік материал жалқы есімдердің жалғаулармен жұмсалыуында — 38,14%, жұрнақтармен жұмсалынуында — 5,09% екендігін көрсетеді.

Ал, бұл қосымшалар қандай формаларда қалай қолданылған?

3-кесте. «Индия очерктеріндегі» жалқы есімдердің жалғаулармен жұмсалыуының құрылымдық формалары

Құрылымдық формалар	Сөзформа саны	Тексттегі кездесу жиілігі	Құрылымдық формалар	Сөзформа саны	Тексттегі кездесу жиілігі
Түбір-септік	214	582	Түбір-көптік	1	11
Түбір-тәуелдік	2	2	Түбір-көптік-септік	2	2
			Түбір-тәуелдік-септік	2	3
			Барлығы	221	600

Егер, 3-кестеге қарасак, бұдан жалқы есімдердің септелетінін, көптелетінін, тәуелденетінін, көптік жалғауын қабылдап барып септелетінін және тәуелденіп барып септелетінін байқаймыз. Бұлардың ішінде сөзформа жағынан да, текстегі кездесу жиілігі жағынан да ең көбі — жай септеулер. Олар 214 сөзформада септік жалғауымен тұрып 582 рет қайталанып қолданылған. Ал мұнымен салыстырғанда басқа формалар жоққа жуық. Демек, бұдан шығатын қорытынды: жай септіктерде қабылдап қолданылу (жұмсалыну) жалқы есімдердің қай тобына болса да тән қасиет.

М. Әуезовтің «Индия очерктеріндегі» жалқы есімдердің қандай грамматикалық формалармен қолданылғандығын көрсететін статистикалық мәліметтер 4-кестеде берілді.

4-кесте. «Индия очерктеріндегі» жалқы есімдердің қосымшалармен қолданылуы

Қосымшалар	Сөзформа саны	Текстегі кездесу жиілігі	Мысалдар	Қосымшалар	Сөзформа саны	Текстегі кездесу жиілігі	Мысалдар
-ның	40	135	Аджентаның	-дан	16	30	Алматыдан
-нің	4	8	Камориннің	-ден	7	17	Гандиден
-дың	9	19	Бабұрдың	-тан	2	2	Хайдарабадтан
-дің	14	41	Гарабаллидің	-тен	2	3	Бенарестен
-тың	9	22	Аббастың	-нан	2	3	Радживаннан
-тің	3	8	Бенарестің	-нен	5	8	Иерусалимнен
-ға	26	61	Аграға	-мен	4	4	Нерумен
-ге	11	28	Бомбейге	-пен	1	1	Сурковпен
-ка	3	7	Мадраска	-сы	1	1	Женевасы
-ке	2	9	Аурацзепке	-ың	1	1	Ғұсманың
-ды	7	7	Алатауды	-дар	1	11	Моғолдар
-ді	7	8	Горькийді	-дарға	1	1	Моғолдарға
-ты	2	2	Блокты	-дармен	1	1	Моғолдармен
-ны	8	24	Будданы	-інің	1	2	Гангінің
-да	21	85	Азияда	-сының	1	1	Индиясының
-де	6	33	Бомбейде	-дай	2	2	Алатаудай
-та	2	6	Мадраста	-лық	2	16	Индиялық
-те	2	8	Бенаресте	-тағы	1	2	Мадрастағы
				-дағы	15	37	Америкадағы
				-дегі	4	17	Бомбейдегі
				-ша	2	3	Европаша
				-ланған	1	1	Европаланған
				-ланғанда	1	1	Европаланғанда
				-лану	1	1	Европалану

Мұнда жалғаулармен қатар кейбір сөз тудыратын жұрнақтардың да бар екенін көресіз. Олар — сын есім тудыратын: *-дай, -лық, -тағы, -дағы, -дегі*; үстеу тудыратын: *-ша*, етістік сөз табын туды-

ратын: *-ланған, -ланғанда, -лану* формалары. Әрине, бұлардың саны текс көлемімен салыстырғанда өте аз. Бірақ, аз болған күнде де жалқы есімдер үшін бұл грамматикалық формалардың жаттығы жоқтай. Аз болған күнде де текст ішінде стильдік ролі бар. Өйткені, олар, біріншіден, табиғаты жағынан тілдік қорға тән тілдің жалпы заңына бағынышты; екіншіден, бұлардың бәрі де бір кездегі жалпы есімдерден жасалған [1, 4-б.].

Қорыта айтқанда, М. О. Әуезовтің «Индия очерктеріндегі» жалқы есімдер материалдары (беріліп отырған жалқы есімдер сөздігін айтып отырмыз) тілді зерттеушілерге, оның ономастика саласымен айналысушыларына зор бай материал бола алары даусыз. Ондағы жалқы есімдерді құрамына қарай талдау, шығу төркінін зерттеу т. б. жайлар бұл мақаланың міндетіне кірген жоқ. Сондықтан ол болашақтың ісі. Әрине, М. О. Әуезовтің «Индия очерктеріндегі» ономастиканы жан-жақты, терең зерттеудің тіл тарихы мен тарихи этнография үшін де зор мәнінің бары анық.

ӘДЕБИЕТТЕР

1. *Жанұзақов Т.* Қазақ есімдерінің тарихы. Алматы, 1971. 218 б.
2. *Әбдірахманов А.* Топонимика және этимология. Алматы, 1975. 208 б.
3. *Жанұзақов Т.* Очерк казахской ономастики. Алма-Ата, 1982. 176 стр.
4. *Тюркская ономастика.* Алма-Ата, 1984. 248 стр.
5. *Қойшыбаев Е.* Қазақстанның жер-су аттары сөздігі. Алматы, 1985. 205 б.
6. *Қазақ ономастикасының мәселелері.* Алматы, 1986. 136 б.
7. *Еспаева К. М.* Әуезовтің «Абай жолы» эпопеясындағы топонимдердің қолданылу аясы // Қазақ ономастикасының мәселелері. Алматы, 1986. 200 б.
8. *Әуезов М. О.* Индия очерктері // Шығармалар, он екі томдық. Екінші том. Алматы, 1967. 2-т. 344—473-б.
9. *Жанпейісов Е. М.* Әуезов «Абай жолы» эпопеясының тілі. Алматы, 1976. 136 б.
10. *Бектаев К. Б., Жұбанов А. Қ., Мырзабеков С., Белботаев А. Б.* М. Әуезовтің «Абай жолы» романының жиілік сөздігі. Алматы, 1979.
11. *Алексеев П. М., Белоусова Е. А., Талицкая Е. Е.* Частотный список собственных имен // Словарь автобиографической терминологии М. Горького. Имена собственные. Л., 1975. С. 102—104.
12. *Алексеев П. М., Белоусова Е. А., Талицкая Е. Е.* Алфавитно-частотный список слов трилогии, вошедших в настоящий выпуск словаря // Там же. С. 97—102.
13. *Якубайтис Т. А., Призуле В. А.* Латышские имена собственные в статистическом аспекте // Известия Академии наук Латвийской ССР. 1982. № 3. (416). С. 53—59.

ЛЕКСИЧЕСКАЯ СВЯЗЬ СЛОВАРЕЙ ПЕРВОЙ И ВТОРОЙ ЧАСТЕЙ РОМАНА «КАН» БАШКИРСКОГО ПИСАТЕЛЯ Д. ЮЛТЫЯ

Лексика художественной литературы базируется на лексике национального языка. Каждый автор, выбирая ему необходимые слова из этого национального языка, посредством их строит словесную ткань произведения: разворачивает события, характеризует героев, создает эмоциональную окрашенность произведения.

В научной литературе существует мнение, что языковое строение художественного произведения обуславливается идейным, тематическим и композиционным строением самого произведения [1]. В романе «Кан» эти три элемента в обеих частях близки друг другу. Роман в целом имеет единую идею и тематику. Композиционное строение первой части близко композиционному строению второй. Исходя из этих предпосылок, нами рассматривается проблема соотношения первой и второй части романа лексически. Выбор объекта исследования определяется тем, что роман «Кан» является уникальным в башкирской литературе. В начале 30-х гг. Д. Юлтыем была опубликована на башкирском и русском языках первая часть романа. В это же время автор подготовил и вторую часть. Однако со смертью автора рукопись второй части романа была утеряна. В 60-е гг. ее нашли, и на башкирский язык ее перевел Ф. Исенгулов. Поэтому представляет определенный интерес рассмотрение лексики обеих частей романа, выявление насколько они близки лексически.

Для анализа лексической связи обеих частей романа нами были получены частотные словари лексем по обеим частям в целом, частотные словари авторской речи по выборкам в 20 тыс. словоупотреблений из обеих частей романа. На основе этих словарей были получены словари слов, встречающихся только в первой и только во второй части, и соответствующих выборках авторской речи, и словари слов, общих для обеих частей и выборок. Далее производился анализ этих словарей. Существуют разные методы сравнения частотных словарей, с целью установления связи между ними. Одним из часто применяемых методов является метод корреляционного анализа рангов слов по сравниваемым частотным словарям [2—4]. Применяется также критерий χ^2 . Ю. А. Тулдавой была предложена своя методика сравнения частотных словарей. Нами для анализа выбрана методика последнего автора.

Получение словарей и все расчеты производились на ЭВМ.

Анализ проводился как по выборкам из авторской речи, так и по всем словоупотреблениям обеих частей в целом. Ниже результаты исследования рассматриваются в отдельности.

Связь словарей-выборок из авторской речи

Анализ словарей-выборок из авторской речи по обеим частям романа дал следующие результаты: в словаре лексем из выборки по первой части число лексем равно 2 622, в словаре выборки из второй части — 3 434 лексемы. Количество слов, встречающихся только в первой выборке, — 1 137. Количество слов, встречающихся только в словаре второй выборки, — 1 949. Слова, общие для словарей обеих выборок, — 1 485 лексем.

Ю. А. Тулдавой была предложена формула для индекса лексической связи словарей двух текстов:

$$R = \frac{C}{V_1 + V_2 - C},$$

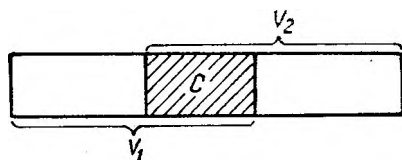


Рис. 1. Соотношение лексем двух текстов

где C — количество общих для словарей слов,

V_1 и V_2 — объемы словарей текстов.

Графически C , V_1 , V_2 изображены на рис. 1.

Индекс R выражает отношение общей части словарей к объему объединенного словаря и принимает значение от нуля до единицы. Когда словари двух текстов не пересекаются, то $R=0$, когда же словари полностью совпадают, то $R=1$.

Вычисление дает результат $R=0,325$.

Доверительные границы для R при доверительной вероятности 0,955 определяются выражением $R \pm 2\delta_R$, где $\delta = \sqrt{\frac{R(1-R)}{V_{1,2}}}$, $V_{1,2}$ — объем объединенного словаря. Формула для δ_R взята из работы Ю. А. Тулдавы [[5]].

Окончательное значение для R вместе с доверительными границами имеет вид $0,325 \pm 0,014$. Это означает, что значение R лежит в пределах от 0,311 до 0,339, т. е. общая часть словарей выборок из авторской речи по обеим частям романа «Кан» составляет 31,1% и 33,9% от объединенного словаря.

Вычисление коэффициента корреляции рангов слов общих обоим словарям выборок по формуле: $\rho = 1 - \frac{6 \sum d_i^2}{n^3 - n}$, где d — разность между рангами слов в выборках, n — число сравниваемых

мых пар слов, дает значение $\rho=0,33$. ρ изменяется от -1 до $+1$. При $\rho=+1$ говорят о полной положительной корреляции, при $\rho=-1$ говорят об отрицательной корреляции, когда же $\rho=0$, то корреляция полностью отсутствует. Вычисленное значение коэффициента корреляции рангов показывает, что между рангами слов, общим обоим словарям выборок, существует положительная корреляция. Это означает, что 33% слов из общих слов являются одинаково высоко-, средне- и низкочастотными в сравниваемых словарях.

Таким образом, можно считать, что Ф. Исенгулов передал в авторском повествовании 1/3 слов из авторского повествования Д. Юлтыя. Причем корреляция рангов этих общих слов в частотном словаре также равна 1/3.

Однако ни индекс лексической связи, ни коэффициент корреляции не дают ответа на вопрос: значимо ли различие двух выборок на уровне словарей статистически, или незначимо. Ю. А. Тулдава предложил метод объединенного словаря с гипотезой об однородности лексики двух текстов и распределении при этом подчастот в объединенном тексте по биномиальному закону [6]. Аналогичные вычисления дают $\chi^2=516,4$, что значительно превышает $\chi_{\text{теор}}^2$ на уровне значимости 0,001 и числе степени $k=1$. Следовательно, есть основание считать, что гипотеза об однородности распределения подчастот не выполняется. То есть выборки из авторских словоупотреблений на уровне словаря статистически разнородны: различие текстов первой и второй части романа «Кан» на уровне частотных словарей авторской речи существенно. В этом проявляется различие стилей авторов, поскольку в выборе слова писателем так же участвует и его индивидуальный стиль [7].

2. Связь полных словарей первой и второй части романа

Анализ словарей по первой и второй части с учетом всей лексики дает следующие результаты: слова, относящиеся только к словарю первой части, — 2010 слов, слова, относящиеся только к словарю второй части романа, — 2494 слов, слова, общие обоим частям — 2436 слов. Индекс лексической связи словарей двух частей $R=0,351$, что несколько больше индекса лексической связи по авторским словоупотреблениям. В словарь общих слов входят все высокочастотные слова в каждом из словарей. Так, из словаря по первой части вошли полностью первые 194 наиболее частотных слова, из словаря по второй части — 86 наименее частотных слов: ул (1309, 821)* —

* В скобках даны частоты по частотным словарям из первой и второй части романа.

«он», без (1298, 470) — «мы», мин (1169, 752) — «я», менэн (735, 443) — «с, вместе», шул (568, 191) — «тот», үз (423, 313) — «свой», бул (471, 344) — «всп. глагол будь», кил (454, 214) — «приходи», һүз (300, 179) — «слово», сық (321, 217) — «выходи», был (300, 354) — «этот», ла (355, 280) — частица, һалдат (313, 205) — «солдат», һугыш (209, 129) — «война», бик (278, 200) — частица, һәм (139, 223) — союз «и», генә (241, 203) — частица и другие [8].

В словарь общих слов вошли также слова и из средних и низкочастотных зон частотных словарей: команда (38, 16) — «команда», командир (32, 24) — «командир», донья (32, 41) — «мир, свет», дошман (22, 34) — «враг», дивизия (18, 15) — «дивизия», ақтар (14, 5) — «выворотить», дауам (13, 20) — «продолжение», тәүзә (7, 9) — «вначале», ишара (6, 4) — «жест», билдә (2, 4) — «метка», йәки (21, 5) — союз «либо» и др.

Из 1 835 одночастотных слов частотного словаря по первой части в объединенный словарь (общий словарь) вошли 1 263 слова, которые составляют 69% одноразовых слов Д. Юлтыя. То есть во второй части в переводе Ф. Исенгулова отражено 2/3 одноразовых слов Д. Юлтыя из первой части, слов, считающихся в лингвостатистике «авторским богатством». Так, во второй части встречаются такие одноразовые слова из первой части: афәт (1, 1) — «беда», батырсылық (1, 3) — «смелость», дүңгәләк (1, 2) — «кольцо», ғәйрәт (1, 1) — «сила», ип (1, 1) — «аккуратность», мөгәмәлә (1, 9) — «обращение», мөхит (1, 3) — «океан», самала (1, 6) — «прикинь», таты (1, 8) — «отведай», ки (1, 1) — частица, вә (1, 1) — союз «и» и другие.

Представляет интерес выяснение, по каким лексико-грамматическим группам происходит максимальное сближение лексики словарей двух частей. С этой целью нами найдены индексы лексической связи для знаменательных частей речи. Данные представлены в выводах:

	V_1	V_2	V_{06}	R
Существ.	1787	1693	903	0.350
Глаголы	1198	1490	739	0.378
Наречия	398	395	165	0.306
Прилагат.	749	825	340	0.276

Как видно из выводов, наибольшее значение индекса лексической связи наблюдается у глаголов, затем у существительных. Наименьшее значение индекса лексической связи у прилагательных. Таким образом, сближение лексики двух частей романа «Кан» — наилучшее среди глаголов.

Впечатляющим является рассмотрение покрываемости текстов обеих частей романа словарем общей части, на основе накопленных частот. Слова общей части словарей покрывают 92% текста первой части и 89% текста второй части. Оба автора как бы составляют словесную ткань произведений из одних и тех же слов. В этом, нам кажется, проявляется стремление Ф. Исенгулова отразить словоиспользование Д. Юлтыя, его стиль на основе анализа первой части романа. Однако, как стиль самого Ф. Исенгулова, так и динамика литературного языка, более чем 30 лет не позволяют достичь полной идентификации стиля Д. Юлтыя. Ф. Исенгулов в своем переводе второй части романа с русского на башкирский язык сумел передать 1/3 словаря Д. Юлтыя. Это не столь много, однако эти слова покрывают более 2/3 текста самого перевода. Ф. Исенгулов сумел передать наиболее частотные, опорные слова из языка Д. Юлтыя. Кроме того, он передал более 1/2 одноразовых слов Д. Юлтыя. Все это приближает язык второй части романа к языку первой части, показывает, что Ф. Исенгулов работал над языком второй части, чтобы передать лексическое своеобразие языка Д. Юлтыя. Однако, как было сказано выше, полной идентификации стиля Ф. Исенгулова и стиля Д. Юлтыя не происходит. Так, исследование авторской речи показывает, что различие в использовании лексики при учете частот статистически значимо.

Анализ словарей обеих частей показывает, что Д. Юлтый и Ф. Исенгулов по-своему подходят к выбору слова. Так, во второй части Ф. Исенгуловым использованы слова: **көллө** — «все», **һәм-мә** — «все», **хақында** — послелог, **шикелле** — послелог, **ауаз** — «звук», **хәтирә** — «память», **әжәл** — «смерть», **ахмак** — «глупый», **шәйлә** — «заметь», **дөмбәслә** — «колоти», **яуапла** — «отвечай», **йәнә** — союз «и», **мәртәбә** — «раз», **ғәзиз** — «дорогой, милый» и другие, которые отсутствуют в первой части. Вместо них Д. Юлтый использовал синонимичные слова: **бары** — «все», **бар** — «все», **тураһында** — послелог, **һымак** — послелог, **истәлек** — «память», **үлем** — «смерть», **байқа** — «заметь», **биргелә** — «колоти», **яуап бир** — «отвечай», **тағы** — союз «и», **рәт** — «раз», **изге** — «дорогой» и другие.

Немалую роль в различении языков двух частей играет процесс развития языка. Так, Д. Юлтыем использованы слова, которые в современном башкирском языке активно не используются, отошли в ряд архаизмов и историзмов: **мәхлүк** — «тварь», **ғәнимәт** — «трофей», **ихтирағсы** — «доброволец», **мөтдәт** — «время», **шәрех** — «комментарий» и другие.

Сложную задачу представляет сохранение вкраплений из русского языка, которые в обилии в речи персонажей у Д. Юлтыя.

В речи тех же персонажей в переводе Ф. Исенгулова их значительно меньше.

Однако, несмотря на эти различия, Ф. Исенгулов использовал ядерную лексику первой части романа. Этот пласт весьма активен во второй части и составляет основу словесной ткани этой части романа «Кан». Тем самым сближается язык второй части с языком первой части романа.

ЛИТЕРАТУРА

1. Бондарева О. Н., Васильева Р. С. О методике сравнения частотных словарей и автоматизации лингвистических работ // Тезисы докладов и сообщений. Л., 18—20 октября 1966 г., С. 21.
2. Гвоздев А. Н. Очерки по стилистике русского языка. М., 1952. С. 335.
3. Ефимов А. Н. Стилистика художественной речи. М., 1957. С. 448.
4. Калинина Е. А. Изучение лексико-статистических закономерностей на основе вероятностного моделирования // Статистика речи. Л., 1968. С. 64—107.
5. Тулдава Ю. А. Об измерении лексической связи текстов на уровне словаря // Вопросы статистической стилистики. Киев, 1974. С. 35—42.
6. Фрумкина Р. М. Статистические методы изучения лексики. М., 1964. С. 115.
7. Dolezel L. O style moderní ceské prosy. Praha, 1960.

С. А. УСКОМБАЕВ

СТАТИСТИЧЕСКОЕ ИССЛЕДОВАНИЕ МНОГОЗНАЧНОСТИ СЛОВ В КАЗАХСКОМ ЯЗЫКЕ

В этой статье проводится сплошной анализ десятичного толкового словаря казахского языка, чтобы определить меры многозначности казахских слов, долю конверсивной омонимии в общей картине многозначности, многозначных слов в каждой части речи, вычислить, сколько в словаре имеется слов с одним, двумя, тремя и более значениями. Надо также определить соотношение количества слов и количества выражаемых ими значений. Такого рода сведения необходимы для решения прикладных задач, например, построения алгоритмов перевода с казахского языка на информационный язык.

Получение таких данных требует выполнения весьма трудоемкой работы по сплошному перебору всего словаря языка и фиксации количества и характера значений каждого отдельного слова.

Изучение многозначности слов ограничивается в нашей работе лишь изолированными словами. Многозначность словосочетаний не рассматривается.

В нашем исследовании регистрировалась многозначность изолированных слов, принадлежащих к следующим частям речи: существительным (З), глаголам (Е), прилагательным (С), и наречиям (У). Всего в десятичном толковом словаре насчитывается 42 709 словарных статей слов, относящихся к этим частям речи, из них 94,3% — слова, имеющие одно значение, и 5,7% многозначных слов.

Из 36 748 слов с одним значением 11 754 (31,9%) — это существительные, 15 370 (41,8%) — глаголы, 8 113 (22,3%) — прилагательные и 1 511 (4%) — наречия.

Из 5 696 словарных статей многозначных слов в 1 086 случаях наблюдается конверсивная омонимия, а в остальных 4 610 словарных статьях и прилагательных однозначные слова примерно равны по встречаемости (табл. 1).

Таблица 1. Встречаемость однозначных и многозначных слов

Часть речи	Однозначные слова		Многозначные слова		Соотношение числа однозначных и многозначных слов
	кол-во	%	кол-во	%	
Существительное	11754	31,9	2208	42	1:0,18
Глагол	15370	41,8	1747	33,26	1:0,11
Прилагательное	8113	22,3	1135	21,7	1:0,13
Наречие	1511	4	162	3,04	1:0,10
Итого	36748	100	5252	100	1:0,14

Надо сказать, что к лексико-грамматической (или консервативной) омонимии относятся все те случаи, когда одно и то же слово в ряде своих значений относится к одной части речи, а в ряде других своих значений — к иной части речи.

Сплошной анализ толкового словаря казахского языка показал, что лексико-грамматическая омонимия встречается в 754 словарных статьях, что составляет 2,23% всего словарного состава.

Количественное соотношение разных типов лексико-грамматической омонимии показано в табл. 2.

Наиболее часто встречаются омонимичные формы существительных и глаголов — 49,8%, омонимия типа «существительное-прилагательное» — 24,6%, «существительное-наречие» — 6,6%,

«прилагательное-наречие» — 3,9, «прилагательное-местоимение» — 0,13%. Таким образом, в сумме два основных типа омонимии — ЗЕ и ЗС составляют 74,64% всех случаев омонимии, а типы конверсивной омонимии ЗУ, ЕУ, СУ составляют в сумме только 13,8%.

Эти закономерности сказываются и на встречаемости более сложных типов омонимии. Омонимия типа ЗЕС встречается чаще других случаев «тройной» омонимии — 3,5%. Все типы омонимии, включающие у (ЗСУ, ЗУЕ) встречаются в сумме в 1% случаев.

Таблица 2. Количественное соотношение разных типов лексико-грамматической омонимии в казахском языке

ТИП ОМОНИМИИ	Общее кол-во случаев	%
Существительное — глагол (ЗЕ)	376	49,8
Существительное — прилагательное (ЗС)	186	24,6
Существительное — наречие (ЗУ)	49	6,5
Существительное — союз (ЗШ)	2	0,26
Существительное — междометие (ЗО)	8	1,1
Существительное — местоимение (ЗД)	2	0,26
Глагол — наречие (ЕУ)	26	3,4
Глагол — междометие (ЕО)	4	0,5
Глагол — союз (ЕШ)	4	0,5
Прилагательное — междометие (СО)	2	0,26
Прилагательное — наречие (СУ)	30	0,9
Прилагательное — союз (СШ)	2	0,26
Прилагательное — местоимение (СД)	1	0,13
Местоимение — глагол (ДЕ)	1	0,13
Существительное — прилагательное — наречие	4	0,5
Существительное — глагол — междометие (ЗЕО)	5	0,7
Существительное — наречие — глагол (ЗУЕ)	4	0,5
Существительное — числительное — глагол (ЗАЕ)	2	0,26
Существительное — прилагательное — глагол (ЗСЕ)	27	3,5
Прилагательное — местоимение — наречие (ЗДУ)	1	0,13
Прилагательное — междометие — глагол (СОЕ)	1	0,13
Глагол — наречие — междометие — (ЕУО)	1	0,13
Существительное — прилагательное — глагол — подражательные слова (ЗЛел)	1	0,13
Существительное — глагол — наречие — союз (ЗЕУШ)	1	0,13
Существительное — глагол — междометие — союз (ЗЕОШ)	1	0,13
Существительное — местоимение — наречие — глагол (ЗДУЕ)	1	0,13
Прилагательное — существительное — союз — глагол (СЗШЕ)	1	0,13
Прилагательное — наречие — существительное — союз (СУЗШ)	1	0,13

Омонимия типа ЗЕСУ не встречается.

Рассмотрим теперь, каковы закономерности количественного

соотношения слов, характеризующиеся различной степенью полисемичности. С увеличением количества значений уменьшается количество полисемичных слов с данным количеством значений. Слов, имеющих три значения, примерно втрое меньше, чем слов, имеющих два значения, а последних в восемь раз меньше, чем слов с одним значением. Эти закономерности отражены графически на рис. 1. На рисунках 2, 3, 4 и 5 показаны соотношения между количеством значений и числом слов, принадлежащих к данной части речи, отдельно для существительных, глаголов, прилагательных и наречий. В них просматриваются общие черты:

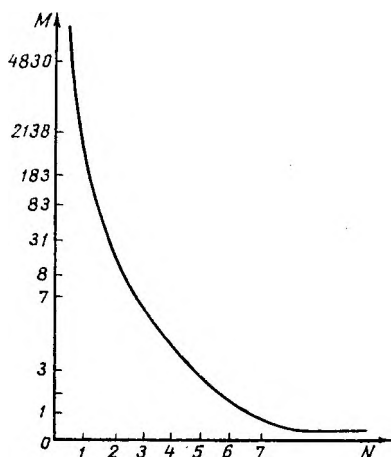


Рис. 1. Зависимость частоты многозначных слов (M) от количества их значений (N)

1 до 5 значений пропорция уменьшения числа разных слов для существительных, глаголов и прилагательных примерно одинакова 0,5—1.

3. В пределах 10—15 значений число многозначных существительных, глаголов и прилагательных достигает минимума.

Анализ графика обнаруживает и некоторые различия между темпом уменьшения числа слов, принадлежащих к разным частям речи, с увеличением количества значений.

Наречия отличаются тем, что пропорции уменьшения числа слов с увеличением количества значений неравномерны для них и равны в интервале 1—7 значений 1316 : 116 : 15 : 5 : 3 : 2 : 1.

У существительных пропорция уменьшения числа слов в интервале 1—5 значений монотонно убывает и каждый раз уменьшается

1. Кривые, выражающие функциональную связь между количеством значений и числом слов, обладающих данным количеством значений, имеют один максимум, соответствующий количеству значений, равному 1. Это означает, что разных слов, имеющих одно значение, всегда больше, чем разных слов, имеющих два, три и более значений.

2. Во всех случаях с увеличением количества значений уменьшается число слов с данным количеством значений. В пределах от

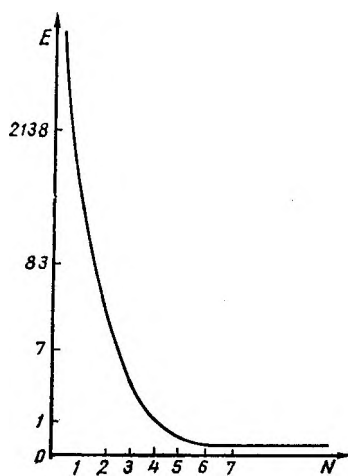
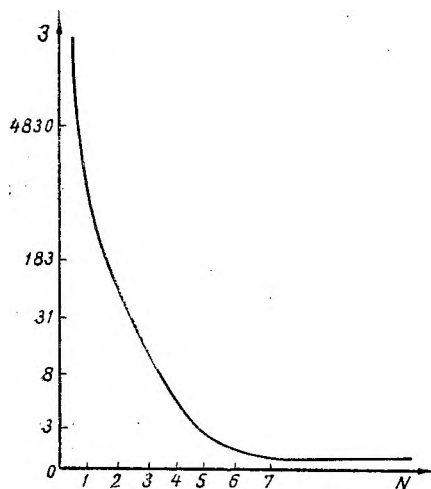


Рис. 2. Зависимость многозначных существительных от количества их значений

Рис. 3. Зависимость частоты многозначных глаголов (E) от количества их значений

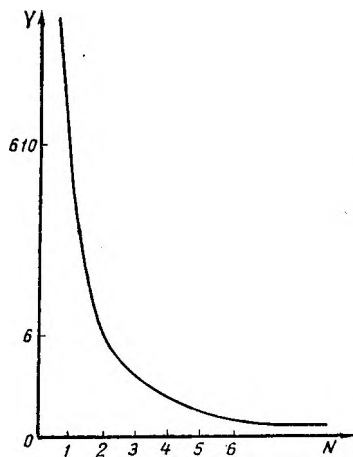
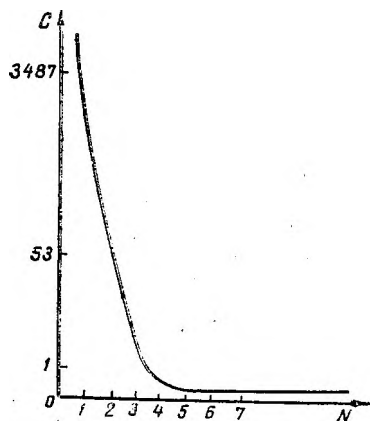


Рис. 4. Зависимость частоты многозначных прилагательных (C) от количества их значений

Рис. 5. Зависимость частоты многозначных наречий (V) от количества их значений

примерно в 8, 3,4, 2,3, 2,2 раза, у глаголов в 14,5, 2,4, 3, 2,5, а у прилагательных в 9,4, 4,6, 3,1, 3,8 раза.

Конкретные величины, характеризующие функциональную связь между количеством значений и числом слов с данным количеством значений, приведены в табл. 3.

Рассматривая связь между количеством значений и числом случаев лексико-грамматической омонимии, можно заметить, что с увеличением количества значений абсолютное число лексико-грамматической омонимии в пределах 2—9 значений монотонно уменьшается в два раза. Наибольшее количество лексико-грамма-

Таблица 3. Соотношение между количеством значений и числом слов с данным количеством значений

Количество значений	Существительные	Глаголы	Прилагательные	Наречия	Всего	Случай лексико-грамматической омонимии	Всего, включая случаи лексико-грамматической омонимии
1	11754	15370	8133	1511	36748	—	—
2	1477	1048	855	132	3512	445	3967
3	423	431	182	17	1053	227	1280
4	180	141	58	6	385	114	499
5	80	56	14	3	154	67	221
6	23	39	11	2	75	41	116
7	9	21	7	1	38	21	59
8	8	7	2	—	17	7	24
9	4	6	4	1	15	12	27
10	3	9	1	—	13	5	18
11	—	1	—	—	1	5	6
12	—	2	—	—	2	2	4
13	—	3	—	—	3	4	7
14	—	1	—	—	1	4	5
15	—	—	—	—	—	1	1
16	—	1	—	—	1	1	2
17	1	—	—	—	1	—	1
18	—	1	—	—	1	3	4
19	—	—	—	—	—	—	—
20	—	—	—	—	—	—	—
21	—	—	—	—	—	—	—
22	—	—	—	—	—	—	—
23	—	—	—	—	—	—	—
24	—	—	—	—	—	—	—
25	—	—	—	—	—	—	—
26	—	—	—	—	—	1	1
24	—	—	—	—	—	—	—
28	—	—	—	—	—	—	—
29	—	—	—	—	—	—	—
30	—	—	—	—	—	1	1

тической омонимии падает на слова, имеющие два значения. В пределах 2—30 значений абсолютное число случаев лексико-грамматической омонимии монотонно убывает.

Интересно сопоставить соотношение лексической многозначности и лексико-грамматической омонимии на одно и то же количество значений. Это соотношение при двух значениях равно 7,7 : 1, при четырех — 3,3 : 1, при пяти — 2,2 : 1, при шести — 1,3 : 1, при семи — 1,8 : 1, при восьми — 2,4 : 1, при девяти — 1,2 : 1, при десяти — 2,6 : 1, при одиннадцати — 0,2 : 1, при двенадцати — 1 : 1, при тринадцати — 0,7 : 1 и т. д.

Таким образом, в пределах от двух до одиннадцати значений число лексико-грамматических омонимов ниже числа слов с лексической многозначностью. От одиннадцати значений наблюдается обратное соотношение, когда случаи лексико-грамматической омонимии преобладают над случаем лексической многозначности. Пределом количества лексических значений многозначного слова является цифра 10. Следовательно, лексико-грамматическая омонимия, лексическая многозначность качественно доминируют в обычном для нее интервале (2—10 значений).

Теперь определим, каково соотношение количества слов и выражаемых ими значений (табл. 4). Для определения того, сколько разных значений охватывает 42 709 словарных статей словаря, надо вычислить сумму произведений числа словарных статей на число выражаемых ими значений. Эти 4 610 словарных статей распределяются по частям речи следующим образом: общее количество многозначных существительных 2 208 (42%), глаголов — 1 747 (33,26%), прилагательных — 1 135 (21,7%), наречий — 162 (3,04%).

Сопоставляя процентные соотношения одно- и многозначных слов, относящихся к разным частям речи, можно заметить, что среди однозначных (соответственно 31,9% и 42%) доля глаголов с одним значением больше, чем доля глаголов с несколькими значениями (41,8 и 33,26%).

Сопоставительные данные о соотношении количества однозначных и многозначных слов в разных частях речи приведены в выводах. Существительные и глаголы как однозначные, так и многозначные, занимают доминирующее положение в общем массиве словарных статей, содержащихся в толковом словаре казахского языка. Их суммарное количество среди однозначных слов составляет 73,7%, а среди многозначных слов — 75,26%, прилагательные и наречия значительно уступают по количеству существительным и прилагательным. Их суммарное количество составляет 26,3% однозначных слов и 25,1% многозначных слов.

Таблица 4. Соотношение количества слов и количества выражаемых ими значений

Кол-во значений	Существительные	Значения существ.	Глаголы	Значения глаголов	Прилагательные	Значения прилагат.	Наречия	Значения наречий
1	11754	11754	15370	15370	8113	8113	1511	1511
2	1477	2954	1048	2096	855	1710	132	264
3	423	1259	431	1293	182	546	17	51
4	180	720	141	564	58	232	6	24
5	80	400	56	280	15	75	3	15
6	23	138	39	234	11	66	2	12
7	9	63	21	147	7	49	1	7
8	8	64	7	56	2	16	—	—
9	4	36	6	54	4	36	2	9
10	3	30	9	90	1	10	—	—
11	—	1	11	—	—	—	—	—
12	—	—	2	24	—	—	—	—
13	—	—	3	39	—	—	—	—
14	—	—	1	14	—	—	—	—
15	—	—	—	—	—	—	—	—
16	—	—	1	16	—	—	—	—
17	1	17	—	—	—	—	—	—
18	—	—	1	18	—	—	—	—
19	—	—	—	—	—	—	—	—
20	—	—	—	—	—	—	—	—
21	—	—	—	—	—	—	—	—
22	—	—	—	—	—	—	—	—

Показательные цифры, указывающие в выводах на соотношение числа однозначных и многозначных слов. В целом в рассматриваемом нами словаре однозначные и многозначные слова представлены в одинаковой степени (соотношение 1 : 0,14). Однако для равных частей речи соотношение числа однозначных и многозначных слов различно. Существительные отличаются некоторым перевесом однозначных слов над многозначными (соответственно 1 : 0,19). У прилагательных наблюдается обратное соотношение — 1 : 0,13. Среди глаголов также преобладают однозначные слова. На каждый многозначный глагол приходится около десяти однозначных (соотношение 1 : 0,11). Наречия в основном однозначны. На каждые десять однозначных наречий приходится одно многозначное (соотношение 1 : 0,1).

Суммарное количество значений, выражаемых всеми существительными, равно 17 446, глаголами — 20 295, прилагательными — 10 853 и наречиями 1 893. В сумме эти четыре цифры составляют 50 487 (все значения, выражаемые словарными статьями четырех частей речи).

