МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РЕСПУБЛИКИ КАЗАХСТАН

Евразийский национальный университет имени Л.Н.Гумилева

Научно-исследовательский институт «Искусственный интеллект»

Алтынбек Шарипбай

ПРОБЛЕМЫ ПЕРЕВОДА КАЗАХСКОЙ ПИСЬМЕННОСТИ НА ЛАТИНСКИЙ АЛФАВИТ

Астана, 2017

НИИ Искусственный интеллект

УДК 81'351+81'352+81'353

ББК Ш1
Ш

Рецезенты:

Қамзабекұлы Д. – доктор филологических наук, профессор, проректор по социально-культурному развитию ЕНУ имени Л.Н.Гумилева.

Ускенбаева Р.К. – доктор технических наук, профессор, проректор по академическим вопросам Международного университета информационных технологий.

Жакупов Ж. - доктор филологических наук, профессор кафедры казахского языкознания Евразийского национального университета им. Л.Н. Гумилева.

Автор:

Ш-25 Шарипбай А.А – директор НИИ «Искусственный интеллект», профессор кафедры «Информатика и Информационная безопасность» ЕНУ имени Л.Н. Гумилева, доктор технических наук, профессор, лауреат государственной премии РК.

Ш-25 Шарипбай А.А. Проблемы перевода казахской письменности на латинский алфавит – Монография, Астана, 2017, -138 С.

РЕКОМЕНДОВАНО
Ученым советом
Евразийского национального университета имени Л.Н.Гумилева

ISBN 978-601-240-913-0

В монографии уточняется звуковая система казахского языка и предлагаются обозначения казахских звуков символами международного фонетического алфавита. Приводятся результаты артикуляционного и акустического анализа звуков казахского языка на основе компьютерного моделирования. Обосновывается необходимость проведения реформы казахского языка. Даются сведения о латинском алфавите и опыте перевода на латинскую графику письменности некоторых тюркских языков. Предлагаются концепция, критерии, задачи и алгоритм перевода казахской письменности на латинский алфавит.

Материалы монографии могут быть полезны широкому кругу читателей, которые проводят как теоретические исследования по казахскому языку, так и практические разработки в области компьютеризации казахского языка и искусственного интеллекта.

1404000000
Ш–25
ISBN 978-601-301-988-8

УДК 81'351+81'352+81'353
ББК Ш1
© Шарипбай А.А., 2017

НИИ Искусственный интеллект

СОДЕРЖАНИЕ

ПРЕДИСЛОВИЕ **6**

ТЕРМИНЫ, ОБОЗНАЧЕНИЯ И СОКРАЩЕНИЯ **13**

ВВЕДЕНИЕ **20**

1. ЗВУКИ ЕСТЕСТВЕННОГО ЯЗЫКА **30**

1.1. Органы речи и органы слуха **30**

1.2. Методы анализа звуков естественных языков **35**

2. ЗВУКОВАЯ СИСТЕМА КАЗАХСКОГО ЯЗЫКА **41**

2.1. Артикуляционный анализ фонем казахского языка **41**

2.1.1. Артикуляция гласных фонем казахского языка 41

2.1.2. Артикуляционная классификация гласных звуков казахского языка 43

2.1.3. Артикуляция согласных звуков казахского языка 54

2.1.4. Артикуляционная классификация согласных звуков казахского языка 60

2.2. Акустический анализ звуков казахского языка **64**

2.2.1. Акустический анализ гласных звуков казахского языка 64

2.2.2. Акустический анализ согласных звуков казахского языка 72

2.2.3. Сравнение казахских звуков со звуками других языков 77

3. КАЗАХСКИЙ АЛФАВИТ **85**

3.1. Перевод казахского алфавита на латиницу **85**

3.1.1. Латинский алфавит и его разновидности 85

3.1.2. Опыт перевода алфавита тюркских языков на латиницу 88

3.1.3. Критерий и задачи перевода казахской письменности на латинский алфавит 96

3.1.4. Правила избавления казахского текста от инородных букв 103

3.1.5. Правила перевода казахской письменности на латиницу 112

ПРИЛОЖЕНИЕ. ГИМН РЕСПУБЛИКИ КАЗАХСТАН **123**

СПИСОК ЛИТЕРАТУРЫ **124**

Список публикаций научной школы **126**

НИИ искусственный интеллект

ПРЕДИСЛОВИЕ

Казахский язык является одним из тюркских языков, имеющий древнюю историю и формировавшийся на протяжении многих веков как и языки родственных тюркских племён, проживавших на территории современного Казахстана [1-4]. Поэтому он, как и любой древний язык, нуждается в исследовании его истории и естественных лингвистических особенностей и возможностей.

Казахский язык является национальным языком казахов, проживающих на территории Республики Казахстан и во многих странах ближнего и дального зарубежья [5]. Поэтому он, как и любой национальный язык, служащий основным средством общения носителей языка, должен изучаться и развиваться для использования его в ежедневной жизни казахского народа.

Казахский язык является государственным языком Республики Казахстан [6]. Поэтому он, как любой государственный язык, должен служить средством общения народов, населяющих территорию Казахстана, поддерживаться государством и использоваться во всех сферах его деятельности, в том числе и во внешних отношениях с другими странами.

Казахстан, интегрируясь в мировое сообщество, должен обеспечивать достаточный (по отношению к мировому уровню) уровень развития государственного языка. Это требование – насущная необходимость при формировании единого культурного и информационного пространства этнических казахов. Поскольку более 15 милионов казахов, живущих в разных странах мира, используют разные алфавиты, основанные на арабице, латинице и кириллице, что мешает их письменному общению.

В своем Послании народу Казахстана от 14.12.2012 Президент страны – Елбасы Н.А. Назарбаев сказал «**Нам необходимо, начиная с 2025 года, приступить к переводу нашего алфавита на латиницу, на латинский алфавит**. Это принципиальный вопрос, который нация должна решить. Ради будущего наших детей мы должны принять такое решение, и это создаст условия для нашей интеграции в мир, лучшего изучения нашими детьми *английского языка и языка Интернета*, и самое главное – это даст толчок модернизации казахского языка. Мы должны провести **модернизацию казахского языка**… .» [7].

НИМ Могущественный интеллект

В своей программной статье «Ориентир на будущее: духовное обновление», опубликованной в республиканской газете «Егемен Қазақстан» Президент Республики Казахстан Н.А.Назарбаев предложил *Проект перехода на латинский алфавит*: «... м**ы должны начинать работы по постепенному переводу казахского языка на латинский алфавит**. … То есть, с 2025 года деловая документация, периодическая печать, учебники – все должны выходить уже на латинском алфавите. Это время уже подходит, и мы должны незамедлительно приступить к этой работе. Правительство должно подготовить график перехода на латиницу. В переходе на латинский алфавит есть глубокий смысл. Это связано с особенностями внедрения современных технологий и коммуникаций, а также с научным и образовательным процессом 21 века. Сегодня дети в стенах школы учат английский язык и без проблем осваивают латинские буквы. Поэтому для молодого поколения не будет никаких проблем. К концу 2017 года в тесном взаимодействии с учёными и всеми представителями общественности необходимо принять единый стандарт нового казахского алфавита и графики. Мы должны в 2018 году приступить к подготовке специалистов для обучения новому алфавиту и начать готовить учебники на латинице. В следующие два года мы должны провести организационные и методические работы... .» [8].

При реализации этого очень важного и своевременного Проекта прежде всего надо учитывать, что:

1) в эпоху глобализации в научной среде многих стран обсуждается проблема использования единого алфавита в глобальном информационном пространстве, в качестве которого предлагается принять английский алфавит, чтобы сократить затраты на поиск и обработку информационных ресурсов, объем которых растет очень быстро. Кроме того, все производимые в мире компьютеры и другие электронные устройства, в том числе и телефоны, изначально поддерживают и будут поддерживать базовый латинский алфавит. Если же алфавит некоторого языка имеет хотя бы одну букву, отличную от буквы базового латинского алфавита, то, чтобы работать на этих аппаратах с алфавитом этого языка, нужно дополнительно создавать шрифты, драйвера, программы сортировки и поиска данных, которые требуют немалых интеллектуальных, временных, трудовых и

НИИ Искусственный интеллект

финансовых затрат. Например, в сети Интернет можно встретить запись латиницей текста на русском языке из-за отсутствия поддержки кириллицы клиентской машиной.

2) в эпоху цифровизации общества многие естественные (английский, французский, русский, японский, китайский и другие) языки изучаются и развиваются с помощью компьютерных программ: созданы многоязыковые электронные словари, мультимедийные вопросно-ответные системы, интеллектуальные системы обучения и оценки знаний, машинные переводчики с одного языка на другой, системы распознавания и синтеза письменной и устной речи и т.д. Основами таких разработок, безусловно, являются *математические* и *компьютерные* модели грамматических правил естественных языков. Подобные проблемы можно ставить и успешно решать и для казахского языка.

Следует отметить, что в ведущих странах (США, Франция, Япония, Россия и др.) эти работы начались более 60-ти лет назад и ориентированы в основном на языки этих стран (английский, французский, японский, русский и др.) языки и выполняются до сих пор коллективами, насчитывающими сотни специалистов в нескольких научных центрах. Эти работы достаточно хорошо финансируются (в объемах десятков миллионов долларов), так как там хорошо понимают значение этих работ, поскольку их результаты служат для будущего страны и, в конечном счете, для создания субъектов и объектов искусственного интеллекта, что является идеалом достижения человеческого разума в области новейших информационных технологий.

В нашей стране работы в этой области начались сравнительно недавно силами малочисленной группы энтузиастов. Но, несмотря на это, уже много сделано по компьютеризации казахского языка:

– разработаны стандарты кодировки букв казахского алфавита для компьютеров, телефонных аппаратов и абонетских станций сотовой связи и соответствующие программы драйверов и распознавателей казахских букв [46-50];

– созданы электронные многоотраслевые терминологические, толковые и орфографические словари [51-61];

НИИ искусственный интеллект

– исследована фонетика, уточнены звуки казахского языка с целью их представления в международном фонетическом алфавите [62-80];

– построены математические модели морфологических правил казахского языка и разработаны программы для автоматического морфологического анализа и синтеза слов и словоформ (морфологический анализатор и синтезатор) [81-111];

– построены математические модели синтаксических правил и разработаны программы для автоматического синтаксического анализа и синтеза словосочетаний и предложений (синтаксический анализатор и синтезатор) [112-126];

– разработаны программы для автоматического семантического анализа и синтеза слов и предложений для интеллектуальных вопросно-ответных систем на казахском языке [127-150];

– разработаны речевые технологии казахского языка (распознавания и синтеза устной казахской речи) [151-180];

– разработаны технологии для автоматизации создания электронных учебников и интеллектуальных систем обучения и оценки знаний на казахском языке [181-222];

– решены проблемы латинизации письменности казахского языка и разработаны программы автоматичекого перевода (конвертор) конвертации казахских текстов с кириллицы на латиницу [223-232].

Причинами такого быстрого получения столь значительных результатов являются высокий уровень квалификации (знания в области компьютерной лингвистики и навыки работы с современными информационными технологиями) исполнителей этих научно-практических исследований и уникальные свойства казахского языка. Дело в том, что казахский язык обладает замечательными свойствами: наличие закона сингармонизма, отсутствия грамматической категории рода (в отличие от русского и других языков), артиклей, префиксов (приставок) и предлогов; подчиненность словообразования строгим правилам добавления аффиксов (суффиксов и окончаний); однозначность аффиксов, конечность числа аффиксов; определеннность структуры простых предложений и т.д

НИИ искусственный интеллект

Некоторые из этих свойств отсутствуют в других тюркских языках, относящихся к огузской группе. Например, закон сингармонизма не действует в турецком, азербайджанском и узбекском языках.

В ходе выполнения этих перечисленных выше работ под научным руководством автора данной монографии были подготовлены и защищены 4 кандидатские диссертации по специальности «05.13.11 – Математическое и программное обеспечение вычислительных машин, систем и сетей» [233-236] и 5 диссертаций PhD по специальности «6D060200 – Информатика» [237-241].

Отдельные результаты этих исследований доложены на многочисленных международных научных конференциях, опубликованы в рейтинговых научных журналах, индексируемых базами Thomson Reuters и Scopus. Список опубликованных работ приведен в Приложении.

Впереди предстоит огромная работа, которая под силу только большому творческому коллективу, в составе которого должны быть десятки специалистов по компьютерной лингвистике. Это дело не одного поколения, так как естественные языки постоянно развиваются и бесконечны по своей природе. В то же время и информационные технологии не стоят на месте.

Известно, что в любом естественном языке устная речь первична, а письменная речь вторична. Это означает, что звуковая система любого естественного языка является основой его письменности. Также известно, что для создания письменности любого естественного языка необходимо сначала придумать буквы (символы) для обозначения его значимых звуков (фонем), объединив которые в одно множество, можно построить алфавит, а затем нужно разработать орфографические нормы (правописание), морфологические и синтаксические правила для записи его слов (словоформ) и словосочетаний (фраз и предложений) буквами алфавита.

Однако на текущий момент нет общепринятого мнения относительно фонетики казахского языка, до сих пор нет стандарта фонетики, где должны быть уточнены казахские звуки и их классификация. Подобного состояния, как в казахского языке, нет ни в одном другом государственном языке. Это связано с тем, что имеющееся учение о фонетике казахского языка унаследовано из

НИИ Искусственный интеллект

русского языка и не отражает точную характеристику казахских звуков, что было показано в работах [29,30]. Поэтому не понятно, что делать с русскими гласными, которые были внедрены в казахский язык без учета его фонетических закономерностей. Один из способов решения этой проблемы - исключение этих звуков из звуковой системы казахского языка, как в азербайджанском языке, и разработка соответствующих орфографических правил.

Структура предлагаемой монографии составлена так, что сначала будут расматриваться проблемы, связанные с артикуляционным и акустическим анализом, уточнением звуковой системы и определением алфавита казахского языка, затем определяются проблемы перевода казахской письменности на латиницу и даются пути их решения.

В монографии уточнена и обоснована звуковая система казахского языка. Приводятся результаты анализа казахских гласных звуков в спектральной области, чтобы определить их формантную структуру, что не было сделано до сих пор. Также приводятся математические модели системы казахских гласных звуков в виде алгебраического выражения и геометрической фигуры, которые более точно разделяют гласные на классы.

Основные результаты, изложенные в монографии, получены в рамках следующих научных проектов, которые выполнялись по грантовому финансированию МОН РК:

1. Создание акустического корпуса казахского языка и уточнение его фонетического строя, представление казахских фонем в международном фонетическом алфавите, 2012-2014 гг.
2. Автоматизация распознавания и порождения письменной и устной речи казахского языка, 2012-2014 гг.
3. Научные, методологические, технологические и методические основы перевода казахской письменности на латиницу, 2013-2015 гг.

Информацию об этих проектах можно найти по адресу:

http://www.e-zerde.kz/zerde/index.php/ru/

http://www.nauka.kz/page.php?page_id=371&lang=3&page=3090

Монография состоит из Введения и трех глав. Во Введении описывается состояние звуковой системы казахского языка и дается обоснование необходимости комплексной реформы казахского языка. В первой главе обсуждаются звуки естественного языка. Во второй главе рассматривается звуковая система казахского языка. Проведен

НИИ искусственный интеллект

формантный анализ гласных звуков казахского языка, показавший, что множество казахских гласных имеет свою геометрическую структуру, которая раскрывает естественную классификацию гласных. Построена математическая модель системы казахских гласных, позволившая установить некоторые лингвистические факты. В третьей главе описывается опыт перевода алфавита тюркских языков на латиницу и решение проблемы перевода казахской письменности на латиницу.

Монография одновременно издалась на казахском и русском языках. Материалы в ней могут быть полезны широкому кругу читателей, которые проводят как теоретические исследования по казахскому языку, так и практические разработки в области компьютеризации казахского языка и искусственного интеллекта.

Автор выражает свою благодарность сотрудникам лаборатории компьютерных наук ЧУ «National Laboratory Astana» PhD Жандосу Есенбаеву и кандидату технических наук Муслиме Карабалаевой за ценные замечания и предложения, которые способствовали уточнению звуковой системы и построению алфавита казахского языка на латинице, известному зарубежному фонетисту профессору Канзасского университета (США) Аллард Йонгману за консультацию при уточнении звуковой системы и обозначении звуков казахского языка в Международном фонетическом алфавите, выражает свою признательность за улучшение качества текста монографии преподавателям кафедры Информатики и информационной безопасности ЕНУ имени Л.Н.Гумилева PhD и кандидату технических наук Гульмире Бекмановой и PhD Асель Мукановой, Бану Ергеш и Марал Амалбекову, а также благодарит своих рецензентов - профессоров Дихан Қамзабекұлы, Раисе Ускенбаевой и Жантас Жакуповау за понимание содержания и поддержку при издании монографии.

Автор: Алтынбек Амирович ШАРИПБАЙ

НИИ Искусственный интеллект

ТЕРМИНЫ, ОБОЗНАЧЕНИЯ И СОКРАЩЕНИЯ

В качестве библиографической базы формирования и уточнения терминов и их определений были использованы источники [12-33].

№	Термин	Определение
1	Аккомодация	Изменение в артикуляции согласных под влиянием соседних гласных и гласных под влиянием согласных.
2	Акустика	Раздел языкознания, который изучает физические характеристики звука.
3	Аллофон	Варианты фонемы.
4	Алфавит	Множество отдельных символов (букв), предназначенных для обозначения звуков некоторого языка.
5	Альвеолы	Ряд бугорков в ротовой полости между верхними передними зубами и передней частью нёба.
6	Артикуляция	Работа, которую совершают органы речи для получения звука и состоящая из трех моментов: *экскурсии*, *выдержки и рекурсии*.
7	Ассимиляция	Влияние одних гласных на другие гласные и согласных на согласные.
8	Аффрикаты	Согласные, представляющие собой слитное сочетание смычного согласного с фрикативным, обычно того же места образования.
9	Буква	Символ (переменная величина), значением которого является звук (постоянная величина).
10	Взрывные	Смычные согласные, при артикуляции которых нёбная занавеска поднята, и

НИИ Искусственный интеллект

		воздух проходит в ротовую полость (в отличие от носовых смычных), а размыкание смычки происходит резко и напоминает взрыв (в отличие от аффрикат).
	Выдержка	Центральный момент артикуляции, когда орган речи произносит звук.
11	Высота звука	Величина, прямо пропорциональная частоте колебаний, единицей измерения которой является герц (одно колебание в секунду).
12	Губное созвучие	Взаимная губная адаптация звуков в составе слога и слова.
13	Дифтонг	Сложный гласный звук, который состоит из двух компонентов.
14	Дифтонгоид	Сложный гласный звук, который в начале имеет незначительный оттенок другого звука.
15	Звук	Минимальная неделимая фонетическая единица речи.
16	Звук гласный	Звук, который образуется с преобладанием голоса над шумами и является слогообразующим.
17	Звук губной	Звук, образуемый с помощью губ.
18	Звук глухой	Звук, который образуется при раскрытом положений голосовых связок или который состоит из одного шума.
19	Звук звонкий	Звук, который образуется при незначительном колебании голосовых связок или шум преобладает над голосом.
20	Звук согласный	Звук, который образуется с преобладанием шумов над голосом и является неслогообразующим.

НИИ Искусственный интеллект

21	Звук сонорный	Звук, который образуется при большем колебании голосовых связок или голос преобладает над шумом.
	Звуковой язык	Язык, передающийся произношением, и принимающийся слушанием.
22	Лингвистический процессор	Компьютерная программа, способная анализировать и синтезировать тексты на естественном или другом языке.
23	Международный фонетический алфавит	Совокупность знаков для обозначения звуков всех языков мира.
24	Монофтонг	Однородный гласный звук.
25	Обертон	Отличная от наименьшей частоты в частотном спектре звука.
26	Округленный звук	Звук, который образуется при округленном положений губ.
27	Органы речи	Полость рта, носовая полость, губы (верхние губы и нижние губы) зубы (верхние зубы и нижние зубы), альвеолы, язык (кончик языка, передняя часть спинки языка, средняя часть спинки языка, задняя часть спинки языка, корень языка), язычок, небё (твердое небо и мягкое небо нёбо), гортань, голосовые связки, легкие.
28	Орфография	Система правил написания слов
29	Орфоэпия	Система правил произношения слов.
30	Открытый звук	Гласный звук, который образуется при низком положении языка в полости рта.
31	Парадигматическое отношение	Отношение, определяющее группировку слов безотносительно к их актуализации.
32	Перцепция	Работа, которую совершает органы слуха, чтобы услышать звук.

НИИ искусственный интеллект

33	Полисемия	Обозначение одним знаком разных предметов (денотатов).
34	Постфикс	Аффикс, стоящий после корня.
35	Префикс	Аффикс, стоящий перед корнем.
36	Правописание	Система правил написания слов и употребления знаков препинания.
37	Просодика	Фонологическая оболочка слога (слова), образованная из звуков.
38	Рекурсия	Заключительный момент артикуляции, когда органы речи завершают произношение.
39	Релевантность	В информационно-поисковых системах – степень соответствия запроса найденным ресурсам.
40	Семантический анализ	Определение смысловых характеристик слов или словосочетаний.
41	Сигнификат	Совокупность некоторых признаков предмета, которые существенны для его правильного означивания или именования.
42	Сила звука	Величина, прямо пропорциональная амплитуде колебаний и измеряется в децибелах.
43	Сингармонизм	Закон в языке, при котором гласные и согласные звуки в составе слога и слова произносятся единым тембром.
44	Сингармоническая просодика	Просодика слога (слова), образованная по закону сингармонизма.
45	Синонимия	Совпадение языковых единиц, имеющих разное написание или звучание, но одинаковый смысл.
46	Синсет	Совокупность синонимов одного понятия
47	Синтаксический	Определение структуры фрагментов

НИИ Искусственный интеллект

	анализ	(предложений) текста.
48	Словоформа	Цепочка фонем, обладающая признаками слова и формально отличающаяся от другой.
49	Слог	Минимальная единица речевого потока, создается единым дыхательным толчком, состоит из одного гласного звука или из сочетания гласного и согласных.
50	Слог закрытый	Слог, который заканчивается на согласный звук.
51	Слог открытый	Слог, который заканчивается на гласный звук.
52	Смычные согласные	Согласные звуки, при артикуляции которых органы речи находятся в таком положении, что поток воздуха из легких полностью блокируется с помощью смычки, создаваемой в полости рта или в гортани.
53	Согласный звук	Звук, при образовании которого выходящий из легких воздух встречает препятствие в ротовой полости.
54	Спектрограмма	Изображение, показывающее зависимость спектральной плотности мощности звукового сигнала от времени.
55	Спирант	Согласный звук, образующийся вследствие трения воздуха, выходящего из легких, о края щели или сужения, образуемого теми или другими органами речи.
56	Тембр	Окраска звука, определяемая количеством, соотношением по высоте и силе, порядком возникновения обертонов, накладывающихся на основной тон.
57	Токен	Часть слова, остающаяся после отсечения

НИИ Искусственный интеллект

		окончания при морфологическом анализе.
58	Тон	Высота
59	Тон основной	Наименьшая частота в частотном спектре звука.
60	Транскрипция	Знаки, с помощью которых обозначаются оттенки звуков.
	Транслитера́ция	Точная передача знаков одной письменности знаками другой письменности.
61	Увула	Язычок, продолжение мягкого неба, с помощью которого произносятся увулярные (картовые) согласные.
62	Узкий звук	Гласный звук, который образуется в полости рта при верхних положениях языка.
63	Фонема	Минимальная смыслоразличительная фонетическая единица языка, которая не имеет самостоятельного лексического или грамматического значения, но служит для различения и отождествления значимых единиц языка (морфем и слов).
64	Фонематическая транскрипция	Знаки в косых скобках, с помощью которых обозначаются лишь фонемы слов или морфем, без учёта их аллофонов.
65	Фонетика	Раздел языкознания, который изучает систему звуков языка.
66	Фонетическая транскрипция	Знаки в квадратной скобке, с помощью которых обозначаются звучание всех звуков без исключения их вариантов.
67	Фонологическая транскрипция	Знаки, с помощью которых обозначается словоразличительные изменения звуков.
68	Фонология	Раздел лингвистики, изучающий структуру звукового строя языка и

НИИ Искусственный интеллект

		функционирование звуков в языковой системе.
69	Форманта	Акустическая характеристика гласных звуков речи, связанная с уровнем частоты голосового тона и образующая тембр звука.
70	Формантный анализ	Процесс определения формант гласных звуков в спектрограмме речевого сигнала, позволяющий на основе их положения в частотной области решать о качестве гласных.
71	Фрикативные согласные	Шумные согласные звуки, при артикуляции которых артикуляторы, не смыкаясь между собой, образуют щель в ротовой полости.
72	Шум	Непериодические (неритмичные) колебания звучащего тела
73	Экскурсия	Начальный момент, когда органы речи готовятся к произношению.

Сокращения и расшифровка

№	Сокращение	Расшифровка
1	LDC	Linguistic Data Consortium
2	TIMIT	Texas Instruments, Inc (TI) and the Massachusetts Institute of Technology (MIT)
3	дБ	Децибел
4	IPA	The International Phonetic Association
5	МФА	Международный Фонетический Алфавит

НИИ Искусственный интеллект

ВВЕДЕНИЕ

Казахская письменность берет свое начало с древнетюркских рунических письменностей VI–X вв. н.э. Один из вариантов древнетюркских алфавитов, знаками которого высечена орхоно-енисейская надпись на надгробном памятнике Кюль-Тегина, впервые был расшифрован (определены значения знаков алфавита) в 1893 году шведским ученым Вильгельмом Томсеном [9].

Близкие к орхоно-енисейским надписям рунические надписи V–III вв. до н.э. были обнаружены в Казахстане в 50 км от г. Алматы, в Иссыкском кургане, в котором был захоронен «Золотой человек». Эти надписи называются «Иссыкское письмо», состоят из 26 знаков, внешне напоминающих орхоно-енисейскую письменность. Поэтому некоторые исследователи признают вероятность того, что потомком Иссыкского письма является более позднее классическое орхоно-енисейское письмо [10].

В связи с распространением ислама в письменности тюркского мира, начиная с IX по XX века н.э. в течение многих веков использовался арабский алфавит. Известно, что на арабском алфавите написана священная книга мусульман – Коран и более тысячи лет создавали свои бессмертные произведения многие ученые, просветители и поэты, чьи труды вошли в мировую сокровищницу науки, образования и культуры.

Однако арабский алфавит, созданный для семитских языков и хорошо приспособленный к требованиям арабского языка, не до конца отражал богатую фонетическую систему тюркских языков: ряд знаков в нем не был нужен тюркским языкам, и, наоборот, немало звуков, имеющиеся в тюркских языках, не находило в нем отражения. В результате появилась необходимость внесения изменений в арабский алфавит, используемый тюркскими языками.

Новый метод тюркского письма на основе арабской графики, называемый «Усуль аль-джадид», был впервые придуман выдающимся крымско-татарским просветилем Исмаилом Гаспралы. Суть его метода заключается в фонетической обработке

НИ... ...ный интеллект

арабского алфавита, которая ставила в соответствие звуки и буквы, в отличие от старого метода «Усуль аль-кадим», предлагавшего слоговое изучение языка, когда отдельные буквы сливались в слоги, а из слогов потом составлялись слова. Новый метод позволил минимизировать недостатки базового арабского произношения и в 2-3 раза сократить срок обучения грамоте [11].

В XX веке казахский язык поменял свой алфавит 3 раза. В первый раз в 1912 году с помощью метода «Усуль аль-джадид» основоположник теории казахского языка Ахмет Байтурсынов перевел казахскую письменность на новый алфавит на основе арабской графики. Для этого он систематизировал и уточнил звуковую систему казахского языка из 28 звуков, из которых 5 гласные фонемы и 19 согласные фонемы, при этом 4 мягкие гласные считались *аллофонами* (на самом деле они являются *фонемами*). Затем он исключил все арабские буквы, обозначающие неказахские звуки и разработал новый алфавит казахского языка, который содержал 24 буквы на основе арабской графики и 1 специальный знак (апостроф). При этом для обозначения мягких гласных использовались диграфы, в которых первыми парами были буквы, обозначающие твердые согласные, а в каждом диграфе второй парой был апостроф. Этим алфавитом в нашей стране пользовались до 1929 года, а казахи, живущие в других странах, в частности, в Китае, Афганистане и Иране, пользуются им и по сей день, добавив еще 3 буквы для обозначения согласных (*вы*), (*фы*) и (*хы*) [12]. Во второй раз в 1929 году вместо алфавита на основе арабицы был принят алфавит на основе латиницы [13], который действовал в нашей стране более 10 лет. В третий раз, в 1940 году, вместо алфавита на основе латиницы был принят алфавит на основе кириллицы. Были добавлены 10 новых букв: две буквы **Ьь** и **Ъъ** для обозначения мягкости и твердости согласных, 8 букв **Ёё, Ии, Уу, Цц, Чч, Щщ, Ээ, Юю, Яя** для обозначения не свойственных казахскому языку русских звуков. Этот алфавит действует до настоящего времени.

НИИ Институт

В таблице 1 показано обозначение фонем (значимых звуков) казахского языка в XX-ом веке, без указанных 10 букв.

Таблица 1. Обозначение фонем казахского языка в XX-ом веке.

№	Название звуков	Обозначение на арабице	Обозначение на латинице	Обозначение на кириллице
1	(а)	ا	A a	А а
2	(ә)	ٱ	Ə ə	Ә ә
3	(бы)	ب	B в	Б б
4	(вы)	ۆ	*Vv*	В в
5	(гі)	گ	G g	Г г
6	(ғы)	ع	Ġ ǵ	Ғ ғ
7	(ды)	د	D d	Д д
8	(е)	ه	E e	Е е
9	(жы)	ج	Č č	Ж ж
10	(зы)	ز	Z z	З з
11	(й)	ي	J j	Й й
12	(кі)	ك	K k	К к
13	(қы)	ق	Q q	Қ қ
14	(ыл)	ل	L l	Л л
15	(мы)	م	M m	М м
16	(ны)	ن	N n	Н н
17	(ыңы)	ڭ	Ŋ ŋ	Ң ң
18	(о)	و	O o	О о
19	(ө)	ۇ	Ö ö	Ө ө
20	(пы)	پ	P p	П п
21	(ыр)	ر	R r	Р р
22	(сы)	س	S s	С с
23	(ты)	ت	T t	Т т
24	(у)	ۋ	V v	У у
25	(ұ)	ۇ	U u	Ұ ұ
26	(ү)	ۇٛ	Y y	Ү ү
27	(фы)	ف	F f	Ф ф
28	(хы)	ح	H h	Х х
29	(шы)	ش	C c	Ш ш

НИИ искусственный интеллект

30	(ы)	ى	Ь ь	Ы ы
31	(і)	ىٔ	І і	І і

До сих пор состав звуковой системы казахского языка является предметом дискуссии. Издаются учебные пособия и научные труды, авторы которых дают разные сведения о количестве и типе казахских звуков. В таблице 2 приводятся некоторые учебные издания по казахскому языку, вышедшие только на казахском языке.

Таблица 2. Учебные издания по казахскому языку.

№	Авторы	Наименование источника	К-во и виды гласных звуков	К-во и виды согласных звуков
1	А. Байтұрсынов	Тіл тағылымы. - Алматы: Ана тілі, 1992, -448 бет.	5 звуков ***а,е,о,у,ы*** 2 полугласные звуки ***у,й***	17 звуков ***б,г,ғ,д,ж,з,к,қ,л,м,н,ң,п,р,с,т,ш***
2	Қ. Жұбанов	Қазақ тілі жөніндегі зерттеулер. - Алматы: Ғылым, 1966,- 362 бет.	7 звуков ***а,е,о,у,ы,ұ,у,ы,й***	12 звуков ***б,г,ғ,д,ж,з, к,қ,п,с,т,ш*** 7 сонорных ***л,н,ң,м,у,й,р***
3	Ғ.Әбуханов	Қазақ тілі. (Пед. училищеге арналған оқулық). - Алматы: Мектеп,1982, -284 бет.	12 звуков ***а,ә,е,и,о,ө, у,ұ,ү,ы,і,э***	25 звуков ***б,в,г,ғ,д,ж,з,й,к,қ,л, м,н,ң,п,р,с,т,ф,х,h, ц,ч,ш,щ***
4	Н.Оралбаева, Ғ.Мадина, А.Әбілқаев	Қазақ тілі (ЖОО. Филология мамандығына арналған оқулық). - Алматы: Мектеп, 1982, -296 бет.	12 звуков ***а,ә,е,и,о,ө, у,ұ,ү,ы,і,э***	26 звуков ***б,в,г,ғ,д,ж,з, й, к,қ,л,м,н,ң, п,р, с,т,у,ф,х, h,ц,ч,ш,щ***
5	І. Кеңесбаев	Қазақ тіл білімі туралы зерттеулер.- Алматы: Ғылым, 1987, -352 бет.	11 звуков монофтонги ***а, ә, е, о, ө, ұ,ү,ы,і,э*** дифтонги ***и,у***	25 звуков ***б,г,д,ж,з,й,к, л,м,н,п,р,с,т, (у),ш,қ,ң,ғ,h, в,ф,х,ц,ч***
6	Ж. Аралбаев	Қазақ фонетикасы бойынша этюдтер. - Алматы:Ғылым, 1988. -144 бет.	11 звуков, монофтонги ***а, ә, е, о, ө, ы, і, ұ, ү, э*** дифтонги ***и,у***	25 звуков ***б,п,м,у,н,л,р, т,д,с,з,ж,ш,ч, ц,й,к,г,қ,ғ,х, ң,в,ф,h***
7	А.Ысқақов, К.Аханов, Б.Кәтенбаева	Қазақ тілі (5-сыныпқа арналған оқулық. 18-ші басылым). - Алматы: Мектеп, 1989, -192 бет.	12 звуков ***а,ә,и,е,э,о,ө, ү,ұ,(у),ы,і***	26 звуков ***б,в,г,ғ,д,ж,з,й, к,қ,л,м,н,ң,п, р,с, т,у,ф,х,h, ш,щ,ц,ч***
8	Р.Әміров, А.Бәкірова	Қазақ тілі (1-сыныпқа арналған	15 звуков ***а, ә, е, е, и, о, ө,у,ұ, ү,***	25 звуков ***б,в,г,ғ,д,ж,з,й,к,қ,л,***