Если не учитывать случаи лексико-грамматической омонимии, то количество значений, выражаемых существительными, будет равно 13 962, глаголов — 17 147, прилагательных — 9 248 и наречий — 1 703. Общая сумма значений без учета лексико-грамматической омонимии будет равна 42 060.

Назовем отношение общего числа значений к общему числу словарных статей «покрывающей способностью» словарных статей. Покрывающая способность словарных статей десятичного словаря казахского языка равна 2. Эта цифра вычислена путем деления суммы значений, выражаемых словами, — 87 235 (из них 50 487, выражаемых многозначными словами и 36 748, выражаемых однозначными словами) на общее число словарных статей — 42 709.

Величина покрывающей способности, вычисленная таким образом, показывает, что в среднем на одно слово, приведенное в словаре, приходится два значения.

К. М. МОЛДАБЕКОВ

ЛИНГВОСТАТИСТИЧЕСКОЕ ИЗУЧЕНИЕ ЗАИМСТВОВАНИЙ ИЗ РУССКОГО ЯЗЫКА В КАЗАХСКИХ ТЕКСТАХ

Определение места русского языка как предмета начальных казахских классов среди других дисциплин и его связи с остальными предметами является одной из актуальных задач методики начальной школы. Вот почему необходимо выявить отношения русского языка с казахским в объеме школьных дисциплин. Это тем более возможно, так как последние имеют традиционную языковую связь, и вледствие этого их лексические структуры взаимно обогащаются.

Современной наукой по частной методике доказана высокая эффективность построения обучения учащихся любому другому языку на базе особенностей русского языка. Использование этой особенности и выявление объема заимствованных в казахском языке слов из русского языка и через русский язык поможет решить нам эту задачу. Для того, чтобы выявить объем заимствованных слов из русского языка и через русский язык в лексике речи казахских детей, мы составили частотные словари текстов из казах-

ской литературы (из журнала «Балдырган» объемом 25 000 словоупотреблений; газеты «Казахстан пионері» — 25 000 словоупотреблений; детской литературы — 50 000 словоупотреблений), и учебников начальных классов. Среди них «Казахский язык» (1, 2, 3 класс) — 43 800 словоупотреблений, «Природоведение» (2, 3 класс) — 31 300 словоупотреблений, «Математика» (1, 2, 3) — 62 550 словоупотреблений, а также номер газеты «Социалистік Қазақстан» («Казахстанская правда») — 10 000 словоупотреблений.

Наблюдение над частотными словарями показало, что в журнале «Балдырган», заимствованных из русского языка и через русский язык словоформ — 1 320, что составляет 5,28% текста, соответственно «Казахстан пионері» — 1 639—7,51%, детской литературы — 1 744,—1,16%, «Казахский язык» — 1 916,—4,37%, «Природоведение» — 1,598,—5,10%, «Математика» — 5 801,—9,42% и «Социалистік Қазақстан» — 2 000,—20%.

Как видно из результатов, заимствованные словоформы чаще всего встречаются в «Математике» (9,42%) и «Социалистік Қазақстан» (20%). Это объясняется тем, что, во-первых, учебник математики переведен с русского на казахский, во-вторых, почти все эти заимствованные словоформы — интернациональные термины, так как газета «Социалистік Қазақстан» знакомит с общественно-политической, социально-экономической жизнью нашей республики и страны.

Если в публикациях «Балдырган» и «Қазақстан пионері», а также в «Қазақском языке» и «Природоведении» чаще всего употреблялись такие слова как пионер, Совет, сумка, сентябрь, вожатый, солдат, отряд, класс, СССР, машина, Ленин, парта, команда, телефон, Москва и др., которые учащиеся часто употребляют в повседневном обиходе, то в газете «Социалистік Қазақстан» — слова: партия, коммунист, коммунистік, бюро, лениншіл, социалистік, экономика, миллион, интернационал, марксизм, ленинизм, НАТО и т. д.

Это показывает, что заимствованные словоформы употреблены в соответствии с общим кругозором читающих. Возникает вопрос: можно ли в сопоставительном изучении казахского и русского языков применять лингвостатистические методы? По нашему мнению, вполне можно использовать частотные словари во время занятий, так как там можно найти даже такие словоформы, частота употребления которых равна только одному. Такие заимствованные слова обычно не входят в словарный запас учащихся.

До сих пор лингвостатистика в языкознании применялась как научно-исследовательский метод. Но, несмотря на это, в казахском языкознании имели место некоторые попытки ее применения в разработке вопросов методики обучения именно в начальных клас-

сах. Например, статья Г. Бегалиева «Эріп, дыбас, буын», написанная в 1935 г. полностью опирается на цифровые данные [1]. По мнению автора, для составления разрезных букварей в первую очередь нужно выявить частоту употребления букв. Для анализа он берет тексты букваря для первого класса казахской школы. По подсчетам, в текстах, представленных в этом учебнике, встречались 506 словоформ. А частота встречаемости букв следующая: а — 17,5% б — 2,2, л — 4,4, р — 5,8% и т. д.

Далее автор анализирует и приводит статистические данные из текстов произведений Беимбета Майлина, учебников для 1—2 классов, по 1 000 словоупотреблений и выявляет частоту встречаемости букв в словах, которые состоят из 3, 4, 5, 9 букв.

На основании статистических данных Г. Бегалиев приходит к следующему:

1. Различная частота употребления слогов не обуславливается различием текстов.

2. Слоги казахского языка устойчиво претерпевают незначительные изменения.

3. В учебниках количество слогов больше, чем в художественной литературе (в первых — 2,8, во второй — 2,4).

4. Очень редко употребляются закрытые слоги со стечением согласных.

Попытка Г. Бегалиева внедрить приемы статистического анализа в практику школьного преподавания была естественной. С помощью лингвистического анализа точно определяются степени (частота) употребления каждой словоформы, слова и виды их изменения по частям речи, а также степени присоединения окончаний (падежных, притяжательных, множественных, личных).

По примеру автора статьи Г. Бегалиева в 3-б классе средней школы № 29 им. В. И. Ленина Бугунского района Чимкентской области мы провели несколько учебно-экспериментальных занятий. На каждом уроке применялись наглядные пособия, на которых были написаны и казахские, и русские слова, понятные учащимся. Обращалось внимание на правильное произношение как казахских, так и заимствованных слов.

Эксперимент проводился в виде игры-соревнования — в течение 5—6 мин. Для этого были специально подготовлены тексты из газеты «Казакстан пионері», журнала «Балдырган», учебников и детских книг, объем которых не превышал 30 словоупотреблений. Надо было подсчитать, сколько в данных текстах существительных из казахского языка, сколько заимствованных. Учащиеся быстро нашли существительные, определили, какие из них казахские, какие — заимствованные. Было организовано чтение учащихся по заранее подготовленным карточкам, в процессе которого учитель

обращал внимание на правильное произношение учащимися заимствованных слов, совместно с ними выявлял в них существительные (или другие части речи, которые необходимо бы проанализировать) и определял частоту их употребления.

В заключении этого урока учитель отметил, что в казахском языке вместе с исконно казахскими словами имеются и заимствованные слова из других языков. Учащиеся переписали в тетради предложения, а учитель проследил, правильно ли они написаны.

Такие виды работы учат детей правильно произносить русские буквы, (так как казахские дети часто путают в, с, б, ч, с, ш, ф, п), повышают грамотность. Учитель по частотным словарям легко может отобрать нужные слова (словоформы) для текста. Таким образом, отборы лексического минимума обеспечивают активное усвоение учащимися программного материала и способствуют разнообразному проведению уроков, развитию навыков речи у детей.

ЛИТЕРАТУРА

Бегалиев Ф. Әріп, дыбыс, буын. Алматы, 1935.

МОРФОЛОГИЯ
МЕН СИНТАКСИС
БОЙЫНША ЗЕРТТЕУЛЕР

ИССЛЕДОВАНИЯ
ПО МОРФОЛОГИИ
И СИНТАКСИСУ

ТЕҚСТ ЛИНГВИСТИКАСЫ ЖАЙЫНДА

Қазіргі кезде әр түрлі тілдерді жан-жақты зерттеудің нәтижесінде өмір талаптарына сәйкес бірнеше тілдік жүйелер іс жүзінде пайдаланылып жүр. Солардың басты-бастыларына мына төмендегілер жатады:

1. Бір тілден екінші тілге аудару жұмысын автоматтандыру жүйесі.

2. Жазба мәліметтерінен автоматты түрде аннотация және рефераттарды (мәліметтердің қысқаша мазмұны) ықшамдап алу жүйесі.

3. Хабарлама (информация) жұмыстарын автоматтандыру жүйесі (қажетті хабарды автоматты түрде іздеу және оны сақтау мәселелері) [1, 5-б.].

Осы аталған тілдік жүйелерді пайдаланудың ойдағыдай шешілуі — сол табиғи тілдің функционалды моделінің бірден-бір дұрыс құрастырылуына тығыз байланысты [2—4].

Бұл жүйелер қай тілде іске асырылса да, олар өздері іштей белгілі бір принципке бағынады. Олар атап айтқанда сөздердің мағынасын ажырату арқылы сөйлемнің мазмұнын ашу, одан соң сөйлемдер жиынтығының мағыналары негізінде тексттің мазмұнын айқындауға саяды. Әрине, бұдан сөйлем мазмұны сөздер мағыналарының қосындысы деп, ал текст мазмұны сөйлемдер мазмұнының жиынтығы деген бірбеткей тұжырым тұмауы керек. Біз текстердің мазмұнын автоматты түрде ашу өте күрделі және арнайы зерттеулерді қажет ететіндігін ғана атап айтпақпыз.

Текстер жайлы бірер сөз. Көптеген ғылым салаларындағы кейбір ұғымдардың тыянақты анықтамасын дәлме-дәл берудің қиындығы сияқты, «текст» ұғымына да тұрақты бір анықтама берілмей келеді. Мұның дәлелі ретінде осыдан он-он бес жылдар бұрын бұл ұғымның ешбір анықтамасыз-ақ қабылданып келгенін айтсақ та

жеткілікті тәрізді [1, 9-б.]. Қазірдің өзінде тілші-мамандардың «текст» ұғымын сипаттайтын анықтамалары әр қилы. Мысалы, «текст дегеніміз — «ұзын сөйлем», «текст — ол бір-бірімен мағыналық қатынаста болатын, кемінде екі сөйлемнен тұратын, сөйлемдер жиынтығы», — деген сияқты анықтамалар беріліп жүр. Басқа да, күрделіленіп, неше түрлі информациялық, логикалық және лингвистикалық ұғымдарды пайдаланып айтып жүрген анықтамаларды да кездестіруге болады. Мысалы, семиотикалық тұрғыдан қарастырғанда «текст — тұтас бір мағыналы кез келген белгілер құрылымы» (знаковая структура), немесе «текст — семантикалық жағынан ұйымдасқан белгілер жүйесі» — деген анықтамалармен түсіндірілген [5, 154-б.]. Ал, логика-лингвистикалық тұрғыдан құрылған анықтамаға тоқтасақ — «текстің» басты белгісі ретінде сызықтық тізбе құрайтын сөйлемдерден тұратындығы, сол жақтық және оң жақтық шекарасы болатындығы, одан соң салыстырмалы толықтық пен байланыстық шарттарының орындалуы тәрізді ерекшеліктер негізге алынады [6, 172-б.].

Ғалымдардың «текст» ұғымы бойынша келтіріп жүрген әр түрлі анықтамаларын қорыта келе, текстің мынадай негізгі қасиеттерін атап кетуге болады. Ол — текстің бір мазмұнды болуы және оның единицаларының өзара грамматикалық, семантикалық әрі логикалық байланысқа түсуі.

Текстерді — байланысты не байланыссыз деп ажыратудың негіздеріне сүйене отыра оның сыр-сипатын ашуға арналған анализ-зерттеулер соңғы жылдарда синтез-зерттеулерге қарағанда көбірек жүргізіліп жүргенін аңғаруға болады.

Текстердің құрылу заңдылықтарын анықтау, әсіресе олардың ішінен ең негізгі деген единицасын айыра білу — өмірдегі оқыту процесінде тәжірибелік мәні зор. Дәлірек айтқанда, оқушылардың аз уақыттың ішінде көлемдірек текстердің мағынасын дұрыс меңгере білу, жаңа программаға сай оқулықтармен оқу құралдарын жазу үстінде аса қажет деуге болады. Сондықтан да әр түрлі текстердің ішкі заңдылықтарын егжей-тегжейіне дейін зерттеу және оларды аша түсу бүгінгі күннің алға қойып отырған талабы десек те артық болмас.

Текст құрылымын зерттеудің қажеттігі, ең әуелі оның негізгі единицаларын анықтап, айыру үшін текстерді сегменттеу ережелерін тағайындап алып, ірі-ірі текстік единицаларға топтау жағын қарастырған жөн. Зерттеуші ғалымдардың ұсыныстары бойынша осындай текст құрайтын ірі единицаларға: сөздер, сөз тіркестері, сөйлемдер, абзацтар, параграфтар, бөлімдер, тараулар, томдарды жатқызуға болады [1, 12-б.].

Өмірдің жеке-жеке объектілердің арасындағы кейбір қарым-қатынастарды айқындау барысында жалаң сөздерге қарағанда сөз

тіркестерінің ролі аса зор екендігі белгілі. Алайда сөз тіркестері әр уақытта сөйлемдер сияқты аяқталған бір ойды білдіретін хабарлама бола бермейтіндігі белгілі. Өмірдің объектілері сапалық және сандық сипаттамаларымен қосарлана отырып, белгілі бір жағдайларға қатысты объектілермен қарым-қатынаста болып отырады. Осындай ситуация құрайтын ерекшеліктер тексте атрибутты, объективті, предикативті, адвербиальды сөз тіркестері түрінде болады. Мысалы, адам баласының жас кезінен-ақ *онай және тұрақты* түрде есінде сақтайтын сөз тіркестері болады: *ақ қар, көк аспан, ана сүті, құс ұшады* т. б. көптеген сөз тіркестері [7, 27-б.]. Міне осындай сөз тіркестері сияқты текст құрылымының элементтері ретінде зат пен оның қасиетін суреттейтін әр түрлі тілдік тәсілдер, соның ішінде, көбінесе, тексте қолдану жиілігі жағынан молырақ кездесетін сөз тіркестерінің түрлері алынып жүр.

Сөз тіркестері өздерінің жасалуы жағынан болсын не құрылымы тұрғысынан болсын құбылмалы шама бола тұрса да, әр авторға қатысты өзіндік айырым қасиеттері де болмай қоймайды. Көпшілік тілдерге тән басқа да ортақ заңдылықтардың кездесе беретіні тәрізді, сөз тіркестерінің қатарындағы ең жиі кездесетіндері екі элементтен тұратын тіркестер екені де жалпы тілге тән заңдылықтар сияқты. Тіркестердің мұндай түрлері, көбінесе, зат пен оның қимыл-әрекетінің арасындағы қарым-қатынастарын айқындайды.

Тіл жүйесінде осы күнге дейін толық шешімі табылмай жүрген мәселелердің ұзын-ырғасы аз емес. Солардың бірі — сөйлемнің жалпы тіл жүйесіндегі алатын орны, екіншісі сөйлемнің семантикалық және синтаксистік жасалу принциптері. Сол сияқты, сөйлемдегі текстің негізгі семантикалық компоненттерінің қатарына жатқызу не жатқызбау мәселесі де әлі өз шешімін таппай келе жатқан жайдың бірі [1, 16-б.].

Текстегі абзац мәселесінің де тілші ғалымдар арасында әлі де көп талас тудырып жүргені мәлім, Бүгінге дейін абзацты текстің қай единицасына жатқызамыз деген сұраққа әр түрлі пікірлер айтылып келді. Оқымыстылардың кейбірі оны композициялық десе, екінші біреулері — логикалық, не семантикалық единица деген топқа қосып жүр. Енді бір зерттеушілердің айтуынша абзац — бір орынға жай ғана топталған сөйлемдердің жиынтығы емес, ол автордың өз ойын оқушының саналы түрде қабылдауына, оны мейгерудің амалдарын жеңілдетудің мақсатына лайықтап құрған сөйлеу практикасында бар текстік единица дейді.

Абзац мәселесін мазмұн тұрғысынан қарағанда көпшілікке мәлім мынадай фактіні атай кетуге болады. Әр текстің тақырыбы іштей бөлек-бөлек тақырыпшаларға, субтемаға және микротемаларға таралады [8, 27-б.]. Тексте автордың ойы шағын-шағын

үзік хабарламалар арқылы беріледі. Ал, осы хабарламалардың бірігіп қосылуынан қомақты хабар пайда болады. Осылайша біріккен хабарламалардың ең шағыны микротема деп аталады да, ал барлық текстің тақырыбын белгілі қатынастық байланыста тұратын бірнеше микротемалардың қосындысы деп есептейді.

Абзацқа енетін әр түрлі текстерден алынған сөйлемдерді байланыссыз бірнеше абзацтардан алынған сөйлемдер тізбегімен салыстыра қарастырғанда төмендегідей ерекшеліктерді байқауға болады: а) абзацқа қатынасты сөйлемдер бір микротеманың не болмаса бір-бірімен тығыз байланысты тұрған бірнеше микротеманың мазмұнын анықтайды; ә) бір абзацтың ішіндегі сөйлемдер арнаулы лексикалық амал-тәсілдер арқылы өзара байланыста болуы керек; б) абзац ішіндегі сөйлемдердің бір-бірімен белгілі синтаксистік байланыста болуы да шарт [9, 14-6].

Сөйлеудің (речь) функционалды-мағыналық түрлеріне қарай тектегі абзацтарды төмендегіше топтастыруға болады: 1) хабарлау абзацтары; 2) сипаттау абзацтары; 3) пікір тудыру абзацтары; 4) сұрақ мәнді абзацтары; 5) ниет, ойды білдіретін абзацтар; 6) аралас абзацтар [10, 20-6.]. Құрылымдық көзқарастан алып қарағанда абзацтың лексикалық тұрғыдан толығы түсуі, синтаксистік және грамматикалық жағынан түрлене түсуі сол абзацтың текстің жалпы құрылымындағы тұрған орнына да байланысты болады. Осы себептен абзацтар: кіріспе (бастапқы), ішкі және қорытынды абзацтар болып үшке саналады.

Шет ел тіл мамандарының да, отандас тілші ғалымдарымыздың да лингвистикалық зерттеулерінің жеке сөйлем көлемінен тысқары шыға қоймағандығын көбірек байқауға болады. Тек соңғы жылдардағы «текст теориясы», «текст лингвистикасы» сияқты атпен шығып жүрген ғылыми ізденістер текстің пайда болу (туу) мәселелерін сөйлем көлемімен ғана шектелмей текст шеңберінде сөз етіп жүр [11, 10-6.].

Жедел дами бастаған текст теориясы көптеген күрделі проблемаларды кеңінен қарастыруға жол ашты. Бұл проблемаларға, қысқаша айтқанда, текстің негізгі единицаларының басын ашу, саралау және олардың өзара байланысының сыр-сипаты мен шарттарын белгілеу және де текстің мазмұнын құрайтын тақырып пен теманың ролін ашу мәселелері жатады.

Жоғарыда сөз болғандай тілші психологтерден кейін көп ұзамай-ақ тілшілердің өздері де текст бойындағы жалпы тақырыптың шағын тақырыпшаларға (подтема, субподтема, микротема) ажыратылатындығын қолдай бастады [12, 25—31-6.].

Зерттеушілер текстің пайда болу (туу) процесінің алгоритмін бір арнаулы жүйелі тәртіп (процедура, диспозиция) ретінде қарастырғанды мақұлдайды. Осы жағдайда ғана сөйлем құрудың әр

түрлі жағдайлары мен жасалу шарттары ескерілуі мүмкін. Әрине, жоғарыда аталған алгоритм құрудың жүйелі тәртібін қатаң бірізді операциялар деп түсінбей, оны таңдау мүмкіншілігі мол ықтималдық ауысулардың негізінде іске асатындығын ескеру қажет [10, 121-б.].

Байланысты текстер жайлы теория, әсіресе сөйлемдердің өзара тіркесу заңдарының синтаксистік тұрғыдан баяндалған нақты ережелері көрсетілген зерттеулер әлі де болса жетіспейді. Жарық көрген еңбектерде текстің «туу» мәселесіне келгенде ғалымдар төмендегідей қорытынды жасайтынын байқауға болады. Текст, тек қана сөйлемдер мағынасымен шектелмейтін, одан да ірілеу текст бөлшектерінің белгілі реттілікті сақтайтын фразадан жоғары (сверхфраза) бөліктермен мағына түзеді [13].

Автордың белгілі бір жайды тұжырымдап айтып жеткізу қажеттігі кезіндегі сөздерді орын-орынымен, рет-ретімен таңдай білу мүмкіншілігінің сырын ашып айту жолдары да оңайға түспейді. Әдетте мамандар сөздер жиынтығын үлкен екі топқа бөледі. Олардың біріншісін мағыналық сөздер тобына жатқызса, ал екіншісін—мәнерлеуіш сөздерге жатқызады [14, 44-б.]. Бірінші топтағы мағыналық сөздердің қалайша сөйлемдердегі синтаксистік құрылысты (конструкция) толықтыра түсу жағын (наполняемость) өз алдына жеке зерттей келе, оқымыстылар ол жағдайдың арнайы логикалық программаның негізінде іске асатындығын және сөйлемде қатынасатын тіл элементтерінің жиілігі мен өзара байланыс күшінің ролін атап айтады.

Автордың текст «жасау» (тудыру) тәжірибесін семантикалық және грамматикалық тұрғыдан зерттеп, сол мақсатқа қалайша тез жетуге болады деген заңды сұрақ тууы мүмкін. Әрине «тез» деген сөз бұл жерде орынды болмауы да мүмкін, бірақ біз әр түрлі мүмкіншіліктермен салыстырғанда әрі тез, әрі дәл түсетін жолды таңдауымыз қажет. Тәжірибе нәтижелеріне қарағанда аталған мақсатқа сатылап жететініміз байқалды [15, 22—26]. Мысалы, сөйлемдегі сөзқолданыстардың атқаратын синтаксистік әрі морфологиялық қызметіне қарай ірі-ірі топтарға бөлуге болар еді. Керекті сөзді таңдап алу үшін оның сөйлемдегі атқаратын функциясын білген жөн. Мысалы, қажетті таңдайтын сөзіміздің сөйлем мүшесі ретінде атқаратын қызметі бастауыш не толықтауыш болу қажет десек, онда ол сөздерді бір зат есімге, не есімдікке жататын сөздер тобынан іздестіргеніміз дұрыс болады. Сол сияқты, кезекті сөздің сөйлемдегі атқаратын қызметі баяндауыш ролінде болса, ондай сөздерді етістік болатын сөздер тобынан, ал анықтауыш сөздерді—сын есім, есімше не болмаса ретті сан есім сөздер арасынан іздеген орынды болады.

Бәрімізге мәлім жайт—ол зат есім, етістік, сын есім және та-

ғы басқа сөз таптарына қатысты сөздердің сан жағынан алғанда өте-мөте мол кездесетіндігі. Егер де әр сөздер тобындағы сөздерді тағы да кішілеу топтарға бөлуге мүмкіндік жасайтын қосымша белгілер табылса, сөз тандау процесі тездетілген болар еді. Мысалы, орыс тілі үшін зат есімнің қосымша белгілері ретінде оны жанды және жансыз деп екі топқа бөліп алу мүмкіншілігі бар. Енді осы топтарды әрі қарайда шағын бөлшектерге бөлуге мүмкіндік беретін басқа да табиғи құбылысының белгілерін іздестіруге болатыны сөзсіз. Осындай семантикалық критерийлердің көмегімен етістік, есімше және т. б. сөздердің де іздестіру зонасын (өрісін) шағындатуға әбден болар еді. Ал, сөздерді көлемі кішірейтілген сөз топтарынан тандап алу мәселесі текстердің стильдік құрамында болатын ықтималдық теориясының мүмкіншіліктеріне негізделеді. Әрине, осындай жолмен тандалып алынған сөз басқа да факторлармен тексерілгені жөн.

Сонымен, ғалымдар адам баласының тексі құру процесінің моделіне сүйене отырып, «қолдан» автоматты түрде текст «тудыру» есебінің алгоритмін құруды сөз еткенде мына жағдайлардың беті ашылғанын жөн көреді: 1) формальды түрде текстің жалпы тақырыбының берілуі; 2) текстің негізгі семантикалық-синтаксистік единицаларының бөлініп алынуы; 3) осы единицаларды байланысты текст түріне біріктіретін ықтималды—детерминделінген ережелерінің берілуі [1, 69-б.]. Осы мақсаттың орындалуы әр түрлі құрылымдағы текстер түрлері мен типтерін арнайы тәсілдермен терең зерттеудің көптен-көп байланысты. Бұл жағдайларда, зерттеуге алынған текстердің саны қанағаттандырарлықтай болуы да шарт.

Біз осы ұсынын отырған кіріспе мақаламызда жалпы текст теориясындағы кейбір алғы шарттарды қазақ оқырмандарына арнап сөз етуді өзімізге мақсат деп білдік.

Қазіргі кездегі партия мен үкіметтің ұлттар тілінің дамуына зор мән беруі тілші ғалымдардың өз ісіне жанаша мағынамен қарауын қажет етеді. Бұлай айтуымыздың бір себебі — техникалық прогрестің жаңа жетістіктерін пайдалана отырып, зерттеу мен және оның нәтижелерін дұрыс тұтыну қажеттігі болады деп есептейміз. Міне осы айтылғандарға байланысты қазақ текстерінің өзіне тән жасалу (туу) моделін құру қазіргі заманға сай үлкен мақсаттардың бірі болмақ.

Болашақта көптеген іс қағаздарының, неше түрлі хабарлардың және де ғылыми-техникалық әдебиеттердің қазақ тілінде басылып шығуы ықтимал. Олай болса қазақ тіліндегі текстерді автоматты түрде — компьютер көмегімен қажеттілікке пайдалану мақсаты тіл зерттеу тәжірибесіндегі жаңа сала — текст теориясымен тығыз шұғылдануды алдын ала бастаған өте орынды деп санаймыз.

Қорытынды ретінде, осы аталған мақсатты іске асыру қажеттігімен қазақ тексін абзац шеңберінде семантикалық-синтаксистік тұрғыдан зерттеу тәжірибесінің бастапқы қадамдары жайлы қысқаша тоқтамақпыз.

Қазақ тексінің құрылымын зерттеуді біз үш түрлі стильді қамтитын текстердің негізінде қарастырмақпыз. Ол — публицистикалық стиль, көркем әдебиет және ғылыми-техникалық стильдерді қамтитын текстер.

Публицистикалық стиль ретінде «Қазақстан мұғалімі» газетіндегі 1986 жылдың қаңтар айы мен 1987 жылдың қаңтар айы аралығындағы мектеп реформасына байланысты: «Совет мектебі жаңа кезеңде», «Алты жасарлар мектепке барады» атты рубрикамен басылып шыққан 35 мақала қарастырылады. Мақаланың авторлары — Қазақстандағы мектеп мұғалімдері, арнаулы газет тілшілері, мұғалімдер білімін жетілдіру институттарының қызметкерлері және тілші ғалымдар. Аталған 35 мақала 503 текстік абзацтан тұрады.

Көркем әдебиеттерге жататын текстер қазақтың белгілі жазушыларының шығармаларынан алынған келесі әңгімелер:

- 1) С. Сейфуллин. Бандыны қуған Хамит, 102 абзац;
- 2) Б. Майлин. Қанды кек, 98 абзац;
- 3) М. Әуезов. Көк серек, 97 абзац;
- 4) С. Мұқанов. Әсия, 47 абзац;
- 5) Ғ. Мүсірепов. Жалон балладасы, 92 абзац;
- 6) С. Ерубәев. Бахыт, 53 абзац.

Осы аталған алты әңгімеде барлығы 500 абзацтан тұратын 20 шағын текстер кездеседі.

Ғылыми-техникалық стильге жататын текст ретінде екі монографияның ішіндегі текстер алынды [16].

Екі монография бойынша алынған текст ұзындығы — 500 абзац.

Сонымен, үш стильге қатысты текстер бойынша 1500 абзац құрылымы зерттелмек.

Тәжірибенің алғашқы қадамдары төмендегі жұмыстарды қамтиды:

1. Аталған текстердегі сөзколданыстарды тиісті сөз табына жатқызып белгілеу.

2. Сөз формаларды реестрлік сөзге келтіріп сегменттеу.

3. Әр текстегі мынадай төрт түрлі синоним сөздерді айырып белгілеу: лексикалық синонимдер, контекстік синонимдер, есімдіктік синонимдер, ассоциативті синонимдер.

4. Бөлек стиль бойынша әр абзацтың алфавитті-жиілік сөздіктерін алу.

5. Стильдер бойынша әр тақырыпшаларға қатысты сөздердің

жалпы жиілігі мен абзацтар санына тиісті жиілігі бөлініп көрсетілген — үлестірімді алфавитті-жиілік сөздікті алу.

6. Үлестірімді сөздік бойынша сөздер жиілігінің кему ретімен түзілген — жиілікті сөздік алу.

7. Негізгі сөздердің потенциалды сөздігін алу.

Электронды есептегіш машина жәрдемімен алынатын осы сөздіктердің материалдарын пайдаланудың негізгі мақсаты әр текстің, әр стилдің негізгі элементтерін, яғни текстің негізгі сөздерін (опорные слова) айырып алу қажеттілігі.

Осы этапта алынатын сөздіктердің материалдарының сапалылығы кейінгі этаптардағы тәжірибелер нәтижелерінің құндылығын арттыратыны сөзсіз.

Сонымен текстердің негізгі сипаттамаларын дұрыс анықтау әрі қарайғы абзац шеңберіндегі семантикалық және синтаксистік зерттеулердің бірден-бір бастамасы болмақ.

ӘДЕБИЕТТЕР

1. *Зубов А. В.* Проблемы порождения текста. Ч. 1. Теория и алгоритмы. Минск, 1987. С. 77.
2. *Виноград Т.* Программа, понимающая естественный язык. М., 1976.
3. *Марчук Ю. Н.* Проблемы машинного перевода. М., 1983.
4. *Невежина З. В.* Структурно-семантическая организация сверхфразового единства в современном английском языке. Киев, 1971.
5. *Russion literature — the Hague Paris, 1972. P. 2.*
6. *Hartmann P.* Zur antropologiachen Fundierung der Sprache // *Symbolae Linguisticae in honorem G. Kurilovich.* Warsszawa, 1965.
7. *Мер-Минасова С. Г.* Словоcочетание в научно-лингвистическом и дидактическом аспекте. М., 1981.
8. *Марчук Ю. Н.* Проблемы лингвистического анализа связного текста (Межфразовый уровень). Таллинн, 1975.
9. *Аграскина С. Н.* Структура абзаца в художественной прозе: (на материале романов Э. Хэмингуэя «И восходит солнце» и «Прощай оружие»). Л., 1972.
10. *Қацнелсон С. Д.* Типология языка и речевое мышление. Л., 1972.
11. *Аспекты общей и частной лингвистической теории текста.* М., 1982. С. 192.
12. *Новиков А. И.* Семантика текста и ее формализация. М., 1983.
13. *Гальперин И. Р.* Текст как объект лингвистических исследований. М., 1981.
14. *Проблемы текстуальной лингвистики.* Киев: Изд-во КГУ., 1983.
15. *Пиотровский Р. Г.* Инженерная лингвистика и теория языка. Л., 1979.
16. *Қалыбеков Т.* Экология және ашық кен. Алматы, 1988; *Көшербаев К.* Минералдар элементінде. Алматы, 1986.

СӨЗ ТАПТАРЫНЫҢ СТАТИСТИКАСЫ

10 томдық «Қазақ тілінің түсіндірме сөздігі» [ҚТТС, 1] — халықтың әлеуметтік, мәдени, рухани өміріндегі аса елеулі құбылыс, лингвистикалық зерттеулер үшін де біршама жинақталған, жүйеленген, мол қазына.

Біз бұл мақалада сөздіктің жалпы редакциясын басқарған Қазақ ССР Ғылым академиясының корреспондент-мүшесі А. Ысқақовтың: «ҚТТС қазақ тіл біліміне жаңадан талай ғылыми зерттеу жүргізуге әрі даңғыл жол ашпақшы, әрі соларға құнарлы материал алуға негіз болатыны айқын. Өйткені бұл сөздіктен (оның нақтылы материалы бойынша) ең алдымен, қазақ тілі лексикасының ая-көлемі шамаланбақшы, сөздердің тұлға-түрпаттары, олардың сала-салалары айқындалмақшы; сөздердің семантикалық сыр-сипаттары, грамматикалық өзгешеліктері, стилистикалық ерекшеліктері бағдарланбақшы; сөздердің қолданылу аясы мен жұмсалыу шеңбері анықталмақшы; сөздердің өзара тіркесу қабілеттерін ашуға мүмкіндік тумақшы, т. т.» [2, 499-б.] — деген пікірін қуаттап, осы бағытта аңғарған, анықтаған кейбір барлау, пайымдауларымызды баяндаймыз.

Сөздік «ең алдымен, қазақ тілі лексикасының ая-көлемін» нақты білуге мүмкіндік береді. Сонау алпысыншы жылдарда ағылшын тілінде 400 мыңнан астам, орыс тілінде 500 мыңға жуық сөз бар [3, 50-б.] деген сияқты жорамал пікірді ұшыратып, қазақ тілінде қанша сөз болуға мүмкін деген ой мазалайтын. Рас, содан бері де орыс тілінде әр түрлі сөздіктер дүниеге келді. Тек солардың ең көлемдісі ғана (17 томдық) жоғарыдағы жорамалдың ширегіне әрең жуықтады. Ал қазақ сөздіктеріндегі сөздердің саны отыз мыңның ортасынан (орфографиялық сөздіктің 2-басылуын айтып отырмыз) ұзап кете қойған жоқ.

«Көп томды түсіндірме сөздіктің маңызы мен мақсаты және ғылыми-теориялық негізі жайында» айтылған алғы сөзде: «...түсіндірме сөздік... қазақ тілінің әрі толық сөздігі, әрі түсіндірме сөздігі есебінде жасалатындықтан... тіліміздегі сөздерді түгел қамтуды... мақсат тұтады» [4, 3-б.] — делінген және мұның өзі бірнеше рет қайталанады. Міне, енді он томдық ҚТТС-тегі реестр сөздер мен фразеологиялық тіркестер де нақты анықталды: «сөздіктің барлық он кітабында 66 мың 931 атау сөз, 24 мың 498 фразеологиялық тіркес [2, 500-б.] қамтылған болып шықты. Бұл аз ба, көп пе? Бір қарағанда аздау болып көрінуі ықтимал. Дұрысы, бұл аз емес. Олай дейтініміз, осы уақытқа дейін қазақ тілінде қырық мыңды да

камтитын сөздік болған жоқ. Оның үстіне әр томдағы сөздің саны да орыс тілінің 17 томдық [5] сөздігіндегіден көп болмаса, аз емес.

Әр түрлі (объективті, субъективті) себептермен сөздікке енбей қалып қойған сөздердің бар екені қазірде белгілі болып отыр. Және ондай сөздердің саны санаулы емес, елеулі деуге болады. Мәселен, біздің анықтауымызша, тек а әрпін тағы да бір жарты мыңдай сөзбен толықтыруға болар еді. Әрине, мәселе тек сөздің санында ғана емес, мәнінде екені белгілі. Алайда санды да елеусіз пәрсе деп қарауға болмас.

Реестр сөздердің түсін түстеп, түрін анықтап алған соң, оларды лексика-семантикалық жақтан айқындау түсіндірме сөздіктің негізгі мән-мағынасы болса керек. Бұл тұрғыдан ҚТТС-тің қол жеткен табысы, өзіндік шабысы басы ашық айқын, дегенмен ара-тұра мүлт кетіп жататын тұстары да жоқ емес. Бұл — өз алдына жан-жақты талдауды қажет ететін күрделі мәселе.

ҚТТС-те реестр сөздердің әрқайсысының сөз табына қатысы көрсетілген. Сөздерді таптастырудың принциптері айқын болғанмен, тілдегі әрбір сөздің грамматикалық табиғатын жазбай тану іс жүзінде аса қиын екенің мойындау керек болады. Соған қарамастан сөздікті жасаушылар әр сөздің тұсына сөз табын көрсетіп отырған. Мұның өзі сөз таптарының ара салмағын анықтауға, олардың тілдегі үлесін нақты білуге мүмкіндік береді.

Статистикалық зерттеулерде жекелеген тексттердегі сөзформа, сөзқолданыстар сөз таптары тұрғысынан айқындалып, ара салмағы ажыратылып баяндалып жүр [6].