НИИ Искусственный интеллект

		оқулық). - Алматы: Мектеп, 1990, -158 бет.	***ы***, ***і***, ***э***, ***ю***, ***я***	***м***,***н***,***ң***,***п***,***р***,***с***,***т***, ***ф***,***х***,***h***,***ц***,***ч***, ***ш***,***щ***
9	С.Рахметова	Қазақ тілін оқыту методикасы (Пед. училищеге арналған оқулық). - Алматы: Ана тілі, 1991, -184 бет.	13 звуков ***а***, ***ә***, ***е***, ***ё***, ***и***, ***о***, ***ө***, ***у***, ***ұ***, ***ү***, ***ы***, ***і***, ***э***	25 звуков ***б***,***в***,***г***,***ғ***,***д***,***ж***, ***з***,***й***,***к***,***қ***,***л***,***м***,***н***,***ң***, ***п***,***р***,***с***,***т***, ***ф***, ***х***,***h***,***ц***,***ч***,***ш***,***щ***
10	М.Қараев	Қазақ тілі. - Алматы: Ана тілі, 1993, -216 бет.	11 монофтонги ***а***, ***е***, ***ы***, ***і***, ***ә***, ***о***, ***ө***, ***ұ***, ***ү*** дифтонгоид ***и***,***у***	25 звуков ***б***,***в***,***г***,***ғ***,***д***,***ж***,***з***,***й***,***к***,***л***, ***м***,***н***,***п***,***р***,***с***,***т***,(***у***),***h***, ***ф***, ***х***, ***ц***,***ч***,***ш***,***қ***,***ң***
11	Н.Оралбаева, Т.Әбдіғалиева, Б.Шалабаев	Практикалық қазақ тілі. - Алматы: Ана тілі, 1993, -272 бет.	12 звуков ***а***, ***ә***, ***е***, ***о***, ***ө***, ***ұ***, ***ү***, ***ы***,***і***, ***э*** дифтонги ***и***,***у***	24 звуков ***б***,***в***,***г***,***ғ***,***д***,***ж***,***зй***,***к***,***қ***, ***м***,***н***,***ң***,***п***,***р***,***с***,***т***,***ф***, ***х***,***h*** ***ц***,***ч***,***щ***
12	И.Маманов	Қазақ тілі. - Алматы: Мектеп, 1989. -176 бет.	15 звуков ***а***, ***ә***, ***е***, ***ё***, ***и***, ***о***, ***ө***, ***у***, ***ұ***, ***ү***, ***ы***, ***і***, ***э***, ***ю***, ***я***	25 звуков ***б***,***в***,***г***,***ғ***,***д***,***ж***,***з***,***й***,***к***,***қ***, ***л***,***м***,***н***,***ң***,***п***,***р***,***с***,***т***,***ф***, ***х***,***h***,***ц***,***ч***,***ш***,***щ***
13	С.Мырзабеков	Қазақ тілінің фонетикасы. - Алматы: Қазақ университеті, 1993, -136 бет.	11 звуков ***а***, ***ә***, ***о***, ***ө***, ***ұ***, ***ү***, ***ы***, ***і***, ***е***, ***у***, ***и***	25 звуков ***б***,***в***,***г***,***ғ***,***д***,***ж***,***з***,***й***,***к***,***л***, ***м***,***н***,***қ***,***ң***,***п***,***р***,***с***,***т***,***у***, ***h***,***ф***,***х***,***ц***,***ч***,***ш***
14	С.М.Исаев	Қазақ тілі. - Алматы: Қайнар, 1993. - 170 бет.	12 звуков ***а***, ***ә***, ***е***, ***и***, ***о***, ***ө***, ***у***, ***ұ***, ***ү***, ***ы***, ***і***, ***э***	26 звуков ***б***,***в***,***г***,***ғ***,***д***,***ж***,***з***,***й***,***к***,***қ***, ***л***,***м***,***н***,***ң***,***п***,***р***,***с***,***т***,***у***, ***ф***,***х***,***h***,***ц***,***ч***,***ш***,***щ***
15	Ш.Бектұров, М.Серғалиев	Қазақ тілі. -Алматы: Білім, 1994, -224 бет.	15 звуков монофтонги ***а***,***ә***,***е***,***о***,***ө***, ***ұ***,***ү***,***ы***,***і*** дифтонги ***и***,***у***, ***э***, ***е***, ***ю***, ***я***	25 звуков ***б***,***в***,***г***,***ғ***,***д***,***ж***,***з***, ***й***,***к***,***л***,***м***,***н***,***ң***,***п***, ***р***,***с***,***т***,(***у***),***h***, ***ф***,***х***,***ц***,***ч***,***ш***,***щ***
16	С.Аманжолов, Ә.Хасенов, И.Ұйықбаев	Қазақ тілі (9-сыныпқа арналған оқулық). - Алматы: Рауан, 1996, -126 бет.	12 звуков ***а***,***ә***,***э***,***е***,***и***,***о***,***ө***,***у***,***ұ***,***ү***,***ы***,***і***	24 звуков ***б***,***в***,***г***,***ғ***,***д***,***ж***,***з***,***й***,***к***,***қ***, ***л***,***м***,***н***,***ң***,***п***,***р***,***с***,***т***, ***ф***,***х***,***h***,***ш***,***ц***,***ч***
17	Ш.Әуелбаев, Ә.Наурызбаева, Р.Ізғұттынова, А.Құлжанова	Әліппе. - Алматы: Ата-мұра, 1997, -160 бет.	15 звуков ***а***,***у***,***о***,***ы***,***ұ***,***е***,***і***,***ә***, ***и***,***ө***,***ү***,***я***,***ю***,***е***,***э***	25 звуков ***б***,***в***,***г***,***ғ***,***д***,***ж***,***з***,***й***,***к***,***қ***, ***л***,***м***,***н***,***ң***,***п***,***р***,***с***,***т***, ***ф***,***х***,***h***,***ш***,***щ***,***ц***,***ч***
18	С.Исаев, К.Назарғалиева	Қазақ тілі (7-сыныпқа арналған оқулық). - Алматы: Рауан, 1997, -192 бет.	12 звуков ***а***,***ә***,***е***,***о***,***ө***, ***ұ***,***ү***,***ы***,***і***,***э***,***и***,***у***	26 звуков ***б***,***в***,***г***,***ғ***,***д***,***ж***,***з***,***й***,***к***,***қ***,***л***, ***м***,***н***,***ң***,***п***,***р***,***с***,***т***,***у***,***ф***,***х***, ***h***,***ц***,***ч***,***ш***,***щ***
19	Жалпы редакциясын басқарған: А.Сейдімбек	Қазақ тілі (әмбебап анықтамалық). - Алматы: Болашақ, 1999, -150 бет.	13 звуков ***а***,***ә***,***е***,***и***,***о***,***ө***,***у***,***ұ***,***ү***, ***ы***,***і***,***ю***,***я***	25 звуков ***б***,***в***,***г***,***ғ***,***д***,***ж***,***з***,***й***,***к***,***қ***,***л***, ***м***,***н***,***ң***,***п***,***р***,***с***,***т***,***ф***,***х***, ***ц***,***ч***,***ш***,***щ***,***h***

НИИ Искусственный интеллект

20	К.Сариева	Қазақ тілі. Оқу жаттығу құралы. - Алматы: Білім, 2000. -136 бет.	15 звуков ***а,ә,е,е,и,о,ө,у,ұ,ү,ы,і,э,ю,я***	25 звуков ***б,в,г,ғ,д,ж,з,й,к,қл,м,н,ң,п,р с,т,фх,h,ц,ч,ш,щ***
21	С.Аманжолова, Ә.Хасенов, И.Ұйықбаев	Қазақ тілі (9-сыныпқа арналған оқулық. 20-шы басылым). - Алматы: Рауан,1996, -126 бет.	12 звуков ***а,ә,е,и,о,ө, у,ұ,ү,ы,і,э***	24 звуков ***б,в,г,ғ,д,ж,з,й,к,қ,л, м,н,ң,п,р,с,т,ф,х,h, ш,ц,ч***
22	Г.Қосымова	Қазақ тілі (5-сыныпқа арналған оқулық). - Алматы: Ата -мұра, 2015, -192 б.	12 звуков ***а,ә,е,и,о,ө, у,ұ,ү,ы,і,э***	26 звуков ***б,в,г,ғ,д,ж,з,й,к,қ,л, м,н,ң,п,р,с,т,у,ф,х, ц,ч,ш,щ h, (у)***
23	А.Ысқақова, К.Аханов, Б.Кәтенбаева	Қазақ тілі (5-сыныпқа арналған оқулық.). - Алматы: Мектеп, 1989, -192 бет.	12 звуков ***а,ә,э,е,и,о,ө, (у),ұ,ү,ы,і,э***	26 звуков ***б,в,г,ғ,д,ж,з,й,к, қ,л,м,н,ң,п,р,с,т, у,ф,х,h,ш,щ, ц,ч***
24	Г.Қосымова, Ж.Даулетбекова	Қазақ тілі (5-сыныпқа арналған оқулық.). - Алматы: Ата-мұра, 2005, -192 бет.	12 звуков ***а,ә,е,и,о,ө, у,ұ,ү,ы,і,э***	26 звуков ***б,в,г,ғ,д,ж,з,й,к, қ,л,м,н,ң,п,р,с,т,у, ф,х,ц,ч,ш,щ h,(у)***
25	Ж.Қ. Балтабаева және басқалары	Қазақ тілі (2-сыныпқа арналған оқулық.). - Алматы: Алматыкітап, 2013, -224 бет.	14 звуков ***а,ә,е,и,о,ө, у,ұ,ү,ы,і,э, ю,я*** қосарлы звуковтар	26 звуков ***б,в,г,ғ,д,ж,з,р,л,м,н, ң,й,(у),к,қ,п ф,с,т,х,ц,ч,ш,щ,h***
26	Ш.Әуелбаев, Ә.Наурызбаева және басқалары	Әліппе. - Алматы: Ата-мұра, 1997, -160 бет.	15 звуков ***а,ә,е,ё,и,о,ө, у,ұ,ү,ы,і,э,ю,я***	25 звуков ***б,в,г,ғ,д, ж,з,й,к,қ,л, м,н,ң,п, р,с,т,ф,х,h, ш,щ,ц,ч***
27	Ш.Әуелбаев, Ә.Наурызбаева, және басқалары	Әліппе. - Алматы: Ата-мұра, 2002, -144 бет.	15 звуков ***а,ә,е,ё,и,о,ө, у,ұ,ү,ы,і,э,ю,я***	26 звуков ***б,в,г,ғ,д,ж, з,й,к,қ,л, м,н,ң, п, р, с,т,у,ф, х,h,ц,ч,ш, щ***

Теперь, чтобы показать наличие противоречия в казахской письменности, рассмотрим 2 примера:

1. Использование в казахском языке гласных звуков (***и***) и (***у***) породило ряд противоречий. Так, например, в современном казахском письме к словам «*би*» - танец, «*ми*» - мозг, «*бу*» - пар, «*гу*» - гул, которые заканчиваются на звуки (***и***) и (***у***), притяжательные окончания в третьем лице добавляются как «*ми+ы*», «*би+і*», «*ту+ы*», «*гу+і*». Они противоречат правилу: “*В казахском языке к словам, заканчивающимся на согласные, присоединяется притягательное окончание в третьем лице «і» или «ы», а к словам, заканчивающимся на гласные – «сі» или «сы»*”.

НИИ Искусственный интеллект

Если в казахском языке звуки (***и***) и (***у***) считать согласными, то в словах «*ми*», «*би*», «*ту*», «*гу*» не окажется ни одного слога. Отсюда возникает два вопроса: “*Правильны ли в казахском языке правила для окончаний*?”, “*Бывают ли в казахском языке слова без единого слога*?”. Чтобы не порождать подобные вопросы, необходимо правильно переписать эти слова с помощью казахских звуков: «*мый*», «*бій*», «*тұу*», «*гүу*». Тогда единицы произношения казахского языка типа корень, слог и морфема будут писаться без искажения, т.е. упомянутые окончания добавятся как «*мый+ы*», «*бій+і*», «*тұу+ы*», «*гүу+і*».

Отсюда можно заключить, что, во-первых, в казахском языке нет гласного звука (***и***), во-вторых, звук (***у***), обозначенный сейчас буквой *Уу,* является согласным и этого звука его нет в русском языке, но он есть в английском языке и обозначается в международном фонетическом алфавите – МФА [31, 32] как [*w*]**.** Звук в русском языке, обозначенный буквой *Уу,* является гласным и в МФА имеет обозначение [*u*], т.е., несмотря на схожесть звучания, нельзя путать между собой согласный звук [*w*] в казахском языке и гласный звук [*u*] в русском языке. Правильное использование указанных звуков в казахском языке будет подтверждать правило: «*В казахском языке слог внутри слова не может начинаться с гласного звука*», т.е. исправленные слова правильно разделятся на слоги в виде «*мы-йы*», «*бі-йі*», «*тұ-уы*», «*гү-уі*» и не нарушат слоговый состав слов.

2. Русские гласные звуки (***ю***) и (***я***) согласно действующей казахской орфографии во многих словах встречаются в сочетании с другими гласными. Например, *аю*, *ою*, *үю*, *ая*, *оя*, *ұя*, что противоречит правилу «*В корневых словах казахского языка не должны встречаться подряд два гласных звуков*». Поэтому вместо этих русских звуков надо использовать сочетания исконно казахских звуков (йұу, йүу) и (йа, йә) соответственно. Тогда мы писали бы: *айұу*, *ойұу*, *үйүу*, *айа*, *ойа*, *ұйа*, что не нарушает ни одного правила казахского языка, в том числе и правила образования слогов и правила добавления морфем. Подобные противоречия отрицательно отражаются на культуре письменной

НИИ искусственный интеллект

речи. В целом, современное состояние казахского языка не позволяет развивать и использовать его как государственный язык. Поэтому необходимо исправить сложившееся положение. Для этого необходимо провести **комплексную реформу казахского языка**. Первоочередными задачами такой реформы являются:

1. Уточнение звуковой системы казахского языка с учетом результатов предыдущих реформ. При этом из заимствованных звуков оставить в системе только те, которые используются в международных непереводимых на казахский язык терминах и не нарушают внутренние закономерности казахского языка.
2. Определить состав нового алфавита на основе латиницы. При этом новый алфавит должен позволить эффективно использовать информационно-коммуникационные технологии.
3. Разработать новые орфоэпические и орфографические нормы казахского языка. При этом заимствованные непереводимые на казахский язык термины должны произноситься и писаться в соответствии с нормами казахского языка.

При уточнении звуковой системы казахского языка следует учесть, что заимстованные согласные звуки (***вы***), (***фы***), (***хы***) и гласные звуки (***и***), (***у***) часто используются для написания фамилий и международных непереводимых на казахский язык терминах. Например: *Вагнер, Хонекер, Форд, Иванов, Хомский, Хинчин, вагон, вектор*, *вето, викторина, вирус, интернет, инвестор, институт, университет, универсам, хлор, хром, хроника, факт, файл, физика, факультет, федерация, филармония, фонетика, формула, функция.* Все названные звуки внедрились в казахский язык так глубоко, что даже у современных носителей казахского языка адаптировались органы речи для их артикуляции (произношения).

Все указанные звуки не входят в состав суффиксов и окончаний, не влияют на образование новых слов и словоформ. Поэтому их надо использовать при составлении нового казахского алфавита. От этого казахский язык не пострадает, наоборот, они помогут сблизить произношение и написание международных терминов к их произношению и написанию в оригинале. Однако, сначала надо применить *лингвистический метод* по разделению

НИИ Искусственный интеллект

звуков языка на *фонемы* и *аллофоны* (см. таблицу 2.2.3.4 в параграфе 2.2.3).

В алфавите любого языка фонемы имеют отдельные графемы, а аллофоны их не имеют, но они обозначаются отдельными символами в МФА и будут присутствовать в транскрипциях других фонетических единиц (морфем, слов и др.) [31, 32].

Для определения нового алфавита казахского языка на основе латинской графики согласные звуки (***вы***), (***фы***), (***хы***) будут считаться фонемами, а гласные звуки (***и***) и (***у***) **–** аллофонами исконно казахских гласных фонем (***і***) и (***ұ***), соответственно.