Жеке шығармада кездесетін сөздерді сөз таптарына қарай топтастыру, олардың проценттік үлес-салмағын айқындау алғаш (17 мыңдай сөзден тұратын) М. Әуезовтің «Абай жолы» романы тілін статистикалық зерттеулерде қолға алып, іске асты [7, 9-б.]. Екі томдық «Қазақ тілінің түсіндірме сөздігінде» (1959, 1961) 18 мыңнан аса сөз бар. Біз оларды сөз табы тұрғысынан анықтадық. Ең негізгі сөз таптары: зат есім, сын есім, етістік және үстеудің осы екі шығармадағы проценттік үлесі мынадай:

Сөз таптары	2 т. ТС	«Абай жолы» романы
Зат есім	46,8	39,85
Сын есім	15,77	24,32
Етістік	33,29	26,25
Үстеу	2,74	4,72

Бұдан сөз саны жағынан жуықтас сөздік пен романда сөз таптарының проценттік көрсеткіші бір-біріне сәйкес келмейтіндігін

аңғару қиын емес. Стыльдің түріне қарай сөз таптары да әр түрлі қолданылатындығы айтпаса да түсінікті.

ҚТТС — тіліміздегі сөздерді мүмкіндігінше толық қамтуды мақсат еткен бірден-бір құнды еңбек. Сөз таптарының мұндағы үлес-салмағы олардың тілдегі объективті қалпын барынша толық таныта алады деп ойлаймыз. Төмендегі кестеден сөз таптарының әр томдағы және тұтас сөздіктегі санын және олардың проценттік үлесін көруге болады.

1 - кесте. ҚТТС-тегі сөз таптарының статистикасы

Сөз табы	1	2	3	4	5	6	7	8	9	10	барлығы	%
Зат есім	1857	2372	1695	2075	1540	1477	2302	2142	1578	1933	18971	28,33
Сын есім	1139	1570	1374	1292	1264	1043	1453	1244	1079	1035	12493	18,66
Етістік	1582	2504	3326	3025	2101	3250	3341	3774	4335	3268	30506	45,58
Үстеу	454	323	328	292	190	160	232	133	170	152	2434	3,64
Сан есім	6	13	26	12	—	1	12	11	5	4	90	0,14
Есімдік	30	19	14	12	18	13	29	8	7	1	151	0,23
Шылау	18	14	25	6	2	7	15	15	14	5	121	0,18
Еліктеу	2	18	33	9	4	27	9	23	15	31	171	0,26
Одағай	52	34	34	6	6	11	33	4	16	19	215	0,32
Әр түрлі	170	265	188	155	120	192	115	298	111	158	1774	2,66
Барлығы	5310	7133	7043	6885	5245	6186	7541	7652	7330	6606	66931	100,0

Қазақ тілінде тоғыз сөз табы бар. Кестеде оныншы болып әр түрлі деген сөз тұр. Мұны былай түсіну керек. Сөздікте кейбір реестр сөздер тіркес түрінде берілген. Мысалы, *Ай-түй*: Ай-түйге қарамады. *Айыз*: Айызы қанды. *Арс*: Арс етті. *Арса*: арса болды т. т. Еліктеу сөздер т. б. көбіне осылай берілген. Сөз табы көрсетілмей қалған сөздер де кездеседі. Қырық шақты сөздің тұсында модаль деген белгі бар. Сондай-ақ көмекші сөз деп көрсетілген үш сөз ұшырайды. Міне, реестрден орын алған осындай сөздерді топтастырып, «әр түрлі» (сөздер) деуге мәжбүр болдық.

Кестеден нені аңғаруға болады? Лексикамыздың сөз таптарына қатысы нақты процентпен көрсетілген. Мұнда зат есім, сын есім, етістіктердің тілдің негізін құрайтыны, олардың үлес-салмағының аса елеулі екені айқын сезіледі.

Әдетте, сөздік құрамның дені зат есім сөздер делінеді. Кестеден байқалатыны, ҚТТС-тегі реестр сөздердің жартысына жуығы етістіктер болып тұр. Дұрысы, бұларды етістік формалар деп қараған жөн болар. Өйткені сөздікте грамматикалық көрсеткіш ретінде қимыл атауы, етістіктердің т. б. түр-түрі алынған. Біз бұларды біріктіріп етістік түрінде алып отырымыз. Реестр сөздердің грам-

матикалық табиғатын дұрыс тану үшін мұны (етістікті) кесте түрінде арнайы көрсеткен дұрыс болар.

Сонда тұсында етістік белгісі қойылған сөздердің саны 8387. Бұл сөздіктегі реестр сөздің 12,51 проценті болады да, етістіктің үлесі азая түседі. Оның үстіне осы етістіктердің қатарында 500-ге жуық *-ыңқыра*, *-іңкіре* формалы сөздер де жүр. Бұл жұрнақтың жалғану мүмкіндігі жоғары екенін ескерсек, оны алу мәселесі ойлануды қажет етсе керек.

2 - кесте. Етістік және оның формалары

Етістік	1	2	3	4	5	6	7	8	9	10	Барлығы	%
Етістік	668	837	691	853	867	908	818	901	947	897	8387	27,49
Қим. ат.	424	904	1806	1474	618	1566	1657	1744	2212	1470	13865	45,45
Өзгел.	258	434	476	399	347	455	493	563	633	566	4588	15,36
Ортақ	79	95	132	114	100	146	129	221	179	150	1345	4,40
Ырықс.	113	145	155	170	120	137	189	286	258	147	1721	5,64
Өздік	41	49	64	26	50	38	55	60	77	49	500	1,63
Барлығы	1583	2495	3327	3023	2102	3251	3342	3775	4335	3239	30506	99,97

ҚТТС-тегі сөз таптарының үлес-салмағын анықтау үстінде бізді реестр сөздердің лексика-грамматикалық, лексика-семантикалық табиғаты, кейбір сөздердің әдеби тілге қатысы көп ойландырды. Мәселен, қимыл атауы формасының реестрден орын алуы қаншалықты заңды. Сөздіктің «түсіндірме» деген атына сай келмейтін, мағынасы ашылмаған қимыл атауы, етіс формаларының саны 22 мыңнан асып, барлық сөздің 33 процентіне жуықтайды. Мұның өзі — ҚТТС-тің үш томы тек осы формаларға арналған цитаталардан тұрады деген сөз. Сондай-ақ басқа сөз таптарының да (сын есім, сан есім т. б.) грамматикалық формалары ұшырайды.

Реестр сөздердің санын көбейтуде сөйлеу тілі (сөйл.), диалект (диал.), жергілікті (жерг.), т. б. сөздердің де үлесі бар. ҚТТС-те сөйлеу тіліне тән 2158, диалект деген белгімен 595 сөз берілген. Санға бар, әдеби тіл дейтін сапаға кереғар бұл сөздер де жарты том болады. Сонда ҚТТС-те лексикалық мағынасы түсіндірілмеген, түсіндірілсе де әдеби тіл талабына сәйкес келе бермейтін сөздердің саны 25 мыңнан асып түседі екен. Мұның өзі тіліміздің лексикалық байлығы, реестр сөздердің жай-күйі туралы әлі де қатты ойлануды қажет етеді.

Қысқасы, біз сөз таптарының ҚТТС-тегі үлес-салмағын қысқа қайырып, кесте түрінде көрсетумен шектелдік, оны терең талдау — алдағы жұмыс.

1. Қазақ тілінің түсіндірме сөздігі. 10 томдық. Алматы, 1974—1986.
2. Қазақ тілінің түсіндірме сөздігі. 10-т. Алматы, 1986.
3. Жуков Дм. Переводчик, историк, поэт. М., 1965.
4. Қазақ тілінің түсіндірме сөздігі. Алматы, I-т. 1974.
5. *Словарь современного русского литературного языка*. М., Т. I—17. 1948—1965.
6. Қазақ тексінің статистикасы. Алматы, 1973.
7. Бектаев К. Б., Жұбанов А. К., Мырзабеков С., Белботаев А. Б. М. Әуезовтің «Абай жолы» романының жиілік сөздігі. Алматы, 1979.

К. Б. БЕКТАЕВ, А. Р. ЗЕКЕНОВА

М. О. ӘУЕЗОВТІҢ ДРАМАЛЫҚ ШЫҒАРМАЛАРЫНДАҒЫ ЗАТ ЕСІМДЕРДІҢ МОРФОЛОГИЯЛЫҚ ЕРЕКШЕЛІКТЕРІ

Қазақ совет әдебиетінің ғана емес, барша совет әдебиетінің лайықты мақтанышы, ойшыл да ғұлама классик жазушы М. О. Әуезов творчествосын зерттеу, оның сөз саптау мәнерінің қыры мен сырын ұға түсу ұрпақтардың барлығының да қасиетті борышы деп ұғынса ешкімге де артық болмас деп білеміз.

Халқымыздың телегей-теңіз бай халық тілін соншалықты игере отырып, тарих пен әдебиет, этнография мәселелеріне де бойлай білген жазушы көркем әдебиеттің барша жанрына да қалам тартқан, кей жанрдың алғашқы негізін салған. Алғаш пьеса жазып, оның қазақ театрының бет пердесін тұңғыш өз драмасының қойылуымен тұсауын да кескен. Ол пьесаның күні бүгінге дейін де қазақ театрлары сахнасынан түспей келе жатуының өзі де М. О. Әуезов пьесаларының өміршеңдігінің дәлелі болса керек.

Біздің зерттеуімізге тиесілі болған жазушының бұрынырақта шыққан 12 томдық шығармаларындағы үш томнан тұратын 22 драмалық шығармалары, оның 4-еуі аударма болса, өзгесі өз төл туындылары. Жуырда ғана 20 томдық шығармалар жинағы басылып шыққанымен, мұнда, да толық драмалық шығармалары қамтылды деуге болмайды.

Алдағы уақытта шығарылуы ойластырылып отырған 50 томдық шығармалар жинағында драмалық шығармалары түгел қамтылар деген сенімдеміз.

М. О. Әуезовтің драмалық шығармаларының тілін зерттеудегі негізгі мақсатымыз — бұл жанр тілін өзге жанр шығармаларының тілімен, атап айтқанда эпопея, әңгімелері, повестері, очерктері, сондай-ақ басқа автордың драмалық шығармаларымен көркем әдебиеттің басқа да салаларымен, яғни балалар әдебиеті, ауыз әдебиеті үлгілерімен, әрі функциональдық стильдің өзге де түрлерімен (публицистика, ғылыми стиль) салыстыру негізінде шығарманың жан-жақты жанрлық, тілдік, стилистикалық ерекшеліктерін көрсететін тілдік құбылыстардың авторға ғана тән өзіндік ерекшеліктерін айқындай түсу. Осылайша жазушының драмалық шығармалары тіліндегі стильдік ерекшеліктерді басқа жанрлармен стильдер сөздіктерімен салыстыра отырып, автордың дәл осы жанрына ғана тән жаңа сөздер жасау үлгісін, морфологиялық, грамматикалық тұлғаларды пайдалану шеберлігін, тіліміздегі бай қорды қаншалықты шебер меңгергендігін, жана мағыналар үстей білу, сөз жасампаздықтағы алдына жан салмайтын шеберлігінің сыры мен қырын ашып, айқындай түсуге талпыныс жасадық. Алғашқы әрекетті сол сөз топтарын индексациялаудан бастадық, олар зат есімді — 3, сын — 6, сан — а, етістік — ч, үстеу — ю, шылау — ш, одағай — о, еліктеуіш — л, модаль, бейтарап, әртарاپ сөздерді — 5, көмекші есімдерді — 1 деп 3 томдық пьесаларды сөз топтарына ажыратып отырып, алфавитті кері алфавитті, ЭВМ-ның көмегімен жиілік сөздіктерді алып, сол сөздіктерді тілге тек ете отырып талдаймыз. Зат есімді сөз табына ажыратуда да бірнеше жік-топтарға бөліп ажыраттық, кісі аттары — 2, жан-жануарлар аттары — ж, географиялық атаулар — г, мемлекет, ру, тайпа аттары — р, бас әрпінен қысқартылған сөздер — с, кітап, оқулық, шығарма, еңбек аттары — б, мәдени-тұрмыстық атаулар — м, астрономиялық атаулар — н.

Пьесалардың жазылу үлгілері бөлек-бөлек, кейбіреуі өлеңмен («Айман—Шолпан», «Бекет», «Қара қыпшақ Қобыланды», т. б.) біреулері қара сөзбен («Бәйбіше, тоқал») енді біреуі аралас — («Еңлік — Кебек», «Қаракөз» т. б.). Алынған сөздікке жүгіне сөйлесек, жалқы есімдер мен атауларды қоспағанда — 18 тәл, 4 аударма пьесаларында барлығы 11814 сөз қолданған, ол сөздер 297000-дай сөз қолданыста жұмсалған, оның ішінде 52647 сөз қолданыс тек авторлық ремаркаға тиесілі де, қалған 245305-і кейіпкер тіліне тән екенін анықтадық.

Ал роман жанрында (4 томдық эпопеяда) барлығы 16983 сөз қолданғаны айқындалған-ды, сонда 466 мындай сөз қолданыста он жеті мындай сөзбен-ақ (әрине ол сөздер грамматикалық тұлға, формалармен құлпырта, өзгертіле қолданылды ғой), соншалықты өлмес мұра қалдырған екен.

Пьеса жанрының өзіне тән дәлелдеуді қажет етпейтін өзгеше-

лігі сол — жалқы есімдердің авторлық ремаркаменен молынан кездесетіндігі. Оған айғақ 2474 сөз (8896 сөз қолданыста) жоғарыда аталған атаулар аттарына тән болды, ал бұл роман жанрында бір мың сөздің үстінде ғана кездескен еді, ал мұнда 2,5 еседей көп қолданылуының өзі осы стильдің өз ерекшелігінің бір қыры десек те болғандай, себебі төрт томдық эпопеяда соншама шытырман оқиғалар мен олардың арасындағы талас-тартыс, әңгіме-дүкен, қарым-қатынастар драмалық қысқа оқиға жүйелеріне ғана арналып жазыла тұрса да, бұған авторлық ремаркадан басқа толып жатқан антропонимикалық атаулар үстеле бергенде, зат есімнің үлесінің көбін осындай жалқы есімдер алатыны байқалды.

Жүргізілген зерттеулер нәтижесінде барлық пьесаларында сөз таптары жағынан ең мол қолданылғаны етістіктен соң зат есім екен. Тіл — тілдердегі зат есім ұғымдар мен түсінік, оқиғалардың, абстракциялық зат атаулары ретінде ең мол пайдаланылатыны, өздеріне тән дербес толық лексикалық мағыналары бар негізгі тірек сөздер екені белгілі. Сондықтан да зат есім сөздері адам қатынас құралы тілдің негізгі қаңқасы іспетті боп келетіні аян. Міне, сондықтан да болар әңгіме мен пьеса жанрларында болмаса, қалған романдары, повестері және очеркінде де М. О. Әуезов зат есімдер тобын басқа сөз таптарынан жоғары жиілікте пайдаланған. Барлық сөз қолданыстың бестен бір бөлігіндей, яғни 245505-тің 56574-і зат есімге тиесілі болып шықты, бұл пьесалардағы әрбір бесінші сөздің заттың атауы болып келетіндігін дәлелсіз-ақ көрсетіп тұр.

Сондай зат есім сөздерді тұлғасы жағынан топтай келгенде жартысына таяуы тек түбір формасында қолданылғаны таңқаларлық жай деп білеміз (24670, 4530), ал енді жалғаулардың түбір әрі туынды зат есімдерге үстелу жайына ойыссақ, онда барлық сөздің 80 проценті септік жалғауларының үлесіне тиеді екен, өйткені мұнда атау септігі тұлғасына тәуелделулі де, жіктелулі де, көптелулі де зат есімдерге ене беретіні мәлім ғой.

Ал егер жалғаулардың үлесіне тоқталатын болсақ, олардың жалғану реттері былайша болады екен: септік жалғаулары — 63,93 процент, тәуелдік жалғаулары — 30,64 процент, көптік жалғаулары — 2,88 процент, жіктік жалғаулары — 2,3 процент. Негізгі сөз байланыстырғыштық қасиет септік жалғауларының ең үнемділігін айқындаса, септіктердің жұмсалыу мүмкіндігі жағынан орындары ауысып кетеді екен: ең алдымен атау, одан соң табыс, барыс, ілік, шығыс, жатыс, көмектес болып кетеді. Септік жалғауларынан кейін пьесалар тіліндегі молынан жұмсалғаны тәуелдік жалғауы екен де, ең төменгі мүмкіндікте қолданылған жалғаулар жіктік және көптік жалғаулары екендігі анықталды. Бұның өзі қолданылуы мен жұмсалыуы жағынан зат есімнің септік жалғауларының

атқаратын тілдік қызметтерінің молдығын тағы да дәлелдей түсіп отыр.

Ал енді осы барлық жалғаулардың көбі бірінің үстіне бірі үстемеленіп жалғанып келіп отыру жайттары жиілігі жағынан төмен болғанымен түрлене қолданылуы жағынан молынан кездесіп отырады. Мәселен, септік жалғауларының ішінде *ды* — жалғауы 1928 жиілікте, *ын* — 1141, *ін* — 840, *ге* — 700, *қа* — 666, *ға* — 826, *ына* — 554, *ді* — 711, *да* — 512 — кездесті. Алайда, қолданылу мүмкіндігі жағынан ең жоғарғы жиілікте ұшырасқан дара күйіндегі тәуелдік жалғаулары болды: *ы* — 1966, *ым* — 1653 сөз қолданыста, *і* — 1452, *ім* — 770, *ың* — 692, *м* — 595 тәрізді болды. Ал енді аз қолданылған көптік жалғауларының ең жоғары жиілікте ұшырасқаны сөзқолданыста 363 рет кездесті.

Агглютинативті тілдің қасиетіне тән жалғамалылықтың өзі нағыз тұрақты қасиет екенін көрсетті, себебі барлық зат есімдерге жалғанған жалғаулардың көпшілігі екі түрлі, үш түрлі, кейде төрт түрлі жалғаулармен қосарланып алмасып келіп отырады екен. Тіліміздегі зат есімдердің пьесалардағы жұмсалуына қарап отырсақ барлық қолданылған сөздердің жартысына жуығын құрайды екен. Егер тұлғасына қарай топтап, жіктей келсек, онда бұл сөз табының 39, 32 проценті тек түбір тұлғасында қолданылғанын байқадық. Мұның өзі ең молынан кездесетін сөздерді жалғаулармен түрлендірмей-ақ көптеп шебер пайдалануға болатынын дәлелдеп отыр. Солай десек те сөз бен сөзді, сөйлем мен сөйлемді, айтайын деген ойды әрлеп, ұғымдырақ, әсерлірек етіп айту үшін аффикстердің атқаратын ролі аса мол екені белгілі. Мұндай жайларды біз тұтас барлық пьесалардағы зат есімдерге тұлғалық талдау жасай келгенде бірден аңғардық, өйткені 31770 сөз қолданыстың барлығы түрліше жалғау формаларымен өзгере қолданылған екен.

Бұрынырақта шыққан мақалада біз тек бір пьесаның өзінде ғана 200-дей жалғау форма варианттарының барлығын ашықтап айтқан едік, ал барлық 22 пьесада сөзқолданыс көлемі бірнеше есе өскенімен де жалғаулармен түрленудің онша көп өсе бермейтіні байқалады.

Филология ғылымдарының кандидаты Ә. Ахабаев публицистика тілін зерттей келе барша 1500000-дай сөзқолданыс ішіндегі зат есімдердің барынша мол мүмкіндікпен түрлене алатын ерекшелігін танытып, 496 түрлі жалғау варианттарымен өзгеретінін айқындап берген еді. Бұндай жағдай көркем әдебиет үлгілерінің ешбір жанрында болмағанына қарағанда публицистика стилінің өзіне ғана тән жалпы халыққа ұғымды, бай әдеби тіл үлгілерін барынша мол пайдаланатындығын дәлелдеп отыр. Міне, бұл, осы публицистика стилінің өзіндік бір үлкен ерекшелігі болып шыққанды.

Өйткені көлемі жағынан бұдан екі есе үлкен М. О. Әуезовтің драмалық шығармаларында 300 000-дай сөзқолданылыс небәрі 435 түрлі жалғау варианттарының барлығын анықтадық.

Ғ. Мүсіреповтің («Ақап сері Ақтоқты») пьесасы мен М. Әуезовтің «Еңлік Кебек» пьесасын салыстыра келгенімізде сондайлық бір зат есімнің түрленуі жағынан басы артық ерекшелену байқала да қойған жоқ еді. Мұндай жағдайдың өзі драма стилі ыңғайы мен тілдік шеберліктері жағынан сошшалықты бір нұқсан келтіре қоятын нәрсе емес екен-ау деген тұжырымға әкеп тірейді.

Бұдан қорытындылауға болатын нәрсе: көркемдігі жоғары жақсы көркем әдебиет шығармаларын жазуда сөзді, әсіресе зат есім сияқты молынан жұмсалатын сөз табын түрлендіре қолдану шығарма тілінің көркемдік дәрежесін анықтайтын бір шкала (деңгей) бола алмайды екен демекпіз.

Бұндай жағдай көркем әдебиет үлгілерінің ешбір басқа жанрларында аз болуына қарағанда публицистика стилінің өзіне ғана тән жалпы халыққа ұғымды да қарапайым, түсінікті халық тілінің үлгілерін барынша мол пайдаланатындығын тағы да бір дәлелдеп отырған тәрізді.

Қазақ тілі стилистикасын зерттеген ғалымдар (Балақаев М. Б., Мұсабекова Ф. М., Е, Жұбанов, Е. Жанпейісов, М. Томанов т. б.) сөздердің стильдік ыңғайда жұмсалуды жайлы айтқанда, олардың сол өз мағынасынан онша көп ауытқымай тек қана бір көп қолданыла бермейтін, кейде көзге еленіп, ерекшеленбейтін жағдайларға не болу сәті деген пікірлерін біз де мақұлдаймыз.

Мысалы үшін, кісі аттарына көптік жалғауларының жалғануы кездесіп отырады, мұның өзі сол адамдардың барлығының атын атап, түсін түстеп жатпай-ақ, ойды тұжырымдап беру, қысқарту мақсатында бірді-екілі аттар аталысымен автордың ықшамдатыңқырау үшін қолданған бір өзіндік стильдік ерекшелігі деп ұғынуға болады, өйткені әр адамның аты бірнешеу емес, бір-ақ ат болатыны белгілі ғой.

Оған мынандай мысалдар келтіруге болады: Абайлар, Кебектер, Сырымдар, т. б. қашанда да драматургия жанрына, стиліне тән нәрсе сөйлеу тілінің басым болатындығы, өйткені драманың оқиға түзілісінің өзі кейіпкерлердің диалогына негізделіп шешіліп отырады ғой. Көбіне қысқа келген кейіпкер диалогтарындағы ой қысқа тұжырымдармен айтылатын сөйлеу мәнеріне байланысты сөздердің түрлену аясы тар болуы заңды болар.

Көлемі пьесалардың жартысындай-ақ публицистикада зат есімнің 496 түрі форма-варианты болса, бұл стильде ол небәрі 435 ғана, алайда бұдан пьеса тіліне нұқсан келді деуге болмайды. Қайта аз да болса мәндеп айтылған деп түйіндеуге болады. Бүткіл зат есімнің жалғаулармен жске үстемеленіп түрленуін мына кес-

теден көруге болады: (р — түбір, с — септік, ж — жіктік, к — көптік, т — тәуелдік жалғауларының белгілері).

Зат есімнің тұлғаларының өзгеруін анықтайтын кесте

Жалғаулардың үстелу рет тәртібі	Жалғау түрлері	Сөз формасы	Сөз қолданысы
Р	—	3199	20589
Р—С	35	3714	9807
Р—Т—С	135	4121	9471
Р—К—С	47	443	657
Р—К—Т—С	85	404	554
		8782	20479
Р—К	6	460	1229
Р—Ж—К	4	11	11
Р—Т—Ж—К	6	6	6
Р—С—Ж—К	1	1	1
		478	1247
Р—Ж	11	311	312
Т—Т—Ж	33	365	555
Р—К—Ж	16	41	47
Р—Т—Т—С—Ж	1	1	1
Р—Т—С—С—Ж	2	2	2
Р—К—Т—Т—Ж	1	1	1
Р—Т—С—Ж	12	17	18
Р—С—Ж	6	10	10
		648	946
Р—К—Т	18	280	459
Р—Т	10	2508	8624
Р—Т—Т	3	4	5
		2792	9088
Р—Т—Д	4	27	47
Р—К—Т—Д		1	
19 Барлығы:	435	15899	52359
түбір тұлғаны алып тастағанда	435	12700	31770

Жанр тіліндегі зат есімдердің түрленуі үлесі жағынан септік жалғауларына тиеді, олардың қолданылу реті алдымен атау тұлғада. Себебі атау септігі тұлғасына көптік, жіктік, тәуелдік жалғауларының формасындағы сөздердің көбі еніп кететіні белгілі. Ал жалпы септік жалғауларының ішіндегі өнімдісі табыс, одан кейін барыс, ілік септіктері, ал қолданылу жиілігі жоғарылығы жағынан алда тұратын тәуелдік жалғауының I, III жақтарындағы сөздер *ы* — 1966, *ым* — 1653, *і* — 1452, *ім* — 770 т. с. с. болып кете барады.

Пьесалардағы жалғау түрлерінің ішінде жалғауларының $\frac{3}{4}$ 302 түрлі форма-вариантта 20479 сөз қолданылса, жіктік жалғаулары-

ның 85 түрі 946 ғана сөзқолданыста, ал 31 форма варианттағы тәуелдік жалғаулары 9088 сөзқолданысты қамтыған. Бұдан байқалатын жағдай: сөзқолданысы жоғарылау жалғаулардың форма варианттары азаяды екен де, керісінше аз қолданылатын сөздердің түрлену, өзгеру қабілеті мол болатыны көрініп тұр.

Бұл да драма жанрының өзіндік ерекшелігінің бір қырын айқындап беріп отыр. Жалғаулардың жалғану тәсілдеріне көз жүгіртсек жалғыз жалғанған жалғау формасындағы сөздердің сөзқолданысы жоғары болғанымен, сөздерге барынша мол жалғану мүмкіндігі бола тұра, сирек қолданылса да барынша көп форма вариантта кездесуі үстемеленіліп жалғанған түрлерінің үлесіне тиіде екен. Әсіресе тәуелдік-жіктік, көптік-септік, көптік-жіктік, көптік-тәуелдік, септік-жіктік, жіктік-көптік, тәуелдік-септік жалғаулары.

Сонымен қоса автордың кейіпкерлерін сөйлету, олардың ойларын әсерлі де ықшамды, эмоционалды түрде жеткізу ыңғайында көптеген зат есімдерді қосарлап, біріктіріп, әдейілеп бұза сөйлету, я болмаса сол күйінде, кейіпкердің білімі мен ой-өрісі дәрсжесі, тіпті ұлтына да қаратып сөйлетіп жіберетіні бар.

М. О. Әуезовтің драмалық шығармаларының тілін әдейі зерттеген көптеген ғалымдарымыз жазушының өлмес мұра боп саналатын драмалық шығармаларының тілін көбіне әдебиеттік ыңғайда зерттей отырса да, жаңағыдай кейіпкерлер аузына саф алтындай сомдалған сөздер салып сөйлете білген деп түйіндеген.

Сондай-ақ, зат есімнің ішінде көптеген, жоғарыда атап көрсеткендей жалқы есім сөздерінің молынан ұшырасқанын баса көрсеткіміз келеді. Адам аттары жайлы сөз болғанда жазушының кісі есімдерін де түрленде отырып атағанын байқауға болады. Бұл ретте өзімсіген, еркелеткен, сый-құрметпен, көңіл-күйі жайына орай атағаны байқалады, оған—*жан, -еке, -тай, -ш, -ша*; формулаларын үстеу арқылы шебер келтіре білген, мысалы: Асюша, Алуаш, Зуратай, Шәкентай, Еңлікжан, Зережан, Құбекжан, Құнеке, Кәке, Дәуке, Рәштай т. б.

Сондай-ақ демеуліктердің *-ай, -ау, -ақ* сияқты түрлерін тіркеп атаудың да диалогтерде біршама стильдік ретте қоланылғанын көрсетеді. Атап айтатын нәрсе, жалпы зат есімдермен қатар, кісі есімдерінің ішінде де, аударма туындылары болғандықтан әрі өз кейіпкерлерінің ұлтына орай басқа ұлт тіліндей адам аттары да мол ұшырасады.

Кейде өз орнымен әдемі ыңғайда есім туғызып айтып жіберетіні де бар, мысалы Көкжал, Көрпебайстан т. б.

Осының бәрі автордың шеберлік мәнерін әрлеп, ұштай тұсу, кейіпкер, оқиға, сюжет, талас-тартыс суреттеулеріндегі жанрлық сөз саптаудағы өз өрнегінің бірі болса керек. Түбір тұлғадағы

адам аттары 820 сөздің 5386 сөзқолданысында кездескен және басқа зат есімдер тәрізді жіктеліп, септеліп, тәуелденуі заңды құбылыс болса, ал енді кісі аттарына көптік жалғауының жалғауы тек қана жазушының осы жанрдағы ең басты стильдік ерекшеліктерінің бірінен саналса керек. Мысалы: *-лар, -тар, -дар, -терің, -ларың, -дер, -лер, -тер, -ңдар, -тері* тәрізді формалар кісі аттарына жалғана қолданылған. Әсіресе, драматургтің жалқы есімдер ішінде стильдік ыңғайда кісі аттарын қолдан жасап жіберу, адамдарға ат қою шеберлігі де кейіпкерлер аузына әдемі салынып отырады: Көзжақсым, Молдажігіт. Бұл атаулардың айтылуының өзі оқиғада, кейіпкерлер арасындағы қарым-қатынастың ұлттық дәстүрдегі үлкеннің атын атамай, ізет сақтау тәрізді, сыйластық, қадірлестікті білдіретін стильдік қызмет атқарып тұрған жалқы есімдер.

Сонымен қатар зат есімнің ішінде көмекшілік қызметін атқара қолданылған біраз сөздер ұшырасады, олар: *алд, арт, маң, қас, жан, аст, ұст* сияқты көмекші есімдер де, негізгі лексикалық есімдер ажыратылуы үшін, бөлек индекс қойылып өз алдына топталынып алынды, сөйтіп өз алдына бөлек сөзтізбе түзілді.

Сондай-ақ, пьесалар сөздігінде жер-су атаулары, астрономиялық атаулар, ру, мемлекет атаулары, жануарлар аттары, бас әріптен қысқартылған сөздер, мәдени-тұрмыстық атаулар мен оқулық, шығарма, газет-журналдар атаулары да индексстермен белгіленіп, бөлек ажыратылды. Бұлардың ішінде әсіресе географиялық атаулардың бәрінен де өнімдірек түрлене (барлығы 1674 сөз қолданыста) қолданылғанын, — 778 сөз формасында ұшырасқанын айтсақта болады. Итжеккен, Жерсібір, Шытырымпар тәрізді жер атауларын өз орнымен қолданғанын аңғардық.

Кейде пьеса текстерінің ішінде бас әріппен берілген ру атаулары, бірде кіші әріппен де беріледі, алайда біз оған қарамастан түгелдей белгілеп өз алдына топтастырудың, олардың саны пьесаларда 144 сөзформадағы 490 қолданыстан тұратын сөздер, оның ішінде барынша көп қолданылған ру аттары тобықты, борсақ, арғын, найман тәрізді атаулар болып келеді. Сондай-ақ жалғаулармен өзгеріп қолданылған 140 сөзформасында 325 сөз бас әріппен қысқарған сөздер болып кездесті.

Енді тоқталатын жәй, зат есімдердің ішінде 8 түрлі ұлт өкілдерінің кейіпкерлерінің қатысуына байланысты, кейде сол ұлттардың кейіпкерлері өз тілінде араластыра сөйлеп жіберетіндігі. Олар орыс, украин, қырғыз, татар, өзбек т. б. тілдеріндегі сөздер. Әрине автор мұндай жайларда әдейі айтқалы отырған ойының, оқиғасының, іс-әрекетінің, солардың мінез құлықтарының қыр-сырын жайып салып, көз алдымызға келтіру мақсатында, өзіндік бір жанрлық әрі стильдік ыңғайда жұмсағанын байқауға болады. Немесе кейіпкерлердің аузына нешәтүрлі күлкілі сөздер салып, кей-

іпкерлердің қалпын, мінез-құлқын дәлірек көрсету мақсатында сөйлетіп жібереді: жан-жананым, Көрпебайстаным, көзелім, т. б. сөздер. Бұлардың бәрінің ешбір артықтығы жоқ, қайта сондай реттілігімен, өрескел көрінбей, құлаққа сол кейіпкерлерден тек қана сондай жайларды күтетіндей күйге түсіреді.

Зат есімнің сөз құрамын қарап отырсақ мұнда сондай-ақ дублет сөздерді де молынан кездестіруге болады екен, олардың да орфографиялық жүйеге келе бермейтін тұлғалары пьесалар тілінде орынды пайдаланылып тұрған іспетті, себебі бұндай жағдай бұл жанрдың бір тілдік өзгешеліктерін айқындай түсіп тұр, әрі сөйлеу тіліне жақындау екенін анадайдан-ақ мезеп тұрғандай.

Мысалы бір газет сөзінің мынадай варианттары қатар жүреді: газит, газета, қазиті; сол тәрізді дұшпан: дұшпан, душпан, дұспан; дозақ, тозақ; тажал, дажал; гаделет, әделет; конфиске, кәмпеске, көнпеске т. с. с. болып 100-ге тарта дублет варианттар бар екен.

Шығарманы оқып отырған оқушы, бұл дублеттерді әдеби нормаға келмесе де өз орнымен автордың өзіндік бір тілдік ерекшеліктерін айқындайтын факторлары ретінде қабылдайды. Әрине мұндай жағдайлар туралы публицистика стилінде, сондай-ақ ғылыми стильде сөз қозғауға болмайтыны стильдік ретте түсінікті нәрсе, алайда драма жанрында автор кейіпкерлерін қалай сөйлетем десе де өзі оның білімі, мінез-құлқы, өзіндік бір ерекшеліктері, оқиғадағы іс-әрекетке тікелей қатысы тіпті көңіл күйіне де байланысты ауызекі сөйлеу тілі үлгісімен де сөйлете алған.

Бұл автордың айтар ойын әрі ұтымды, әрі ұғымды, әрі сенімді әрлендіру, сондай бір белгісіздеу, көзге қораштау сөйлеу тілі сөздері арқылы да тамаша стильдік өз мөнерін көрсеткен-ау деген ойға тірелесін.

Сонымен, айтайын дегеніміз: М. О. Әуезов драмалық шығармаларын осындай қарапайым, сөйлеу тілінің — халық тілінің мол байлығына сүңгі отырып, барынша әрлендіре, өзгерте, кейде жаңа мағына үстей де орынды пайдалана алған, жай бір қарадүрсін тәрізді сөздермен-ақ, құбылта отырып қандай тамаша сценалық образдар туғызғанына танқалған үстіне тандана бересін. Біз мұнда тек зат есімдердің ғана драмалық шығармалардағы сөздердің кездесу жиілігі, жалғаулармен және жұрнақтармен құлпыра түрленуі, олардың түр-түрлеріне, үстемелену жағдайларына біраз тоқталу арқылы, автордың өз төл туындылары мен аударма туындыларындағы ең негізгі деген сөз табының сөз қолданыстарына тоқталғанда, айтарымыз: пьесалардағы әр айтылған сөздің алтыдан бірі зат есімге тиесілі және автордың өзіндік ерекшеліктерінің көбі де осы сөз табындағы негізгі лексикалық мағыналы сөздерге байланысты гүлдене, құлпыра шыққан деген тұжырымға келеміз. Жаңа сөз жасау шеберлігі бола тұрып-ақ, автор мәлім сөздердің

өзін аффикстермен өзгерту, дублеттермен беру, я болмаса әдейі бұзып айтқызу арқылы-ақ талай балама ойлар мен кейіпкерлерін сомдай білген деп түйеміз. Басқа стильдегі зат есім сөздерімен салыстыра келсек, көлемі жөнінен салыстыруға боларлық публицистика (Ақабаев Ә.), роман жанрларында («Абай жолы») ең жиі қолданылған сөз табы зат есім болатын болса, драма жанрында егістікті молынша пайдаланғанын айттық. Мұның өзі де автордың алдымен айтар ойын ойланып, толғанып, әсерлеп-әрлеп барып айтатын қасиетін ұқтыра түскендей әсер аласың.

Соныменен, автор әр кейіпкерін өз дәуірінің, заманының тілімен сөйлетіп, сол сөйлету үстінде кейіпкерлер арқылы өзінің де шебер сөз зергері екендігін аша түсіп, дәлелдеп бергендей. Тілін кесек сомдау арқылы жазушы өзінің де өзіндік мәнері, стилі, сөз зергері екендігін өзі көрсетіп бере алған деген қорытынды жасаймыз.