Переход гласных звуков (***и***) и (***у***) из статуса фонемы в действующем алфавите на основе кириллицы в статус аллофона в новом алфавите на основе латиницы обьясняется тем, что согласно МФА произношение гласного звук (***и***) – [i] может варьироваться от русского гласного звука (***и***) до казахского гласного звука (***і***) -[ɪ] и, аналогично, произношение звука (***у***) - [u] может варьироваться от русского гласного звука (***у***) до казахского гласного звука (ұ) - [ʊ] (здесь в круглой скобке – наименование звука, а в квадратной скобке – его обозначение в МФА) [31, 32]. Этот ***феномен,*** когда два звука могут одинаково встречаться в определенном контексте, при этом, замена одного из них на другой не порождает новое слово или предложение, называется ***Свободное варьирование*** [33]. Поэтому звуки (***и***) и (***у***) не будут рассматриваться при анализе звуковой системы казахского языка (проводимого в главах 2.1 и 2.2) и как аллофоны не будут обозначаться отдельными буквами в новом алфавите. Такая практика имеется во многих (английский, французский, турецкий, узбекский и других) языках.

Таким образом, можно считать, что в реформе казахского языка будут участвовать **31** фонема, из них **9** гласных фонем и **2** аллофоны: (а) - [ɑ], (ә)-[æ], (е) - [e], (о) - [ɔ], (ө) - [ɵ], (ұ) - [ʊ, u], (ү) - [ʏ], (ы) - [ɤ], (і) - [ɪ, i]; **22** согласные фонемы: (бы) - [b], (вы) - [v], (ғы) - [ɣ], (гі) - [g], (ды) - (кі) - [k], (қы) - [q], (жы) - [ʒ], (зы) - [z], (й) - [j], (ыл) - [l], (мы) - [m], (ны) - [n], (ыңы) - [ŋ], (пы) - [p], (ыр) - [r], (сы) - [s], (ты) - [t], (ў) - [w], (фы) - [f], (хы) - [h], (шы) - [ʃ],

НИУ Искусственный интеллект

которые должны обозначаться отдельными графемами (буквами) в алфавите.

НИИ Искусственный интеллект

1. ЗВУКИ ЕСТЕСТВЕННОГО ЯЗЫКА

1.1. Органы речи и органы слуха

Любой звук естественного языка является волной, которая распространяется в упругой среде, обладающей определённым спектром, интенсивностью и диапазоном.

Звуки естественного языка произносятся органами речи человека (артикуляция), распространяются через воздух (акустика) и воспринимаются (распознаются) органами слуха (перцепция) [20].

Органы речи человека показаны на рисунке 1.1.

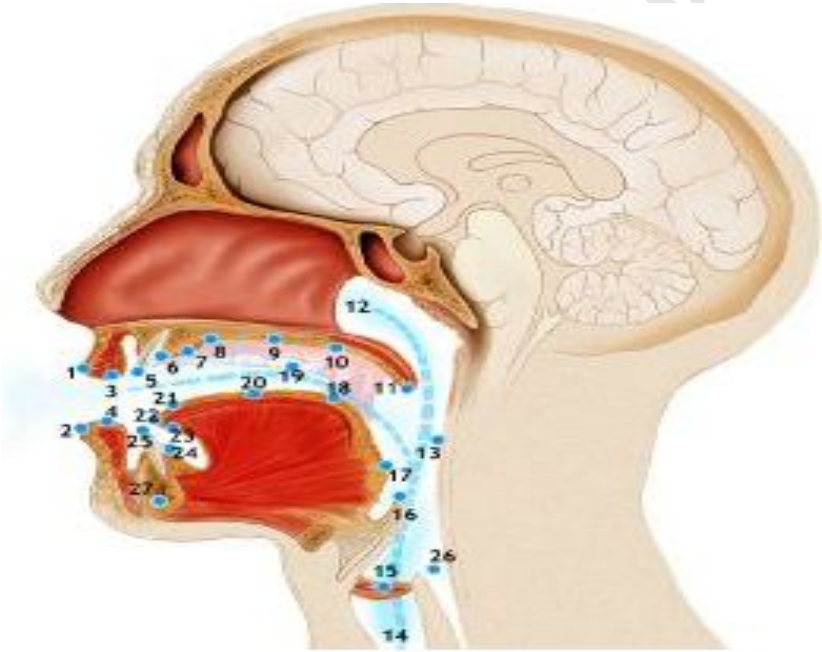


Рисунок 1.1. Органы речи человека.

1. Верхняя внешняя губа; 2. Нижняя внешняя губа; 3. Верхняя внутренняя губа; 4. Нижняя внутренняя губа; 5. Передние верхние зубы; 6. Передняя часть альвеолы; 7. Задняя часть альвеолы; 8. Передняя часть твёрдого нёба; 9. Основная поверхность твёрдого нёба; 10. Мягкое нёбо; 11. Язычок гортани; 12. Носовая полость; 13. Задняя стенка глотки; 14. Гортань, переходящая в трахею, бронхи и лёгкие; 15. Голосовые связки; 16. Надгортанная часть или область глотки; 17. Корень языка; 18. Задняя поверхность языка; 19. Коренные зубы; 20. Основная поверхность языка; 21. Верхняя часть кончика языка; 22. Кончик языка; 23. Нижняя часть кончика языка; 24. Соединение дёсен с передними зубами; 25. Передние

НИИ ИСТС Минтеллект

зубы нижней челюсти; 26. Гортань глотки, переходящая в пищевод; 27. Нижняя челюсть.

Органы речи делятся на *активные* (*подвижные*) и *пассивные* (*неподвижные*). Активные органы речи способны менять объём и форму речевого тракта и создавать в нём препятствия для выдыхаемого воздуха. К ним относятся нижняя челюсть (*maxilla*), язык (*lingua*), губы (*labiis*), язычок (*uvula*), голосовые связки (*vocalis funibus*), нёбо (*palatum molle*). Пассивные (*неподвижные*) органы речи лишены способности изменения формы речевого тракта. К пассивным органам речи относятся верхняя челюсть (*maxillae*), зубы (*dentium*), альвеолы (*alveoli*) и твёрдое нёбо (*palatum durum*), мягкое нёбо (*palatum molle*), носовая полость (*cavitas nasi*).

Работа органов речи в получении звука называется артикуляцией. Артикуляция состоит их трех моментов: *экскурсии*, *выдержки*, *рекурсии*.

Экскурсия является начальным моментом, когда органы речи готовятся к произношению. Выдержка является центральным моментом, когда осуществляется произношение звука и его звучание. Рекурсия является заключительным моментом, когда органы речи завершают произношение.

Все звуки делятся на гласные и согласные в зависимости от того, какие органы речи и как участвуют в его образовании. При образовании гласных звуков участвуют только язык и губы, которые не создают преграды, и воздух свободно проходит в полость рта. При образовании согласных звуков участвуют не только язык и губы, но и нёбо, которые создают преграды. Поэтому гласные звуки образуются только голосом, а согласные звуки - сочетанием голоса и шума; либо только шумом.

Язык очень гибок и подвижен, может занимать различные положения в ротовой полости для образования различных звуков. Язык условно подразделяется на 3 части: *передняя часть* (*front part*), *средняя часть* (*middle part*) и *задняя часть* (*back part*). В состоянии покоя передняя часть спинки языка располагается

НИИ Могучий интеллект

против альвеолярной дуги (*alveolar ridge*) на твёрдом нёбе, средняя часть – против твёрдого нёба (*hard palate*), задняя часть – против мягкого нёба (*soft palate*). Поверхность всего языка называется дорсальной поверхностью (*dorsal surface*), поверхность его передней части – предорсальной поверхностью (*predorsal*).

Голосовые связки (*vocal cords*) образуются из двух мускульных эластичных складок, которые могут сближаться и одновременно натягиваться или раздвигаться и расслабляться. Пространство между ними называется голосовой щелью (*glottis*). Над гортанью располагаются полость зева, или фарингальная полость (*pharyngealcavity*), ротовая полость (*mouth cavity*) и носовая полость (*nasal cavity*). Их объём влияет на качество звуков. Из фарингальной полости выдыхаемый воздух может выходить через ротовую полость, если мягкое нёбо (*soft palate*) поднято, или через носовую полость, если мягкое нёбо опущено. В первом случае образуются ротовые (*oral*) звуки, во втором – носовые (*nasal*) звуки различной высоты и тональности.

Когда человек дышит, его легкие разжимаются (вдох) и сжимаются (выдох). При сжатии легких воздушная струя проходит через гортань и попадает в ротовую полость, если увула не закрывает проход, или в носовую полость, если увула закрывает проход. При этом возможны случаи:

1) образуются гласные звуки, если голосовые связки сдвинуты, и воздушная струя проходит, трогая голосовые связки, вызывая их гармонические, периодические, равномерные звуковые колебания самой разной частоты, представляющие музыкальный звук – тон;

2) образуются согласные звуки, если голосовые связки раздвинуты, и воздушная струя проходит, не трогая голосовые связки, вызывая их негармонические, непериодические, неравномерные звуковые колебания, представляющие немузыкальный звук – шум (глухие) или сочетание шума и тона (звонкие).

НИИ искусственный интеллект

Тон и шум являются сложным звуком и состоят из множества составляющих, а ротовая и носовая полость являются резонатором и формируют качество тона и шума, усиливая его составляющие.

Сила звуковых колебаний измеряется относительно порога слышимости в *децибелах* (дБ). Частота звука, которую воспринимает орган слуха человка как высоту звука, измеряется в герцах. Органы слуха человека могут воспринимать звуковые волны длиной от 1,6 см до 20 м, что соответствует диапазону 16 – 20 000 Гц (20 кГц).

Органы слуха человека преобразуют звуковое колебание воздуха в электрические импульсы и состоят из трех основных частей: *наружного уха*, *среднего уха* и *внутреннего уха*.

Органы слуха человека показаны на рисунке 1.2.

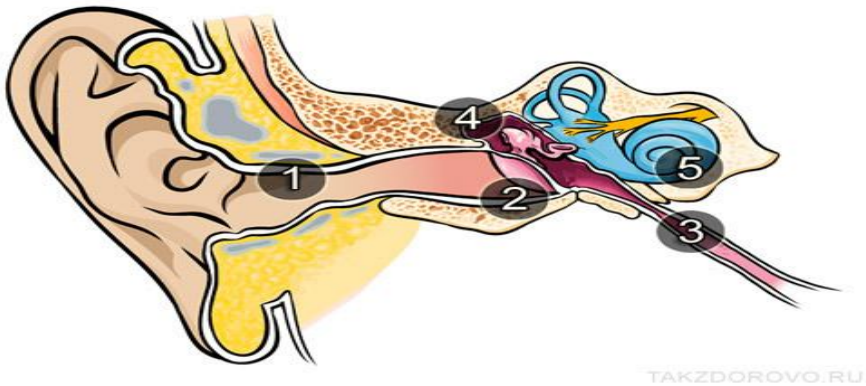


Рисунок 1.2. Органы слуха человека.
1 - ушная раковина и наружный слуховой проход; 2 - барабанная перепонка; 3 - евстахиева труба; 4 – слуховые косточки (молоточек, наковальня и стремечко); 5 - внутреннее ухо.

Наружное ухо предназначено для улавливания и фокусировки звуковых волн, состоит из ушной раковины и наружного слухового прохода. Наружное ухо не имеет никакого отношения к преобразованию звуковых колебаний воздуха в электрические импульсы.

Среднее ухо состоит из барабанной перепонки, слуховых косточек (молоточек, наковальня, стремечко) и евстахиевой трубы. Барабанная перепонка вибрирует от звуковой волны, а слуховые косточки передают звуковых колебаний от барабанной перепонки в

область внутреннего уха. Евстахиева труба соединена с носоглоткой и выравнивает давление в среднем ухе.

Внутреннее ухо состоит из двух частей: вестибулярного аппарата и улитки, которая превращает вибрацию в нервные импульсы, передавемые к слуховым центрам мозга, т.е. преобразует звуковых колебаний в электрические импульсы и передает их на слуховые нервы.

НИИ Искусственный интеллект

1.2. Методы анализа звуков естественных языков

В этом параграфе рассматриваются методы анализа звуков естественных языков, которые описаны в источниках [12-33]. Эти методы будут применены в главах 2.1 и 2.2 для проведения артикуляционного анализа и акустического анализа звуков казахского языка.

В естественных языках наименьшими единицами речи являются звуки, благодаря которым мы произносим, слышим и различаем слова, т.е. они выполняют смыслоразличительную функцию. С помощью звуков естественных языков можно превратить одно слово в другое, при изменении одного звука меняется смысл слова.

В зависимости от точки зрения на изучаемый объект наука о звуках речи представлена соответствующими разделами:

1. *Артикуляционная фонетика* посвящена описанию и классификации звуков речи с точки зрения работы органов артикуляции.

2. *Акустическая фонетика* исследует физические особенности звуков речи, доступных для восприятия органами слуха.

3. *Перцептивная фонетика* выявляет особенности восприятия звуков слушающим и опознавание звуковых единиц.

Анализ звуков любого естественного языка может быть артикуляционным (по функции органов речи), акустическим (по произношению звуков) или перцептивным (по восприятию звуков).

Артикуляционный анализ языка определяет органы речи и их движение, участвующие в образовании его звуков и устанавливает артикуляционные признаки:

– *место образования звуков;*

– *способ образования звуков;*

– *участие голоса и шума в образовании звуков*.

Как сказано выше в параграфе 1.1, звуки естественного языка в зависимости от того, какие органы речи и как участвуют в их образовании, делятся на *гласные звуки* и *согласные*, имеющие разные артикуляционные признаки.

НИИ искусственный интеллект

При образовании гласных звуков выходящая из легких струя воздуха должна пройти через гортань и вызвать колебание сближенных голосовых связок, находящихся в нем.

По месту образования гласные звуки могут подразделяться на:

– альвеолярные (переднеязычные);
– среднеязычные;
– увулярные (заднеязычные);
– губные (округленные);
– негубные (неокругленные).

По способу образования гласные звуки могут в зависимости от вертикального движения языка подразделяться на:

– верхнеязычного подъема;
– среднеязычного подъема;
– нижнеязычного подъема.

При образовании согласных звуков на пути воздушного потока создается преграда органами речи и могут быть следующие события: блокирование речевых органов, сильное или слабое ограничение речевых органов, перенаправление воздушного потока через носовые органы, происходит неполное колебание голосовых связок (*сонорные согласные* и *шумные согласные*) или колебание отсутствует (*глухие согласные*).

По месту образования согласные звуки могут подразделяться на:

– *переднеязычно-зубные (предлингво-дентальные);*
– *переднеязычно-переднёбные (предлингво-предпалатальные);*
– *среднеязычно-твердонёбные (среднелингво-палатальные);*
– *заднеязычно-твердонёбные (постлингво-палатальные);*
– *заднеязычно-мягконёбные (постлингво-велярныые);*
– *губно-губной (билабиальные);*
– *губно-зубные(лаби-дентальные).*

Здесь на первом месте указан активный орган речи, на втором месте – пассивный орган речи.

По способу образования согласные звуки могут подразделяться на *смычные* и *щелевые*. При произношении смычных согласных в ротовой полости происходит полный смык

НИИ искусственный интеллект

органов речи, и воздушная струя, шумно преоделевая смык, находит выход наружу. Для смычных согласных звуков выдержка вкючает момент полного прекращения протекания воздушной струи через ротовую полость.

Смычные согласные могут подразделяться на: *взрывные, аффрикаты, носовые.*

При артикуляции взрывных нёбная занавеска поднята, и воздух проходит в ротовую полость, а размыкание смыков происходит резко и напоминает взрыв. При артикуляции аффрикат размыкание смыков происходит не резко. При артикуляции носовых небная занавеска опущена и воздух выходит через нос.

При произношении щелевых согласных в ротовой полости происходит сближение активного и пассивного органов речи, в результате чего воздушная струя, проходя через образовавшуюся щель, производит шум (трения - фрикацию).

Щели по форме различаются *круглощелевые* и *плоскощелевые*, а по локализации подразделяюся на *срединные* и *боковые*.

Акустический анализ языка различает в звуке такие акустические характеристики, как:

– *звучность звука;*

– *высота звука;*

– *сила звука;*

– *длительность звука;*

– *тембр звука*.

Звучность звука зависит от соотношения тона (периодического колебания) и шума (непериодического колебания). По звучности звуки делятся на вокальные и невокальные. Вокальными звуками являются гласные, состоящие только из тона, и сонорные согласные, которые состоят из тона и шума, но при этом тон преобладает над шумом. Невокальными звуками являются шумные согласные звуки, в которых шум преобладает над тоном (звонкие согласные) или тон вообще отсутствует (глухие согласные).

Высота звука зависит от частоты колебаний (числа колебаний в единицу времени). По высоте звуки делятся на высокие и низкие.

НИИ искусственный интеллект

Чем больше частота колебаний источника звука, тем выше издаваемый им звук. К высоким относятся гласные звуки переднего ряда, а также переднеязычные и среднеязычные согласные звуки. К низким относятся все остальные гласные и согласные звуки.

Сила звука зависит от амплитуды звуковых колебаний: чем больше амплитуда колебаний, тем сильнее (громче) звук. По силе звуки делятся на консонантные и неконсонантные звуки. Консонантные звуки – это слабые (тихие) звуки, к ним относятся все согласные, а неконсонантные звуки – это сильные (громкие) звуки, к ним относятся все гласные звуки.