А. Б. БЕЛБОТАЕВ

СӨЗФОРМАЛАРЫНДАҒЫ СӨЗ ТАПТАРЫНЫҢ СТАТИСТИКАЛЫҚ СИПАТТАМАСЫ (ҚАЗАҚ ТЕКСТЕРІНЕН ТҮЗІЛГЕН ЖИІЛІК СӨЗДІКТЕР НЕГІЗІНДЕ)

Жалпы тіл статистикасында сөз таптары мәселесінің сөз етілуі бұрыннан да бар дәстүр. Дегенмен де бұл құбылыстар түркі тілдерінде, оның ішінде қазақ тіл білімінде соңғы жылдардың жемісі деуіміз керек. Өйткені, қазақ тілі мәселелерін карауда зерттеу үшін алынған әр түрдегі (жанрдағы, стильдегі) қазақ текстері статистикалық тұрғыдан сипатталып, мүмкіндігінше ол текстерге лингвистикалық талдау жасалынады. Мұндай кезде ең алғаш анықталатын жай — белгілі бір текстегі сөз таптарының мөлшері, олардың өз ішіндегі ара салмағы.

Жалпы, қазақ тіліндегі сөз таптары мәселесін, оның қазақ текстеріндегі көрініс-мөлшерін, өз ішіндегі ара салмағын, ол бойынша алатын реттік орын тәртібін сөз еткенде, біз бұл жайларды қай кезде де үш тұрғыдан: «түбір сөз» («сөз»), «сөзформа», «сөзқолданыс» тұрғыларынан талдап, бағалағанды жөн деп білеміз. Өйткені, бұл — «сөз», «сөзформа», «сөзқолданыс» едіншалары —

тілдің бір-бірімен тығыз байланыстағы тұрақты элементтері. Солай болғанымен де, олардың жеке-жеке өзіндік ерекшеліктері де бар. Мәселен, кез-келген текстегі «сөз» — «сөзформадан», «сөзформа» — «сөзқолданыстан» қай кезде де аз болады [1, 4-б.]. Міне осындай заңдылықтар бізді бұл тілдік единицалардың қазақ текстеріндегі көрінісіне сөз таптары тұрғысынан жеке-жеке статистикалық талдау жасауға итермеледі.

Біздің сөз еткелі отырғанымыз — сөзформаларындағы сөз таптарының көрінісі жайлы.

Сөздер текст ішінде «әр түрлі тәсілдермен түрленіп, алуан түрлі формаларға ие болады» [2, 111-б.]. «Сөзформа» біздің қолдануымызда, «дербес немесе көмекші мағыналы сөздің түбір тұлғасы мен оған жалғанған форма тудырушы морфемалардың саналуан көріністері» [3, 24—25-б.]. Бұл сөзформалар тілдік қарым-қатынаста әр түрлі жиілікте қолданылып, жұмсалынады.

Қазақ тіл білімінде қазақ тілі мәселелерін зерттеудің негізінде қазақ текстерінен алынған сөзформалардың жиілік сөздіктері біршама бар. Олардың ішінде де әр түрлі тілдік стильдерден алынғандары да, әр түрлі тілдік жанрлардан алынғандары да, оқулық текстерінен түзілгендері де т. т. кездеседі.

Бұл жиілік сөздіктер туралы, ондағы сөзформалар жайлы тіл білімінің лингвистикалық статистика саласының қазақ тіл біліміне келуінен бері сөз болып келеді десек артық айтқандық болмайды. Өйткені, қандай текстен болмасын ЭВМ-дер арқылы әлгі текстің негізінде жиілік сөздік алар болсақ, бізге машина ең алдымен сөзформаларының сөздігін алып берген болар еді.

«Сөзформа» единицасы — грамматикамызда бұрыннан белгілі болғанымен де лингвостатистикадағы сөзформалар сөздігі, оның негізінде сөз таптарының көрінісі қазақ тіл білімінде енді-енді сөз болып отыр. Айталық, қазірге дейінгі қазақ тілі зерттеулерінде сөзформалардағы сөз таптарының статистикасы тек бір-екі әдебиеттерде ғана келтірілген. Онда да М. О. Әуезовтің «Абай жолы» роман-эпопеясындағы статистикасы ғана [1, 4—5-б.; 4, 9-б.].

Сөзформаларындағы сөз таптарының статистикасы мәселесі қазақ текстерінің негізінде (бірнеше текст көлемінде) біздің жұмыста алғаш рет қаралып отыр. Ал, қазақ тілінің сөзформаларындағы сөз таптарына лингвостатистикалық талдау жасау үшін тілдік материал болып ЭВМ-нің көмегімен қазақ текстерінен құрастырылған мына төмендегідей жиілік сөздіктер пайдаланылды (1-кестеге қараңыз).

Бұл кестеде 11 әдебиеттен түзілген жиілік сөздіктердің тізімі және ол сөздіктердің әрқайсысындағы сөзформа саны берілген. Оған зер сала қарасақ екі түрлі жайды байқаймыз: біріншісі —

І - кесте. ЭВМ-нің көмегімен құрастырылған
жиілік сөздіктер

Сөздіктің аты	Сөзформа саны	Құрастырушы авторлар:
М. Әуезовтің «Абай жолы» романының жиілік сөздігі. Алматы, 1979.	61424	Бектаев Қ. Б. Жұбанов А. Қ. Мырзабеков С. Белботаев А. Б. Бектаев Қ. Б. Белботаев А. Б.
М. Әуезовтің «Өскен өркен» романының жиілік сөздігі (қолжазба)	20610	Белботаев А. Б.
М. Әуезовтің повестері тілінің жиілік сөздігі (қолжазба)	6680	Белботаев А. Б.
М. Әуезовтің «Асыл нәсілдер» әңгімесінің жиілік сөздігі (қолжазба)	5544	•
М. Әуезовтің «Әйел жолы» әңгімесінің жиілік сөздігі (қолжазба)	9008	•
М. Әуезовтің «Шатқалаң» әңгімесінің жиілік сөздігі (қолжазба)	4174	•
М. Әуезовтің әңгімелері тілінің жиілік сөздігі (қолжазба)	23595	•
М. Әуезовтің «Индия очерктері» тілінің жиілік сөздігі (қолжазба)	12038	•
М. Әуезовтің очерктері тілінің жиілік сөздігі (қолжазба)	9132	•
М. Әуезовтің «Қаракөз» пьесасының жиілік сөздігі (қолжазба)	4376	Зекенова А. Р.
«Математика» тексінің жиілік сөздігі (қолжазба)	3393	Белботаев А. Б.

жиілік сөздік алынған текстер екі түрлі тілдік стильден (көркем әдебиет және ғылыми-техникалық), сондай-ақ көркем әдебиет стилінің бес түрлі жанрларынан (роман, повесть, очерк, әңгіме, пьеса) алынғандығы болса, екіншісі — көркем әдебиет стиліне жататын он жиілік сөздіктің бәрінің де текстерінің авторы бір жазушы (М. О. Әуезов) екендігі.

Әрине, сөзформаларының жиілік сөздіктерін бұлайша әр түрлі тілдік стильдерден, әртүрлі тілдік жанрлардан алуымыз да кездейсоқтық емес. Өйткені, алынған текстер неғұрлым әр түрлі болған сайын олардағы сөз таптарының көріністері де, сандық сипаттары да соғұрлым жалпы тілдік негізде айқындала бермек.

Егер 1-кестеде келтірілген сөздіктердегі сөзформаларын әр текст бойынша өз ішінде сөз таптарына бөліп, ондағы әр сөз табындағы сандық көрсеткіштерін өз тексіндегі сөзформаларының жалпы санына қарай шағып, проценттік үлестерін шығарар болсақ, бұл жағдай 2-кестедегідей көрінер еді (2-кестеге қараңыз).

Егер 2-кестеге қарасақ, бірінші аңғарылатын жай — сөзформаларының сөздігінде зат есім, етістік, сын есім сөз таптарындағы

2-кесте. Сөзформаларындағы сөз таптарының статистикасы

Сөз таптары	«АЖ» роман %	«ӨӨ» роман %	«Повестер» % ₀	«АН» әңгіме % ₀	«ӘЖ» әңгіме % ₀	«Ш» әңгіме % ₀	«Әңгі- мелер» % ₀	«ИО» очерк % ₀	«Очерктер % ₀	«Қаракөз» пьеса % ₀	«Математика» % ₀
Зат есім	46.55	45.78	39.40	41.29	44.84	36.39	42.78	55.26	52.81	38.02	44.77
Етістік	38,35	35.55	37.12	37.41	35.86	39.65	38.41	26.36	26.84	40.12	29.76
	84.90	81.31	76.62	78.70	80,70	76.04	81.19	81.62	79.65	78.14	74.53
Сын есім	9.37	11.23	11.48	11.07	10.15	10.28	10.73	11.81	12.11	9.06	12.18
	94.27	92,56	88,00	89.77	90,85	85.32	91.92	93.43	9176	87,20	83.71
Үстеу	2.12	2.38	4.19	2.87	2.35	3.59	2.44	2.32	2.62	3.87	2.86
Есімдік	1.53	2.38	3.22	3,73	3.62	5.41	2.66	1.85	2.64	4.66	5.00
Сан есім	0.60	0.84	1.38	1.35	0.85	0.91	0.87	0.96	1.14	1.11	1,91
Одағай	0.42	0.33	0.22	0.22	0.52	0.86	0.42	0.03	0.05	0,59	—
Модаль сөздер	0.39	0.63	0.63	0.77	0.68	0.60	0.53	0.43	0.60	1.41	0.92
Еліктеуіш с-р	0.34	0.21	0.90	0.20	0.16	0.86	0.46	0.09	0.15	0,04	—
Шылау	0.32	0.67	1.45	1.08	0.96	1.44	0.69	0.83	1.03	1.05	2.60
	5.72	7.44	11.99	10.22	9.15	13.67	8.07	6.56	8.23	12.74	13.29
проценті	99.99	100,00	99.99	99.99	99,99	99.99	99.99	99.99	99.99	99.94	100.00
Барлығы: сөзформа саны	61424	20610	6680	5540	9008	4174	23595	12038	9132	4676	3397

Ескерту. «Асыл нәсілдер»—«АН», «Әйел жолы»—«ӘЖ», «Шатқалаң — «Ш», «Индия очерктері» — «ИО» деп белгіленеді.

сөзформаларының сандық мөлшері ең көп проценттерді құрайды (86, 32—94,27% аралығында). Бірақ, сөзформаларының сөздігінде негізгі сөз таптары — зат есім мен етістік. Бұлар барлық текте орташа алғанда 74,53—84,90% аралығында. Бұл жерден байқайтынмыз біріншіден, зат есім мен етістік сөз таптарының тіліміздегі негізгі сөз таптары екендігі, екіншіден — екі сөз табының грамматикалық қарым-қатынасқа көп түсетіндігі, сөйтіп ол кезде тектегі сөзформалары сөз тудырушы және сөз түрлендіруші грамматикалық формаларда өнімді қолданылатындығын танытады.

Сөзформаларындағы сын есім туралы сөз ерекше. Өйткені, бұл сөз табы сөзформаларының жеке тектерінде (сөздіктерінде) болсын, не жалпы тектерде болсын форма тудыратын (не өзгертетін) грамматикалық аффикстермен қолданылуы жағынан зат есім мен етістіктен кейінгі үшінші орынды иемденеді. Оның сөзформаларындағы үлестік мөлшері, жалпы тектер (сөздіктер) бойынша 9,06—12,18% аралығында. Бұл құбылыстың тек қазақ текстеріне ғана емес, «қазіргі түркі тілдеріне» де ортақ болуы ғажап емес (5, 78-б.).

Сөзформалары сөздіктеріндегі бұл айтылған үш сөз табынан басқа сөз таптарының (үстеу, есімдік, сан есім, одағай, шылау, модаль, еліктеуіш сөздер) мөлшері, бар-жоғы 5,72—13,29% аралығында.

Егер сөзформаларындағы сөз таптарының мөлшерін, өз ішіндегі ара салмағын жалпы тектер бойынша қарайтын болсақ, олардың сандық мөлшер жағынан реттік орындары төмендегідей болады:

1. Зат есім: 36,39—55,26% аралығында;
2. Етістік: 26,36—40,12%;
3. Сын есім: 9,06—12,18%;
4. Есімдік: 1,53—5,4%;
5. Үстеу: 2,12—4,19%;
6. Шылау: 0,32—2,60%;
7. Сан есім: 0,60—1,91%;
8. Модаль сөздер: 0,39—1,41%;
9. Еліктеуіш сөздер: 0,04—0,90%;
10. Одағай сөздер: 0,03—0,86%.

Сөзформаларындағы сөз таптарының статистикалық ара салмағы бойынша алатын реттік орын тәртібі туралы не айтуға болады?

Тіліміздің сөзформаларындағы сөз таптарының реттік орын тәртібі мәселесі әр түрлі тектер көлемінде алғаш рет сөз болып отыр.

Сөзформаларындағы сөз таптарының реттік орын тәртібін анықтау мақсатымен, сөзформаларындағы сөз таптарының реттік орын

3-кесте. Сөзформаларындағы сөз таптарының статистикалық

Реттік орны	«АЖ» (роман)		«ӨӨ» (роман)		«Повестер»		«АН» (әңгіме)		«ӘЖ» (әңгіме)	
	сөз табы	%	сөз табы	%	сөз табы	%	сөз табы	%	сөз табы	%
1	Зат.	43.55	Зат.	45.76	Зат.	39.40	Зат.	41.29	Зат.	44.84
2	Ет.	38.35	Ет.	35.55	Ет.	37.12	Ет.	37.41	Ет.	35.86
3	Сын.	9.37	Сын.	11.23	Сын.	11.48	Сын.	11.07	Сын.	10.15
4	Үст.	2.12	Үст.	2.38	Үст.	4.19	Ес.	3.73	Ес.	3.62
5	Ес.	1.53	Ес.	2.38	Ес.	3.22	Үст.	2.87	Үст.	2,35
9	Сан.	0.60	Сан.	0.84	Шыл.	1.45	Сан.	1.35	Шыл.	0.96
7	Од.	0.42	Шыл.	0.67	Сан.	1.38	Шыл.	1.03	Сан.	0.85
8	Мод.	0.39	Мод.	0.63	Ел.	0.90	Мод.	0.77	Мод.	0.68
9	Ел.	0.34	Од.	0.33	Мод.	0.63	Од.	0.22	Од.	0.52
10	Шыл.	0.32	Ел.	0.21	Од.	0.22	Ел.	0.20	Ел.	0,17
Барлық %		99.99		100.00		99.99		99.99		99.99
Барлық сөзформа		61 424		20 610		6 680		5 544		9 008

тәртібін көрсетуге көмектесетін кесте мен схема жасалды (3-кесте мен оның схемасына қараңыз).

Бұл 3-кесте мен схеманың көмегімен және бұларды 2-кестемен салыстыра отырып, әзірге мынадай жайларды айтуға болады:

1. 2-кесте бойынша сөзформаларындағы зат есімді сандық мөлшер жағынан бірінші орынға қойсақ, 3-кесте мен схема біздің бұл қорытындымыздың әлі нақты еместігін көрсетеді. Осы сияқты екінші орындағы етістік сөз табы туралы да осындай қорытынды жасауға болады. Бірақ, сөзформаларында — бірінші, екінші орындарды осы екі сөз табының бөлісетіндігіне дау жоқ. Бұл анық жай.

Ал, осындағы осы екі сөз табының сөзформалар бойынша проценттік мөлшерінің молдығын, бірақ әлі де болса реттік орындарының тұрақсыздығын қалай түсіндіруге болады?

1. Зат есім де, етістік те қашаннан атауыш сөз таптарының ішіндегі ең үлкен сүбелі сөз таптары. Бұл жайды бұдан бұрынғы зерттеулер де дәлелдеп жүр.

2. Бұл екі сөз табы да формалық жақтан өзгеруде өздеріне тән өзіндік ерекшеліктері бар. Атап айтқанда, зат есім жалғаулардың барлық түрлерімен өзгере алады, ал етістіктің болса зат есімге қарағанда түрлену мүмкіндігі әлдеқайда мол. Өйткені, олар жалғаулармен қатар арнаулы жұрнақтармен де түрленеді [1].

3. Етістік сөзформаларының бірінші орында көрінуі — көркем әдебиет стилінің тек әңгіме («Шатқалаң» — 39,65%), пьеса («Қаракөз» — 40,12%) жанрларынан байқалады.

реттік орын тәртібі

(әңгіме) «Ш»		«Әңгімелер»		(очерк) «ИО»		«Очерктер»		«Қаракөз» (пьеса)		«Математика»	
сөз табы	%	сөз табы	%	сөз табы	%	сөз табы	%	сөз табы	%	сөз табы	%
Ет.	39.65	Зат.	42.78	Зат.	55.26	Зат.	52.81	Ет.	40.12	Зат.	44.77
Зат.	39.39	Ет.	38.41	Ет.	26.36	Ет.	26.84	Зат.	38.02	Ет.	29.76
Сын.	10.28	Сын.	10.73	Сын.	11.81	Сын.	12.11	Сын.	9.06	Сын.	12.18
Ес.	5.41	Ес.	2.66	Үст.	2.32	Ес.	2.64	Ес.	4.66	Ес.	5.00
Үст.	3.59	Үст.	2.44	Ет.	1.85	Үст.	2.62	Үст.	3.87	Үст.	2.86
Шыл.	1.44	Сан.	0.87	Сан.	0.95	Сан.	1.14	Мод.	1.41	Шыл.	2.60
Сан.	0.91	Шыл.	0.69	Шыл.	0.83	Шыл.	1.03	Сан.	1.11	Сан.	1.91
Од.	0.86	Мод.	0.53	Мод.	0.48	Мод.	0.60	Шыл.	1.06	Мод.	0.92
Ел.	0.86	Ел.	0.46	Ел.	0.09	Ел.	0.15	Од.	0.59	Од.	—
Мод.	0.60	Од.	0.42	Од.	0,03	Од.	0,05	Ел.	0.04	Ел.	—
99.99		99.99		99.99		99.99		99,94		100.00	
4 174		23 595		12 038		9 132		4 676		3 396	

Бұл жайдың біздіңше үш түрлі себебі бар сияқты:

Бірінші — жанрлық ерекшелік болуы мүмкін. Бірақ, бұлай дейтін болсақ, зерттеуге алынған текстерде бұдан басқа да үш әңгіме («АН», «ӘЖ», «Әңгімелер») бар. Олардағы жағдай — әдеттегідей. Ал, бұл біздің бірінші болжалымыздың қисынсыз екендігін дәлелдейді.

Екінші — текстер көлеміне байланысты болуы мүмкін. Бұлай дейтін себебіміз, басқа текстермен салыстырғанда бұл екі сөздіктің сөзформа көлемі аз (қараңыз: «Шатқалаң» — 4174 сөзформа, «Қаракөзде» — 4676).

Үшінші — авторлық стиль де (ерекшелік) болуы мүмкін.

Қалай дегенмен де бұл құбылыстың «күпиясын» нақтылай ашу үшін әлі де зерттей түсу қажет.

2. Сөзформаларындағы сын есімдердің орны сандық мөлшері жағынан да (9,06—12,08%), текстердегі схемалық көрінісі жағынан да орнықты, тиянақты — үшінші орын.

3. Сөзформаларындағы есімдік пен үстеу сөз таптарының реттік орын тәртібі жайы да сырттай қарағанда зат есім мен етістіктердің жайыма көп ұқсас: бұл екі сөз табының қайсысын алғашқы, қайсысын соңғы орынға қою керек екенін қазірге ашып атау қиын. Өйткені, 2-кестеде берілген сөзформалары сөздіктеріндегі сөз таптарының статистикасына қарасақ есімдікті — төртінші (1,53—5,4%), үстеуді — бесінші (2,12—4,19%) орынға қояр едік. Бірақ 3-кесте мен оның схемасы біздің бұл шешіміміздің әлі де дәл

Реттік орны	„АЖ“ (ро- ман)	„ӨӨ“ (ро- ман)	„Пове- стер“	„АН“ (әңгі- ме)	„ӘЖ“ (әңгі- ме)	„Ш“ (әңгі- ме)	„Әңгі- мелер“	„Ю“ (очерк)	„Очерк- тер“	„Қара- көз (пьеса)	„Мате- матика“	
1	зат.	зат.	зат.	зат.	зат.	ет.	зат.	зат.	зат.	ет.	зат.	● Зат есім
2	ет.	ет.	ет.	ет.	ет.	зат.	ет.	ет.	ет.	зат.	ет.	● Етістік
③	сын.	сын.	сын.	сын.	сын.	сын.	сын.	сын.	сын.	сын.	сын.	● Сын есім
4	үст.	үст.	үст.	ес.	ес.	ес.	ес.	үст.	ес.	ес.	ес.	● Есімдік
5	ес.	ес.	ес.	үст.	үст.	үст.	үст.	ес.	үст.	үст.	үст.	● Үстеу
6	сан.	сан.	шыл.	сан.	шыл.	шыл.	сан.	сан.	сан.	мод.	шыл.	○ Шылау
7	од.	шыл.	сан.	шыл.	сан.	сан.	шыл.	шыл.	шыл.	сан.	сан.	○ Сан есім
8	мод.	мод.	ел.	мод.	мод.	од.	мод.	мод.	мод.	шыл.	мод.	○ Модаль сөздер
9	ел.	од.	мод.	од.	од.	ел.	ел.	ел.	ел.	од.	од.	○ Одағай сөздер
10	шыл.	ел.	од.	ел.	ел.	мод.	од.	од.	од.	ел.	ел.	○ Еліктеуіш сөздер
Барлық сөзфор- ма	61424	20610	6680	5544	9008	4174	23595	12038	9132	4676	3396	

Сөзформалардағы сөз таптарының статистикалық
реттік орын тәртібінің қазақ текстеріндегі
схемалық көрінісі.

еместігін көрсетеді. Мұнда үстеу — роман, повесть, очерк («ИО») жанрларында төртінші орындарда да, басқа жанрларда (өңгіме, очерк («очерктер»), пьеса мен «математика» тексінде) бесінші орында. Ал, есімдік сөз табы — роман, повесть, очерк («ИО») жанрларында бесінші орындарды көрсетеді де, басқа тектерде төртінші орындарда тұр.

Біздіңше, екі сөз табының арасындағы бұл құбылыс ұзақ мерзімдік процесс сияқты. Өйткені, біріншіден, екеуінің де проценттік көрсеткіштері (сандық мөлшер-шамасы) бір-біріне жуық та (есімдік: 1,53—5,41; үстеу: 2,12—4,19), басқа сөз таптарынан бөлектеу, көп алшақ; екіншіден — екеуі қай кезде де осы орындардың (төртінші, бесінші) несі. Мүмкін, бұл тілдің дамуындағы тарихи процестің нәтижесі де болар.

4. Қазақ текстерінің сөзформаларындағы біз атап көрсеткен сөз таптарынан басқа — шылау, сан есім, еліктеуіш, одағай, модаль сөздер сияқты сөз таптары бар. Олардың реттік орын төртіптері жайлы қазір нақтылы пікір айта қоюға әлі ерте сияқты.

Біздіңше мұның екі түрлі себебі бар: бірінші себебі — бұл сөз таптарының әлі де болса қазақ текстерінде (жазба әдебиетінде) бір қалыпты жүйеде (мысалы, сын есімдей) көрінбеуінен, түрленіп жұмсалмауынан да болуы мүмкін болса, екінші себебі — текстің не жанрлық, не авторлық стиліне байланысты болуы да мүмкін. Тек, тілдік стильдерге байланысты емес. Бұл анық. Өйткені, сөзформаларынан алынған текстердегі тек «математика» тексін есептемесек, басқа текстердің бәрі бір тілдік стильге — көркем әдебиет стиліне жатады емес пе? Осының өзінде (бір тілдік стильде болса да) жоғарыдағы сөз таптарының көрінісінде реттік заңдылық, не тұрақтылық байқалмайды. Ал мұндай реттік заңдылықты, не тұрақтылықты әлі де болса зат есім, етістік, үстеу, есімдік сөз таптарынан күтуге болады. Шамасы, бұл — уақытпен шешілетін, тілдің, оның ішінде жазбаша әдебиеттің дамуына байланысты көрінетін құбылыс болса керек.

ӘДЕБИЕТТЕР

1. Мырзабеков С. Статистико-лингвистический анализ структуры глагола современного казахского языка (по роману М. О. Луэзова «Абай жолы»): Автореф. дис. ... канд. филол. наук. Алма-Ата, 1973. 32 с.

2. Аханов К. Грамматика теориясының негіздері. Оқулық. Алматы, 1972. 216 б.

3. Ахабаев Ә., Бектаев К. Б. Текст единицалары және олардың сипаттамалары. Кітапта: Қазақ тексінің статистикасы. Алматы, 1974. 23—48-б.

4. Бектаев К. Б., Жубанов А. Қ., Мырзабеков С., Белботаев А. Б. М. Әуезовтің «Абай жолы» романының жиілік сөздігі. Алматы, 1979. 336 б.

5. Нұрмаханова Ә. Н. Түркі тілдерінің салыстырмалы грамматикасы. Алматы, 1971.

АВТОМАТИЧЕСКИЙ АНАЛИЗ МОРФОЛОГИИ КАЗАХСКОЙ СЛОВОФОРМЫ

Автоматический анализ морфологии слова является одним из наиболее важных и необходимых этапов лингвистической операции. Путем автоматического морфологического разбора при обработке отдельно взятых словоформ обеспечивается получение первичных грамматических характеристик и отделение морфемной основы словоформ от аффиксальных морфем. Это необходимо для последующих этапов распознавания словесной величины.

В системах автоматической обработки текстов (например, в автоматической поисковой системе) морфологический анализ применяется для установления фактов парадигматических соотносительностей слов и выявления связующих грамматических показателей слов и их сочетаний, необходимых для исследования на синтаксическом уровне. Опыт моделирования лексики и морфологии некоторых индоевропейских языков [1—3] показал, что в связи с инженерно-лингвистическим моделированием и автоматическим переводом текстов возникает необходимость создания машинной морфологии, «опирающихся на нетрадиционное выделение так называемых машинных флексий и создание словаря машинных основ (квазиоснов)» [4, 5]. В результате автоматического анализа морфемного состава можно получить первичные грамматические признаки только по предписанному грамматическому коду.

Автоматическое приведение словоформы к основной (словарной) форме является сложной задачей из-за большого количества фоно-морфологических особенностей лексики, представляющих национальные черты тюркских языков, в частности казахского.

Задачи автоматического морфологического анализа казахских текстов выдвигают необходимость накопления и хранения максимально полной информации об отдельном слове (графические, фонетические характеристики и даже семантические, парадигматические и социолингвистические).

Наши исследования, направленные для решения задачи лемматизации, проводятся на материалах десятилетнего толкового словаря казахского языка, алфавитно-частотных, обратных словарей, составленных на ЭВМ по отдельным жанрам казахских письменных текстов различного объема. Изучение и выявление внутреннего строения казахской словоформы и всевозможных правил порождения с помощью аффиксов и основ казахских слов, является одной из важных задач при автоматическом морфологическом ана-

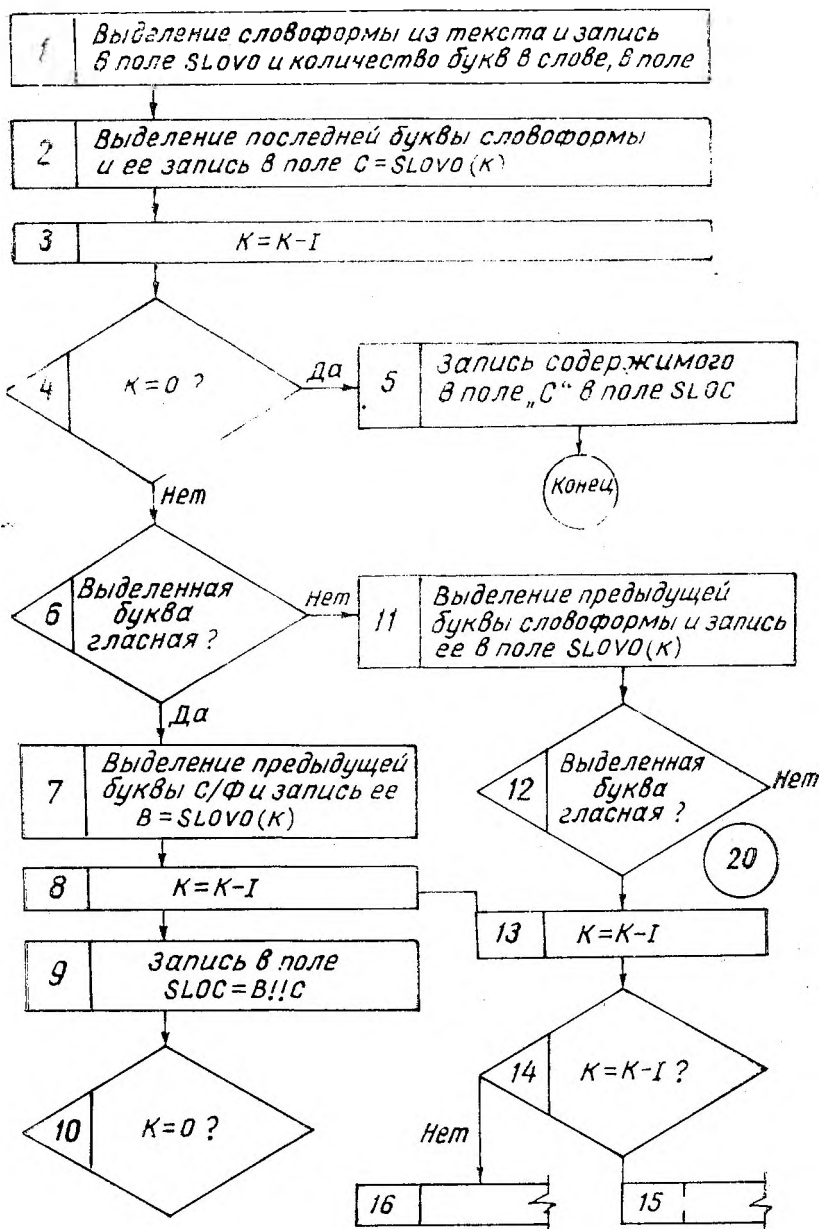
лизе казахского текста. На базе этих исследований разрабатываются алгоритмы автоматического слогаделения и автоматического морфологического анализа. Ранее полученные А. Ахабаевым, С. Мирзабековым и А. Джубановым статистические данные по морфологической структуре казахских существительных, глаголов, прилагательных, местоимений [6] послужили материалом для выявления и изучения закономерностей аффиксации словоформ по отдельным частям речи. В алгоритме автоматического морфологического анализа мы попытались учесть случаи соответствия словоформ, соотносящихся с основной формой слова, при котором возникает немало трудностей. Например, слова «бал» (мёд), «балға» (молоток) в словосочетаниях «балға мал» (мокни в мёд) и «ұлкен балға» (большой молоток) затрудняют автоматическое выделение основы «бал» или «балға». Поэтому такие омонимичные слова задаются как исключения в виде отдельной таблицы. Для ускорения скорости поиска в нашем алгоритме также указывается длина основы в слогах, намного упрощающая процесс отделения основ от казахских словоформ. Кроме словаря основ и словоформ, в алгоритме используется также словарь сочетаний словообразующих и формообразующих аффиксов. При этом аффиксы, присоединяемые к существительным и образующие существительные от других классов слов, составляют одну группу, а аффиксы, присоединяемые к существительным и образующие глаголы, — в другую группу и т. д. Словарь сочетаний аффиксов можно использовать и для установления парадигматических эквивалентов слов.

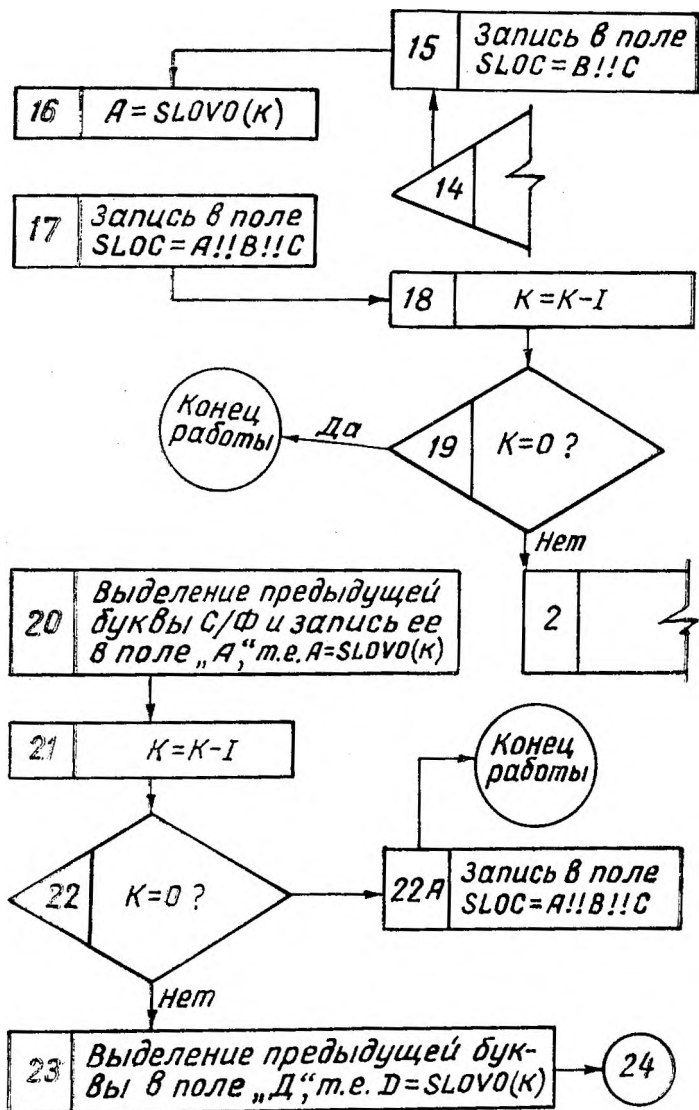
В казахском языке только у существительных количество формообразующих аффиксов достигает 500, а многие казахские глаголы имеют до 1000 различных форм [7]. В связи с этим возникает необходимость проводить классификацию частей речи и грамматических категорий путем машинной обработки морфологических данных для их формализации.

В тюркских языках части речи традиционно классифицируются в основном с точки зрения их смыслового содержания слова (лексического), т. е. значения принадлежности к той или иной системе словообразования или словозменения, а также синтаксических функций.

На основе этих признаков в тюркских языках первичные и вторичные морфемы делятся на три группы [8]:

1. Корневые морфемы, выражающие имя, т. е. название предмета или статического признака предмета;
2. Корневые морфемы, выражающие глагол, т. е. обозначающие действие — динамический признак;
3. Корневые служебные морфемы, выражающие грамматические отношения слов в составе словосочетаний и предложений.





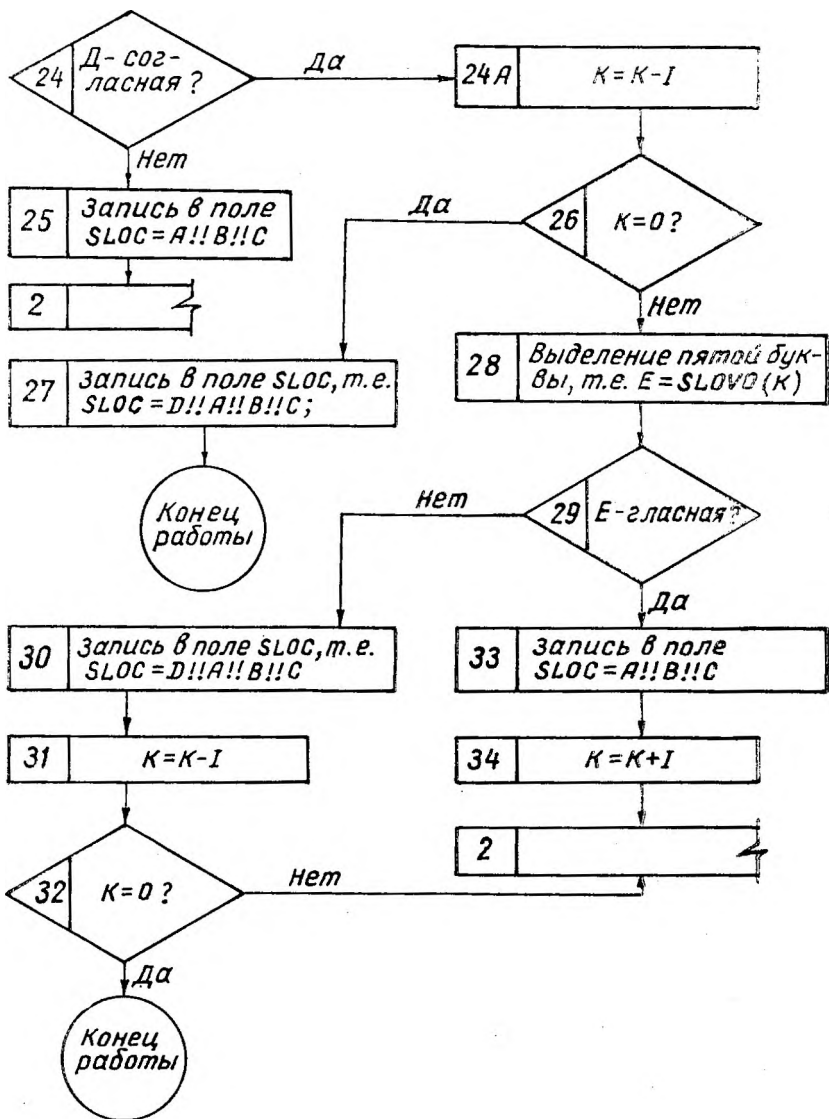


Рис. 1. Блок-схема алгоритма автоматического слогаделения

К первой группе относятся имена существительные, имена прилагательные, наречия, числительные, междометия и мимемы, местоимения, ко второй группе — глаголы и к третьей — частицы, послелоги, союзы, модальные слова.

Слова, относящиеся, к первым двум группам, имеют особые, отличные друг от друга системы аффиксального словособразования. Однако формально некоторые омоаффиксы неразличимы (омографы). Омонимичные слова могут иметь аффиксы, присоединяемые к корневым морфемам как первого, так и второго типа. Всю совокупность корневых морфем можно подразделить также на два больших класса: самостоятельные, или знаменательные; несамостоятельные, или служебные.

Кроме того, в тюркских языках существуют группы слов, которые на морфологическом уровне, т. е. вне контекста, нельзя отнести к определенной части речи (омонимичные слова).

Для изучения природы казахского языка необходимо хорошо представлять систему слоговых структур. Казахская письменность, как и многие письменности индоевропейских языков, основана на звуковой системе записи. Несмотря на это, в казахском языке больше всего преобладает слоговой оттенок. Поэтому знания морфологической структуры языка сопряжены с умением сегментировать слова казахского языка на слоги. Еще в 1934 г. Х. Жубанов в статье «Буын жігін қалай табуға болады?» (Как определить границы слогов? [9]) на базе выделенных им шести свойств слогов предложил своеобразный алгоритм определения границ слогов в казахских словах (словоформах). Ниже перечислены шесть свойств, характерных для казахского языка, когда:

не существует слога без участия гласного звука;

в одном слоге не может быть более одного гласного;

изолированный гласный звук может образовать слог и самостоятельно;

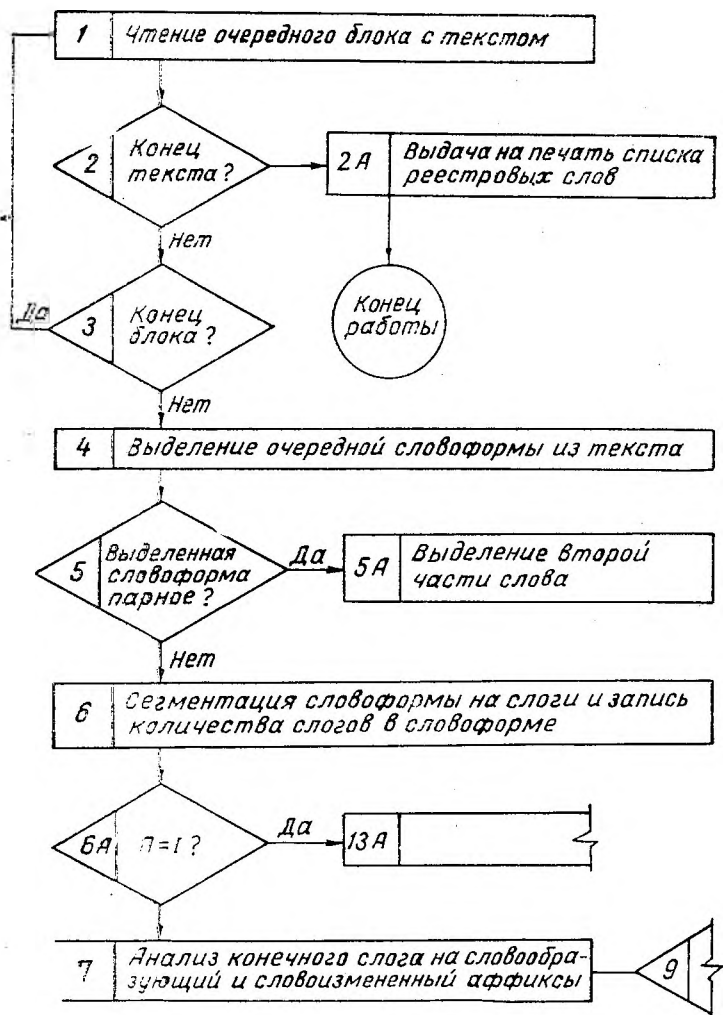
в одном слоге участвует не более трех согласных;

сочетание двух согласных в одном слоге возможно только после гласного звука; причем в обязательной композите: сонор (звонкий) + глухой;

внутри слова не встречается слог, начинающийся с гласного звука. В середине или конце слова слог обязательно начинается с инициального звука.

Сегментацию слов на слоги в казахском языке профессор Х. Жубанов предлагает начать с конца слова, причем конечному гласному должен предшествовать согласный звук, который означает начало границ между слогами.

Количество звуков в слогах собственно казахских словоформ определяется в пределах от одного до четырех.



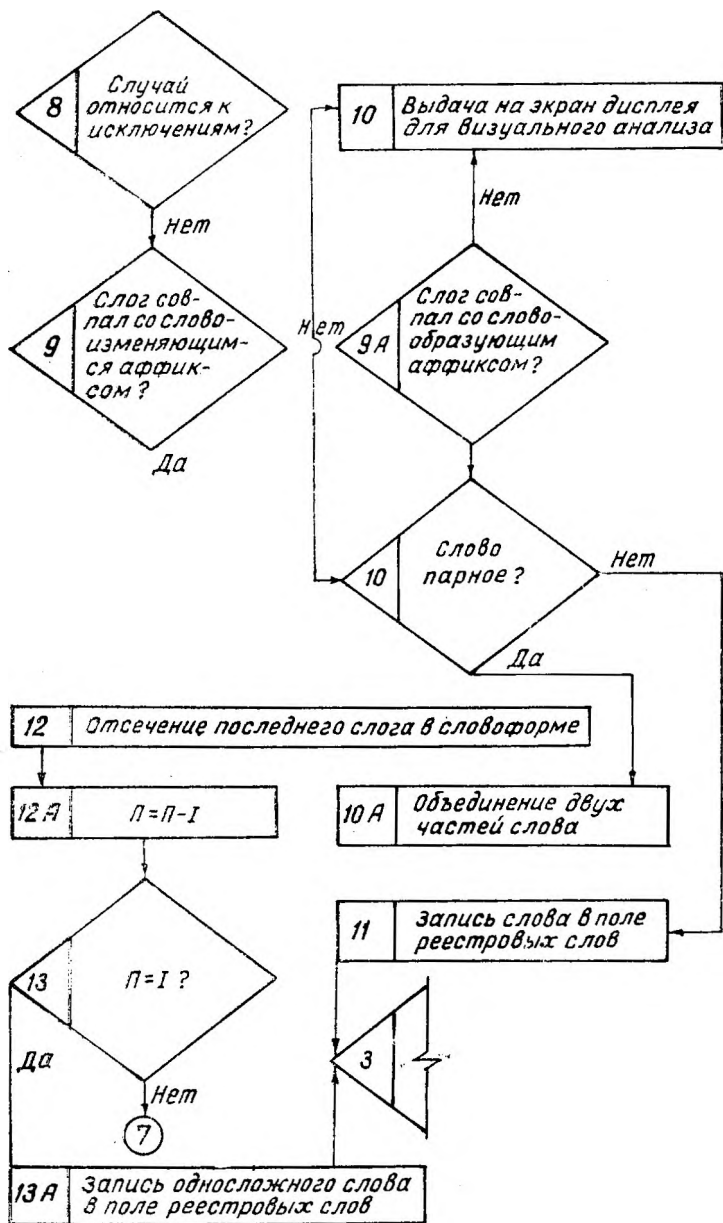


Рис. 2. Блок-схема алгоритма автоматического морфологического анализа

Х. Жубанов выделяет шесть вариантов слогов по разновидности структур звуков: Г; СГ; ГС; СГС; СГСС.

Основываясь на этом, построим алгоритм слогаделения казахской словоформы, блок схемы его ниже (см. блок-схему 1).

Автоматический морфологический анализ казахской словоформы, а именно приведение любой словоформы текста к реестровой форме мы проводим согласно слоговой структуре казахского слова.

В блок-схеме 2 в укрупненном виде приводится алгоритм выше-названной задачи. Решение ее основано на распознавании словообразующих и словоизменяющих аффиксов среди слоговой структуры конкретного слога. Анализ проводится с конечных слогов слова, путем сравнения с заранее заданным перечнем названных аффиксов, производится отсечение конечных словоизменяющих аффиксов. При этом учитываются случаи омонимии аффиксов. Некоторые конкретные ситуации заданы в виде таблиц исключений. Те случаи, которые не подчиняются нашим правилам сегментации, выдаются на экран дисплея для последующей визуальной корректировки. Такие нетипичные случаи фиксируются в таблице исключений. Данный алгоритм реализован на разных, различающихся по стилю и типам, текстах.

ЛИТЕРАТУРА

1. Статистика речи. Л., 1969. С. 259.
2. Статистика речи и автоматический анализ. Л., 1971 г. С. 240.
3. Зорев М. Г. Машинная морфология и ее применение в немецко-русском автоматическом словаре // СТРААТ. Л., 1973. С. 270.
4. Коваль Н. А. Опыт построения словаря основ и таблиц окончания украинского глагола для анализа при машинном переводе // Применение новых методов в изучении языка. Вып. 1. Днепропетровск, 1969. С. 245.
5. Баскаков Н. А. Морфологическая структура слова и части речи в тюркских языках // Советское востоковедение. 1957. № 1. С. 72—85.
6. Статистика казахского текста. Алма-Ата, 1973. С. 49—136, 231.
7. Там же. С. 39.
8. Там же.
9. Жубанов Х. Исследования по казахскому языку Алма-Ата, 1965.

ЛЕКСИЧЕСКАЯ СТРУКТУРА ИМЕН СУЩЕСТВИТЕЛЬНЫХ В ЯЗЫКЕ ДРАМАТУРГИИ

В нашей статье мы рассматриваем характеристики некоторых своеобразных черт жанрово-стилистической жизни слов в пьесах основателя казахской драматургии М. О. Ауэзова, а также количественное соотношение и качественное своеобразие текста и с лексико-стилистической стороны. Особенности языка пьес М. О. Ауэзова ранее исследовались Х. М. Нурмуқановым [1], но в основном в литературоведческом аспекте.

В статье приводится статистическое сопоставление употребления лексики имен существительных с различными лексическими пластами внутри пьес. Поскольку в исследовании языка казахской драматургии статистический метод применяется впервые, то сама методика исследования является мало разработанной. Опорой и моделью текста драматургии явились полученные на ЭВМ алфавитные и обратно-алфавитные частотные словари, которые составляют по 3 томам, (по изданному в 1967—1969 гг. полному собранию сочинений писателя) около 297 000 словоупотреблений и 18 000 словоформ, 11 592 слов. Путем сопоставления разных жанров произведений ученые старались выявить особенности языка драмы с применением статистического метода. В качестве единицы алфавитного, обратно-алфавитного или частотного словаря выступает словоформа. Для статистического исследования лексики надо решить и морфологические задачи, осуществить индексацию частей речи. В своей речи мы опирались на грамматику современного казахского языка, а также на труды ученых С. К. Кенесбаева, Г. Г. Мусабаева, А. И. Искакова, М. Б. Балакаева [2].

Четкое разграничивание при индексации позволило устранить конверсионную омонимию; например, слово *ак* — прилагательное, дающее понятие белизны, цвета, а дополнительный смысловой оттенок обозначает движение — *течь*, наконец, понятие молочных продуктов и т. д. Вышеназванные словари являются той статистической моделью, с помощью которой описывается распределение вероятностей отдельных лексических единиц, где указана статистическая характеристика вместе с указанием частей речи.

В лингвостатистике сплошным изучением пользуются тогда, когда велика генеральная совокупность, и тем более необходимо учесть все употребления интересующих нас лингвистических единиц. Так бывает при статистическом описании языка писателя или при исследовании языка отдельного художественного произведения.

Известно, что язык драматургических произведений сложен и своеобразен, потому что здесь взаимодействуют и сталкиваются два фактора — устная разговорная речь, индивидуальность и мастерство повествователя. В этой статье мы рассматриваем только одну из частей речи — имя существительное и применение его с отдельными лексическими пластами, слова иноязычного происхождения, слитные, а также парные непонятные слова в речи персонажей. Одной такой частью речи являются слова русского происхождения, использованные в переводных, а также в авторских пьесах для стилизации речи персонажей. Так, в пьесах 876 слов русского происхождения звучало в устах героев, это были в основном социально-политическая лексика и бытовые термины. К таким словам относятся: адмирал, генерал, прокурор, судья, герцог, герцогиня, поэзия, проза и т. д. Автор использует такие слова без перевода на казахский язык, потому что некоторые из этих слов не имеют аналогичного перевода, а некоторые не переводит специально, такие употребления имеют чисто стилистическую окраску. К этим словам, как к исконно казахским, присоединяются казахские аффиксы. Всего 876 слов, измененных аффиксами, использованы в 3540 словоупотреблениях. Интересным является то, что некоторые русские слова в казахском языке передаются в искаженной форме сознательно. В речи персонажей слово гимназия употреблено в трех разновидностях, как: кемназия, кемпажна, кемнажид. Подобных слов насчитывается 158, измененных грамматическими формами в 224 словоупотреблениях, слова следовательно — седуэтл, седуэтель; ошибка — эшебке, эшепке; газета — кэзит, газита, казит, газати; книга — кнега, кнежке, конфетка — кэмпеске, кэнпеске, кемпес; институт — уністет; пристав — бірстап; бочка — мошке, бошке; существительное — сошеветний; охрана — ахран и т. п. иногда образуя парные слова: ушоный-пушоный; учитель-мучитель и т. д. Автор языком персонажей называет такие слова, показывая типы героев в разных ситуациях, в разное время развития личности.

Имена существительные в языке драматургии М. О. Ауэзова очень разнообразны и по форме. Использованы парные слова в словоформах — 440, около 660 словоупотреблений: арба-сока, айран-шалап, акпан-кантар, бай-кулак, бай-жуан, бай-манап, биторе, жуан-феодал, феодалдык-хандык и т. д. Новаторством писателя здесь является то, что он обогащает состав парных слов, образуя одну часть из исконно казахских, другую из русских компонентов.

Слитные слова использованы мало, они встречаются в 229 словоупотреблениях, как — асказан (желудок), алашорда — название организации, дунбасы — городничий, желбау — одна из веревок

юрты; пманшарт — название религиозной книги, калынмал — выкуп за невесту, шекара — граница и т. д. Здесь основу создают только исконно казахские слова, кроме некоторых слов тива шекара — по происхождению вроде калькированного перевода. Небольшая часть слов относится к религиозным, культовым слоям слов, которые в настоящее время редко можно услышать из уст героев, верующих в бога: кудай — бог, дозак — ад, пайгамбар — апостол, намаз — молитва, ораза — пост, бейіш — рай и т. д.

В пьесах М. О. Ауэзова живут персонажи восьми национальностей, поэтому встречаются в словаре слова и на татарском, уйгурском, узбекском языках. Большая доля из них относится к татарскому языку, например: башым — моя голова, аял — женщина, мэсляк — мозг, кыз — девушка, шатлык — радость и т. п. Насчитывается 63 слова в 76 словоупотреблениях. Таким образом, небольшая часть существительных относится к чисто устаревшей, просторечной лексике, как ока, найза, сауга, зындан, айбалта, часть к политико-административной лексике — ояз, болыс, султан, старшын, болыстык, кул, хан, патша, мырза, малай, отчасти к просторечным словам: бетім-ау, батыр-ау, салдакы, бишара, жәдігей, былшыл мәнэс сарамас шыракайым, каралдым, дойдірім и т. д. Частотность употребления таких лексем невысока, многие употребляются по одному разу, но каждая лексическая единица имеет свое место при стилизации речи персонажей.

Особо выделим мастерство писателя в словотворчестве — словообразовании. Герои М. О. Ауэзова говорят языком своего времени, своей эпохи. О применении некоторых, редко употребляемых и непонятных слов, можно сказать следующее: о словах алкаганы, шарпы, нор, байшоғел, текпілі, батыр, эреке, кеден, кужандау, сойы, дулау, уйысу, нами была написана статья к 90-летию М. О. Ауэзова в 1987 г. [3].

О словах малоупотребительных и непонятных, которых не оказалось в десятитомном толковом словаре казахского языка, они также указывают на индивидуальные особенности стиля писателя.

Особо можно выделить мастерство драматурга в словообразовании, уместное использование слов в другом смысле. Здесь речь идет о словах, кажущихся непонятными. Но они высоко стилизованы тем, что показывают большое мастерство индивидуального стиля писателя. Их немного (по употребительности), и они очень низкочастотны (около ста словоупотреблений): кеп и дат — слово; кәр — синоним слов қаһар; берен — кымбат мата, мылтык, батырлар сауыты — дорогой материал, вид ружья, снаряжения батыра; самар — сөзге шешен — оратор; сойы — тегі — происхождение чего-либо, кого-либо; зұлмат — жауыз, алапат — негодяй, монта-

ны мінез — характер, кажушийся тихим; аяшым — очень ласковое обращение к женщине; тэу ету — поклонение чему-нибудь священикому; кер — кедергі — кесір — препятствие; бал бала — дөушка, сладкая как мед; сал нөпір — сделать многое.

Некоторые слова требуют толкования обстоятельного: тоғам алу, әдеп — әркен, сөкек қыз, еменде семен, нысапы; өмірдің ал гүлі, шәңгі

Автор умело использовал исконные казахские слова для передачи политико-административных терминов, например: Ұлы совет — жоғарғы совет — Верховный Совет; кіндік — орталық — центр; партияласу — делится на партию; кеңес — совет. Подобное употребление, естественно передавало стилистическую окраску лексики того времени. На вид «недоступные» слова — такие как козетім, жанан, жан — жананым, қачқай и т. д. становятся понятными только в контексте, хотя не имеют «своего лица» как лексические единицы, так как автор использует их только в качестве стилистического средства.

Все драматические произведения М. О. Ауэзова характеризуются наличием различных лексических пластов, что является жанрово-стилистической особенностью языка драматургии писателя.

ЛИТЕРАТУРА

1. Нұрмұқанов Х. М. М. Ауэзов драмалық шығармаларының эмоционалды-экспрессивті лексикасы. Автореферат канд. дисс. Алма-Ата, 1969. С. 27.
2. Кеңесбаев Е. К., Мұсабаев Ф. Г. Қазіргі қазақ тілі. Алма-Ата, 1962; Ысқақов А. Ы. Қазіргі қазақ тілі. Алма-Ата, 1964; Балақаев М. Б., Қордабаев Т. Қазіргі қазақ тілі грамматикасы. Алма-Ата, 1961.
3. Мухтар Ауэзов и современная литература, Алма-Ата, 1989. С. 222—227.

ҚОСЫМША
ПРИЛОЖЕНИЕ

М. О. ӘУЕЗОВ ШЫҒАРМАЛАРЫНДАҒЫ (ДРАМАТУРГИЯ, ПОВЕСТЬ) ЕТІСТІК СӨЗДЕРДІҢ ҮЛЕСТІРІМДІК АЛФАВИТТИ-ЖИІЛІК СӨЗДІГІ

Морфологияның қарастыратын негізгі мәселелерінің бірі — сөз таптары. Сондықтан да, ол тіл білімінде басқа мәселелерге қарағанда әлдеқайда ертерек сөз бола бастады. Олардың тұрақталып айқындалуына, бір тәртіпке түсірілуіне, сондай-ақ олардың грамматикадан алатын орны, ролі, маңызы туралы көптеген белгілі тілші ғалымдарымыз творчестволық еңбек етті. Әсіресе, «бұл мәселеде, яғни қазақ тіліндегі сөз таптарының тұрақталып айқындалуы және олардың ұнасымды бір тәртіпке түсірілуіне ерекше үлес қосқан ғалымдарымыз Қ. Жұбанов, Н. Сауранбаев, А. Ысқақовтар болды» [1, 15-б.; 2, 77-б.]. Дегенмен де, қазіргі қазақ тіліндегі сөз таптары мәселесі — әлі де күн тәртібінен түсе қойған жоқ. Керісінше, морфологияның әлі де болса зерттеліне түсетін түйінді тақырыптарының біріне айналып отыр. Өйткені, «сөз таптары жайындағы мәселелер — тілдің грамматикалық құрылысының, сөздің грамматикалық құрылымының әрі кіндік қазығы, әрі негізгі арқауы» [3, 132-б.].

Соңғы кездегі статистикалық зерттеулерде жекелеген текстер негізінде, ондағы сөзформа, сөзқолданыстар — сөз таптары тұрғысынан айқындалып, олардың ара салмағы сандық мөлшермен көрсетіліп жүр [4, 9-б.; 5, 50-б.; 6, 14-б.]. Бұлардың қай-қайсысында да зат есім, сын есім, етістіктер — қаралынған текстердің негізін құрайтыны, олардың бұл текстер үшін үлес-салмағының аса елсулі екені айқын сезіледі. Оны мына төмендегі кестеден де анық көруге болады:

Текстердегі сөз таптары статистикасы

Сөз табы	«Публицистика» Ә. Ахабаев о/о	«Абай жолы» (роман) о/о	«Балалар әдебиеті» К. Молда- беков о/о
Зат есім	45.19	35.36	37.77
Етістік	21.86	31.44	35.84
Сын есім	12.65	9.35	7.00
Барлығы:	79.70	76.15	70.61
Сөзқолда- ны саны:	150000	455691	100000

Ескерту: Кестеде келтірілген мәліметтер «М. Әуезовтің «Абай жолы» романының жиілік сөздігінен» [1979, 9-б.], Ә. Ахабаевтың [1971, 9-б.], Қ. Молдабековтің [1985, 14-б.] жұмыстарынан алынды.

Қазақ тіліндегі грамматикалық сөз таптарының ішіндегі ең көлемді де, күрделісі — етістік. Сондықтан да болар, етістіктің «лексикалық құрамы көлемді, грамматикалық құрылымы аса күрделі сөз табы» дейміз [7, 5-б.]. Оның тіліміздің грамматикасында алатын орны да, атқаратын қызметі де өте үлкен. Ал, сандық мөлшер жағынан алып қарайтын болсақ, басқа сөз таптарымен салыстырғандағы ара салмағының қандай екенін жоғарыда келтірілген кестеден-ақ көріп отырмыз. Онда ол — зат есімнен кейінгі екінші орынды немденіп тұр.

Етістік сөз табының осындай қасиеттерін ескере отырып сіздердің назарларыңызға М. О. Әуезов шығармаларының ішінен драматургиясымен повесінен түзілген етістік сөздердің алфавитті—жиілік сөздігін ұсынып отырмыз. Бұл — ұлы классик жазушының драматургиясы [8] мен повесіндегі [9] қолданылған етістік сөздердің қолданылу жиілігін салыстыра көрсетуге арналған сөздік.

Жазушының драматургиясындағы (текстеріндегі) қолданылған барша сөздер — «кейіпкер сөзі», «автор сөзі» делініп жеке-жеке бөлініп қаралады да, бұлардан бұл сөздікке «кейіпкер сөзі» ғана алынды. Драмалық шығармалардағы «кейіпкер сөзінің» көлемі — 245 мың сөзқолданыс болса, «повестің» көлемі — 21307 сөзқолданыстан тұрады.

Сөздік екі бағанамен берілген: бірінші бағанада — реестрлік сөздер берілсе, екіншіде — текстердегі [8, 9] қолдану жиіліктері көлбеу сызық [/] арқылы бөлініп берілген. Көлбеу сызықтың алдындағы сан — «драматургия» тексіндегі, сонындағы сан — «повесть» тексіндегі етістік сөздердің кездесу жиілігін білдіреді.

Екі не одан көп текстердегі лингвистикалық единицаларды статистикалық әдіспен салыстыра зерттеу кезінде, текст көлемдерінің шамалас болғаны өте-мөте қажет. Ал, олай болмаған жағдайда қатынас жиілігінің (относительная частота) көмегімен де салыстыру әдісін қолдана беруге болады (әрине, бұл кезде қаншалықты қате кету мүмкіншілігін ескерген жөн).

Айталық, біз қарастырып отырған төмендегі сөздікке алынған текстердің көлемдері де бірдей емес, яғни «драматургия» тексі (245 мың сөзқолданыс) «повесть» тексінен (21307 сөзқолданыс) 11.4 есе үлкен. Не істеу керек?

Сөздікте екі текстегі етістік сөздердің кездесу жиілігі салыстырылып берілген. Айталық, «ағыз» етістігі «драматургияда» 10

рет кездесіп қолданылған болса, «повесте» бұл сөз — 3 рет кездескен. Бұған қарап «ағыз» етістігін «драматургияда» көп, «повесте» аз қолданылады екен деген пікір тұмауы керек. Өйткені, бұл жерде ең алдымен «ағыз» сөзінің әр тектегі қолданыстың жиілігін анықтауымыз қажет. Ол үшін әр тектегі кездескен жиіліктің (абсолютті жиіліктің) сол текстің көлемін білдіретін санға қатынасын табуымыз керек. Мысал үшін «ағыз» етістігінің әр текте қатысты жиілігін анықтап көрейік. Ол үшін 10 санын 245000-ға бөлсек «драматургияға», 3 санын 21307-ге бөлсек «повестке» қатысты жиіліктері шығады:

(«драматургия») — $10 : 245000 = 0,00004$,

(«повесть») — $3 : 21307 = 0,00014$,

Демек, бұл жерде «ағыз» сөзінің «драматургияға» қарағанда «повесте» әлдеқайда жиі қолданылатынын (9 еседей) байқаймыз.

Сөзіміз дәлелдірек бола түсуі үшін тағы бір мысал келтірейік: «Ал» етістігі қазақ текстерінде ең жиі қолданылатын сөздердің бірі (4, 313-б.; 10, 348-б.). Ол «драматургияда» — 2398 рет, «пьесада» — 320 рет кездескен. Сұрақ: бұл сөз қай текте жиі қолданылған?

(«драматургия») — $2398 : 245000 = 0,009$,

(«повесть») — $320 : 21307 = 0,015$.

Демек, «ал» етістігі «повесте» «драматургияға» қарағанда жиірек (2 еседей) қолданылған.

Сондай-ақ, сөздікте бір, екі рет кездесіп, бірақ бірінде бар, бірінде жоқ сөздер (ағыл, ажыра, ажырай т. б.) кездеседі. Әрине, мұндай сөздер сол өзі бар текст үшін маңызды болғанымен, бізге пәлендей сандық сипаттама бере алмайды. Ал, екі текстің бірінде аса көп кездесіп, екіншісінде тіпті кездеспейуі (мыс.: аз, айны, алсын әкел, боп, қып, секіре-секіре, сүйіп, тақ істе, т. б.) мүмкін. Әрине, ол сөз бар текстердің ерекшелігін білдіретіні анық.

«Драматургиядағы» «кейіпкер сөзі» 11814 реестрлік сөзден тұрады. Мұның 5068-і етістік сөздер де (42,9%), олар текте 69009 рет қайталанылып қолданылған. Яғни, жазушы драматургиясындағы «кейіпкер сөзінің» 25,4% камтиды. Ал, жазушының «Қараш-қараш оқиғасы» повесінің тексі — 21307 сөзқолданыстан тұрады. Мұнда 3752 реестрлік сөз бар. Мұның 7171 сөзқолданыстан (33,65%) тұратын 1265 реестрлік сөзі (33,45%) етістік сөздер.

Жалпы, бұл ұсынылып отырған сөздік — екі жанр лексикасын, олардың қолданысын, көрінісін анықтауға, сондай-ақ, әр түрлі мақсаттағы сөздік жасаушы лексикографтар мен жазушы тілін зерттейтін ғалымдар мен ізденушілерге өте қажетті, құнды тілдік материал бола алады деген ойдамыз.

Сөз	Абсолютті жиілік	Сөз	Абсолютті жиілік	Сөз	Абсолютті жиілік
абайла	9/0*	айналасокта	1/0	албырт	6/0
абыржы	6/0	айналыс	3/0	алғыз	19/5
ағалас	0/1	айны	31/0	алғызыс	0/1
ағар	10/1	айныт	2/0	алда	109/4
ағарт	7/0	айрыл	108/7	алдамас	1/0
ағыз	10/3	айрылыс	13/0	алдан	26/1
ағызба	3/0	айт	2863/99	алдандыр	1/0
ағыл	1/1	айт-айт	1/0	алдансын	1/0
ағыт	3/0	айтакта	4/0	алдаңқыра	1/0
ағытыл	2/0	айткыз	77/2	алдарман	1/0
адакта	3/0	айтса-айтпады	1/0	алдат	3/0
адақыл	1/0	айтта	1/0	алдаусырат	1/0
адамшыла	1/0	айтты-айт- пады	7/0	алдыр	19/1
адас	69/3	айттыр	0/1	алдырт	3/0
адаспан	1/0	айтыл	54/2	алжас	1/0
адастыр	13/6	айтып-айтып	2/0	алжы	5/0
адинайтес	2/0	айтыс	39/4	алкылда	0/1
адында	2/0	айық	28/2	алкындыр	1/0
адырай	0/1	айықтыр	0/1	алмасыр	2/0
адырақал	1/0	айыпта	7/0	алмасын	4/0
адырақалғыр	2/0	айыр	82/2	алсын	41/0
адыранда	1/0	айырбаста	2/0	алын	34/3
ажарлан	0/1	айырт	3/0	алынбасын	1/0
ажуала	1/0	айырыл	7/2	алып-соқ- кыз	0/1
ажыра	2/0	ақ	107/13	алып-алып	1/0
ажырай	0/1	ақмақтан	1/0	алынқашты	1/0
ажырас	1/0	ақпақұлақтан	1/0	алып-ұшып	0/1
аз	53/0	ақса	3/0	алыс	115/6
азай	3/0	ақсат	1/0	алыстат	1/0
азайма	1/0	ақсит	0/1	алыстыр	4/0
азайт	3/0	ақта	62/2	алшаңда	3/0
азғыр	15/0	ақтал	16/1	амалда	2/0
азғырыл	1/0	ақтар	13/1	аманда	12/0
аздыр	10/0	ақтармала	0/1	амандас	14/5
азына	4/3	ақтармалат	0/1	аңди	1/0
азынат	1/0	ақтарт	3/0	аңжала	1/0
айбында	2/0	ақтарыл	1/0	аңтұр	3/0
айғайла	14/3	ақылдас	27/2	аңтас	1/1
айда	57/14	ақыр	26/2	аңқта	15/4
айдал	9/0	ақырында	1/0	аңықтал	2/0
айдат	0/1	ақырындат	0/1	аңсар	25/0
айқайла	8/1	ал	2398/320	аңсарт	3/0
айқас	14/1	алабұрт	2/0	аңда	27/0
айқында	1/0	алағыз	0/2	аңды	44/0
айқындал	0/1	алаңда	2/0	аңдыс	5/0
айлан	0/1	алаңдат	0/2	аңқаусы	5/0
айландыр	1/0	алар	3/1	аңқы	7/0
аймала	1/1	аларт	2/0	ақса	14/0
айманда	1/0	аласта	8/0	аңыр	2/0
айнал	76/26	аласұр	8/1	аңыра	9/0
айналдыр	15/4				

анырат	2/0	асыктыр	5/0	ауыздан	1/0
ап	67/1	асыл	14/2	ауыздандыр	1/0
апар	75/9	асылдап	3/0	ауыр	18/3
апарат	4/0	асылдандыр	3/0	ауырла	9/0
апкел	3/0	асылдансын	1/0	ауырмасын	1/0
апкет	5/0	асын	8/1	ауырт	1/0
апта	2/0	асынсын	1/0	ауыс	15/7
аптык	3/0	асын-сасып	2/0	ауыстыр	8/0
араздас	1/1	асып-тасып	1/0	аш	178/7
арала	12/9	асыр	23/0	ашы	21/2
аралас	30/6	асыра	28/1	ашык	0/2
аралат	2/1	ат	162/18	ашыл	77/10
аралатсын	3/0	ата	126/8	ашын	6/2
аралаттыр	12/0	атал	26/4	ашындыр	0/1
арамкат	1/0	аталсын	1/0	ашыр	6/0
арамда	0/1	атама	1/0	ашыктыр	4/0
арамсын	0/1	атап	12/4	ашыс	9/0
аранда	4/0	атандыр	1/0	ашулан	29/1
арандас	4/0	атап-атап	1/0	ашуландыр	2/0
арандат	3/0	атат	2/0	аюсы	1/0
арашала	6/2	атар	1/0	ая	193/10
арба	2/2	атас	0/1	аязда	1/0
арбас	0/1	атастыр	1/0	аякта	8/0
ардакта	4/0	аткар	7/4	аяктал	3/0
арестовать	2/0	аткарыс	1/0	аяктал	1/0
арзы	1/0	аткы	1/0	аяла	7/2
аркала	0/2	аткыз	5/0	аялда	5/0
арна	26/5	атқыла	3/0	аян	24/0
арнал	17/4	атқыр	15/0	аянда	1/5
арналсын	4/0	ашкуылда	1/0	аяс	1/1
арпалыс	3/0	атпасын	1/0	әдеміле	5/0
арсылда	2/0	атсын	22/0	әзілде	2/0
арт	27/4	атта	18/1	әзілдес	1/0
артыл	1/0	аттан	36/9	әзірейлім- сіп	1/0
аруакта	1/0	аттан-аттан	0/1	әзірле	24/0
арыл	48/11	аттандыр	10/5	әзірлен	33/0
арылма	2/0	аттансын	2/0	әзірлес	2/0
арылт	5/1	аттат	1/0	әзірлет	3/0
арыл-талған	1/0	атыл	11/1	әйгіле	1/0
аршы	4/2	атыс	14/1	әйгілен	3/0
ас	182/20	атып-атып	1/0	әйтесіз	1/0
аса	31/1	ау	37/3	әйт	1/0
асат	2/0	ауғас	2/0	әйтсе	1/0
аскакта	12/0	ауғастыр	1/0	әйтсін	1/0
аскактат	1/0	аудар	15/3	әйтті	1/0
аскыз	1/1	аударт	1/0	әйтірге	6/0
аскында	3/0	аударыл	1/0	әке	40/2
аскындат	1/0	аукаттан	1/0	әкел	170/0
аспанда	1/0	аула	3/1	әкет	62/15
аспандат	0/1	аулакта	2/0	әкпел	5/1
аспасын	1/0	аулан	1/0	әкпет	3/0
асын	7/0	ауна	4/5	әкімсі	1/0
аста	4/0	аунаксы	0/1	әлектен	1/0
астыр	3/0	аунат	1/0	әлпште	4/0
асык	70/7	аусарлан	3/0		

әлсіре	9/1	балпанда	1/0	баяула	1/0
әлсірет	1/0	балат	1/0	бәсекелес	2/0
әндет	19/0	балаусала	1/0	бәсендет	0/1
әнгімелео	2/0	балбыра	2/0	без	12/2
әнкілде	1/0	балкы	2/0	бездір	1/0
әпер	38/1	балкыт	2/0	безе	18/1
әперт	1/0	балтала	1/0	безен	2/0
әпкел	30/0	бапта	3/0	безер	3/0
әпкет	11/0	бар	987/76	бейімде	1/0
әулнесі	2/0	барғыз	1/0	бейімдел	1/0
әулнесіт	1/0	барла	10/0	бекте	14,0
әулік	4/0	барсын	20/0	бекі	1/0
әуреле	2/0	барымтала	0/1	бекін	35/15
әурелен	2/0	барыңқыра	1/0	бекіс	2/0
әшкереле	6/0	барып-барып	2/0	бекіт	11/5
әшкерелен	2/0	барыс	3/0	бекіттір	1/0
әшкерелет	1/0	бас	209/35	белгіле	5/2
бағала	9/0	басқар	22/2	белгілен	4/0
бағалаң	1/0	басқарсын	1/0	белдес	4/0
бағып-қағып	1/0	басқарыс	1/0	белсен	3/0
бағын	18/1	басқыз	4/3	бер	1696/156
бағынбасын	1/0	баспала	2/4	бергіз	13/0
бағындыр	1/0	бассал	1/0	бере-бере	1/0
бағыттал	1/0	бассын	3/0	бермәймән	1/0
бадырайт	2/0	баста	237/38	берікте	2/0
бажырай	10/0	бастал	32/7	беріктен	1/0
бажырайт	1/0	басталсын	10/0	берін	34/1
байқа	111/13	басталсын	2/0	беріп-беріп	2/0
байқаған-		бастама	1/0	беріс	12/0
тоқыған	1/0	бастамасын	1/0	бетте	30/4
байқал	3/4	бастасын	6/0	беттес	1/0
байқал-бай-	0/1	бастат	1/0	беттестір	1/0
кап		бастыр	12/3	биле	57/2
байқар	1/0	бастырма-		билеп-төсте	3/0
байқас	1/0	лат	1/0	билет	8/0
байқат	2/1	бастырт	0/1	бипаздан	1/0
байла	98/8	басыл	21/2	бит	16/0
байлап	10/0	басымда	0/1	биікте	1/0
байланыс	6/0	басын	9/0	биязыла	1/0
байлас	11/0	басынт	1/0	бозда	2/0
байласын	1/0	бат	69/7	бойла	5/0
байлат	8/0	баталас	1/0	бойұсын	2/0
байпаңда	1/0	батпақта	1/0	боқта	1/6
байсалдан	1/0	батпасын	1/0	боқтап-	
байы	6/4	батсын	3/0	боқтап	0/1
байызда	2/0	батыр	12/7	бол	6234/496
бақ	99/3	баулы	3/0	бола-бола	1/0
бақсын	1/0	баура	11/0	болғай-ак	2/0
бақтыр	3/0	бауызда	19/0	болды-болды	1/0
бақыла	4/0	бауыздал	2/0	болдыр	3/1
бақылда	2/0	бауыздат	2/0	болжа	24/3
бақыр	16/1	бауырла	3/0	болсын	548/0
бақырай	2/0	бауырымда	1/0	болыс	34/4
бала	16/0	башымыңи	1/0	болсын	1/0
балала	6/0	баянда	2/1	бомбала	1/0