Длительность (долгота) звука зависит от его продолжительности во времени. По длительности выделяются долгие (длительные) и краткие звуки.

Тембр звука зависит от совокупности колебаний голосовых складок и результата работы резонаторов, является индивидуальным качеством его акустических признаков.

Акустическая характеристика звуков естественного языка базируется на экспериментальных данных электроакустики. Она основана как на собственно частотных характеристиках формант, так и на расположении формант в спектре звука.

Для проведения акустического анализа звуков естественного языка можно использовать следующие акустические признаки:

1. Признак *звонкость-глухость* основан на оценке наличия или отсутствия гармонических колебаний в низких частотах, позволяет разделять все звуки на звонкие и глухие. К звонким звукам относят все вокальные - гласные, сонорные - согласные и звонкие - шумные. К глухим относят невокальные - глухие шумные согласные.

2. Признак *прерывность-непрерывность* основан на характеристике звуковой волны в спектрограмме. Прерывные звуки на спектрограммах характеризуются резко оборванным краем волны, а у непрерывных звуков существуют переходные тоны.

НИ ОТБЕННЫЙ ИНТЕЛЛЕКТ

3. Признак *назальность-неназальность* основан на оценке участия в образовании звука носового резонатора.

4. Признак *резкость-нерезкость* основан на оценке интенсивности шума. При образовании резких звуков в ротовой полости возникают турбулентные воздушные потоки, когда часть воздуха идет поперек общего потока, возмущая его.

Перцептивный анализ языка осуществляется на основе экспериментов, проводимых в специальных звукоизолированных помещениях с целью слуховой идентификации его звуковых единиц. При этом со смыслоразличительной стороны согласные обладают большими различительными особенностями, нежели гласные: по согласным можно угадать слово, а по гласным – нельзя.

В качестве анализируемых звуковых сегментов используются начитанные носителем естественного языка на диктофон или искусственно синтезированные звуки, слоги, слова или фразы анализируемого языка или искусственные слова (логотомы), не имеющие значения в данном языке.

К методам перцептивных исследований относятся *восприятие*, *различение* и *сравнение* предъявляемых звуковых сегментов:

1) восприятие звука, слога, слова или фразы анализируемого языка;

2) восприятие энергетической реализации слога (восприятие ударного и заударного слогов внутри каждого отдельно взятого слова по громкости);

3) восприятие временной реализации слога (восприятие ударного и заударного гласных по долготе внутри каждого отдельно взятого слова и восприятие ударных гласных по долготе), а также выявление характера примыкания согласного к предшествующему ударному гласному.

Таким образом, звуков естественного языка можно отличать друг от друга артикуляционной, акустической и перцептивной классификацией. Кроме указанных классификаций, есть и функциональная (смыслоразличительная) классификация

смысловых звуков - фонем, которая заключается в том, что звук может различать звуковые оболочки слов и их форм.

НИИ Искусственный интеллект

2. ЗВУКОВАЯ СИСТЕМА КАЗАХСКОГО ЯЗЫКА

2.1. Артикуляционный анализ фонем казахского языка

2.1.1. Артикуляция гласных фонем казахского языка

Во введении сказано, что в казахском языке имеются **9** гласных фонем: (а) - [ɑ], (ә) - [æ], (е) - [е], (о) - [ɔ], (ө) - [ɵ], (ұ) - [ʊ], (ү) - [ʏ], (ы) - [ɤ], (і) - [ɪ].

Артикуляция этих гласных будет описан, исходя из их классификации, рассмотренной в параграфе 2.2.1.

Гласный звук (а) - [ɑ] образуется при открытой челюсти, нейтральном положении губ, низком положении языка и отодвинутом назад языке. Следовательно, звук (а) - [ɑ] по положению челюсти является открытым, по положению губ является неокругленным, по вертикальному положению языка является нижним, а по горизонтальному положению - заднеязычным гласным звуком. Артикуляционные признаки звука [ɑ]: *открытый*, *нижний*, *заднеязычный*, *неокругленный*.

Гласный звук (ә) - [æ] образуется при открытой челюсти, нейтральном положении губ, ненапряженно-нижнем положении языка и отодвинутом вперед языке. Следовательно, звук (ә) - [æ] по положению челюсти является открытым, по положению губ является неокругленным, по вертикальному положению языка является ненапряженным нижним, а по горизонтальному положению - переднеязычным гласным. Артикуляционные признаки звука [æ]: *открытый*, *ненапряженно-нижний*, *переднеязычный*, *неокругленный.*

Гласный звук (е) - [е] образуется при полуоткрытой челюсти, нейтральном положении губ, среднем положении языка и отодвинутом вперед языке. Следовательно, звук (е) - [е] по положению челюсти является полуоткрытым, по положению губ является неокругленным, по вертикальному положению языка является средненижним, а по горизонтальному положению - переднеязычным. Артикуляционные признаки звука [е]: *полуоткрытый*, *средненижний*, *переднеязычный*, *неокругленный*.

НИИ Интеллектуальный интеллект

Гласный звук (о) - [ɔ] образуется при полуоткрытой челюсти, округленных губах, средненижнем положении языка и отодвинутом назад языке. Следовательно, звук (о) - [ɔ] по положению челюсти является полуоткрытым, по положению губ является округленным, по вертикальному положению языка является средненижним, а по горизонтальному положению - заднеязычным гласным. Артикуляционные признаки звука [ɔ]: *полуоткрытый*, *средненижний*, *заднеязычный*, *округленный*.

Гласный звук (ө) - [ө] образуется при полуоткрытой челюсти, округленных губах, средневерхнем положении языка и отодвинутом вперед языке. Следовательно, звук (ө) - [ө] по положению челюсти является полуоткрытым, по положению губ является округленным, по вертикальному положению языка является средневерхним, а по горизонтальному положению - переднеязычным гласным. Артикуляционные признаки звука [ө]: *полуоткрытый*, *средневерхний*, *переднеязычный*, *округленный*.

Гласный звук (ұ) - [ʊ] образуется при закрытой челюсти, округленных губах, ненапряженно-верхнем положении языка и отодвинутом назад языке. Следовательно, звук (ұ) - [ʊ] по положению челюсти является закрытым, по положению губ является округленным, по вертикальному положению языка является ненапряженноверхним, а по горизонтальному положению - заднеязычным гласным звуком. Артикуляционные признаки звука [ʊ]: *закрытый*, *ненапряженноверхний*. *заднеязычный*, *округленный*.

Гласный звук (ү) - [ʏ] образуется при закрытой челюсти, округленных губах, ненапряженно-верхнем положении языка и отодвинутом вперед языке. Следовательно, звук (ү) - [ʏ] по положению челюсти является закрытым, по положению губ является округленным, по вертикальному положению языка является ненапряженно-верхнем, а по горизонтальному положению - переднеязычным гласным звуком. Артикуляционные признаки звука [ʏ]: *закрытый*, *ненапряженно-верхний*, *переднеязычный*, *округленный*.

НИИ Искусственный интеллект

Гласный звук (ы) - [ɤ] образуется при закрытой челюсти, нейтральном положении губ, верхнем положении языка и отодвинутом назад языке. Следовательно, звук (ы) - [ɤ] по положению челюсти является закрытым, по положению губ является неокругленным, по вертикальному положению языка является верхним, а по горизонтальному положению - заднеязычным гласным. Артикуляционные признаки звука [ɤ]: *закрытый*, *верхний заднеязычный*, *неокругленный*.

Гласный звук (і) - [ɪ] образуется при закрытой челюсти, нейтральном положении губ, ненапряженно-верхнем положении языка и отодвинутом вперед языке. Следовательно, звук (і) - [ɪ] по положению челюсти является закрытым, по положению губ является неокругленным, по вертикальному положению языка является ненапряженно-верхнем, по горизонтальному положению является переднеязычным. Артикуляционные признаки звука [ɪ]: *закрытый*, *ненапряженно-верхний*, *переднеязычный*, *неокругленный*.

2.1.2. Артикуляционная классификация гласных звуков казахского языка

Во многих источниках по фонетике казахского языка, например, в [15-18] для артикуляционной классификации казахских гласных звуков выделяют 3 бинарных (имеющих по 2) признака:

Таблица 2.1.2.1. Три бинарные артикуляционных признака гласных звуков казахского языка.

№	Названия признаков	Значения признаков	
1	Горизонтальное положение языка	Заднеязычное	Переднеязычное
2	Положение челюсти	Открытое	Закрытое
3	Положение губ	Округленное	Неокругленное

Признак «Положение языка» можно более точнее назвать как «Горизонтальное положение языка» со значениями «Заднеязычное»

НИИ Искусственный интеллект

и «Переднеязычное», поскольку имеется признак «Вертикальное положение языка», значения «Твердое» и «Мягкое» относятся к акустическому признаку, а не к артикуляционному.

Кроме того, по принципу Дирихле [21] для корректного и полного описания характеристик 9-ти объектов трех бинарных (двузначных) признаков недостаточно, поскольку каждый признак имеет по 2 значения и максимальное количество различимых между собой признаков равно $2^3 = 8$. Однако традиционная классификкация 9-ти казахских гласных звуков, показанная в таблице 2.1.2.2, основана на 3-х бинарных признаках.

Таблица 2.1.2.2. Традиционная классификация казахских гласных звуков на основе трех бинарных признаков.

Гласные звуки		**Артикуляционные признаки**					
		Положение челюсти		Положение губ		Положение языка	
Наименование	Обозначение в международном фонетическом алфавите (МФА - IPA)	Открытый	Закрытый	Неокругленный	Округленный	Твердый	Мягкий
(а)	[ɑ]	+	–	+	–	+	–
(ә)	**[æ]**	**+**	**–**	**+**	**–**	**–**	**+**
(о)	[ɔ]	+	–	–	+	+	–
(ө)	[ɵ]	+	–	+	–	–	+
(ұ)	[ʊ]	–	+	–	+	+	–
(ү)	[ʏ]	–	+	–	+	–	+
(ы)	[ɤ]	–	+	+	–	+	–
(і)	[ɪ]	–	+	+	–	–	+
(е)	**[е]**	**+**	**–**	**+**	**–**	**–**	**+**

Из таблицы 2.1.2.2 видно, что в традиционной классификации артикуляционные признаки гласных звуков (ә) и (е) полностью совпадают. Это означает, что традиционная классификация не

НИИ Искусственный интеллект

способна различить эти два гласных звука. Следовательно, должен быть проведен более тщательный анализ артикуляционных и акустических характеристик системы казахских гласных звуков. Для этого можно предложить два способа в порядке их появления:

1. Оставить количество признаков без изменения, но каждому из этих признаков добавляется еще по одному значению. Например, в книге [15] на стр. 38 добавлено одно значение «Полуоткрытое» только для признака «Положение челюсти».

2. Добавить еще один признак относительно вертикального положения языка с двумя значениями.

Способ 1. Классификация казахских гласных звуков на основе трех тернарных (имеющих по 3) признаков с добавлением новых значений - *среднеязычные*, *полуокругленные*, *полуоткрытые* [22], как это показано в таблице 2.1.2.3.

Таблица 2.1.2.3. Три тернарных артикуляционных признака.

Гласные звуки		**Артикуляционные признаки**								
		Положение челюсти			Положение губ			Положение языка		
Наименование	Обозначение в МФА(IPA)	Открытые	**Полуоткрытые**	Закрытые	Неокругленные	**Полуокругленные**	Округленные	Заднеязычные	**Среднеязычные**	Переднеязычные
(а)	[ɑ]	+	-	-	+	-	-	+	-	-
(ə)	[æ]	+	-	-	+	-	-	-	-	+
(о)	[ɔ]	+	-	-	-	-	+	+	-	-
(ө)	[ɵ]	+	-	-	-	-	+	-	-	+
(ұ)	[ʊ]	-	-	+	-	-	+	+	-	-
(ү)	[ʏ]	-	-	+	-	-	+	-	-	+
(ы)	[ɤ]	-	-	+	+	-	-	+	-	-
(і)	[ɪ]	-	-	+	+	-	-	-	-	+

НИИ Искусственный интеллект

(е)	[е]	-	+	-	-	+	-	-	+	-

На основе этих трех тернарных признаков (таблица 2.1.2.3) можно построить геометрическую модель системы гласных звуков казахского языка в виде двух пирамид с совмещенными между собой вершинами, которая показана на рисунке 2.1.2.1.

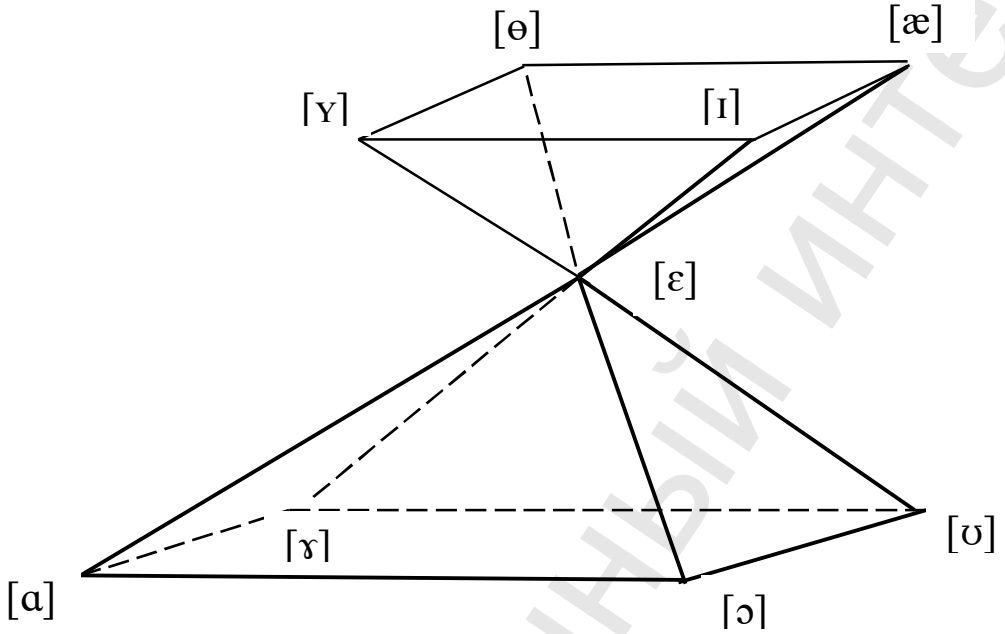


Рисунок 2.1.2.1. Геометрическая модель системы казахских гласных на основе трех тернарных артикуляционных признаков.

Здесь плоскость [ɑ] [ɔ] [ʊ] [ɤ] представляет заднеязычные (твердые) гласные звуки (а) (о) (ұ) (ы), плоскость [æ] [ɵ] [ʏ] [ɪ] – переднеязычные (мягкие) гласные звуки (ə) (ɵ) (ү) (і), плоскость [ɑ] [ɔ] [ɵ] [æ] – открытые гласные звуки (а) (о) (ɵ) (ə), плоскость [ɤ] [ʊ] [ʏ] [ɪ] – закрытые гласные звуки (ы) (ұ) (ү) (і), плоскость [ɔ] [ʊ] [ʏ] [ɵ] – округленные гласные звуки (о) (ұ) (ү) (ɵ), плоскость [ɑ] [ɤ] [ɪ] [æ] – неокругленные гласные (а) (ы) (і) (ə), а вершина [е] представляет особый звук, имеющий признаки: среднеязычный полуоткрытый (полузакрытый) и полуокругленный гласный звук (е).

На рисунке 2.2.1 показано, что каждый казахский гласный звук, кроме звука [е], имеет три признака и стоит на пересечении трех плоскостей в трехмерном пространстве. Каждая плоскоскость представляет определенный класс звуков, к которому может

НИИ Искусственный интеллект

принадлежать тот или иной гласный звук. Поскольку гласный звук [е] стоит на пересечении 4-х плоскостей, то этот звук имеет 4 варианта произношения в зависиости от наличия в контексте произносимого слова звуков: [ɑ], [æ], [ɔ], [ɵ], [ʊ], [ʏ], [ɤ] и [ɪ]. Мы провели эксперимент, микшируя осциллограммы заднеязычных (твердых) звуков [ɑ], [ɔ], [ʊ], [ɤ] с осциллограммой звука [е] и обнаружили такие свойства, представленные на рисунке 2.1.2.2:

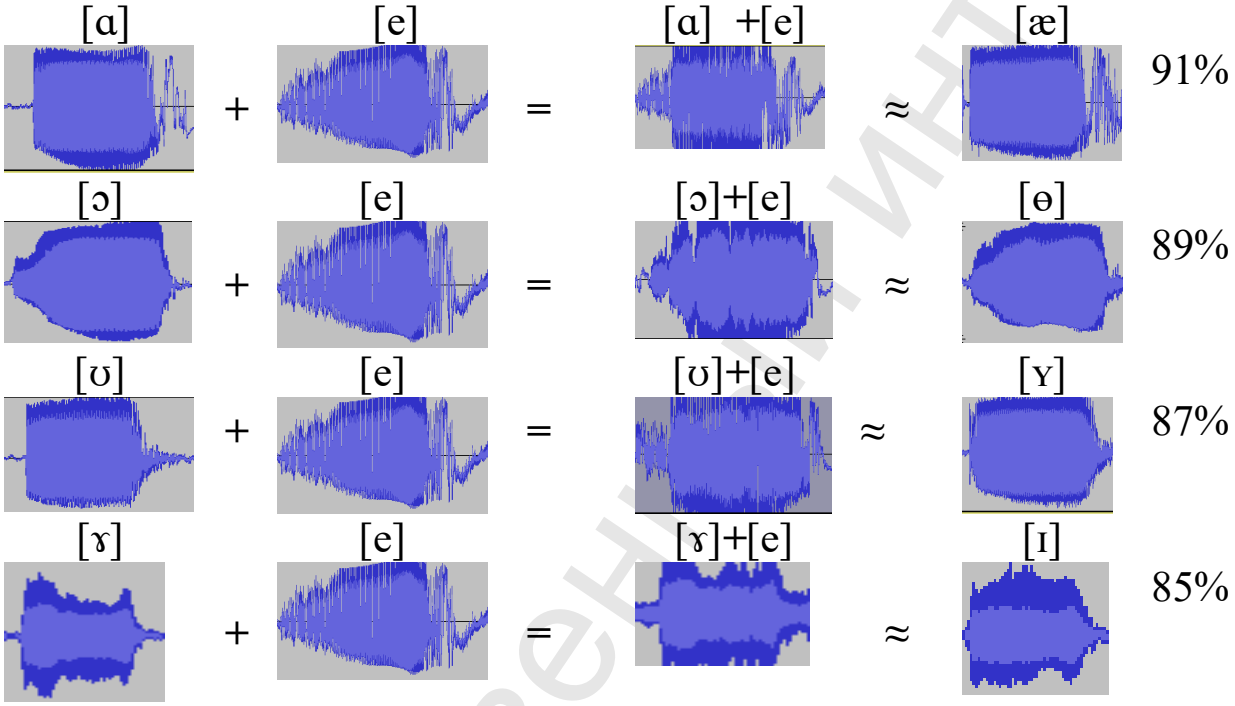


Рисунок 2.1.2.2. Свойства гласных звуков казахского языка.

Классификация казахских гласных на основе трех тернарных артикуляционных признака показана в таблице 2.1.2.4.

Таблица 2.1.2.4 Классификация казахских гласных звуков на основе трех тернарных артикуляционных признаков.

Положение языка	**Положение языка**				
	Заднеязычные		**Среднеязычные**	**Переднеязычные**	
	Неокругленные	Округленные	Полуокругленные	Неокругленные	Округленные
Открытые	[ɑ]	[ɔ]		[æ]	[ɵ]
Полуоткрытые			[е]		
Закрытые	[ɤ]	[ʊ]		[ɪ]	[ʏ]

Следует отметить, что выше указанные квалификации не соответствуют классификации звуков в МФА [31, 32]. Поэтому они не будут использоваться на практике.

НИИ искусственного интеллект

Теперь построим артикуляционную классификацию гласных звуков казахского языка на основе 4 бинарных признаков в соответствии с артикуляционными признаками звуков в МФА.

Способ 2. Классификация на основе добавления нового, четвертого признака относительно *вертикального положения языка* с двумя значениями: *верхние, нижние.*

В МФА термин «верхние» называется «close», а «нижние» – «open», подразумевая близость при подъеме языка к небу [31].

Обоснование введения четвертого признака дается в разделе 2.2.1 при рассмотрении акустического анализа казахских гласных звуков на основе компьютерного анализа, используемого для многих естестественных языков [31, 32].

В таблице 2.1.2.5 показаны четыре бинарных артикуляционных признака казахских гласных звуков.

Таблица 2.1.2.5. Четыре бинарные артикуляционных признака казахских гласных звуков.

Гласные звуки		**Артикуляционные признаки**							
Наименование	Обозначение в МФА	Положение языка				Положение губ		Положение челюсти	
		Вертикальное		Горизонтальное					
		Нижние	Верхние	Заднеязычные	Переднеязычные	Неокругленные	Округленные	Открытые	Закрытые
(а)	[ɑ]	+	–	+	–	+	–	+	–
(ə)	[æ]	+	–	–	+	+	–	+	–
(ы)	[ɤ]	+	–	+	–	+	–	–	+
(і)	[ɪ]	+	–	–	+	+	–	–	+
(ұ)	[ʊ]	–	+	+	–	–	+	–	+
(ү)	[ʏ]	–	+	–	+	–	+	–	+

НИИ искусственный интеллект

(о)	[ɔ]	–	+	+	–	–	+	+	–
(ө)	[ө]	–	+	–	+	–	+	+	–
(е)	[е]	–	+	–	+	+	–	+	–

Теперь, используя Булеву алгебру [24], мы построим математическую модель системы гласных звуков [25] в виде алгебраического выражения на основе использования значений 4-х артикуляционных бинарных признака, показанных в 2.1.2.5.

Обозначим 4 артикуляционных признака *Вертикальное положение языка* и *Горизонтальное положение*, *Положение губ*, *Положение челюсти*, *языка* казахских гласных звуков через логические переменные x_1, x_2, x_3 и x_4, соответственно. Эти переменные принимают только значения 1 для истины и 0 для лжи:

– если язык находятся в верхнем положении, то $x_1 = 1$, в противном случае $x_1 = 0$;

– если язык находится в заднем положении, то $x_2 = 1$, в противном случае $x_2 = 0$;

– если губы находится в округленном положении, то $x_3 = 1$, в противном случае $x_3 = 0$;

– если челюсть находится в открытом положении, то $x_4 = 1$, в противном случае $x_4 = 0$.

На основании установленных значений для переменных можно построить Булеву модель системы казахских гласных звуков, представленной в таблице 2.1.2.6.

Таблица 2.1.2.6. Булева модель системы казахских гласных.

Наименование звука	**Обозначение в МФА**	x_1	x_2	x_3	x_4
(а)	[ɑ]	0	1	0	1
(ә)	[æ]	0	0	0	1
(ы)	[ɤ]	0	1	0	0
(і)	[ɪ]	0	0	0	0
(ұ)	[ʊ]	1	1	1	0
(ү)	[ʏ]	1	0	1	0
(о)	[ɔ]	1	1	1	1

НИИ Искусственный интеллект

(ө)	[ө]	1	0	1	1
(е)	[е]	1	0	0	1

Далее для каждого из 9 гласных звуков, выписывая признаки в виде их конъюнкции и объединяя эти конъюнкции, получим следующую дизъюнктивную нормальную форму из 9 членов:

$$\left\{\begin{matrix}(\overline{x}_1 \wedge x_2 \wedge \overline{x}_3 \wedge x_4) \vee (\overline{x}_1 \wedge \overline{x}_2 \wedge \overline{x}_3 \wedge x_4) \vee (\overline{x}_1 \wedge x_2 \wedge \overline{x}_3 \wedge \overline{x}_4) \vee \\ (\overline{x}_1 \wedge \overline{x}_2 \wedge \overline{x}_3 \wedge \overline{x}_4) \vee (x_1 \wedge x_2 \wedge x_3 \wedge \overline{x}_4) \vee (x_1 \wedge \overline{x}_2 \wedge x_3 \wedge \overline{x}_4) \vee \\ (x_1 \wedge x_2 \wedge x_3 \wedge x_4) \vee (x_1 \wedge \overline{x}_2 \wedge x_3 \wedge x_4) \vee (x_1 \wedge \overline{x}_2 \wedge \overline{x}_3 \wedge x_4)\end{matrix}\right\} \quad (2.1)$$

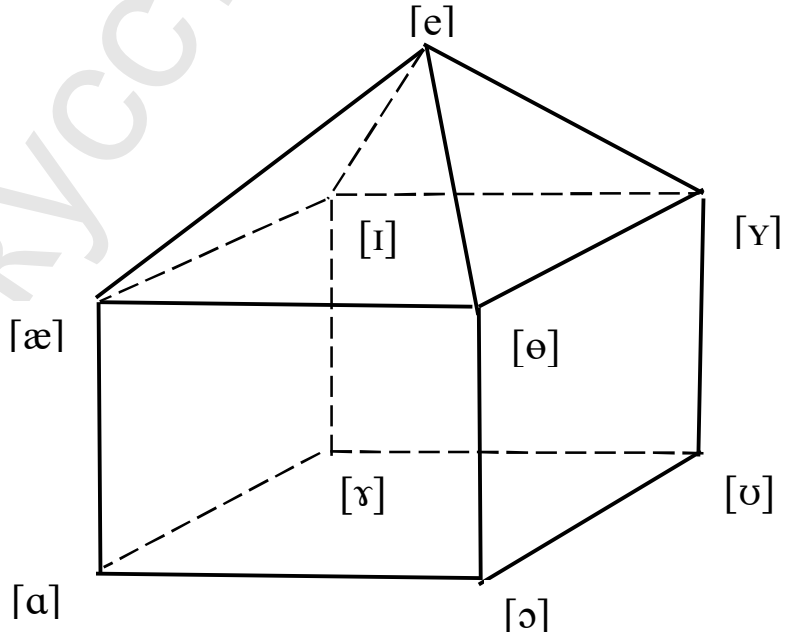
Применяя аксиомы Булевой алгебры, получим упрощенное выражение:

$$(x_1 \wedge x_3) \vee (\overline{x}_1 \wedge \overline{x}_3) \vee (x_1 \wedge \overline{x}_2 \wedge \overline{x}_3 \wedge x_4) \quad (2.2)$$

Выражение (2.2) будем называть *функцией принадлежности*, которая характеризует систему казахских гласных в теореме, приведенной ниже.

Теорема принадлежности. Гласный звук λ принадлежит системе казахских гласных тогда и только тогда, когда его артикуляционные признаки x_1, x_2, x_3 и x_4, определенные выше, удовлетворяют дизъюнктивной нормальной форме $(x_1 \wedge x_3) \vee (\overline{x}_1 \wedge \overline{x}_3) \vee (x_1 \wedge \overline{x}_2 \wedge \overline{x}_3 \wedge x_4)$.

На основе этих четырых артикуляционных признаков можно построить геометрическую модель системы гласных звуков казахского языка, показанной на рисунке 2.1.2.3 [25].



НИИ Искусственный интеллект

Рисунок 2.1.2.3. Геометрическая модель системы казахских гласных на основе 4-х бинарных артикуляционных признаков.

Здесь плоскость [ɑ] [ɔ] [ʊ] [ɤ] представляет заднеязычные (твердые) гласные звуки (а) (о) (ұ) (ы), плоскость [æ] [ɵ] [ʏ] [ɪ] – переднеязычные (мягкие) гласные звуки (ә) (ө) (ү) (і), плоскость [ɑ] [ɔ] [ɵ] [æ] – открытые гласные звуки (а) (о) (ө) (ә), плоскость [ɤ] [ʊ] [ʏ] [ɪ] – закрытые гласные звуки (ы) (ұ) (ү) (і), плоскость [ɔ] [ʊ] [ʏ] [ɵ] – округленные гласные звуки (о) (ұ) (ү) (ө), плоскость [ɑ] [ɤ] [ɪ] [æ] – неокругленные гласные (а) (ы) (і) (ә), а вершина [е] представляет особый звук (е).

Можно возразить, что, предложив четыре артикуляционных признака для классификации гласных звуков, геометрическая модель также должна быть описана в четырехмерном пространстве. Однако, как видно из таблицы 2.1.2.5, все гласные, кроме [е], описываются тремя признаками и только для отличия [е] от остальных был привлечен четвертый признак. Поэтому, концептуально гласная [е] была расположена вверху, чтобы выделить ее особенный статус. Другой причиной того, почему [е] расположена вверху (а не в середине, как это было показано на рисунке 2.1.2.1), является то, что [е] – наиболее периферийная гласная в смысле первой и второй форманты (см. рисунок 2.2.1.4). Этот факт также наблюдается в возрастании второй форманты для гласных [а], [ә], [е] (таблица 2.2.1.2), соответствуя изменению положения языка от более задней гласной [а] к более передней [е] через среднюю гласную [ә]. Аналогичное поведение второй форманты наблюдается и для троек {[о], [ө], [е]}, {[ы], [і], [е]} и {[ұ], [ү], [е]}. Наоборот, для тех же троек первая форманта уменьшается, отражая увеличение положения языка при переходе от более низкой гласной к [е] через среднюю гласную.

Из данной геометрической модели можно установить следующие лингвистические факты (аксиомы):

1. Никакие вершины (гласные звуки) из нижней и верхней плоскостей, включая [е], не могут оказаться в исконно казахском слове. Этот факт называется небной (палатальной) гармонией гласных или сингармонизмом.

2. В казахском языке нет исконно казахского слова, содержащего более трех различных гласных звуков, и если оно

НИИ искусственный интеллект

содержит три гласных звука, то это мягкие гласные, один из которых обязательно [е].

В геометрической модели на рисунке на рисунке 2.1.2.2 это можно интерпретировать так: никакие 3 гласные нижней или верхней плоскостей не встречаются в казахском слове. Слово, содержащее 3 различные гласные, включает вершины плоскости, проходящей через [е] и любые 2 вершины верхней плоскости.

Данный факт был проверен на небольшом словаре и на данный момент является лишь гипотезой.

Теперь можно привести классификацию казахских гласных звуков на основе четырех бинарных артикуляционных признака (по горизонтальному положению языка, по вертикальному положению языка, по положению челюсти, по положению губ), показанную в таблице 2.1.2.4.

Таблица 2.1.2.4. Классификация казахских гласных звуков на основе четырех бинарных артикуляционных признаков.

На этой таблице сверху показаны признаки по горизонтальному положению языка, слева - признаки по положению челюсти, справа - признаки по вертикальному положению языка, а признаки по положению губ учтены при

расположении двух звуков в одной строчке – слева неокругленные, справа – округленные.

В следующем параграфе будет детально описываться артикуляция казахских согласных звуков с указанием артикуляционных признаков.

НИИ Искусственный интеллект

2.1.3. Артикуляция согласных звуков казахского языка

Во введении уточнено, что в казахском языке имеются **22** согласные фонемы: (бы) - [b], (вы) - [v], (ғы) - [ɣ], (гі) - [g], (ды) - [d], (кі) - [k], (қы) - [q], (жы) - [ʒ], (зы) - [z], (й) - [j], (ыл) - [l], (мы) - [m], (ны) - [n], (ыңы) - [ŋ], (пы) - [p], (ыр) - [r], (сы) - [s], (ты) - [t], (ұ) - [w], (фы) - [f], (хы) - [h], (шы) - [ʃ].

Ниже будут даны уточненные описания артикуляции этих звуков с указанием их места и способа образования, а также участия голоса.

Звук (бы) - [b] образуется обеими губами, следовательно, по месту образования является губно-губным (билабиальным) согласным. При образовании звука [b] нижние и верхние губы смыкаются и резко размыкаются с сильным толчком воздуха, образуя взрыв, а голосовые связки производят частичное колебание, следовательно, по способу образования он является смычно-взрывным (эксплозивным) согласным, а по участию голосовых связок – звонким согласным. Артикуляционные признаки звука [b]: *губно-губной*, *смычно-взрывной*, *звонкий.*

Звук (вы) - [v] образуется нижней губой и верхними зубами, следовательно, по месту образования является губно-зубным (лабиодентальным) согласным. При образовании звука [v] верхняя губа приподнята, верхние зубы обнажены, нижняя губа соприкасается с верхними зубами, кончик языка и верхние альвеолы сближаются, образуя плоскую щель, а голосовые связки производят частичное колебание, следовательно, по способу образования он является плоскощелевым согласным (спирантом), а по участию голосовых связок – звонким согласным. Артикуляционные признаки звука [v]: *губно-зубной*, *плоскощелевой*, *звонкий*.

Звук (ғы) - [ɣ] образуется задней частью спинки языка и веляром (мягким нёбом), следовательно, по месту образования является заднеязычно-велярным согласным. При образовании звука [ɣ] задняя часть спинки языка и язычок смыкаются и размыкается,

НИИ искусственного интеллект

образуя круглую щель, а голосовые связки производят частичное колебание, следовательно, по способу образования он является круглощелевым (фрикативным) согласным (спирантом), а по участию голосовых связок – звонким согласным. Артикуляционные признаки звука [ɣ]: *заднеязычно-велярный*, *щелевой*, *звонкий*.

Звук (гі) - [g] образуется задней частью спинки языка и веляром (мягким нёбом), следовательно, по месту образования является заднеязычно-велярным согласным. При образовании звука [g] задняя часть спинки языка и мягкое нёбо смыкаются и струя воздуха мгновенно преодолевает смычку, образуя взрыв, а голосовые связки производят частичное колебание, следовательно, по способу образования он является смычно-взрывным согласным, а по участию голосовых связок звонким согласным. Артикуляционные признаки звука [g]: *заднеязычно-велярный*, *смычно-взрывной*, *звонкий*.