боп	0/32	бұқтыр	1/0	бүтінде	1/1
божы	0/1	бұла	1/0	бүтіндел	1/1
бора	3/0	бұландат	2/0	быксы	2/0
борандат	1/0	бұлға	1/0	былға	14/1
борсы	2/0	бұлғакта	1/0	былған	6/0
бос	16/0	бұлда	5/0	былжырат	1/0
боса	43/0	бұлдан	6/0	былық	3/0
босан	3/0	бұлдант	1/0	былықтыр	5/0
босансы	2/0	бұлдыра	2/1	былшылда	3/0
босат	63/4	бұлдырат	6/0	быршы	1/0
босаттыр	2/0	бұлжытпа	7/0	бытыра	5/1
бостыр	0/1	бұлында	2/1	біл	1308/39
бошал	1/0	бұр	33/2	біледі	1/0
боя	3/0	бұра	6/0	білгей	3/0
бояу	2/2	бұрал	1/4	білгіз	1/0
боялсын	11/0	бұрала	3/0	білдір	65/13
бөге	32/0	бұрғыз	2/1	білдірсін	2/0
бөгел	32/8	бұркат	1/0	біле	43/9
бөгесін	1/0	бұркылда	1/2	білеулендір	
бөк	1/0	бұрма	5/0	білмеген	0/1
бөксер	1/0	бұрмалаңда	1/0	көрмеген	1/0
бөл	20/3	бұрти	3/0	білмегенсі	2/0
бөле	3/0	бұрыл	17/7	білсәң	1/0
бөлен	1/0	бұрып-бұ-		білсін	52/0
бөлмесін	1/0	рып	0/1	білтеле	4/0
бөліп	8/2	бұтақта	1/0	білушімең	3/0
бөліп-бөліп	0/1	бұтарлас	1/0	білін	12/5
бөліп-жа-		бұтты-батты	1/0	біліс	2/2
рып	2/0	бүгіл	5/1	біртінде	1/1
бөлшекте	2/0	бүйде	1/0	бірік	27/0
бөрліктір	1/0	бүйт	65/5	біт	173/20
бөрік	1/0	бүк	15/0	бітеулен	2/0
бес	2/0	бүксі	1/0	бітсін	4/0
бөтенсі	1/0	бүксін	1/0	бітір	53/2
бу	16/1	бүкте	4/3	бітіс	2/0
бұдырат	1/0	бүктел	4/1	важничает	2/0
була	1/0	бүктетіл	1/0	величать	1/0
булды	1/0	бүктүс	1/0	вспомни	1/0
булық	5/1	бүкши	0/1	встать	1/0
бусан	2/1	бүлдір	35/0	глядит	2/0
бусын	1/0	бүлдірт	2/0	глядя	1/0
буын	6/1	бүлкілде	0/1	говорить	5/0
буындыр	2/1	бүлкілдет	0/1	гөй-гөйле	2/0
бұз	145/4	бүлін	19/0	гөр	5/0
бұзаула	2/0	бүлініс	0/1	гүлде	2/0
бұздыр	1/0	бүр	9/0	гүлдендір	1/0
бұзыл	35/3	бүрк	1/0	газапта	1/0
бұзыс	0/1	бүрке	9/0	гайыпта	4/0
бұйдала	1/0	бүркемеле	1/0	ғыл	95/0
бұйыр	140/2	бүркен	3/0	дабылдат	1/0
бұйырт	1/1	бүркес	1/0	дабырай	1/0
бұйрыл	2/0	бүрсенде	0/1	дабырла	3/1
бұк	14/3	бүріс	1/0	дабыста	2/1
бұқпантайла	1,0	бүрістір	1/0	давай	2/0

дағдар	4/0	дүрди	0/1	ентелет	4/0
дағдылан	0/1	дүрсе	2/1	еңбекте	2/0
дайында	9/2	дүркіре	1/1	еңкей	9/1
дайындал	5/1	дүрсілде	0/1	еңкейіңкіре	0/1
дайындат	2/0	дүрілдет	1/0	еңсер	3/1
далдала	2/0	дүңкілде	1/0	еңсеріп-	
далой	2/0	дірілде	2/3	жапырып	0/1
дамы	4/0	с	7025/252	еңіре	19,0
дандайсы	1/0	еге	5/0	еңірет	3/0
даноста	1/0	егес	29/1	епте	2/3
даңғырла	1/0	егізде	1/0	ептеп-ептеп	0/1
даңғырлат	1/0	егіл	5/0	ер	146/5
данктыр	1/0	ез	10,0	ергіз	1/0
дары	5/0	езбеле	1/0	ерегес	1/0
дарыт	5/0	езгіле	1/0	ерегіс	4/0
даула	1/0	ездір	1/0	ерсуілде	2/0
даулас	10/1	езіл	14/0	ержет	5/1
даурық	15/2	езуле	1/0	еркеле	4/0
дауылдат	2/0	ек	30/1	еркелет	4/0
дауыста	3/2	екпінде	1/4	еркінде	1/0
дәлде	5/2	екше	3/0	еркінсі	2/0
дәлелде	3/0	екшел	2/0	ернеуле	0/2
дәмелен	1/0	екілен	2/2	ерт	37/2
дәмет	5/0	елбіре	2/0	ертте	0/3
дәнде	1/0	еле	19/1	ерттел	0/1
дәнін	1/1	елгіз	0/1	ерік	8/0
дәріле	1/0	елен	1/0	ерін	5/0
де	4712/349	елен-ескер	1/0	еріт	7/0
дегенсі	0/1	елер	1/0	ес	16/0
дегіз	8/0	елерт	1/0	есей	1/0
дедектет	1/0	елесте	0/1	еселеп-есе-	
демал	1/0	елеусізден	0/2	леп	1/0
деме	10/2	елжіре	7/0	есеңгіре	1/0
денде	1/2	елжірет	2/0	есенте	18/0
дерттен	2/0	елнендет	2/0	есептес	1/0
дес	16/16	елті	4/0	ескер	16/1
диәрмін	1/0	елікте	2/0	ескерт	6/1
дигән	3/0	еліктір	1/2	ескертіл	2/0
дидарлас	1/0	елір	8/3	ескеріл	1/0
доғар	44/0	елірт	1/0	ескір	4/0
долдан	2/2	ем	513/4	есті	270/11
долой	1/0	емде	16/0	естіл	30/7
домала	3/2	емдел	2/0	естіп-біліп	1/0
домалат	2/0	емексіт	5/0	естіп-көріп	1/0
домбық	1/0	емен	40/0	естірт	12/0
домбытпала	1/0	еміз	2/0	есір	7/0
доноста	1/0	емін	2/0	есірке	6/1
дорбала	1/0	емірен	2/0	есіркеп-мү-	
дөңгелет	1/0	емірент	1/0	сірке	1/0
дудыра	1/0	ен	13,0	есіт	330/4
дула	2/0	енгіз	3/0	ет	9/35
дұрыста	5/0	енде	1/1	еткіз	14/3
дүңки	0/1	ендет	0/1	етпетте	0/1
дүңкілде	0/1	ентеле	2/0	етіл	7/0

етіс	4/2	жақында	20/14	жанып-	
жабық	5/0	жақындас	3/0	жанып	1/0
жабықтыр	1/0	жақындас-	0/1	жанып-	
жабыл	30/9	тыр		сөнген	1/0
жабырка	0/1	жақындат	2/0	жаныс	1/5
жабыс	50/5	жала	8/1	жаныш	0/1
жабыстыр	1/0	жалала	2/0	жанышта	1/0
жағала	19/0	жаламасын	1/0	жаншы	3/0
жағалас	12/0	жалаң	6/1	жаншып-	
жағалат	6/0	жалаңашта	3/0	қырып	0/1
жағыл	0/1	жалаңда	0/1	жаңғыр	0/1
жағын	9/1	жалас	7/0	жаңғырт	7/1
жағыстыр	1/1	жалбақта	1/0	жаңғырық	1,0
жадағайлат	2/0	жалбарын	3/0	жаңсақтат	1/0
жадыра	5/1	жалбырла	1/0	жаңыл	41/2
жадырат	1/0	жалғас	11/2	жаңылдыр	1/0
жаз	249/7	жалғызсыра	0/1	жаңылыс	2/0
жазала	21/3	жалда	2/0	жап	78/2
жазғыз	0/1	жалдан	3/1	жапкыз	6/0
жазғыр	25/0	жалма	3/2	жапсарлас	1,0
жазда	13/0	жалли	1/1	жапсар	3/0
жаздыр	2/1	жалпиған-		жаптыр	4,0
жасын	1/0	талтнған	1/0	жапыр	3/1
жазыл	43/1	жалпит	2/0	жапырлас	1/1
жазыс	1/0	жалпылда	1/0	жапырыл	5/0
жай	42/1	жалтар	12/7	жар	33/5
жайға	1/1	жалтарт	1/0	жара	131/10
жайғас	8/1	жалтыра	1/1	жарал	9/0
жайғастыр	2/0	жалық	2/0	жарала	12/0
жайдақта	3/0	жалын	40/0	жаралан	2/0
жайдыр	1/0	жалында	17,0	жаран	1/0
жайқа	2/0	жалындыр	0/1	жарас	37/2
жайқал	2/2	жама	5/0	жарасын	1/0
жайла	22/5	жамал	5/0	жарастыр	3/0
жайлама-		жаманда	8/1	жарат	23/3
сын	1/0	жамандат	1/0	жарғыла	2/0
жайлан	3/3	жамап-жас-		жарғылас	2/0
жайлансын	1/0	қап	1/0	жарияла	17/1
жайлас	2/0	жамбаста	5/1	жариялан	1/0
жайна	13/1	жамыл	18/3	жаркылда	2/1
жайнаңда	1/0	жамылт	0/1	жаркыра	16/2
жайнат	3/0	жамырат	1/0	жарылқа	3/0
жайра	10/0	жан	43/1	жармас	8/1
жайрағыр	4/0	жаңа	1/0	жартысын	1/0
жайраңда	1/0	жанас	8/4	жартыла	1/0
жайрат	12/1	жанасала	0/1	жарыл	19/4
жайыл	18/10	жанасың-		жарылға	5/4
жайынсы	0/2	қыра	0/1	сын	17/0
жақ	78/3	жандарал-		жарылқа	2/0
жақаура	1/0	сы	1/0	жарылқа-	
жақаурат	1/0	сандыр	2/0	сын	2/0
жақсыла	7/1	жанжалдас	2/0	жарылма-	
жақта	11/1	жансың	6/0	сын	2/0
жақтыр	8/1	жаны	11/1		

жарылсын	1/0	жеккөр	1/0	жиып	4/0
жары	17/0	жеккіз	1/0	жиыр	1/0
жарыс	11/3	жектір	1/0	жиыстыр	1/0
жарыса-жарыса	1/0	жел	1/2	жиіле	0/2
жарыт	5/0	желбіре	1/1	жиіркент	1/0
жаса	165/11	желкеле	1/0	жоғал	75/5
жасал	8/0	желкілде	1/0	жоғалма-сын	1/0
жасан	5/1	желпі	1/1	жоғалсын	4/0
жасаңсыт	1/0	желпін	2/1	жоғалт	20/1
жасар	3/6	желік	8/0	жаяула	0/2
жасарт	1/0	желіктір	1/0	жой	63/0
жасартсын	1/0	жеме	13/0	жойсын	1/0
жасас	4/0	жемесін	1/0	жойыл	9/0
жасасын	11/0	жентекес	2/0	жойылсын	3/0
жасат	2/1	жең	101/7	жоқта	16/1
жасаура	0/1	жеңгіз	1/1	жоқтат	5/0
жаска	2/0	жеңдір	1/0	жола	12/1
жасқан	15/1	жеңілдет	48/1	жоласын	1/0
жасқандыр	0/1	жеркен	1/1	жолат	6/2
жастаң	2/1	жерле	2/0	жолаушыла	4/0
жасы	22,1	жерлесін	4/1	жолда	3/0
жасыр	48/3	жерсін	1/0	жолык	17/0
жасырып	9/2	жері	1/0	жорамалда	3/0
жасырып-сын	1/0	жерін	1/0	жорғала	2/2
жасырыс	2/0	жес	0/1	жорт	13/2
жасыт	12/0	жесін	5/0	жоры	2/0
жат	413/136	жет	272/22	жорыт	2/5
жага-жас-тана	0/1	жетекте	9/3	жосы	2/2
жаткыз	7/3	жетектесін	1/0	жосылт	3/0
жатын	6/0	жетектет	5/1	жосыт	2/0
жатта	3/0	жетеле	5/2	жотала	1/0
жаттык	2/2	жеткіз	65/8	жөде	5/0
жатыр	345/28	жеткізсін	3/0	жөпе	1/0
жатырқа	1/0	жетсін	29/0	жөнел	88/12
жау	41/4	жетіл	7/0	жөнелмесін	1/0
жауаптас	2/2	жетілдірсін	1/0	жөнелт	16/1
жаугыз	2/0	жетімсіре	0/1	жөнелтсін	1/0
жаудыр	5/0	жетімсірет	1/1	жөнеліс	1/1
жаула	0/2	жетімсіт	1/1	жөңкіл	1/2
жаулаң	1/0	жетіс	6/1	жөңкілт	1,1
жаулас	11/0	жетістір	2/0	жөтел	0/1
жаура	2/0	жи	108/13	жөткір	1/0
жауық	4/0	жиғыз	4/0	жу	14/3
жаяула	3/0	жиді	3/0	жуанда	1/0
жәрдемдес	3/0	жина	21/4	жуансы	0/1
же	121/10	жинал	19/4	жуасы	1/0
жебе	2/0	жиналсын	1/0	жуасыт	2/0
жебеле	3/0	жирен	18/2	жуса	3/0
жеделдет	2/0	жирент	3/0	жусат	2/0
жей-жей	1/0	жисып	3/0	жуып	1/0
жск	8/0	жиыл	43/3	жуып-жай	1/0
		жиылсын	2/0	жуыт	3/0
		жиылыс	1/0	жұбан	2/0

жұбапта	1/0	жылжи-		зорайт	1/0
жұбат	8/2	жылжи	1/0	зорла	11/0
жұғыс	1/0	жылжы	5/2	зорықтыр	1/0
жұқ	6/1	жылми	2/0	зула	4/1
жұқала	1/0	жылмит	1/0	зымыра	1/0
жұқтыр	2/2	жылпылда	0/1	и	12/0
жұл	24,7	жылпында	1/0	иді	4/0
жұлқы	1/1	жылушыра	0/1	некте	5/0
жұлқыла	0/1	жылы	2/0	немде	1/0
жұлқын	6/1	жылын	7/1	немден	1/0
жұлқыс	2/0	жылыс	6/0	изе	0/2
жұлсын	1/0	жылыстан-	0/1	нкемде	3/0
жұлын	3/0	қыра		нлан	17/0
жұлыс	3/0	жылыт	1/2	нле	2/0
жұм	9/1	жымдас	1/0	нмен	5/0
жұмбақта	6/0	жыми	3/3	нреңде	1/0
жұмдыр	1/0	жындан	16/0	нтер	0/1
жұмса	34/13	жыр	7/0	нтермеле	0/1
жұмсал	2/0	жырақтат	1/0	нтеріңкіре	0/1
жұмсалсын	1/0	жырла	15/0	нтшіле	3/3
жұмсар	2/0	жырт	13/1	йоғал	1/0
жұмсарт	3/1	жыртыл	1/0	ні	6/0
жұмсаат	3/0	жырыл	4/1	ніл	2/0
жұмыр	4/0	жырыс	1/0	нілт	2/0
жұмырт-		жібер	338/60	нілтірес	1/0
кала	2/0	жібермесін	1/0	нір	0/1
жұт	17/5	жіберсін	1/0	ніріл	0/1
жұта	4/0	жіберіл	6/0	ніске	8/1
жұтат	1/0	жіберілсін	1/0	қапұт	2/0
жұтсын	2/0	жіберіс	1/0	қадірле	3/0
жұтыл	0/3	жібі	7/1	кәмпескеле	1/0
жұтын	0/1	жібіт	3/0	кәмпескелен	1/0
жүгін	4/0	жікте	1/0	кәукілде	1/0
жүгір	75/4	жіктен	1/0	кебірсі	0/1
жүгір-жүгір	1/0	жіңішкеле	0/1	кез	15/3
жүгірт	2/2	жіңішкер	3/3	кезгіз	0/1
жүде	25/2	заберите	1/0	кезде	1/0
жүдет	10/0	завоет	1/0	кезден	1/0
жүз	2/0	заплагет	1/0	кездес	66/4
жүкте	4/0	запони	1/0	кездестір	1/1
жүнде	3/0	зарла	10/0	кезе	5/0
жүнжі	1/0	зарлат	4/1	кезен	3/0
жүр	1397/124	зарлық	3/0	кезер	1/0
жүргіз	25/1	зарықтыр	3/1	кезерт	1/0
жүр-жүр	1/0	застучит	1/0	кезік	15/1
жыбырла	3/0	заулат	3/0	кейі	4/0
жығыл	46/10	зашумит	2/0	кейінде	1/0
жығылсын	1/0	зәрлен	0/1	кейт	2/0
жый	0/1	звенит	2/0	кеке	31/0
жық	31/4	звонит	3/0	кекес	1/0
жыла	164/11	зекі	1/3	кекет	1/3
жылай-		зертте	20/0	кекжекте	1/0
жылай	1/10	знать	5/0	кекжи	0/1
жылат	34/1	зовут	1/0	кексіні	1/0

кектен	2/1	кетір	28/1	көр	1450/87
кел	2264/233	кетіріл	3/0	көрген-	
келген-кеткен	2/0	кетіс	14/1	баққан	1/0
келді-кетті	1/1	көш	81/5	көргіз	1/0
келер-келмес	1/0	кешік	33/0	көре-көре	1/0
кел-кел	2/0	кешіктір	1/0	көркей	3/0
келер-келер	0/1	кешір	28/4	көркейт	1/0
келтір	58/5	ки	63/6	көрсет	146/20
келіп-келіп	2/1	кигіз	9/0	көріп	193/33
келіп-кеткен	1/0	кидір	2/0	көріп-біліп	3/0
келіс	28/1	кил	2/0	көріп-көріп	0/1
келістір	6/0	килік	23/0	көріп-танып	1/0
кемсеңде	0/1	кимеле	2/0	көріс	37/0
кемсіп	3/1	кит	1/0	көрістір	2/0
кемсіт	2/0	киіп	20/3	көсемсі	1/0
кемі	4/2	киіндір	4/0	көсі	1/0
кеміңкіре	0/1	конфескеле	9/0	көсіл	2/0
кемір	7/0	конфиске-		көтер	113/15
кеміт	5/0	лен	3/0	көтере-кө-	
кене	6/0	көбей	14/5	терс	1/0
кенел	1/0	көбейт	6/1	көтерт	1/0
кенелт	2/0	көбіктен	4/0	көтеріл	16/9
кенжеле	2/0	көгендет	1/0	көтерілсін	1/0
кеңейт	1/0	көгер	4/1	көтеріңкіре	0/1
кеңес	9/1	көгерт	6/0	көтеріс	1/0
кеңкілде	2/0	көзде	13/3	көтібарла	1/0
кені	1/0	көзден	1/0	кеш	71/1
кеп	258/10	көздес	0/1	көрсенші	0/1
кептел	1/1	көздестір	2/0	көшір	19/2
кептес	2/0	көзс	1/1	көшіріл	1/0
кептір	2/0	көзеп-көзеп	2/0	крой	1/0
кер	16/2	көздестір	2/0	крутить	1/0
кергі	1/0	көйт	1/0	кур	1/0
кердеңде	2/0	көкп	1/0	күбін	1/0
керне	12/2	көксе	11/0	күбірле	1/0
кернейлет	1/0	көкте	4/0	күбірлес	1/0
керт	0/1	көкі	6/0	күдіктендір	1/1
керіл	2/0	көлгірсі	3/0	күжірей	0/1
керіс	3/0	көлдененде	1/2	күзе	6/1
кес	88/8	көлденендет	1/2	күзет	10/5
кесекте	0/1	көм	16/1	күзетіл	1/0
кесес	5/0	көмектес	3/0	күй	64/5
кес-кес	1/0	көміл	0/1	күйбеңде	0/1
кес-кесте	2/0	көн	143/9	күйдір	18/0
кескіз	0/1	көндір	19/0	күйеушіле	2/0
кескіле	1/0	көнші	6/0	күйзел	3/0
кескілес	12/1	көңілден	3/1	күйзелт	6/0
кестеле	1/0	көңілдендір	1/0	күйлен	4/0
кестір	1/0	көңсін	1/0	күйрег	3/0
кесіл	11/5	көпте	2/0	күйсе	0/1
кесіп-кесіп	1/1	көптеп-		күйш	14/0
кесіп-піш	1/1	көмекте		күйшідір	1/0
кесіс	1/1	көптір		күйш-жан	1/0
кет	1416/131	көпір		күйш-ортен	1/0

күл	104/10	қаға-қаға	1/0	қақа	2/0
күлгірсі	4/0	қағын	6/0	қақай	1/0
күлдір	1/3	қағыс	7/3	қақал	4/0
күлкілде	1/0	қағыстыр	1/1	қақпа	4/0
күлімсіре	8/1	қағыт	3/3	қақпайла	2/0
күліңкіре	0/1	қада	5/1	қақпала	2/0
күліс	0/4	қадағала	1/0	қақса	2/0
күмілжі	1/0	қадал	22/15	қақсат	3/0
күнде	1/0	қадалт	0/12	қақсын	1/0
күнелт	1/1	қадалма	2/0	қақта	1/0
күңкілде	1/0	кадірле	10/0	қақтық	2/4
күнірен	12/0	қажа	5/0	қақтықпа	2/0
күнірент	8/0	қажала	2/0	қақтықтыр	1/0
күпін	1/0	қажы	41/0	қақтыр	2/1
күргейле	3/0	қажып-		қақшит	1/0
күре	1/0	талып	2/0	қақылда	1/0
күренде	0/1	қажыт	8/0	қақырын	1/1
күрес	1/0	қаз	27/0	қал	1788/120
күркіре	3/0	қазал	1/0	қала	34/2
күрмел	2/0	қазбала	2/0	қалан	1/0
күрсін	8/6	қаздыр	3/0	қалас	2/0
күрсіндір	0/1	қазыс	2/0	қалбақта	0/1
күт	127/12	қайғыр	4/0	қалғы	16/5
күткіз	1/0	қаймық	12/0	қалдыр	50/5
күттір	4/3	қайна	22/10	қалжақта	3/0
күтіл	1/0	қайнама	2/0	қалжыңда	3/0
күтін	2/0	қайнас	1/0	қалжыра	4/1
күтірле	1/0	қайнат	6/3	қалқақта	1/0
күтіс	1/0	қайра	3/1	Қалқы	4/1
күшей	2/0	қайранда		қалмай-	
күшейт	5/1	қайрап-	1/0	талмай	0/1
күшікте	1/0	қайрап	1/0	қайырсын	30/0
кідір	4/0	қайрат	1/0	қалтыра	10/2
кіжін	2/2	қайраттан	1/1	қалтырат	0/1
кінала	24/1	қайрыл	9/1	қалыс	3/1
кіналас	4/0	қайсарлан	0/1	қалшылда	1/0
кінәластыр	1/0	қайт	473/22	қама	23/4
кір	179/22	қайтала	4/0	қамал	9/0
кіргіз	20/7	қайтар	39/2	қамалт	1/0
кіржи	0/1	қайтарт	1/0	қамат	2/0
кірде	2/2	қайтарыл	2/0	қамда	7/0
кіріс	48/2	қайт-қайт	2/0	қамдан	24/5
кірістір	2/0	қайтсын	15/0	қамдат	4/0
кісенде	4/0	қайтыс	1/1	қамты	6/0
кісімсі	2/0	қайшылан	1/0	қамшыла	3/0
кісіне	3/2	қайшылан-		қамшылан	0/1
кісінес	0/1	дыр	0/1	қамшылап-	
кішірей	2/0	қайыр	22/1	қамшылап	0/1
қабатта	0/1	қайырсын	1/1	қамық	4/0
қабаттал	0/1	қайырт	2/0	қан	8/2
қабылда	13/1	қайрыл	3/0	қана	9/1
Қабырғала	0/1	қайыс	8/0	қанағат-	
қабыс	2/0	қайыстыр	1/0	тан	2/0
		қақ	43/8		

қанағат- тандыр	1/0	қарғап- сілеп	1/1	қауіптен	1/0
қанат	2/0	қарғы	0/1	қаш	188/27
қанаттан	2/0	қарлық	1/0	қашқалақ- та	2/0
қанаттас	1/0	қарма	7/0	қашқан- пысқан	1/0
қанда	1/0	қармала	1/0	қашпа	3/0
қандат	1/0	қарман	0/2	қашып- лысып	2/0
қандыр	6/1	қарсыла	1/1	қашыр	9/0
қанжарлас	1/0	қартай	8/1	қи	60/2
қансыра	2/1	қарулан	5/0	қиқанда	1/0
қантала	3/4	қарулан- дыр	1/0	қиқылда	1/0
қанғақса	1/0	қарусыз- дандыр	2/0	қиға	1/0
қанғы	13/2	қарыздан	2/0	қиғалда	8/0
қанғыр	6/2	қарын	1/0	қиғна	41/1
қанғыра	1/0	қарыс	11/1	қиғнал	34/2
қанғырақ- тә	1/0	қарыста	2/0	қиғрат	3/0
қанғырла	1/0	қасар	4/0	қиғралап	0/1
қанғырт	2/2	қасарыс	2/0	қиғсай	1/0
қанғыт	2/1	қасас	1/0	қиғсайт	2/0
қанқақса	1/0	қаси-қаси	2/0	қиғсайың- қыра	0/1
қаңқи	1/0	қасқай	3/0	қиғсындыр	1/0
қаңсыла	1/0	қасқайт	1/0	қиғыл	6/1
қаңыра	3/0	қасқар	1/0	қиғында	5/0
қаңырат	2/0	қаста	1/0	қиғыс	1/2
қан	92/3	қасы	1/0	қиғяла	3/2
қапалан	1/0	қасыт	3/0	қиғялда	2/0
қапсын	1/0	қат	71/8	қиғянатта	2/0
қапта	43/5	қатай	5/1	қобалжы	8/2
қапта-	1/0	қаталас	3/1	қобалжыт	1/0
қаптал	0/1	қатар	2/1	қоз	2/0
қапталда	0/4	қатарлан	1/0	қозға	12/1
қаптал	1/0	қатарлас	3/2	қозғал	22/9
қара	303/53	қателес	19/0	қозғалақ- та	0/1
қарабас	1/0	қатнас	3/2	қозғалыс	0/1
қарай	51/6	қатнастыр	1/0	қозда	4/1
қарай-	1/0	қатпа	11/1	қоздат	2/0
қарайлас	2/0	қатсын	3/0	қоздыр	5/2
қарақта	2/0	қатта	1/1	қой	791/37
қарал	7/0	қаттама	1/0	қойғыз	9/0
қарала	6/3	қатулан	4/1	қойдыр	1/0
қаралас	1/0	қатынас	1/0	қой-қой	1/0
қаран	3/0	қатып- сөніп	1/0	қойсын	26/0
қарас	19/3	қатыр	11/1	қойыл	12/0
қарастыр	4/1	қатыс	9/1	қойыс	0/1
қарат	7/0	қаудырла	1/0	қоқай	0/1
қарауылда	0/3	қаула	2/0	қоқалап	1/0
қарауыт	0/4	қаумала	1/0	қоксыт	1/0
қарбыт	0/1	қауса	1/1	қолбақта	1/0
қарға	22/0	қаусат	3/0	қолда	4/0
қарғама	3/0	қаусырмала	2/1	қолдан	11/0

қомақта	1/0	қуандыр	1/0	құстыр	4/0
қемдан	0/1	қуант	17/1	құтқар	59/6
қомсын	3/0	қуаныс	1/0	құттықта	19/1
қомда	1/0	қуар	52/5	құтыл	65/11
қон	34/8	қуғыз	0/2	құтыр	53/4
қонақта	0/1	қудала	2/1	құтырт	10/1
қондыр	8/0	қудыр	2/0	құш	17/0
қонжи	1/0	қура	6/0	құшақта	11/5
қоңыл-		қусыр	3/0	құшақтас	5/1
тақсы	2/0	қусырыл	0/1	құшырлан	1/0
қоңыраулат	1/0	қутында	0/1	қыбырла	7/2
қоңырлат	1/0	куыл	5/0	қыбырлат	1/0
қоңырсы	6/0	қуыр	2/1	қыдыр	7/2
қоңырсыт	1/0	құбыжық-		қыдырт	1/0
қопанда	4/0	тапдыр	0/1	қыз	46/0
қопар	1/0	құбыр	11/2	қызар	5/3
қопарыл	0/1	құдалас	0/1	қыздыр	4/1
қораздан	1/2	құй	30/1	қызық	55/5
қорала	0/1	құйыл	3/1	қызықта	10/0
қоралай-		құла	10/11	қызықтыр	6/0
қоралай	0/1	құлазы	5/1	қызылдан	0/1
қорбаңда	1/0	құлай-		қыл	651/71
қорға	62/0	құлай	0/1	қылғыз	0/1
қорғал	2/0	құлақтан-		қылдыр	4/1
қорғала	7/0	дыр	2/0	қылжақта	1/0
қорған	8/1	құлат	23/3	қылжи	2/0
қорғанда	1/0	құлашта	1/0	қылқа	1/0
қорғат	2/0	құлда	0/1	қылқи	2/0
қорғашта	2/1	құлдан	0/1	қылқылда	1/0
қорқып-		құлдила	1/1	қылпы	1/0
пысып	1/0	құлдилат	1/1	қылтит	1/0
қорқыт	20/5	құлдырат	1/1	қылтында	1/0
қорла	23/1	құлпыр	1/0	қылымсы	4/0
қорлан	1/2	құлшын	4/1	қылыс	0/1
қорли	1/0	құмарт	1/0	қылыштас	1/0
қорша	6/0	құндызда	0/1	қымбатта	0/1
қоры	4/0	құнтта	1/0	қымсын	2/0
қорық	108/9	құнтит	1/0	қымтан	2/1
қорылда	1/0	құпта	7/0	қынжыл	5/0
қорыт	12/0	құр	60/0	қыңқылда	1/0
қорытыл	1/0	құра	18/0	қыңырай	2/0
қос	131/15	құрал	6/0	қыңыратқы	4/0
қосақта	2/0	құрас	1/0	қып	0/17
қоста	51/4	құрастыр	1/0	қыпылда	3/0
қостат	2/0	құрғат	1/0	қыр	56/1
қостыр	3/0	құрметте	2/1	қырғыз	3/0
қосыл	70/6	құрсан	3/0	қырдыр	1/0
қосметте	0/1	құрсын	22/0	қырқ	3/0
қошқыл-		құрт	47/2	қырқыл	4/1
дандыр	0/1	құры	95/5	қырқыс	4/0
қоштас	15/0	құрықта	1/0	қырлан	0/1
қу	141/10	құрыл	5/1	қырсық	5/0
қуала	4/3	құс	1/0	қырсықтан	1/0
қуан	67/3	құса	6/0	қырық	2/0

кырыл	17/0	марды	1/0	мүдiр	1/0
кырында	1/0	масайра	3/1	мүдiрт	1/0
кырып-жой	1/0	масайрат	1/0	мүжi	1/0
кырып-өрте	1/0	масатсы	1/0	мүйзде	3/0
кырыс	1/0	маскарала	1/0	мүнкi	3/0
кыс	19/11	мата	10/0	мүшеле	0/1
кыскарт	21/2	маужыра	1/0	мыгымда	1/0
кыста	4/0	мауык	1/0	мыжыма	1/0
кыстат	3/0	мәдехала	1/0	мызгы	5/0
кыстыр	4/1	мәймеңкеле	2/0	мыкта	11/7
кысыл	38/3	мәлiмде	1/2	мыкшында	1/0
кысылма	1/0	мәлiмден	1/0	мылжала	1/0
кысыра	1/0	мәлiмдесiн	1/0	мыңгыра	1/0
кышы	3/0	мәпеле	1/0	мыңкылда	1/0
лайла	3/0	мегде	1/1	мін	56/26
лақ	3/0	мегзес	1/0	мінбе	1/0
лақтыр	5/1	мекенде	3/1	мінгiз	13/5
ласта	2/0	мелши	0/1	міндiр	0/1
лаула	1/1	менсiн	11/0	міне	16/0
ләйлі	1/0	меншіктен	1/0	мінес	1/0
ләнетте	1/0	менгер	6/0	мини	3/0
ләпiлде	0/1	меңдет	1/0	мінсiн	2/0
ләпiр	2/0	мензе	1/0	мініс	0/2
лоблы	4/0	мертiк	3/0	мінгестiр	0/1
лупiлде	1/0	мертiктiр	1/0	мінгiрле	1/0
лык-лыкта	0/1	мешай	1/0	неғыл	5/0
лыкылда	2/0	миаула	1/0	назалан	3/0
лыпы	0/1	мобиля-		найзала	1/0
любить	1/0	зацияла	2/0	наказать	2/0
мадакта	4/0	мойы	2/0	накұрыс-	
мазақта	4/3	мойында	5/1	тан	1/0
мазала	1/0	мойындат	1/1	налы	6/0
мазда	0/2	мойынсұн	2/0	налын	1/0
маздат	0/1	молай	4/2	намыстан	1/1
майла	1/0	молайма	1/0	нан	15/0
майыс	12/0	молайт	3/1	нандыр	1/1
мақта	22/2	молык	0/1	наныцкыра	1/0
мақтал	1/1	молықтыр	1/0	написать	1/0
мақтан	15/0	молчать	1/0	нарла	1/0
мақтандыр	2/0	монтансы	1/0	научить	2/0
мақтасык	2/0	мөлдiре	2/0	нолетте	1/0
мақұлда	8/0	мөлтiлде	1/0	неғыл	18/1
мақұлдат	1/0	мөлшерле	0/1	неселес	1/0
малдан	3/0	мөңкi	2/0	немкұра	1/0
малма	1/0	мөңкiр	1/0	несметь	1/0
малты	0/1	мөңiре	1/0	нешли	3/0
малтық	2/1	мұзда	2/0	нобайла	0/1
малынт	1/0	мұзданкыра	0/1	ноктала	1/1
мандыма	1/0	мұздат	2/1	нұска	11/3
мандыт	1/0	мұқал	1/0	нығай	3/0
маңайлас	1/1	мұқат	1/0	нығайт	1/0
маңғаздан	1/0	мұнарт	0/1	нығызда	1/2
маңыра	2/0	мұсдала	1/0	нығыт	1/0
мардамсы	2/0	мұтыл	1/0	обалсын	1/0