Звук (ды) - [d] образуется кончиком языка и верхними альвеолами, следовательно, по месту образования является переднеязычно-альвеолярным согласным. При образовании звука [d] кончик языка смыкается с верхними зубам и альвеолами, струя воздуха мгновенно преодолевает смычку, образуя взрыв, а голосовые связки производят частичное колебание, следовательно, по способу образования он является смычно-взрывным (эксплозивным) согласным, а по участию голосовых связок – звонким согласным. Артикуляционные признаки звука [d]: *переднеязычно-альвеолярный*, *смычно-взрывной*, *звонкий*.

Звук (жы) - [ʒ] образуется кончиком языка и задней частью верхней альвеолы, следовательно, по месту образования является переднеязычно-постальвеолярным согласным. При образовании звука [ʒ] кончик языка и верхние альвеолы сближаются, образуя круглую щель, а голосовые связки производят частичное колебание, следовательно, по способу образования он является круглощелевым (апикальным) согласным (в начале слов – аффриката), а по участию голосовых связок – *звонким* согласным. Артикуляционные признаки звука [ʒ]: *переднеязычно-*

НИИ Искусственный интеллект

постальвеолярный, *круглощелевой* (в начале слова *аффриката*), *звонкий*.

Звук (зы) - [z] образуется кончиком языка и верхними альвеолами, следовательно, по месту образования является переднеязычно-альвеолярным согласным. При образовании звука [z] кончик языка и верхние альвеолы сближаются, образуя плоскую щель, а голосовые связки производят частичное колебание, следовательно, по способу образования он является плоскощелевым (однофокусным) согласным (спирантом), а по участию голосовых связок – звонким согласным. Артикуляционные признаки звука [z]: *переднеязычно-альвеолярный*, *плоскощелевой*, *звонкий*.

Согласный звук (й) - [j] образуется средней частью спинки языка и палатальным (твердым нёбом), следовательно, по месту образования является среднеязычно-твердонёбным согласным. При образовании звука [j] средняя часть спинки языка и твердое нёбо сближаются, образуя щель, а голосовые связки производят большие колебания, но без образования шума, следовательно, по способу образования он является плоскощелевым (фрикативным) согласным (аппроксимантом), а по участию голосовых связок – сонорным согласным. Артикуляционные признаки звука [j]: *среднеязычно-палатальный*, *плоскощелевой*, *сонорный.*

Согласный звук (кі) - [k] образуется задней частью спинки языка и мягким нёбом, следовательно,по месту образования является заднеязычно-мягконёбным (велярным) согласным. При образовании звука [k] задняя часть спинки языка и мягкое нёбо смыкаются, и струя воздуха мгновенно преодолевает смычку, образуя взрыв, а голосовые связки остаются нейтральными, следовательно, по способу образования он является смычно-взрывным согласным, а по участию голосовых связок – шумным глухим согласным. Артикуляционные признаки звука [k]: *заднеязычно-велярный*, *смычно-взрывной*, *шумный глухой.*

Согласный звук (қы) - [q] образуется задней частью спинки языка и язычком (увулой), следовательно, по месту образования

НИИ Искусственный интеллект

является заднеязычно-увулярным согласным. При образовании звука [q] задняя часть языка и язычок смыкаются, и струя воздуха мгновенно преодолевает смычку, образуя взрыв, а голосовые связки остаются нейтральными, следовательно, по способу образования он является смычно-взрывным согласным, а по участию голосовых связок – глухим согласным. Артикуляционные признаки звука [q]: *заднеязычно-увулярный*, *смычно-взрывной*, *глухой.*

Согласный звук (ыл) - [l] образуется кончиком языка и верхними зубами, следовательно, по месту образования является переднеязычно-зубным согласным. При образовании звука [l] кончик языка и альвеолы смыкаются, бока языка с зубами сближаются, образуя боковую щель, а голосовые связки производят большие колебания, следовательно, по способу образования он является смычно-боковым согласным, а по участию голосовых связок – сонорным согласным. Артикуляционные признаки звука [l]: *переднеязычно-зубной*, *смычно-боковой*, *сонорный.*

Согласный звук (мы) - [m] образуется губами и по месту образования является губно-губным (билабиальным) согласным. При образовании звука [m] губы смыкаются, и воздух проходит в обход, в носовую полость, а голосовые связки производят большие колебания, следовательно, следовательно, по способу образования он является смычно-носовым согласным, а по участию голосовых связок – сонорным согласным. Артикуляционные признаки звука [m]: *губно-губной*, *смычно-носовой*, *сонорный*.

Согласный звук (ны) - [n] образуется передней частью спинки языка и верхними зубами, следовательно, по месту образования является переднеязычно-зубным согласным. При образовании звука [n] передняя часть спинки языка и верхние альвеолы смыкаются, и воздух проходит в обход, в носовую полость, а голосовые связки производят большие колебания, следовательно, по способу образования он является смычно-носовым согласным, а по участию голосовых связок – сонорным согласным. Артикуляционные

НИИ Искусственный интеллект

признаки звука [n]: *переднеязычно-зубной*, *смычно-носовой*, *сонорный* (сонант).

Согласный звук (ыҥ) - [ŋ] образуется средней частью спинки языка и мягкой частью нёба, следовательно, по месту образования является среднеязычно-мягконёбным (велярным) согласным. При образовании звука [ŋ] средняя часть спинки языка и мягкая часть нёба смыкаются, и воздух проходит в обход, в носовую полость, а голосовые связки производят большие колебания, следовательно, по способу образования он является смычно-носовым согласным, а по участию голосовых связок – сонорным согласным. Артикуляционные признаки звука [ŋ]: *среднеязычно-велярный*, *смычно-носовой*, *сонорный*.

Согласный звук (пы) - [p] образуется обеими губами, следовательно, по месту образования является губно-губным (билабиальным) согласным. При образовании звука [p] губы смыкаются, и струя воздуха мгновенно преодолевает смычку, образуя взрыв, а голосовые связки остаются нейтральными, следовательно, по способу образования он является смычно-взрывным (эксплозивным) согласным, а по участию голосовых связок – глухим согласным. Артикуляционные признаки звука [p]: *губно-губной*, *смычно-взрывной*, *глухой*.

Согласный звук (ыр) - [r] образуется передней частью спинки языка и верхними альвеолами, следовательно, по месту образования является переднеязычно-альвеолярным согласным. При образовании звука [r] передняя часть спинки языка и верхние альвеолы то смыкаются, то сближаются, образуя поочередно смычку и узкую щель, а голосовые связки производят большие колебания, следовательно, по способу образования он является смычно-дрожащим согласным, а по участию голосовых связок – сонорным согласным. Артикуляционные признаки звука [r]: *переднеязычно-альвеолярный*, *смычно-дрожащий*, *сонорный.*

Согласный звук (сы) - [s] образуется кончиком языка и верхними альвеолами, следовательно, по месту образования является переднеязычно-альвеолярным согласным. При

НИИ искусственный интеллект

образовании звука [s] кончик языка и альвеолы сближаются, образуя плоскую щель, и струя воздуха проходит через проход, суженный преградой, а голосовые связки остаются нейтральными, следовательно, по способу образования он является плоскощелевым (фрикативным) согласным (спирантом), а по участию голосовых связок – глухим согласным. Артикуляционные признаки звука [s]: *переднеязычно-альвеолярной*, *плоскощелевой*, *глухой.*

Согласный звук (ты) - [t] образуется передней частью спинки языка и верхними зубами, следовательно, по месту образования является переднеязычно-зубным согласным. При образовании звука [t] кончик языка и передние зубы смыкаются, и струя воздуха мгновенно преодолевает смычку, образуя взрыв, голосовые связки остаются нейтральными, следовательно, по способу образования он является смычно-взрывным согласным, а по участию голосовых связок – глухим согласным. Артикуляционные признаки звука [t]: *переднеязычно-зубной*, *смычно-взрывной*, *глухой.*

Согласный звук (ў) - [w] образуется губами и мягким нёбом следовательно, по месту образования является губно-мягконёбным (лабиовелярным) согласным. При образовании звука [w] спинка языка приближается к мягкому нёбу, образуя круглую щель, губы принимают округлую форму, делая звук лабиализованным, а голосовые связки производят большие колебания, следовательно, по способу образования он является круглощелевым (апикальным) согласным, а по участию голосовых связок – сонорным согласным (аппроксимантом). Артикуляционные признаки звука [w]: *губно-мягконёбный*, *круглощелевой*, *сонорный.*

Согласный звук (фы) - [f] образуется нижней губой и верхними зубами, следовательно, по месту образования является губно-зубным (лабиодентальным) согласным. При образовании звука [f] нижняя губа и верхние зубы сближаются, образуя плоскую щель, и струя воздуха проходит через проход, суженный преградой, а голосовые связки остаются нейтральными, следовательно, по способу образования он является плоскощелевым (фрикативным)

НИИ искусственный интеллект

согласным (спирантом), а по участию голосовых связок – глухим согласным. Артикуляционные признаки звука [f]: *губно-зубной*, *плоскощелевой*, *глухой*.

Согласный звук (хы) - [x] образуется задней частью спинки языка и мягким нёбом, следовательно, по месту образования является заднеязычно-мягконёбным (велярным) согласным. При образовании звука [x] задняя часть спинки языка и мягкое нёбо сближаются, образуя плоскую щель, и струя воздуха проходит через проход, суженный преградой, а голосовые связки остаются нейтральными, следовательно, по способу образования он является плоскощелевым согласным, а по участию голосовых связок – глухим согласным. Артикуляционные признаки звука [x]: *заднеязычно-велярный*, *плоскощелевой*, *глухой*.

Согласный звук (шы) - [ʃ] образуется передней частью языка и задней частью верхней альвеолы (постальвеолы), является переднеязычно-постальвеолярным согласным. При образовании звука [ʃ] передняя часть спинки языка и верхние альвеолы сближаются, образуя плоскую щель, и струя воздуха проходит через проход, суженный преградой, а голосовые связки активизируются, следовательно, по способу образования он является плоскощелевым (фрикативным) согласным, а по участию голосовых связок – глухим шумным согласным. Артикуляционные признаки звука [ʃ]: *переднеязычно-постальвеолярный*, *плоскощелевой*, *шумный*, *глухой*.

2.1.4. Артикуляционная классификация согласных звуков казахского языка

В этом параграфе дается артикуляционная классификация согласных звуков казахского языка.

Во многих естественных языках характеристика согласных звуков состоит из 5 основных признаков:

1) место образования шума;

2) способ образования шума;

3) уровень шума (сонорность / шумность);

НИИ Искусственный интеллект

4) глухость / звонкость;

5) твёрдость / мягкость.

В таблице 2.1.4.1 представлены артикуляционные признаки всех согласных звуков казахского языка, артикуляция которых была описана в 2.1.3.

В таблице 2.1.4.2 приведена артикуляционная классификация этих согласных звуков в зависимости от артикуляционных признаков, указанных в таблице 2.1.4.1.

НИИ Искусственный интеллект

Таблица 2.1.4.1. Артикуляционные признаки казахских согласных звуков.

Название согласных звуков	Обозначение согласных звуков в МФА	**Место образования**								**Способ образования**						**Участие голоса**		
		Губные		Язычные						Смычные				Щелевые				
		Губно-губные	Губно-зубные	Переднеязычные			Среднеязычные	Заднеязычные		Смычно-взрывные	Смычно-носовые	Смычно-боковые	Смычно-дрожащие	Плоскощелевые	Круглощелевые	Глухие	Звонкие	Сонорные
				Зубные	Альвеолярные	Постальвеолярные	Твердонёбные	Мягконёбные	Увулярные									
(бы)	[b]	+	–	–	–	–	–	–	–	+	–	–	–	–	–	–	+	–
(вы)	[v]	–	–	+	–	–	–	–	–	–	–	–	–	+	–	–	+	–
(ғы)	[ɣ]	–	–	–	–	–	–	+	–	–	–	–	–	–	+	–	+	–
(гі)	[g]	–	–	–	–	–	–	+	–	+	–	–	–	–	–	–	+	–
(ды)	[d]	–	–	–	+	–	–	–	–	+	–	–	–	–	–	–	+	–
(кі)	[k]	–	–	–	–	–	–	+	–	+	–	–	–	–	–	+	–	–
(қы)	[q]	–	–	–	–	–	–	–	+	+	–	–	–	–	–	+	–	–
(жы)	[ʒ]	–	–	–	+	+	–	–	–	–	–	–	–	–	+	–	+	–
(зы)	[z]	–	–	–	+	–	–	–	–	–	–	–	–	+	–	–	+	–
(й)	[j]	–	–	–	–	–	+	–	–	–	–	–	–	+	–	–	–	+
(ыл)	[l]	–	–	–	+	–	–	–	–	–	–	+	–	–	–	–	–	+
(мы)	[m]	+	–	–	–	–	–	–	–	–	+	–	–	–	–	–	–	+
(ны)	[n]	–	–	+	–	–	–	–	–	–	+	–	–	–	–	–	–	+
(ың)	[ŋ]	–	–	–	–	–	+	–	–	–	+	–	–	–	–	–	–	+
(пы)	[p]	+	–	–	–	–	–	–	–	+	–	–	–	–	–	+	–	–
(ыр)	[r]	–	–	–	+	–	–	–	–	–	–	–	+	–	–	–	–	+
(сы)	[s]	–	–	–	+	–	–	–	–	–	–	–	–	–	+	+	–	–
(ты)	[t]	–	–	+	–	–	–	–	–	+	–	–	–	–	–	+	–	–
(ұ)	[w]	+	–	–	–	–	–	–	–	–	–	–	–	–	+	–	–	+
(фы)	[f]	–	+	–	–	–	–	–	–	–	–	–	–	+	–	+	–	–
(хы)	[x]	–	–	–	–	–	–	+	–	–	–	–	–	+	–	+	–	–
(шы)	[ʃ]	–	–	–	+	+	–	–	–	–	–	–	–	+	–	+	–	–

НИМ Искусственный интеллект

Таблица 2.1.4.2. Артикуляционная классификация согласных звуков казахского языка.

Способ образования		Место образования								Участие голоса
		Губные		Язычные						
		Губно-губные	Губно-зубные	Переднеязычные			Среднеязычные	Заднеязычные		
				Зубные	Альвеолярные	Постальвеолярные	Твердонёбные (Палатальные)	Мягконёбные (Велярные)	Увулярные	
Смычные	Взрывные (эксплозивные)	[p]		[t]				[k]	[q]	Глухие
		[b]			[d]			[g]		Звонкие
	Назальные (Носовые)	[m]		[n]				[ŋ]		Сонорные
	Боковые			[l]						Сонорные
	Дрожащие				[r]					Сонорные
Щелевые	Плоскощелевые (Фрикативные)		[f]		[s]	[ʃ]		[x]		Глухие
			[v]		[z]					Звонкие
							[j]			Сонорные
	Круглощелевые (Апикальные)									Глухие
						[ʒ]		[ɣ]		Звонкие
		[w]								Сонорные

НИИ Искусственный интеллект

2.2. Акустический анализ звуков казахского языка

2.2.1. Акустический анализ гласных звуков казахского языка

Общепринятым способом акустического анализа гласных звуков естественных языков является *формантный анализ*, который нацелен на выявление определенных формант в спектрограмме (двумерной диаграмме, в которой горизонтальная ось представляет время, а вертикальная ось - частоту) речевого сигнала и решает вопрос о качестве гласных звуков.

Форманта получается с помощью спектрографа. Количественно форманты соответствуют максимумам в речевом спектре (рисунок 2.2.1.1) и на спектрограммах обычно появляются в виде темных горизонтальных полос (рисунок 2.2.1.2).

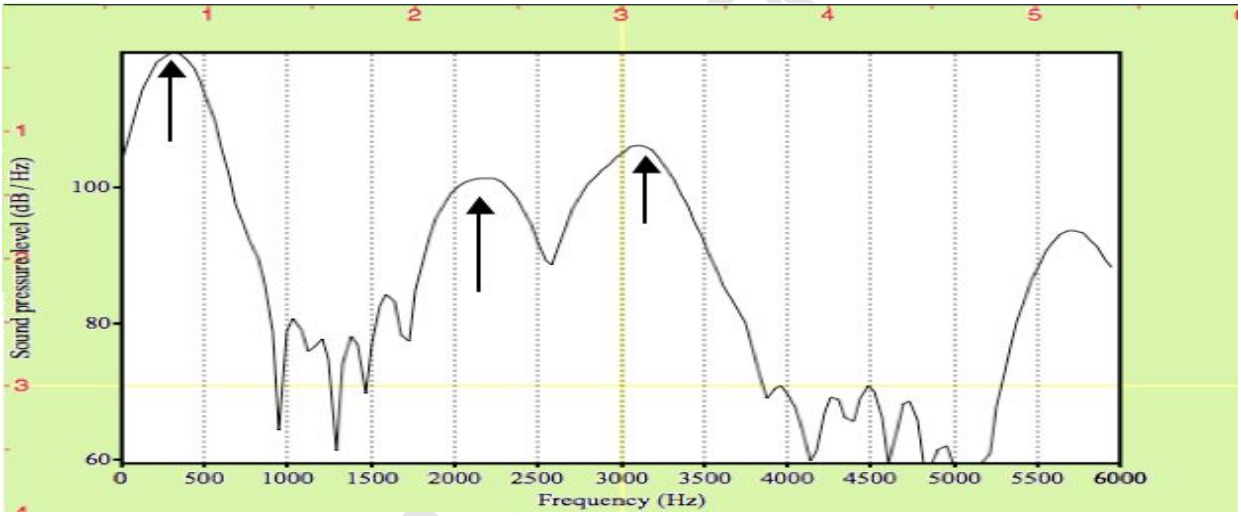


Рисунок 2.2.1.1. Форманты в спектре речевого сигнала.