обижать	1/3	оп	2/0	өйде	3/0
обый	1/0	опыр	0/1	өкпеле	13/3
оз	36/1	ор	2/1	өксі	1/0
озсын	1/3	ора	24/5	өксіп-	
ой	13/1	орағыт	1/3	өксіп	1/1
ойғар	1/3	орайла	1/3	өкте	8/1
ойла	273/34	орал	37/3	өктеме	1/0
ойлама	1/3	оралт	0/1	өктет	1/0
ойлан	67/9	оран	1/2	өкшеле	1/1
ойлап-		орна	16/0	өкін	22/1
ойлап	1/1	орналас	9/0	өкінбе	8/0
ойлас	7/3	орналастыр	8/0	өкіндір	1/1
ойлат	4/1	орнат	12/1	өкір	4/2
ойлатпа	3/3	орнатыс	1/0	өл	359/19
ойна	65/4	орнык	1/1	өлдім-	
ойнай-		орныктыр	1/2	талдым	2/0
ойнай	4/3	орыл	0/1	өледет	12/0
ойнақта	1/3	орында	42/2	өлерменден	1/3
ойнакшы	2/1	орындал	13/1	өлмесін	1/0
ойнакылан	0/1	орындalma	2/3	өлсін	6/0
ойнас	1/0	орындат	2/1	өлтір	163/3
ойнат	1/1	орындатпа	1/0	өлтірт	11/0
ойранда	2/0	орышала	2/0	өлше	5/3
ойсыра	1/0	орышалас	1/3	өлшестір	1/0
ойсырат	1/3	орыткы	2/0	өліп-тал	2/0
ойыл	1/3	ос	1/3	өліп-тіріл	0/1
ойыс	6/3	оскорблять	3/3	өліс	11/0
ойыстыр	3/3	осқыр	0/2	өн	8/1
оқтал	4/1	оснақта	18/0	өнді	1/3
оқы	228/11	осы	5/0	өндір	8/0
оқыған-		осыл	1/0	өңгер	1/1
тоқыған	1/3	осыр	2/3	өн	3/0
оқымын	1/3	ота	2/3	өр	22/3
оқып-оқып	1/0	отама	1/0	өрбі	3/1
оқыран	1/2	отарла	0/1	өркні	5/1
оқысын	2/3	отта	11/1	өркенде	1/0
оқыт	20/4	оттас	1/0	өркендет	1/0
оқытсын	1/0	отык	2/3	өркештен	0/1
олқысын	5/3	отыр	723/135	өрле	1/7
омағат	1/3	отырғыз	7/1	өрлесін	1/0
омакап	2/0	отырсын	6/0	өрлет	2/0
омыраула	1/1	отырт	4/1	өрмеле	0/1
он	6/0	отырып-		өрте	49/2
онайлат	1/3	отырып	0/1	өртен	28/1
онал	3/1	отырыс	0/3	өртенсін	1/0
онашала	2/0	оян	20/5	өршелен	7/0
онба	22/3	оянба	2/0	өрте	0/2
онғар	2/0	оянсын	1/0	өріл	5/0
онғарыл	1/0	оят	11/3	өрісте	1/0
онғарылма	1/0	өзгер	19/4	өс	86/11
онда	6/0	өзгерт	13/0	өсекте	1/0
ондал	2/3	өзеуре	6/3	өспесін	1/0
ондыр	2/3	өйт	29/3	өссін	7/0
онсын	2/0	өйт-бүйт	0/1	өсти-өсти	2/0

өсір	39/3	пысықсы	0/2	сандыра	1/3
өсірсін	1/3	пысылда	1/3	сандырақта	15/3
өсіт	1/0	ніс	27/3	санғы	1/3
өт	186/43	нісір	0/2	саптаc	4/3
өте	16/3	ниш	5/0	сапыр	2/0
өтер-өтпес	0/1	разоружить	1/0	сапырылыс	0/1
өтесін	2/0	рақаттан	2/0	сарала	1/3
өткен-		рәнжіт	1/0	сарғай	7/1
кеткен	2/3	ренжі	19/0	сарғайт	1/0
өткер	1/3	ренжіме	2/0	сарк	2/0
өткіз	21/5	ренжіт	8/0	саркыл	2/0
өткіздір	1/3	ретте	5/0	сарқыра	1/0
өткізсін	1/3	саба	11/3	сарна	8/0
өткізілсін	1/0	сабап-		сарғылда	1/1
өтсін	12/3	сабап	1/0	сары	14/7
өтіл	3/3	сабал	1/0	сарыл	5/3
өтін	51/3	сабыл	5/0	сарында	2/0
өтіс	1/3	сабылт	1/0	сас	21/0
өш	29/2	сағын	26/2	састыр	2/0
өшсін	15/3	сазар	1 4	сасы	5/0
өштес	1/0	сай	12/2	сасыт	1/0
өшік	2/3	сайла	17/1	сат	60/1
өшір	28/3	сайлан	5/0	сатқыз	1/1
өшірт	1/0	сайлат	2/2	сатыл	10/1
найдалан	11/2	сайра	6/2	сатырға	1/0
нанала	2/3	сайранда	3/0	сау	5/1
паналат	2/3	сайрандат	1/0	саугала	0/1
парла	1/3	сайыс	2/3	саудала	3/1
партиялас	1/3	сақта	93/5	саудалас	5/0
пең	1/3	сақтан	13/2	саудалат	1/0
пер	1/3	сақтандыр	3/0	саула	1/0
перімсі	1/3	сақылда	3/0	саулат	1/0
періп-	0/1	сал	475/33	сауык	0/1
періп	2/3	салбыра	6/3	саябырла	1/2
печатать	2/3	салбырат	0/1	сәлемдес	1/0
писать	2/3	салғастыр	2/3	сәнде	2/3
пли	2/3	салғыз	5/2	сез	75/13
погоди	2/0	салғылас	4/3	сезбегенсі	0/1
подводить	1/0	салдыр	3/0	сезген-	
поддержи-	1/0	салдыра	1/3	білген	1/0
вать		салдырлас	1/3	сездір	9/2
подумать	2/3	салқында	1/0	сезіктен	5/1
поздрав-	2/3	салын	5/3	сезіл	3/1
лять		салып-		сезіп	1/3
показать	2/0	салып	1/3	сейіл	3/0
помолись	2/3	салыс	11/2	секенде	1/0
помнит	1/3	салыстыр	6/1	секи	0/1
помогайт	1/3	самға	2/0	сектен	3/0
помогать	1/3	самсы	1/0	секилден	3/0
прикрепить	2/3	сана	42/5	секір	6/7
пропеть	2/0	санай-		секіре-	
пұрткөлде	0/1	санай	1/3	секіре	0/7
пырылда	0/1	санас	5/3	секірт	1/0
пыс	3/3	сандал	17/3	селдіре	1/0
пысыр	3/1				

сем	4/0	сойыл	5/2	сұқ	9/0
семіасін	0/1	сойыс	1/2	сұла	0/1
семір	1/2	соқ	123/19	сұлат	1/0
семірт	3/0	соққыла	2/0	сұңқылда	1/0
сен	113/15	соқ-соқ	4/0	сұңқылда	1/0
сендел	2/0	соқтық	15/2	сұра	184/10
сенделт	1/0	соқтыр	3/1	сұрал	3/0
сендір	3/0	соқтықтыр	0/1	сұрас	1/0
сел	3/0	сол	32/0	сұрастыр	1/1
селтес	1/0	солғында	1/0	сұрат	4/1
селтір	1/0	солкылда	1/0	сұркиялан	1/0
селтірт	2/0	сор	2/0	сұрлан	0/3
сергі	5/4	сорай	2/0	сұрыпта	1/0
серей	2/2	сорғала	1/3	сұрыптал	1/0
өерейт	4/0	сорғалат	1/0	сүз	6/1
серме	12/0	сорла	5/0	сүзіс	3/0
сермел	1/1	сорлат	9/0	сүй	160/3
серпі	3/0	соттал	3/0	сүйгіз	1/0
серпіл	4/2	соттас	1/0	сүйде	4/0
серттес	0/1	сөге-		сүйдім-	
серік	0/1	жаманда	1/0	күйдім	0/1
серіктес	1/0	сөгіл	1/0	сүйдір	1/0
серіп	13/0	сөздес	1/0	сүйе	23/1
сескен	15/11	сөйде	4/0	сүйемелде	1/0
сескендір	0/1	сөйле	194/23	сүйен	17/2
сиғыз	1/0	сөйлеп-	1/1	сүйеніс	2/0
сиқырлан	3/0	сөйлеп		сүйес	4/0
сипа	8/2	сөйлес	189/3	сүйке	0/2
силай-силай	1/0	сөйлестір	6/0	сүйкектет	1/0
сипала	2/0	сөйлет	12/0	сүйре	10/1
сипан	1/0	сөйт	71/9	сүйренде	2/0
сипаңқыра	0/1	сөк	28/1	сүйрет	9/4
сипат	1/0	сөлменде	0/1	сүйретіл	1/3
сихырлан	3/0	сөн	25/1	сүйретің-	
сиын	1/0	стой	5/0	кіре	1/0
сиыркүйым-		суал	8/0	сүйсін	16/10
шақта	1/0	суалт	1/0	сүйсіндір	4/0
сияқтан	3/21	суар	9/0	сүйт	75/0
сияқтандыр	0/1	суарыл	1/0	сүйте-	
сказать	3/0	суарын	1/0	сүйте	1/0
склонять	1/0	суйди	3/0	сүйін	3/0
слушать	1/0	сула	1/0	сүйіндір	4/0
собирать	2/0	суманда	3/0	сүйіс	9/0
советовать	1/0	суретте	2/0	сүменде	0/1
соғыл	3/0	сусылда	1/0	сүнгіт	2/1
соғыс	27/2	сусында	1/0	сүр	3/1
соғыстыр	0/1	сусында	1/0	сүргіз	2/0
соз	26/3	сусы	1/0	сүрт	4/0
созыл	11/5	суы	2/0	сүрін	2/0
созып-		суыр	10/4	сыба	1/1
созып	1/0	суырыл	7/0	сыбырла	4/3
сой	16/9	суыт	4/3	сыбырлас	4/0
сойғыз	1/1	сұғыл	0/1	сығала	1/1
сойқақта	0/1	сұғыл	2/0	сытымда	0/3

сыдырт	1/0	сіркеле	0/1	тамашала	3/3
сыз	0/1	табала	6/0	тамсан	1/0
сызда	0/1	табанда	5/0	тамыз	1/0
сыздан	0/4	табандат	0/1	тан	9/1
сыздықтат	0/1	табандас	1/1	тандыр	4/0
сызыл	3/0	табжыл	3/0	танғы	1/1
сый	18/2	табыл	49/2	тантыра	1/0
сыйғыз	2/0	табын	23/0	таны	109/7
сыйлас	0/1	табындыр	1/0	таныл	10/2
сыйын	7/0	табыс	48/3	таныс	18/0
сыйыс	1/1	табыстыр	4/0	тапыстыр	3/0
сык	2/0	тағайында	1/0	таныт	34/1
сыка	4/0	тағыл	1/0	таң	0/1
сыкпырт	1/0	тазала	3/0	таңбала	2/0
сыкситың-		тазарт	4/0	таңғал	10/0
қыра	0/1	тай	15/1	таңғалдыр	1/1
сықшылықта	0/2	тайғанас	1/0	таңда	46/3
сықырла	3/0	тайғыз	2/0	таңдан	6/1
сыл	0/1	тайдыр	13/0	таңдат	1/0
сыла	4/0	тәккі	6/0	таңғал	2/0
сылан	1/0	тайлық	2/0	таңыл	2/0
сылап-сипа	1/0	тайпал	0/1	таңырка	5/1
сылты	2/0	тайранда	2/0	тап	482/3
сын	9/1	тайсал	3/0	тапа	1/0
сына	68/1	тайталас	4/0	тапжыл	10/0
сынал	10/0	тайын	1/0	тапкыз	1/0
сынал-міне	1/0	тайынба	1/0	тапсыр	74/2
сынас	9/0	так	45/0	тапсырыл	4/0
сынат	4/0	тақа	23/6	тапта	14/0
сындыр	12/1	тақал	1/0	таптал	1/0
сынықсы	1/0	таққыз	1/0	таптыр	4/0
сыңғырла	1/0	тақылда	6/0	тара	31/3
сынқылда	1/0	тақымда	0/1	тарай-	
сыпыр	5/2	тал	34/2	тарай	1/0
сыпырып-		тала	21/1	тарал	1/0
снырып	0/1	талап-		тарап-	
сыр ыт	1/0	талап	1/0	тарап	1/0
сырқыра	2/0	талаптан	1/0	тарас	7/1
сырла	1/0	талас	45/2	тарат	14/1
сырлас	4/0	таластыр	1/0	тарби	1/0
сыртылда	0/1	талат	6/0	тарка	6/0
сытыл	3/4	талға	9/1	таркат	8/0
сытырла	0/1	талда	3/0	тарпақтас	1/0
сілейт	5/0	талдыр	13/1	тарпы	6/0
сілкін	3/3	талқанда	7/0	тарпылда	1/0
сілкінт	2/0	талқандал	2/0	тарсылда	2/0
сілте	18/5	талқыла	2/0	тарт	260/36
сілтес	0/1	талқылан	3/0	тартқыз	2/0
сілік	6/0	талмаура	1/0	тартпа	1/0
сімір	1/3	талпын	6/4	тарттыр	3/0
сің	7/1	талтай	0/1	тартыл	5/0
сіңбір	1/0	талти	1/0	тартын	13/1
сіңір	3/0	там	2/1	тартын-	
сірес	2/1	таммақтан	1/0	кертiл	1/0

тартыл-		геңсел	1/0	толғат	1/0
тартыл	0/1	геңселдір	1/0	толқы	12/2
тарғыс	4/3	геңселт	1/0	толқыт	3/0
тарыл	3/1	ген	27/4	толтыр	8/1
тарылт	1/1	гененде	1/0	толықсы	2/0
тас	2/0	генкіле	2/2	томсар	2/6
тасала	0/1	генсін	1/2	томырыл	1/0
таскыз	1/0	гср	15/2	тона	6/0
тасқында	2/0	гербе	1/0	тонал	2/0
таста	149/25	гербет	3/0	тонас	1/0
тастаған-		герге	7/2	тоң	5/2
талмаған	1/0	тергел	5/0	тоңазы	0/4
тастал	3/0	тергіз	1/0	тоңдыр	2/0
тастас	0/1	теренде	0/1	тоңқылда	1/0
тасы	14/2	герендет	1/0	топайшылат	1/0
тасырла	5/2	терле	3/0	топта	1/1
тасырлат	2/0	терлеп-		топтан	1/2
тасыт	1/0	геншіп	0/1	топтаңқыра	0/1
тағ	13/0	терлет	2/0	топшыла	2/1
таткыз	1/0	тес	4/0	топылдат	1/0
тагулас	2/1	тесіл	2/1	топырла	1/0
таты	12/1	тесірей	0/1	топырлас	1/0
татыма	1/0	ти	183/23	торал	2/0
таусыл	7/4	тигіз	25/1	торла	2/0
тауыс	5/0	тие	2/0	торы	15/1
тая	8/0	тиер-тимес	0/1	торық	2/0
таян	4/0	тиет	14/0	торықтыр	1/0
тәлтіректе	0/1	тиыл	5/0	тос	148/5
тәлімісі	1/0	тиіс	14/1	тосансы	1/0
тәрбиеле	10/0	тиянақта	2/0	тоса-тоса	1/0
тәрізден	6/2	тиянақтат	1/0	тотит	1/0
тәріздендір	0/1	тобарсы	3/0	тоятта	2/0
тәуірлен	2/0	тоғыс	1/1	төбелес	6/0
тәуірсі	1/0	тоғыт	6/0	төгіл	6/1
тебін	0/3	тоз	26/1	төз	3/0
тебірен	1/0	тоздыр	3/1	төйтпейле	3/0
тебірент	1/0	той	21/0	төк	44/0
тегереп	1/0	тойғыз	9/0	төлде	0/1
теже	3/1	тойла	12/0	төле	25/1
тездет	3/1	тойтар	5/0	төлен	2/0
тексер	24/2	тойтарыл	2/0	төлет	1/0
тексеріл	1/0	тойын	0/1	төменде	5/0
тептір	1/0	тоқайлас	2/0	төмендет	1/0
телменде	1/0	тоқта	192/6	төн	5/1
телмір	3/3	тоқтал	11/0	төңкер	5/0
телмірт	0/1	тоқтас	4/0	төңкеріл	2/4
телі	11/0	тоқтат	102/1	төңіректе	1/0
теңітпе	8/0	тоқы	9/1	төпе	0/1
теңірет	2/0	тоқырат	1/0	төпеле	1/0
теңе	6/0	тол	25/4	төсе	18/1
тенел	1/0	толаста	1/0	төссел	2/1
теңелт	1/0	толға	9/0	төсен	0/1
теңес	2/0	толған	22/0	төсте	2/0
теңестір	2/0	толғандыр	1/0	төте	1/0

төгеле	0/1	түксн	1/1	тырайт	1/0
требовать	1/0	түкір	6/1	тырбан	2/0
ту	193/6	түле	7/0	тыржит	4/1
туғыз	11/0	түнделет	2/0	тырмыс	3/1
тудыр	12/1	түне	5/0	тырна	2/0
тудырып-		түнер	4/4	тырнақта	1/0
өсір	0/1	түніл	22/0	тырнала	1/0
тула	12/2	түңілдір	2/0	тырсылда	1/0
тулат	3/1	түңілт	2/0	тыртыс	2/1
тура	1/0	түптет	1/0	тырыс	14/1
турала	1/0	түр	4/1	тышкырт	2/0
туралат	1/0	түрегел	8/3	тігіл	2/0
турамышыла	1/0	түрегелгіз	1/0	тіз	19/0
туу	3/0	түрлен	3/1	тізеле	1/0
туып-өскен	2/0	түрлент	1/1	тізелес	1/0
туыс	2/0	түрт	15/1	тізіл	2/0
туыстыр	1/0	түрткіле	0/1	тік	41/5
тұзда	3/1	түртпекте	1/0	тіки	0/2
тұқырт	2/3	түртін	2/0	тіккіз	2/0
түн	2/0	түршік	7/0	тіксің	1/2
түнжыра	4/1	түс	261/111	тіктір	2/0
түншык	2/1	түсе-түсе	0/1	тікірей	1/0
түншықла	2/0	түсін	104/3	тіл	6/1
түншықтыр	8/0	түсіндір	6/0	тіле	1/0
тұр	1067/81	түсініс	4/0	тілгіле	1/0
тұрақта	3/0	түсір	39/10	тілде	6/0
тұратұр	1/0	түшкір	3/0	тілдес	2/0
тұрғыз	13/0	түшкірт	1/0	тілдестір	2/0
тұрыс	3/0	түшіркен	2/0	тіле	183/18
тұса	4/0	тығыл	21/0	тілеме	1/0
тұсат-		тығып-		тілен	4/1
тұсап	0/1	сүйреп	0/1	тілес	4/0
тұспала	1/0	тығыс	1/0	тілет	3/0
тұспалда	2/1	түй	45/1	тілін	7/0
тұт	48/1	түйыл	12/4	тімтіл	2/0
тұтан	1/0	тык	68/9	тінт	8/2
тұтас	6/0	тыккышта	1/0	тіре	1/1
тұтат	0/2	тықпала	1/0	тірел	7/2
тұтыл	3/1	тыксыр	1/0	тірес	3/0
түгенде	2/0	тыкырт	1/0	тірке	2/0
түгендел	1/0	тыкыршы	1/0	тіркел	3/0
түгендес	1/0	тын	24/2	тіркес	2/0
түгесіл	1/0	тындыр	3/0	тіркет	1/0
түзе	21/2	тыншыт	0/1	тірліктес	1/0
түзел	19/0	тынык	5/1	тіріл	5/0
түзеліңкіре	0/1	тынышта	3/0	тірілт	2/0
түзет	5/0	тыныштал	6/0	тісе	1/0
түй	22/5	тыныштан		тісте	4/0
түйгіз	0/1	дыр	2/0	тістен	0/1
түйне	1/0	тында	180/9	тіс-тір-	
түйре	3/0	тындама	1/0	кекте	1/0
түйіл	10/2	тындас	1/0	тітірет	1/0
түйін	3/0	тындат	1/0	уағызда	1/0
түйіс	3/1	тыпырла	1/0	уайымда	3/0

уат	4/0	ұлықтыр	1/0	ұялма	1/0
уәделес	2/0	ұлып-ұлып	1/0	ұялт	9/0
убирать	1/0	ұлыт	1/0	үгітте	10/0
уволить	2/0	ұмыт	105/1	үде	6/0
укала	1/0	ұмытыл	32/6	үдет	3/0
уқалан	2/0	ұмытылма	1/0	үдеттір	1/0
ула	7/0	ұмытсын	1/0	үз	23/5
улан	6/1	ұмыттыр	2/0	үздік	1/0
улат	5/0	ұна	26/0	үздір	3/0
улап-шула	1/1	ұнат	19/0	үзіл	20/0
уле	1/0	ұнатпа	1/0	үзіліп-	
улық	1/0	үр	110/7	талған	1/0
уметь	1/0	ұрғыз	2/0	үзіп-үзіп	1/0
урала	3/0	ұрла	17/1	үй	2/0
успокаи-	2/0	ұрлан	0/1	үйде	2/0
вать		ұрма	1/0	үйлен	25/1
учить	2/0	ұрсын	14/0	үйленбе	1/0
уыл	1/0	ұрсып-		үйлес	7/0
уылда	1/0	соғып	0/1	үйрен	16/1
уылжы	1/0	ұрсыс	2/0	үйренбе	1/0
уыста	1/0	ұртта	2/0	үйреніс	1/0
ұғыл	1/0	ұрттат	1/0	үйрет	36/1
ұғын	13/2	ұрын	11/0	үйретпе	1/0
ұғындыр	3/2	ұрындыр	1/0	үйт	17/0
ұғынсын	1/0	ұрынт	1/0	үйтіл	1/0
ұғыныс	5/0	ұрыс	47/2	үйіл	4/2
ұғыс	21/0	ұрыстыр	1/0	үйіп-	
ұғыстыр	1/0	ұрысып-		төгіп	1/1
ұжымдас	2/0	талас	1/0	үйір	2/0
ұза	13/3	ұса	3/0	үйірле	1/0
ұзарт	2/0	ұста	179/26	үйіріл	2/1
ұзат	17/3	ұстал	20/2	үйіріліп-	
ұйғар	10/0	ұстама	2/0	айналып	0/1
ұйқас	5/0	ұстан	0/1	үк	1/0
ұйқысыра	0/1	ұстас	2/10	үлгер	3/0
ұйлық	5/0	ұстаспа	2/0	үлгір	2/0
ұйпа	5/0	ұстат	13/2	үлес	2/1
ұйы	4/1	ұсын	37/1	үлестір	1/0
ұйық	1/0	ұсыныл	1/0	үлкей	3/1
ұйықта	23/5	ұт	6/0	үміттен	1/0
ұйымдас	4/0	ұтал	1/0	үнде	56/29
ұйыстыр	6/0	ұтқыз	2/0	үнемде	2/0
ұйыт	1/0	ұтыл	2/0	үңіл	5/3
үк	75/7	үш	144/4	үңірей	0/2
үкпағансы	0/1	үшқында	2/0	үр	7/2
үкса	20/2	үшсын	1/0	үргіз	2/0
үксастыр	3/0	үштас	1/0	үрейлен	2/0
үксын	2/0	үшықты	0/1	үркіңкіре	0/1
үктыр	8/0	үшыр	32/3	үркіт	10/1
үлас	3/0	үшыра	5/3	үрле	1/0
Үласпа	2/0	үшырас	2/0	үрме	1/0
үлғай	5/2	үшырат	6/3	үрпі	2/4
үлғайт	3/0	үял	24/0	үрік	19/10
үлы	9/0	үяла	7/0	үрікпе	2/0

үсте	1/1	шалыныс	1/0	шеш	60/0
үсті	5/0	шамала	3/0	шештір	1/1
үсі	4/0	шамалас	1/1	шешіл	7/2
үсін	2/0	шамдан	2/0	шешін	4/0
үшкір	1/0	шан	1/0	шиелен	0/1
үштел	2/0	шанды	2/0	шиеленіс	3/0
фунсиндеу	1/0	шандырла	1/0	шиеленіс-	
хабарла	4/0	шаншы	3/1	тір	2/0
хабарлан-		шаншыл	1/3	шира	2/0
дыр	7/0	шаныш	2/0	ширык	3/1
хабарлас	2/0	шаңдат	1/0	шиыр	3/0
харакәтта	1/0	шап	195/20	шоқын	2/0
характерле	1/0	шапкыз	2/0	шоқындыр	1/0
хаталас	1/0	шапқыла	4/2	шоқыт	1/0
хатта	2/0	шаптық	2/0	шол	18/4
хаттал	1/0	шаптыр	3/0	шолжында	1/0
хисаплан	2/0	шапшанда	0/1	шом	2/1
ходотайла	1/0	шапшы	7/3	шомыл	1/0
хорла	1/0	шапшыт	1/1	шоңқит	1/0
читать	2/0	шарла	7/1	шошанда	1/0
шабақта	2/1	шарпы	2/0	шошы	33/6
шабалан	1/0	шарпыс	3/0	шошын	7/3
шабанда	1/1	шарша	7/4	шошыт	13/3
шабыл	3/0	шарықта	5/0	шөк	5/0
шабын	2/0	шат	5/0	шөктір	1/0
шабыс	5/0	шатас	10/0	шөлдет	1/0
шағылыс	0/1	шатастыр	4/1	шөлірке	1/0
шағылыс-		шатақаяқта	2/0	шуақта	1/0
тыр	2/0	шапкыз	1/0	шудалан	1/0
шай	1/3	шапшаң	4/0	шула	16/1
шайка	8/0	шатыл	2/0	шулат	4/0
шайкал	3/0	шатына	1/1	шуылда	1/0
шайқас	4/0	шатырла	2/0	шұба	1/2
шайлық	2/0	шатырлас	2/0	шұбат	4/0
шайлықпа	1/0	шатыс	1/0	шұбыр	4/0
шайлықтыр	1/0	шаш	24/1	шұбырт	8/2
шайлаң	5/0	шашыл	8/0	шұбырыс	0/1
шайнас	1/0	шашыра	1/7	шұғылдан	3/0
шайыл	1/0	шәлдірек-		шұқы	3/0
шак	30/3	те	1/0	шұқырай	1/0
шақта	2/0	шоуілде	2/0	шұлғы	1/3
шақтыр	1/0	шегеле	1/1	шұнтит	1/0
шақыр	139/6	шегер	3/0	шұңқырай	0/1
шақырт	44/2	шегін	9/0	шұрқыра	0/1
шақырыл	6/1	шегіне-		шұрқырас	1/0
шақырыс	0/1	шегіне	1/0	шұрылда	1/0
шал	46/4	шек	46/0	шұбалән	1/0
шалдықтыр	0/1	шекіс	7/0	шүй	3/1
шалдыр	2/0	шендес	1/0	шүйселе	1/0
шалжиыс	1/1	шене	3/0	шүйір	2/0
шалкай	3/0	шенгелде	1/0	шүйіркелес	1/0
шалқы	17/0	шерлен	1/0	шығар	338,31
шалутта	2/0	шерт	1/1	шығарт	12/0
шалық	2/0	шерткіз	1/0	шығарыл	0/1

шығында	5/1	шірітпе	1/0	іздетпе	1/0
шығындат	1/0	ығыс	2/2	ізерле	2/0
шығынқыра	1/0	ызалан	3/3	іл	19/3
шығыс	0/4	ызғарлаң	1/0	ілгеріле	0/1
шыда	92/3	ызғыт	1/0	ілгіз	1/0
шыдай-		ызылда	1/0	ілдір	1/1
шыдай	1/0	ық	11/2	ілес	0/1
шыдат	6/0	ықпа	1/0	ілестір	1/0
шыжылда	1/1	ықта	5/0	ілме	1/0
шыжылдат	1/0	ықтыр	2/0	ілік	2/1
шық	671/76	ымда	2/0	іліктес	0/1
шыққыз	1/0	ынтық	4/0	ілін	21/8
шықпа	5/0	ынты-		іліндір	1/1
шықсын	24/0	мақтас	1/0	ірік	24/5
шыла	1/0	ыңғайла	0/2	ірі	2/0
шылдыра	1/0	ыңғайлаң	2/0	ірік	18/3
шымышы	1/0	ыңқи	1/0	ірікте	4/1
шында	12/0	ыңқылда	1/0	іріле	0/1
шыныс	3/0	ыныран	3/1	іріп-шірі	1/0
шыңғыр	1/0	ыңырсы	2/1	іріт	1/0
шында	1/0	ырғал	0/3	іс	1/1
шындал	1/0	ырғы	4/0	істе	300/0
шыңқылда	1/0	ырғыт	1/0	істел	15/0
шырайлаң	1/0	ыржақта	2/0	істелме	1/0
шырағала	2/0	ыржаңда	2/0	істес	2/0
шырқа	9/1	ырзала	2/3	істет	11/0
шырқат	2/0	ырзалас	1/0	ісін	2/0
шырқыра	1/0	ыршы	1/1	ісір	0/1
шырыа	16/0	ыршыт	1/0	іш	76/7
шырмал	5/0	ысқыр	5/0	ішкіз	7/0
шыт	1/0	ыстықта	1/0	ішле	2/0
шытпа	1/0	ысыл	1/0	ішпе	2/1
шытың	1/2	ысылма	1/0	ішсін	3/0
шытырла	1/0	ытқы	0/3	ішіп-жеп	1/0
шімірген	1/0	ышқын	0/1	электр-	
шірен	1/1	ізде	125/13	лендір	1/0
шірі	5/2	ізден	30/1		
шіріме	2/0	іздес	1/0		
шіріт	2/2	іздет	3/2		

* Келтірілген сандардың біріншісі — драматургия, екіншісі — повесті көрсетеді.

ӘДЕБИЕТТЕР

1. Хасенова А., Ибатов А. Қазақ тілі морфологиясын зерттеудегі табыстар мен болашақ міндеттер // Қазақ тілі грамматикасы бойынша зерттеулер. Алматы, 1975.
2. Ибатов А. Қазақ тілі морфологиясының зерттелуі // Развитие казахского советского языкознания. Алма-Ата, 1980.
3. Ысқақов А. Қазіргі қазақ тілі. Морфология. Алматы, 1974.
4. Бектаев Қ. Б., Жұбанов А. Қ., Мырзабеков С., Белботаев А. Б. М. Әуезовтің «Абай жолы» романының жиілік сөздігі. Алматы, 1979.

5. *Амбаев Ә., Бектаев К. Б.* Зат есім сөздердің морфологиялық құрылымына статистикалық талдау // Қазақ тексінің статистикасы. Алматы, 1973.
6. *Молдабеков Қ.* Лингвостатистические исследования казахских текстов для младших школьников (на материале текстов учебников начальных классов и литературы для детей). Дис. ... канд. филол. наук. Алма-Ата, 1985.
7. *Хасенова А.* Етістіктің лексика-грамматикалық сипаты. Алматы, 1971.
8. *Әуезов М. О.* 12 томдық шығармалар жинағы. Алматы, 1969. 8, 9, 10-т.
9. *Әуезов М. О.* Қараш-қараш оқиғасы. 12 томдық шығармалар жинағы. Алматы, 1967. 1-т. 238—310-б.
10. *Қазақ тексінің статистикасы.* Алматы, 1973.

М. Әуезовтің «Индия очерктеріндегі» жалқы есімдердің түсіндірме сөздігі

Аббас (3)	Қісінің аты. Ходжа — Ахмед Аббас — кино — суреттердің сценарийін жазушы.
Агра (11)	Қаланың аты, Индияда.
Аджента (7)	Қыстақтың аты. Хайдарабаттың қасында
Азия (7)	Осы аттас қыстақтың атымен аталған тарихи ескерткіш /храм/.
Аджента ескерткіші (1)	Материк. Дүниенің жер шарындағы ең ірі бөлігі.
Айвангар (1)	Қісінің аты. Тас жазуларын шешуші
Айя-София мешіті (1)	Мешіт. Дүние жүзіндегі жеті ғажайып архитектура ескерткіштерінің бірі. Стамбулда.
Акбар (7)	Қісі аты. Хумаюның баласы, Бабур патшаның немересі.
Ақсақ Темір (3)	Қісінің аты. Орта Азиялық қолбасшы, өмір /1370—1405/. Монғолдың түріктенген Барлас тайпасынан шыққан Тарағай Бектің баласы.
Аладин (3)	Қісінің аты. Патша.
Алатау (3)	Таудың аты.
Алексей Сурков (25)	Қісінің аты. Ақын.
Али-Сардар Жафри (6)	Қісінің аты. Ақын.
Алихан (1)	Қісінің аты.
Алматы (6)	Қаланың аты. Қазақстанның астанасы.
Альгамбра сарайы (1)	Кордово кафедратындағы сарайдың аты. Дүние жүзіндегі жеті ғажайып архитектура ескерткіштерінің бірі.
Альмара (1)	Ағаштың аты.
Альпы (3)	Таудың аты.
Америка (13)	Материк. Батыс жарты шардағы екі құрылықтан тұратын дүние бөлегі. Атлант және Тынық мұхиттарының аралығында.
Англия (15)	Мемлекет. Ұлыбританияның саяси,

«Англия король ханымының алқасы» (1)
Антей (2)

Апенин аралы (1)
Арабстан теңізі (4)

Арабстан шөлі (2)

Аравали (1)
«Ауамахал» (1)

Аураңгабад (1)
Аураңзеп (4)

Афина (1)

Африка (3)
Ашока (8)

Әулие Марк соборы (1)

Әулие Петр соборы (1)
Бабұр (11)

«Бабұрнама» (1)

Бағиран (2)
Балига (2)

Бангалор (13)
Банжора (1)
Батыс Германия (1)
Батыс Европа (2)
Бахадұршах (2)
Бенарес (30)
Бенгал қойнауы (3)

Бенгал уәлаяты (2)

Бенгали штаты (4)
Бенгал шығанағы (4)
Бизитти (1)

әкімшілік бөлігі, оның тарихи ұйытқысы.

Бомбейдегі теңіз жағалауының атауы

Кісі аты. Ежелгі грек мифологиясындағы алып.

Жер аты, Италияда.

Индия мұхитының бір қойнауындағы теңіз. Жағасында Бомбей қаласы бар.

Африкадағы шөл. Сахара шөлінің солтүстік шығыс жалғасы.

Таудың аты.

Сарайдың аты (Дворец ветров — Желді сарай дегенді білдіреді.

Қаланың аты.

Шахжиханның баласы, Жихангердің немересі. Акбардың шөберегі.

Бұл жерде Афины қаласын айтып отыр. Грецияның астанасы. Грек мифологиясындағы басты құдайлардың бірі — қыз құдайдың атына байланысты қойылған.

Материк.

Кісінің аты. Индияның ұлы патшаларының бірі.

Собор. Дүние жүзіндегі жеті ғажайып архитектура ескерткіштерінің бірі. Венецияда.

«Санжиованни» храмы. Римде.

Бабұр /Бабур/ Захир ад-Дин Мухаммед — бұрынғы Ферғана әмірі. Үндістан билеушісі, қолбасшы.

Бабұр патша қалдырған тарихи мемуар. Әрі тарихи, әрі әдеби шығарма. Орта Азия халықтарының көрнекті әдеби ескерткіш — мұрасы.

Ғажайып ертегілердегі сәнді бақ.

Кісінің аты. Индия — совет қоғамының бастығы /сол кездегі/.

Қаланың аты. Индияда.

Қаланың аты. Индияда.

Орталық Европадағы Мемлекет.

Географиялық аймақ.

Кісінің аты. Патша

Қаланың аты. Индияда

Үндістан мен Үндіқытай түбектерінің арасы. Үнді мұхитында

Бенгал шығанағындағы өлке, ел-жұрт.

Бенгали штаты.

Үнді мұхитындағы шығанақ. Үндістан мен Үндіқытай түбектерінің арасында.

Кісінің аты. Жазушы.

- Бирма (1)
- Бкара каналы (1)
- Блок (1)
- Бомбей (57)
- Борджия (1)
- Ботичелли (2)
- «Бөтен адам» (1)
- Брама (5)
- «Брахманчарилар» мектебі
- Бриндаван (1)
- Британия (3)
- Будда (50)
- Бупалсинг (1)
- Вагора (1)
- Валатхоло (7)
- Варгеца паркі (1)
- Василий Блаженный шіркеуі (1)
- Ватикан (7)
- Венгрия (1)
- Венеция (2)
- Виктор Эммуэул ескерткіші (2)
- Висн (2)
- Внуково аэродромы (1)
- Ганг (9)
- Гануман (4)
- Гарибальди ескерткіші (6)
- «Гастхауз» (3)
- Германия (3)
- Геродот (1)
- Горький (1)
- Мемлекет. Оңтүстік — Шығыс Азияда Үндіқытай түбегінің батысында орналасқан. Каналдың аты, онда Нангап плотинасы бар. Кісінің аты. Блок Александр Александрович — орыс ақыны. Қаланың аты, Индияда. Кісінің аты. Кісінің аты. Италиян суретшісі. Индия жазушысы Бизиттидің операсы. Жаратушы құдайдың аты. Индуизм дінінің болашақ гүлама жреңтерін әзірлейтін мемлекет қарамағындағы мектеп. Индияның Читоргар қаласында. Демалыс орны, парктің аты. Англияның ескі аттарының бірі. Құдайдың аты. Будда дінінде нирвана жағдайына, нағыз әулиелікке жеткен жан иесі. Кісінің аты. Махараджа. Өзеннің аты. Оңтүстік Индияның қарт жазушысы — ақыны. демалыс орны. Дүние жүзіндегі Жеті ғажайып архитектура ескерткіштерінің бірі. Москвада. Өз алдына мемлекет болған қала. Католик шіркеуінің басшысы — папаның ордасы. Орталық Европадағы мемлекет. Италияның солтүстік-шығысындағы қала. Адриат теңізінің жағалаында. Италия королі Виктор Эммуэулге қойылған ескерткіш. Кісінің аты. Проза шебері, қысқа әңгімелер авторы. Аэродромның аты. Москвада. Өзен, дария аты. Әулие су саналады. «Рамаяна» дастанындағы маймылдардың бастығы. Италияның халық батырына қойылған ескерткіш. Каморин мүйісіндегі үкімет қонақжайының аты. Мемлекет. Европада екінші дүние жүзілік соғыстың аяғына дейін өмір сүрген мемлекет. Ежелгі грек тарихшысы. Максим Горький — пролетарлардың

Греция (5)
Грифиц (1)
Гувер плотинасы (1)
Грута заманы (1)
Густав Ваза ескерткіші (1)

Гайса (8)
Гұмаюн (2)
Гұсман (4)
Дания (1)
Даңқ мұнарасы (1)

Дели (56)
Дели университеті (4)

Джавахарлал Неру (2)

Дикенс (1)

Дрезден (2)

«Дрезденская галерея» (3)

Дюран (3)
Дюссельдорф (2)
Европа (37)
Египет (1)

Ескі Дели (2)
«Жазғы сарайлар» (1)

Жайпур (2)
«Жалпының жоспары» (2)

Жамбыл (2)

Жаңа Дели (Нью Дели) (2)
«Жаңа заң» (2)
Женева (6)

Женева көлі (1)
«Женіс мұнарасы» (1)

ұлы жазушысы. Социалистік реализм әдебиетінің негізін қалаушы. Ежелгі грек мемлекеттерінің жерінің жалпы атауы. Кісінің аты. Хайдарабодтағы Аджепта және Эллора ескерткіштерін зерттеуші: голым. Плотинаның аты. Индияда. Солтүстік Үндістандағы Магадха мемлекетін билеушілер заманын /4—6 ғғ./ айтады. Швед королінің құрметіне қойылған ескерткіш. Швецияда. Кісінің аты. Пайгамбарлардың бірі. Бабұрдың баласы. Белгілі «төрт шариярдың» бірі. Батыс Европадағы мемлекет. Мұнараның аты. Читторгар қаласында. Қаланың аты. Үндістанның астанасы. Индиядағы ең үлкен оқу орындарының бірі. Үндістанның саяси және мемлекет қайраткері. Кісінің аты. Ағылшынның көрнекті реалист жазушысы. Қаланың аты. Герман Демократиялық республикасында. Дүние жүзіндегі ең таңдаулы көркемсуреттер жинақталған орындардың бірі. ГДР-дің Дрезден қаласында. Кісінің аты. Англия оқымыстысы. Қаланың аты. Батыс Германияда. Материк. Көне Мысыр. Ніл өзенінің төменгі ағысында, солтүстік-шығыс Африкада болған ертедегі мемлекет. Дели қаласының ескі бөлегі. Көл үстіндегі көне ескерткіштер тобы. Удайпурда. Қаланың аты. Индияда. Индия өкіметінің үлкен шаруашылық үйірмелерінің бірі. Қоғамның аты /«Общинный проект»/. Кісінің аты. Жамбыл Жабаев — қазақ халық поэзиясының алыбы. Делидің жаңа бөлегі. Жамбыл Жабаевтың жыры. Қаланың аты. Швейцарияның астанасы. Швейцариядағы келдің аты. Мұнараның аты, Читторгар қаласында

- Жихангер (1)
- Жумна өзені (2)
- Жұмамешіт (5)
Захириддин Бабұр (1)
Иерусалим (2)
- Индира Ганди (1)
- Индия (369)
Индия мұхиты (5)
- Индия — Совет достығы (10)
Индия теңізі (1)
Иран (8)
- Иран бағы (2)
- Иса (1)
- Италия (7)
Италия жарым аралы (1)
«Кабалиширатем» (1)
Кавери (1)
Каир (6)
- Калькутта (24)
Каморин мүйісі (8)
- Карачи (4)
Карбузье (1)
- XII Қарл мүсіні (2)
Кашмир (3)
Кербаланың шөлі (2)
Киплинг (1)
- Колизей қалдығы (2)
- Копенгаген (2)
Кордово халифаты (2)
- «Корслевское общество» (1)
- Патша. Акбардың баласы. Хумаюнның немересі. Бабұрдың шөбересі. Өзеннің аты. Делидің ортасымен ағады. Мешіттің аты. Шахжиһан салдырған. Патша. Ақсақ Темірдің немересі. Қала. Ескі аты Құдыс, Палестинада. Үндістанның саяси және мемлекет қайраткері. Джавахарлал Нерудің қызы. Үндістан. Оңтүстік Азиядағы мемлекет. Негізінен Үндістан түбегінде орналасқан. Астанасы — Дели. Үнді мұхиты. Көлемі жөнінде жер шарындағы үшінші (Тынық және Атлант мұхиттарынан кейінгі, мұхит. Қоғамның аты. Үнді мұхиты. Мемлекет. Азияның оңтүстік-батыс бөлегінде 1035 ж. дейін Персия деп аталынған. Ертегідегі еститін «Иран бағы» — «Бағиран». Кісінің аты. «Тәжмахолды» салушы шебер. Мемлекет. Еуропаның оңтүстік бөлігінде. Аралдың аты. Италияда. Храмының аты. Өзеннің аты. Қаланың аты. Ескі аты Мысыр. Египет Араб Республикасының астанасы. Қаланың аты. Индияның оңтүстігіндегі Индия мұхитына мольна сұғылып тұрған жер шегі. Қаланың аты. Пакистанда. Синд провинциясының әкімшілік орталығы. Кісінің аты. Француз архитекторы. Белгілі Швед королы — XII Қарлге қойылған ескерткіш. Швецияда. Қаланың аты. Солтүстік Үндістанда. Қала. Данияның астанасы. Құмның аты. Географиялық аймақ. Киплинг Джозеф Редьярд — ағылшын жазушысы. Көне заманнан қалған ескерткіш орны. Қала. Данияның астанасы. Пиреней түбегінде 929 жылдан 1031 жылға дейін өмір сүрген феодалдық мемлекет. Қоғамның аты. Англияда.

Қочин (6)
«Көркемсөз қоғамы» (1)

Крит (2)
Крит аралы (4)

Кришан Чандр (4)

Кришна (8)
«Кришнарджсагар» (1)
Кеенофонт (1)
«Қуатүлнслам» (1)
«Кульпаттандалам» (1)
Кумба ескерткіші (1)
Кунгар мырза (2)

Кіші Алматы биігі (1)

Қазақстан (1)

Құддүс (2)

Құтбиддин (1)

«Құтбминар» мұнарасы (3)

Қызыл сарай (Редфорт) (3)
Қытай (2)
Лакшмивилас (1)
Лахну мектебі (1)
Лахор (2)

Лахор дарбазасы (1)

Лахор мектебі (1)
Лев Толстой (2)
Ленинград (2)
Личн (1)

Луи Жаколино (2)
Мадрас (27)
Майсур (10)

Майсур тоғайы (3)
Майсур уэляты (1)

«Мақшәр күні» (3)

Малавар (1)
Мариям-ана (3)

Қаланың аты.
Индиядағы ақын-жазушылар қоғамының аты.

Таудың аты.
Крит тауына байланысты қойылған аралдың аты.

Кісінің аты. Индия қайраткері, халық басшысы.

Көгертуші құдайдың аты.
Қонақ үйдің (Отель) аты. Майсурда.

Ежелгі грек жазушысы, тарихшы.
Мешіт, Индияда.

Индия өкіметінің үлкен шаруашылық үйірмелерінің бірі.

Раджа Махмуд сұлтанды жеңген жауынгерге қойылған ескерткіш.
Нангал құрылысының бас инженері, конгрессистер партиясының мүшесі.

Алатаудың Алматы тұсындағы бір бөлігінің аты.

Қазақ Советтік Социалистік Республикасы.
ССРО құрамындағы одақтас республикалардың бірі. Астанасы — Алматы.

Қаланың аты. Жаңа аты — Иерусалим. Пакистанда.

Кісінің аты. Индия патшаларының бірі.
Мұнара, Құтбиддин патша салдырған. Делиде.

Шахжіһан салдырған патша сарайы.
Қытай Халық Республикасы.

Сарай аты. Удайпурда.
Суретшілер мектебі. Индияда.

Қаланың аты. Пакистанның солтүстік шығысында.
Лахор қаласына кіретін негізгі қақпаның аты.

Суретшілер мектебі. Индияда.
Орыс жазушысы.

Қаланың аты. Советтер Одағында.
Біздің жаққа мәлім емес, мол бұтақты, үлкен жеміс ағашы.

Кісінің аты. Ағылшын жазушысы.
Қаланың аты. Оңтүстік Үндістанда.

Тамилнад штатының әкімшілік орталығы.
Қаланың аты. Оңтүстік Үндістанда.

Майсур қаласының маңындағы тоғай.
Оңтүстік Үндістандағы ел-жұрт. Араб теңізі жағалауында.
Микельанджелонның туындысы — «Страшный суд».

Таудың аты.
Ғайса пайғамбардың анасы.

- Маркс (3) Кісінің аты. Дүние жүзіндегі жұмысшы табының кемеңгер көсемі, ұлы ұстазы, халықаралық революциялық қозғалыстың шебер стратегі. Терең ойшыл теоретик, ғылыми коммунизмнің негізін салушы.
- Мауңбет (1) Кісінің аты. Лорд Мауңбет — Англияның вице-королі.
- «Махабалипурам» (1) Храмның аты. VII ғасыр ескерткіші. Шива құдайына арналған.
- «Махабхарата» (3) Үнді халқының әрі дін кітабы, әрі ғажайып дастаны. Санскритте «Ұлы Бхарата туралы аңыз».
- Махамана (1) Кісінің аты. Патша.
- Махин (1) Бомбейге қарасты бір ауданның аты.
- Маяковский (1) Орыс совет ақыны.
- Мейвар (1) Кісінің аты. Король.
- Менон (1) Кісінің аты. Индияның (сол кездегі) Одақтағы елшісі.
- Микельанджело (7) Кісінің аты. Микельанджело Буанарроти — итальян мүсіншісі, кескіндемекші, архитектор әрі ақыи.
- Милан (1) Қаланың аты. Италияда. Ломбардия облысының және Милан провинциясының әкімшілік орталығы.
- Мирай (1) Храмның аты. Бір сұлу қыз атынан қойылған.
- Моғол мектебі (2) Суретшілер мектебі. Индияда.
- Моғол паркi (2) Парктің аты. Халық демалатын демалыс орны.
- Моғол сарай (1) 1707 жылы өлген Ауранзеппке дейін «Ұлы моғолдар нәсілі» деп аталатын патшалар (Бабур, Хұмаюн, Акбар, Жихангер, Шахжихан, Ауранзепп т. б.) тұрған сарайдың аты.
- Монблан (4) Шыңның аты. Батыс Альпідегі тау массиві.
- Москва (8) Қаланың аты. Советтік Социалистік Республикалар Одағының астанасы.
- Мундассери (1) Кісінің аты. Революцияшыл бағыттағы жазушы.
- Мұмтазбегім (1) Кісінің аты. Шахжихан патшаның сүйікті ханымы.
- Мухаммед (4) Пайғамбар. Ислам дінінің негізін салған діни уағызшы.
- Мүлк Раджинан (11) Кісінің аты. Индия жазушысы.
- «Мың бір түн» (1) Араб әдебиетінің орта ғасырдағы классикалық ескерткіші. Ол үш жүздей аңыз-әңгімелерден тұрады.
- Мысыр (4) Қаланың аты. Каир қаласының ескі аты. Египет Араб Республикасының астанасы.
- Нагар (1) Кісінің аты. Удайлурдың жергілікті қайраткері.

- Наггал плотинасы (2)
- Наггалстрой (1)
- Непал (1)
- Неру ханым (1)
- Нобель (1)
- Норвегия (2)
- Оксфорд университеті (1)
- Октавиан Август (1)
- Оңтүстік Индия (1)
- Орта теңіз (2)
- Осло (2)
- «Осман» (1)
- Осман Алихан (1)
- «Осман» университеті (1)
- Ост-Индия компаниясы (1)
- Пакистан (5)
- Памнадай (1)
- Парфенон (1)
- Пелопоннес жарым аралы (1)
- Пенджап (4)
- Пенджап уәлаяты (2)
- Пенджу (1)
- Петр—I (5)
- Пиллай (1)
- Пичула (1)
- Подбанабапурам (1)
- Полтава (1)
- Прасад (6)
- Пенджаптың бес өзенінің бірі — Сетледж өзеніне салынған құрылыс. Жазушының Индияға барған кезінде жүріп жатқан плотина құрылысының аты.
- Оңтүстік Азиядағы мемлекет. Жер шарындағы ең биік Гималай тауының орталық бөлігіне орналасқан. Джавахарлал Нерудің туған аясы. Нобель сыйлығы — Дүниежүзілік сыйлықтың аты.
- Солтүстік Европадағы мемлекет. Астанасы — Осло.
- Ұлыбританияның байырғы және ең ірі жоғары оқу орындарының бірі. Оксфорд қаласында.
- Рим императоры Юлий Цезарьдің немересі.
- Географиялық аймақ.
- Теңіздің аты.
- Қаланың аты. Норвегияның астанасы.
- Хайдарабаттағы аурухананың аты.
- Кісінің аты. Хайдарабаттың әміршісі (әкімі).
- Хайдарабаттағы жоғары оқу орны.
- Компанияның аты.
- Мемлекет. Оңтүстік Азияда.
- Сарайдың аты. Шахзаданың жанын қорғап қалған бір құтуші әйел атынан қойылған.
- Храмның аты. Афина құдайына арнап салынған. Дүние жүзіндегі жеті таңғайып архитектура ескерткіштерінің бірі. Афиныда.
- Аралдың аты. Грецияда.
- Солтүстік-Батыс Индиядағы штаттың аты.
- Географиялық аймақ. Пенджап елжұрты.
- Демалыс бағының аты. Осы аттас қыстақтың атынан қойылған.
- Орыс патшасы. Россия мемлекетінің императоры, мемлекеттік қайраткер, қолбасшы және дипломат.
- Кісінің аты. Зерттеуші-ғалым, журналист, ақын.
- Көлдің аты, Удайпур қаласының маңында.
- Тривандрум маңындағы ескі сарайлардың бірі, XIV ғасырға жатады.
- Қаланың аты. Украинада. Полтава облысының орталығы.
- Кісінің аты. Прасад Раджендра — Индияның президенті, саяси және мемлекет қайраткері.

- Претап (1)
Пушкин (2)
- «Пьета» скульптурасы (2)
- «Пілдер үңгірі» (1)
Рабиндранат Тагор (4)
- Равана (1)
- Рагнер (1)
Рагнер (1)
- Раджагопалачария (1)
- Раджама (1)
Раджа Махмут (1)
Раджендр Прасад (3)
- Раджестан (5)
- Раджестан уәлаяты (1)
- Раджинан (5)
Радж Капур (3)
- Раджпур Бенгал мектебі (2)
- Радхакришнан (3)
- Рама (7)
- Рамаманрам (1)
- «Рамаяна» (4)
- Рамен (8)
- Рангун (1)
- Рафаэль (4)
- Рачпромук (2)
- Реймский собор (1)
- Рим (35)
- Кісінің аты. Махараджа. Александр Сергеевич — орыс жазушысы. Жаңа орыс әдебиетінің негізін қалаушысы.
- Атақты Микельанджело қолынан шыққан туынды.
Таудың аты. Бомбейдің жанында. Үнді әдебиетінің классигі, прозаик, сыншы, ақын, публицист.
«Рамаяна» дастанының кейіпкері, жын (демон).
Кісінің аты. Индия жазушысы.
Қаланың аты. Жазушы Рагнер атындағы қала.
Раджагопалачария Чакраварти — Индияның саяси және қоғамдық қайраткері. Көпке мәлім қарт дипломат.
Кісінің аты. Әйел ақын.
Кісінің аты. Орта ғасырдағы сұлтан.
Кісінің аты. Индия президенттерінің бірі.
Солтүстік Батыс Индиядағы штат. Орталығы Жайпур қаласы.
Географиялық аймақ. Раджестан елжұрты.
Кісінің аты.
Кісінің аты. Индиядағы белгілі кино қайраткері.
Мектептің аты. Индиядағы 16—19 ғғ. болған негізгі мектептердің бірі. Суретшілер мектебі.
Радхакришнан Сарвепалли — Индияның көрнекті саяси және мемлекет басшыларының бірі. Вице-президент, ғалым-философ.
«Рамаяна» дастанының кейіпкері. Дін ұғымы бойынша құдай сипатты зат.
Кісінің аты. Революцияшыл бағыттағы жазушы.
Ежелгі Индияның эпикалық шығармаларының бірі.
Кісінің аты. Индияның атақты ғалымы, профессор.
Қаланың аты. Бирман Социалистік Республикасы Одағының астанасы.
Итальян живописец және архитекторы.
Кісінің аты. Махараджа. Майсурдың әкімі.
Собор. Дүние жүзіндегі жеті ғажайып архитектура ескерткіштерінің бірі. Францияда.
Қаланың аты. Ежелгі көне қала. Италияның астанасы.

- Россия (4)
- Самарқан (1)
- «Санжиованни» храмы (1)
- «Санипаты» (1)
- «Сан Пьетро Инквинколыи» (3)
- Санта-Мария храмы (1)
- Сәбит (1)
- «Сегіз піл сарайы» (3)
- Сетледж (3)
- Сидхарта Будда (2)
- Сикстинская копелла (3)
- Сикстинская мадонна (3)
- Симмла (2)
- Сита (6)
- Сиям (1)
- Скандинавия компаниясы (1)
- Совет Одағы (2)
- Советтік Шығыс (2)
- Стамбул (1)
- Стокгольм (11)
- Субраманьян Шастри (1)
- Сурат (1)
- Суэц каналы (2)
- «Тажмахал» (6)
- Талғар биігі (1)
- «Тажмахал» (11)
- Тәжікстан (1)
- Ұлы Октябрь Социалистік революциясына дейін өмір сүрген орыс мемлекеті
- Қаланың аты. Самарқанд облысының орталығы. Өзбек ССР-інде.
- Әулие Петрдің атақты соборы, Римде.
- Индия өкіметінің үлкен шаруашылық үйірмелерінің бірі.
- Храмның аты. Шынжырлаулы Петр храмы. Римде.
- Храмның аты, Римде.
- Кісінің аты. Сәбит Мұханов — қазақ совет жазушысы. Қоғам қайраткері, Қазақ совет әдебиетінің негізін салушылардың бірі.
- Сарайдың аты. Удайшурда.
- Өзеннің аты. Пакистанның, Индияның және Қытайдың үстімен ағады.
- Құдайдың аты. Будда дінінде Нирвана жағдайына, нағыз әулиелікке жеткен жан несі делінеді.
- Сикстин копелласы. Италия қайта өрлеу дәуірінің Рим Ватиканындағы тамаша ескерткіші.
- Сурет өнерінің құдайындай саналатын Рафаэль қолымен жасалған атақты ғажайып шығарма.
- Қаланың аты. Солтүстік Индияда.
- «Рамаьяна» дастанының героинясы, Раманың сүйікті сұлу жас ғашығы.
- Мемлекеттің аты.
- Компанияның аты.
- Советтік Социалистік Республикалар Одағы. Астанасы — Москва.
- Географиялық аймақ.
- Қаланың аты. Турияның ең ері өндірісті және мәдени орталығы.
- Қаланың аты. Швецияның астанасы.
- Кісінің аты. Ақын.
- Қаланың аты.
- Қапалдың аты, Аравияда.
- Хайдарабаттағы қопақ үйдің аты.
- Алматының бір шетіндегі Алатау шоқыларының бірі.
- «Сарайлар тәжі». Әрі мешіт, әрі күмбез. Дүние жүзіндегі жеті ғажайып архитектура ескерткіштерінің бірі.
- Шахжихан патша өзінің сүйікті ханымы Мұмтазбегімге арнап салдырған.
- Тәжік Советтік Социалистік республикасы. Орта Азияның оңтүстігіндегі

«Теософическое общество» (1)

Тибет (1)

Токио (1)

Траванкор (3)

Тривандрум (9)

Тіршілік мәнісі туралы» (1)

Удайпур (21)

«Үнсіз мұнара» (4)

Фатей (1)

Фейнмор Купер (1)

Фергюссон (1)

Фергана (1)

Фердауси (1)

Флоренция (1)

Форум қалдығы (3)

Фостер (1)

Франция (2)

Хайдарабад (12)

Хайдарабадхауз (2)

Хельсенки (2)

Херингер (1)

Ходжа-Ахмед Аббас (3)

Хумаюн (1)

Читоргар (4)

Чандигар (1)

Чандлер (1)

Чандогарх (1)

«Чандра Махал» (1)

«Чарминнар» (1)

Цейлон аралы (5)

«Чоури» (1)

Шантиникетан университеті (1)

одақтас республика. Астанасы — Душанбе.

Бениарес қаласындағы бір қоғамның аты.

Таудың аты.

Қаланың аты. Жапонияның астанасы.

Қаланың аты.

Қаланың аты.

1932 ж. Англия оқымыстысы Дюран жазған кітап (шығарма).

Қаланың аты, Индияда. Малавр тауының етегіндегі мұнараның аты.

Көлдің аты. Удайпурда.

Кісінің аты. Ағылшын жазушысы.

Кісінің аты. Хайдарабаттағы Аджента және Элора ескерткіштерін зерттеуші ғалым.

Қаланың аты, Өзбекстанда. Облыс орталығы.

Кісінің аты. Шығыстың ұлы шайыры (ақыны).

Қаланың аты. Италияда.

Көне заманнан қалған көне ескерткіш орны.

Кісінің аты. Ульям Фостер — американ және халықаралық жұмысшы қозғалысының қайраткері.

Франция Республикасы. Батыс Европадағы ірі капиталистік мемлекеттердің бірі.

Қаланың аты, Индияда.

Сарай, жатын жай. Делиде.

Қаланың аты, Швейцария астанасы.

Хайдарабадтағы Аджента және Элора ескерткіштерін зерттеуші.

Кісінің аты. Кино-суреттердің сценарийін жазушы.

Кісінің аты. Патша. Бабұр патшаның баласы.

Қаланың аты. Раджестанның орталығы.

Қаланың аты.

Кісінің аты. Ағылшын ғалымы.

Қаланың аты.

Сарайдың аты. (Лунный дворец) - Ай сәулелі сарай. Жайпурда.

Мешіттің аты. Хайдарабадта. XVI ғасыр туындысы.

Аралдың аты.

Храмның аты. XI ғасыр туындысы. Читоргарда.

Индиядағы жоғары оқу орындарының бірі.

Шахжихан (9)

Швейцария (3)

Швеция (1)

«Шемали» (1)

Шемали (1)

Шива (14)

Шивалик (1)

«Шиваның жұмағы» (1)

Шолохов (1)

Эллора (7)

Эллора ескерткіші (1)

Эмпиनावудранат (1)

Югославия (1)

Юлий Цезарь (1)

Япония (1)

Ұлы моғолдар патшасының бірі. Жи-хангердің баласы. Акбардың немересі, Хумаюнның шөбересі.

Швейцария Конфедерациясы. Орталық Европадағы мемлекет.

Швеция корольдығы. Солтүстік Европадағы мемлекет.

Пьесаның аты.

«Шемали» пьесасындағы героиняның аты.

Тәңірінің аты. Күйретуші құдай.

Таудың аты. Индияда.

Храмның аты. Эллора қыстағында.

Кісінің аты. Михаил Александрович — орыс совет жазушысы.

Қыстақтың аты. Хайдарабаттың қасында.

Осы аттас қыстақтың атымен аталынатын тарихи ескерткіш (храм).

Кісінің аты. Революцияшыл бағыттағы жазушы. Валатхолоның баласы.

Югославия Социалистік Федерациялық Республикасы. Европаның оңтүстік бөлегіндегі социалистік мемлекет.

Юлий Цезарь Гай — Римнің мемлекеттік және саяси қайраткері. Қолбасшы, жазушы.

Жапония. Шығыс Азиядағы мемлекет.

**М. О. ӘУЕЗОВТИҢ «ИНДИЯ ОЧЕРКТЕРІНДЕГІ»
КІСІ ЕСІМДЕРІНІҢ ЖИІЛІК СӨЗДІГІ.**

кісі есімдері	кездесу жиілігі	кісі есімдері	кездесу жиілігі
Аббас	3	Иса	1
Айвангар	1	Қарбузье	1
Ақбар	7	Киплинг	1
Ақсак Темір	3	Кришан Чандр	4
Аладли	3	Кришна	8
Алексей Сурков	25	Ксенофонт	1
Али-Сардар Жафри	6	Кунгар мырза	2
Алихан	1	Құтбилдин	1
Антей	2	Лев Толстой	2
Аураңзеп	4	Луи Жаколино	2
Ашока	8	Мариям-ана	3
Бабұр	11	Маркс	3
Баха дұршах	2	Маунбет	1
Бизитти	1	Махамана	1
Балига	2	Маяковский	1
Блок	1	Мейвар	1
Борджия	1	Менон	1
Ботичелли	2	Микельанджело	1
Брама	5	Мундассери	1
Будда	50	Мұхаммед	4
Бупалсинг	1	Мүлк Раджинан	1
Валатхоло	7	Нагар	1
Виси	2	Неру ханым	1
Гануман	4	Октавион Август	1
Геродот	1	Осман Алихан	1
Горький	1	Петр-1	5
Грифиц	1	Пиллай	1
Ғайса	8	Прасад	6
Ғумаюн	2	Претап	1
Ғұсман	4	Пушкин	2
Джавахарлал Неру	2	Рабиндрант Тагор	4
Дикенс	1	Равана	1
Дюран	3	Рагнер	1
Жамбыл	2	Раджагопалачария	1
Жихангер	1	Раджама	1
Захириддин Бабұр	1	Раджа Махмут	1
Индира Ганди	1		

Раджендр Прасад	3	Фергюссон	1
Раджинан	5	Фердауси	1
Радж Капур	3	Фостер	1
Рархакришнан	3	Херингер	1
Рама	7	Ходжа-Ахмед Аббас	3
Рамамирам	1	Хумаюн	1
Рамен	8	Чандлер	1
Рафаэль	4	Шахжихан	9
Рачпромук	2	Шемали	1
Сэбит	1	Шива	14
Сидхарта Будда	2	Шолохов	1
Сита	6	Эмпинавудранат	1
Субраманян Шастри	1	Юлий Цезарь	1
Фенимор Купер	1		

**М. О. ӘУЕЗОВТІҢ «ИНДИЯ ОЧЕРКТЕРІНДЕГІ»
ГЕОГРАФИЯЛЫҚ АТАУЛАРДЫҢ ЖИЛІК СӨЗДІГІ**

географиялық атаулар	кездесу жиілігі	географиялық атаулар	кездесу жиілігі
Агра	11	Дрезден	2
Аджента	7	Дюссельдорф	2
Азия	7	Европа	37
Алатау	3	Египет	1
Алматы	6	Ескі Дели	2
Альпы	3	Жайпур	2
Америка	13	Жаңа Дели	2
Англия	15	Женева	6
Апенин (аралы)	1	Женева (көлі)	1
Арабстан (теңізі)	4	Жумна (өзені)	2
Арабстан (шөлі)	2	Иерусалим	2
Аравали	1	Индия	369
Аурангабад	1	Индия (мұхиты)	5
Бангалор	13	Индия (теңізі)	1
Афина	1	Иран	8
Африка	3	Италия	7
Банжора	1	Италия жарым (аралы)	1
Батыс Германия	1	Каверн	1
Батыс Европа	2	Қаир	6
Бенарес	30	Калькутта	24
Бенгал (қойнауы)	3	Каморин (мүйісі)	8
Бенгал (уәлаяты)	2	Қарачи	4
Бенгали (штаты)	4	Қашмир	3
Бирма	1	Қербаланың (шөлі)	2
Бқара (каналы)	1	Копенгаген	2
Бомбей	57	Қордово (халифаты)	2
Британия	3	Кочин	6
Вагера	1	Крит	2
Ватикан	7	Крит (аралы)	4
Венгрия	1	Кіші Алматы (биігі)	1
Венция	2	Қазақстан	1
Ганг	9	Құдұс	2
Германия	3	Қытай	2
Греция	5	Лахор	2
Дания	1	Ленинград	2
Дели	56	Мадрас	27

Майсур	10	Совет Одағы	2
Майсур (тоғайы)	3	Советтік Шығыс	3
Малавар	1	Стамбул	1
Махнн	1	Стокгольм	1
Милан	1	Сурат	1
Монблан	4	Суэц (каналы)	2
Москва	8	Талғар (биігі)	1
Мысыр	4	Тәжікстан	1
Непал	1	Тибет	1
Норвегия	2	Токио	1
Оңтүстік Индия	1	Травинкор	3
Орта (теңіз)	2	Тривандрум	9
Осло	2	Удайпур	21
Пакистан	5	Фатей (көлі)	1
Пеллепоннес жарым (аралы)	1	Ферғана	1
Пенджап	4	Флоренция	1
Пенджап (уәлаяты)	2	Франция	2
Пичула (көлі)	1	Хайдарабат	12
Полтава	1	Хельсинки	2
Рагнер	1	Цейлон (аралы)	5
Раджестан	5	Чандигар	1
Раджестан (уәлаяты)	1	Чандогарх	1
Рангун	1	Читоргар	1
Рим	35	Швейцария	3
Россия	4	Швеция	1
Самарқан	1	Шивалик	1
Сетледж	3	Югославия	1
Симмла	2	Япония	1
Сиям	1		

МАЗМҰНЫ

Алғы сөз 5

Түркі тілдерінің машиналық қорын жасау және статистикалық зерттеу мәселелері

Е. Қ. Жұбанов. Лингвистикалық зерттеулерді математикаландыру идеяларының бастау көздері (тілдік элементтерді сандық тұрғыдан талдау) 12

А. Б. Белботаев, А. Қ. Жұбанов, З. А. Сиразитдинов. Түркі тілдеріндегі ықтималды-статистикалық әдістер негізінде жүргізілген лингвистикалық зерттеулер туралы 21

А. Қ. Жұбанов. Түркі тілдерінің машиналық қорын жасаудың алғышарттары 31

Фонетика бойынша зерттеулер

Ә. Ж. Жүнісбеков. Орфография мен транскрипция жайында бірер сөз 44

Л. Гржебичек. Абай Құнанбаев поэзиясындағы дауысты дыбыстар үндестігінің эстетикалық қызметі 49

А. Е. Жұбанова. Поэтикалық тексті статолингвистикалық тұрғыдан зерттеу 53

Д. А. Молдалиева. Қазақ текстеріндегі дауысты және дауыссыз дыбыстардың энтропиясы 57

З. М. Базарбаева. Интонация және оның семантикамен байланысы жөнінде 63

Ә. Ж. Жүнісбеков, Қ. Қ. Алдабергенова. Қазақ тіліндегі сингармодыбыстарға статистикалық талдау 71

Лексика бойынша зерттеулер

М. Қ. Айымбетов, Қ. Б. Бектаев. Қарақалпақ, өзбек және қазақ публицистикасының лексикалық жақындығы туралы лингвостатистикалық зерттеулер 76

Б. А. Мұратова. Орысша-қазақша сөздіктегі жұрнақты зат есімдердің лексикографиялық тәсілмен берілуі 90

А. Б. Белботаев, М. О. Әуезовтің «Индия очерктеріндегі» жалқы есімдер статистикасы	94
З. А. Сиразятдинов. Башқұрт жазушысы Д. Юлтыйдың «Қан» романының бірінші және екінші бөлімдерінің сөздіктерінің арасындағы лексикалық байланыс	100
С. А. Өскенбаев. Қазақ тіліндегі көп мағыналы сөздерді статистикалық зерттеу	105
К. Молдабеков. Қазақ текстеріндегі орыс тілінен енген сөздерді лингвостатистикалық тұрғыда зерттеу	113

Морфология мен синтаксис бойынша зерттеулер

А. Қ. Жұбанов. Текст лингвистикасы жайында	118
С. Мырзабеков. Сөз таптарының статистикасы	126
А. Р. Зекенова, Қ. Б. Бектаев, М. О. Әуезовтің драмалық шығармаларындағы зат есімдердің морфологиялық ерекшеліктері	130
А. Б. Белботаев. Сөзформаларындағы сөз таптарының статистикалық сипаттамасы (қазақ текстерінен түзілген жиілік сөздіктер негізінде)	139
С. А. Өскенбаев, А. Қ. Жұбанов. Қазақ сөзформасын автоматты түрде морфологиялық талдау	148
А. Р. Зекенова. Драмалық шығармалар тіліндегі зат есімнің лексикалық құрылымы	157

Қосымша

А. Б. Белботаев, А. Р. Зекенова, М. О. Әуезов шығармаларындағы (драматургия, повесть) етістік сөздердің үлестірімдік алфавитті-жиілік сөздігі	162
А. Б. Белботаев, М. О. Әуезовтің. «Индия очерктеріндегі» жалқы есімдердің түсіндірме сөздігі	188

СОДЕРЖАНИЕ

Предисловие	8
-----------------------	---

Вопросы статистического исследования и создания машинного фонда тюркских языков

Е. Х. Жубанов. У истоков идей математизации лингвистических исследований (количественный анализ элементов речи)	12
А. Б. Белботаев, А. Х. Джубанов, З. А. Сиразитдинов. О лингвистических исследованиях с применением вероятностно-статистических методов в тюркских языках	21
А. Х. Джубанов. Предпосылки создания машинного фонда тюркских языков	31

Исследования по фонетике

А. Д. Джунисбеков. Замечания по орфографии и транскрипции	
Л. Гржебичек. Эстетические функции гармонии гласных в поэзии	
Абая Кунанбаева	44
А. Е. Жубанова. Статолингвистическое изучение казахского поэтического языка	49
Д. А. Молдалиева. Энтропия гласных и согласных букв казахского текста	53
З. М. Базарбаева. К вопросу об интонации и ее связи с семантикой (на материале казахского языка)	57
А. Д. Джунисбеков, С. Алдабергенова. Статистический анализ сингармозвуков в казахском языке	63

Исследования по лексике

М. К. Айымбетов, К. Б. Бектаев. Лингвостатистические исследования лексической близости каракалпакского, казахского и узбекского подъязыков публицистики	76
Б. А. Муратова. Лексикографическое оформление суффиксальных существительных в русско-казахском словаре	90
А. Б. Белботаев. Статистика собственных имен «Индийских очерков» М. О. Ауэзова	94

З. А. Сиразитдинов. Лексическая связь словарей первой и второй части романа «Кан» башкирского писателя Д. Юлъя	100
С. А. Ускомбаев. Статистическое исследование многозначности слов в казахском языке	105
К. М. Молдабаев. Лингвостатистическое изучение заимствований из русского языка в казахских текстах	113

Исследования по морфологии и синтаксису

А. Х. Джубанов. К вопросу о лингвистике текста	118
С. М. Мырзабаев. Статистика частей речи	126
К. Б. Бектаев, А. Р. Зекенова. Морфологические особенности употребления существительного в драмах М. О. Ауэзова	130
А. Б. Белботаев. Статистическая характеристика частей речи в словоформах (на основе материалов частотных словарей казахских текстов)	139
С. А. Ускомбаев, А. Х. Джубанов. Автоматический анализ морфологии казахской словоформы	148
А. Р. Зекенова. Лексическая структура имен существительных в языке драматургии	157

Приложение

А. Б. Белботаев, А. Р. Зекенова. Распределительный алфавитно-частотный словарь глагольных слов в произведениях М. О. Ауэзова (драматургия, повесть)	162
А. Б. Белботаев. Словарь собственных имен «Индийских очерков» М. О. Ауэзова	188

Научное издание

СТАТИСТИКА КАЗАХСКОГО ТЕКСТА

*Утверждено к печати Ученым советом
Института языкознания
Академии наук Казахской ССР*

Рецензенты: доктора филологических наук
А. Б. Болганбаев, М. С. Сергалиев

Зав. редакцией *Дробжева Л. И., Беркимбаев Т. Б.*
Редакторы *Сатова Н. Р., Аманбекова Р. Ш., Сергиенко Т.*
Художественный редактор *В. Н. Афуксениди*
Оформление художника *Н. Ф. Чурсина*
Технический редактор *В. К. Горячкина*
Корректоры *Кинасова З. А., Каймулдина С. О.*

ИБ № 3148

Сдано в набор 26.10.89. Подписано в печать 3.07.90. УГ10105.
Формат 60×84¹/₁₆. Бум. тип. № 1. Литературная гарнитура. Высокая печать.
Усл. п. л. 12,09. Усл. кр.-отт. 12,38. Уч.-изд. л. 12,79. Тираж 800.
Заказ 258. Цена 2 р. 80 к.

Издательство «Гылым» Казахской ССР
480100, Алма-Ата, ул. Пушкина, 111/113
Типография издательства «Гылым» Казахской ССР
480021, Алма-Ата, ул. Шевченко, 28

2

00

[Redacted]

[Redacted]

[Redacted]

[Redacted]