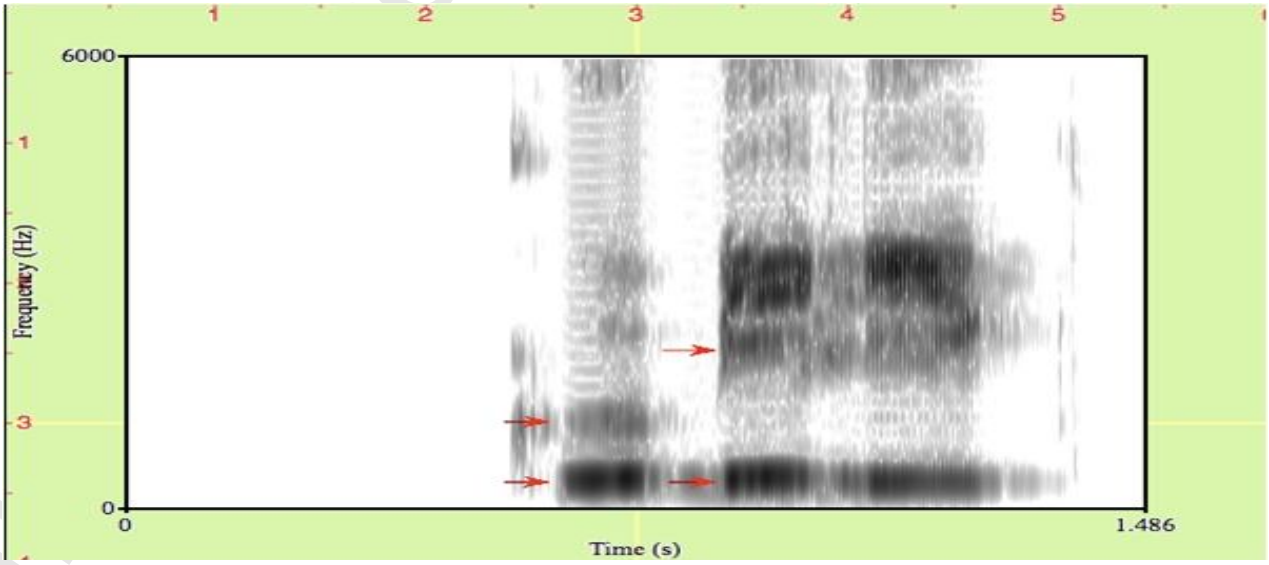


Рисунок 2.2.1.2. Форманты в спектрограмме речевого сигнала.

Обычно первые две форманты несут достаточно информации о гласных. В частности, первая форманта F1 отражает положение языка в вертикальном направлении, а вторая форманта F2 отражает положение языка в горизонтальном направлении [26, 27].

Ниже приводятся результаты эксперимента по формантному анализу гласных звуков казахского языка, проведенного в [28]. В экспериментах определены первые три форманты, но при анализе использовались только первые две. Подход основан на работе, проделанной в [29] только применительно к казахскому языку

В данном эксперименте участвовало 7 мужчин в возрасте от 25 до 30 лет. Каждый из них произносил по три раза гласные. При этом если звук звучал не отчетливо, то он произносился в контексте слова. Таким образом, всего имелся 21 пример каждого гласного звука. Анализ проводился с использованием программного обеспечения Praat [30]. Спектрограммы получены, используя 5 мс окном Хэмминга с шагом 2.5 мс и частотным шагом в 20 Гц.

Результаты эксперимента для первых трех формант представлены в таблице 2.2.1.2.

Таблица 2.2.1.2. Средние значения формантных частот.

Наименование	**Обозначение в МФА – IPA**	**F1**, Гц	**F2**, Гц	**F3**, Гц
(ə)	[æ]	685	1840	2795
(ө)	[ɵ]	357	1329	3214
(і)	[ɪ]	500	1683	2667
(ү)	[ʏ]	370	1710	2520
(а)	[ɑ]	738	1200	2913
(о)	[ɔ]	388	775	2950
(ұ)	[ʊ]	530	1170	2880
(ы)	[ɤ]	417	850	3283
(е)	[e]	313	2200	2925

НИИ Искусственный интеллект

Данные в таблице 2.2.1.2 не являются абсолютными величинами, но усредненными. Они раскрывают типичные закономерности и взаимосвязь между формантами каждой гласной.

На рисунке 2.2.1.3 показаны форманты казахских гласных.

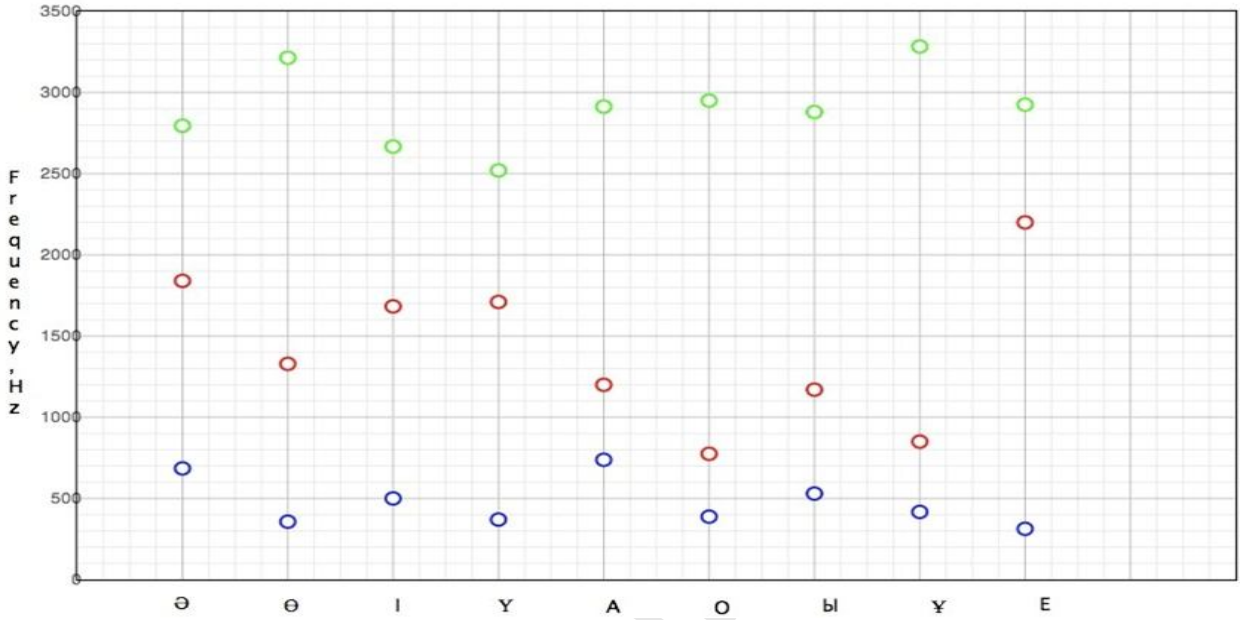


Рисунок 2.2.1.3. Форманты казахских гласных звуков,
F1 – синяя, F2 – красная, F3 – зеленая.

Распределение гласных звуков в пространстве F1-F2 показано на Рисунке 2.2.1.4.

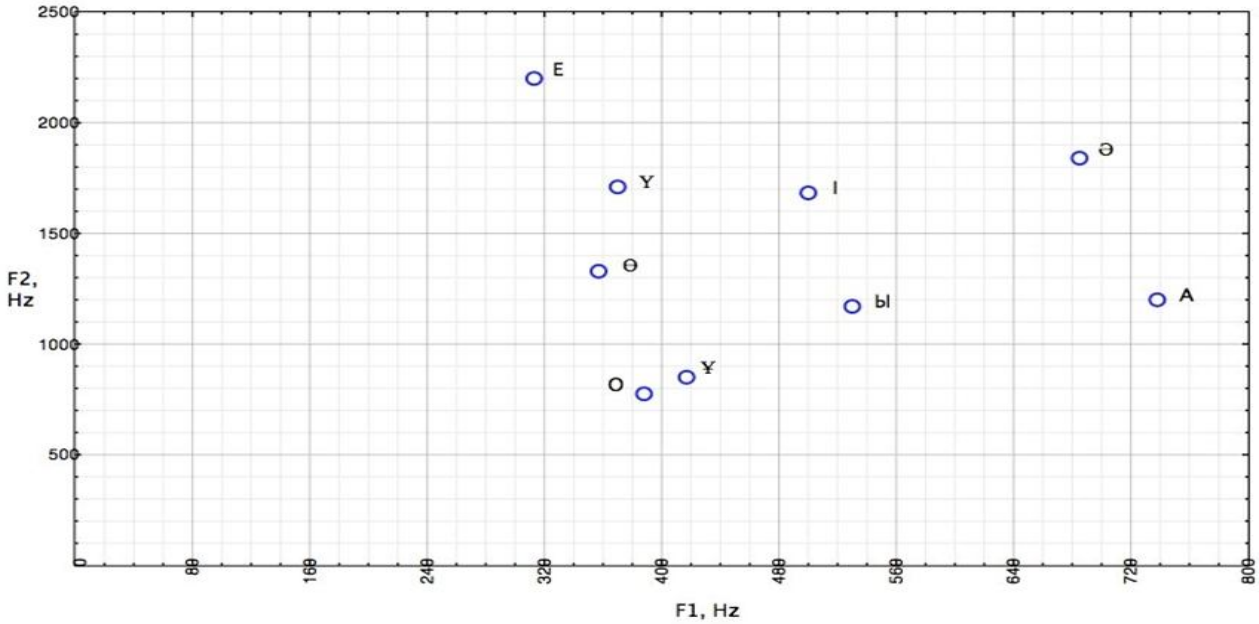


Рисунок 2.2.1.4. Гласные звуки в пространстве F1-F2

Данные из таблицы 2.2.1.2 позволяют разделить акустическую классификацию казахских гласных звуков на 2 класса по двум акустическим признакам – *мягкие* и *твердые*, что соответствует артикуляционным признакам – *переднеязычные* и *заднеязычные.*

Классификация мягких и твердых гласных казахских звуков показана на рисунке 2.2.1.5.

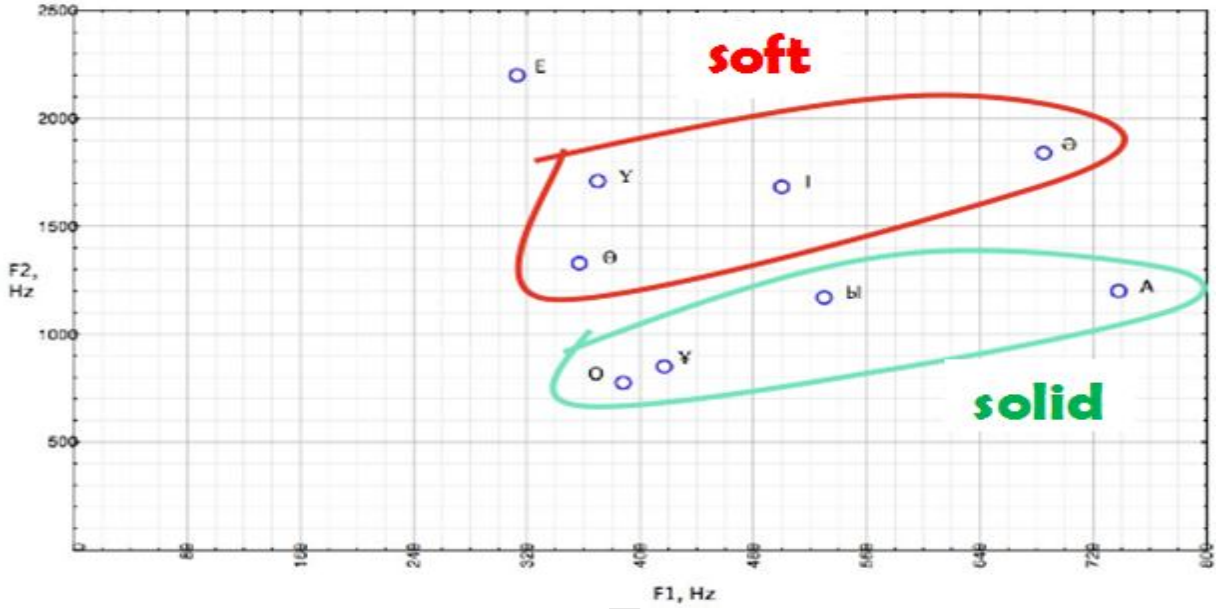


Рисунок 2.2.1.5. Классификация мягких и твердых гласных.

Здесь большое значение форманты F2 соответствует мягким гласным, а малое значение соответствует твердым гласным звукам.

Классификация высоких и низких гласных казахских звуков показана на рисунке 2.2.1.6.

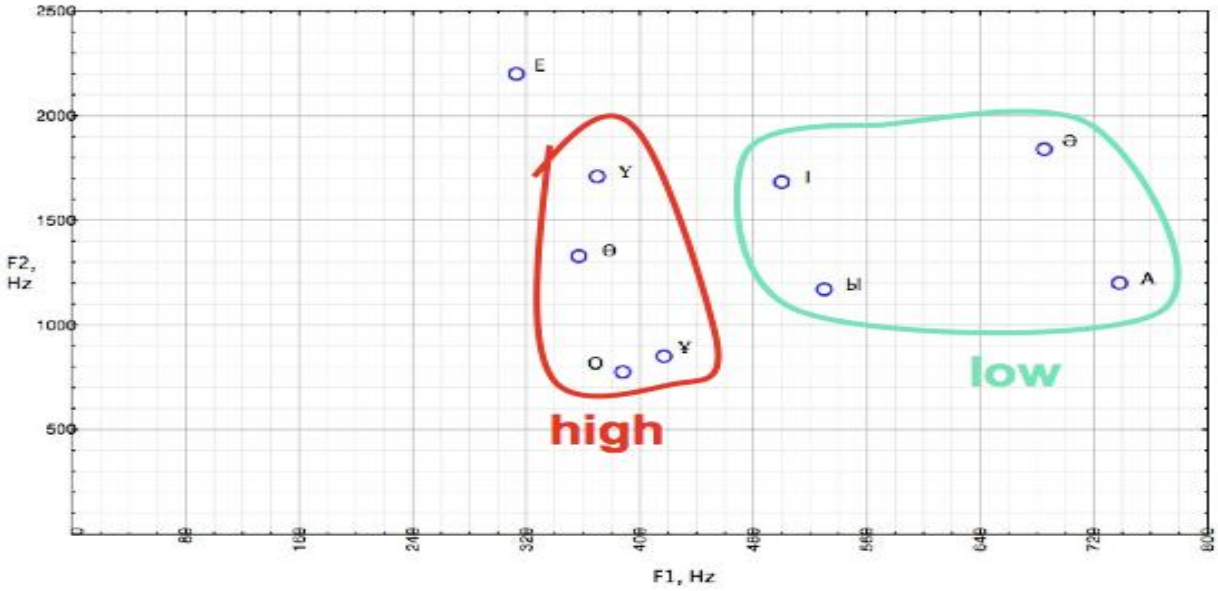


Рисунок 2.2.1.6. Классификация высоких и низких гласных.

На рисунке 2.2.1.6 показано, что значения форманты высоких (округленных) гласных звуков О - [ɔ], Ұ - [ʊ], Ө - [ɵ], Ү- [ʏ] меньше, чем у низких (неокругленных) гласных Ы - [ɤ], А - [ɑ], І - [ɪ], Ә - [æ].

Действительно, первая форманта F1 обратно пропорциональна высоте языка: чем меньше значение F1, тем выше положение языка.

Хотя гласные Ы - [ɤ] и І - [ɪ] образуют свой кластер - заднеязычные и переднеязычные соответственно, но мы предпочитаем относить их вместе с гласными А - [ɑ] и Ә - [æ] в один класс – нижние гласные.

Классификация гласных относительно положения челюсти не так очевидна. Хотя известно, что положение челюсти влияет на высоту языка и форму губ, увидеть этот эффект в имеющихся данных непросто. Традиционно открытыми гласными по положению челюсти являются А - [ɑ], Ә - [æ], О - [ɔ], Ө - [ɵ], а закрытыми гласными являются Ы - [ɤ], І - [ɪ], Ұ - [ʊ], Ү- [ʏ].

Классификация открытых и закрытых гласных по отношению к положению челюсти показана рисунке 2.2.1.7, где гласные различаются по положению челюсти.

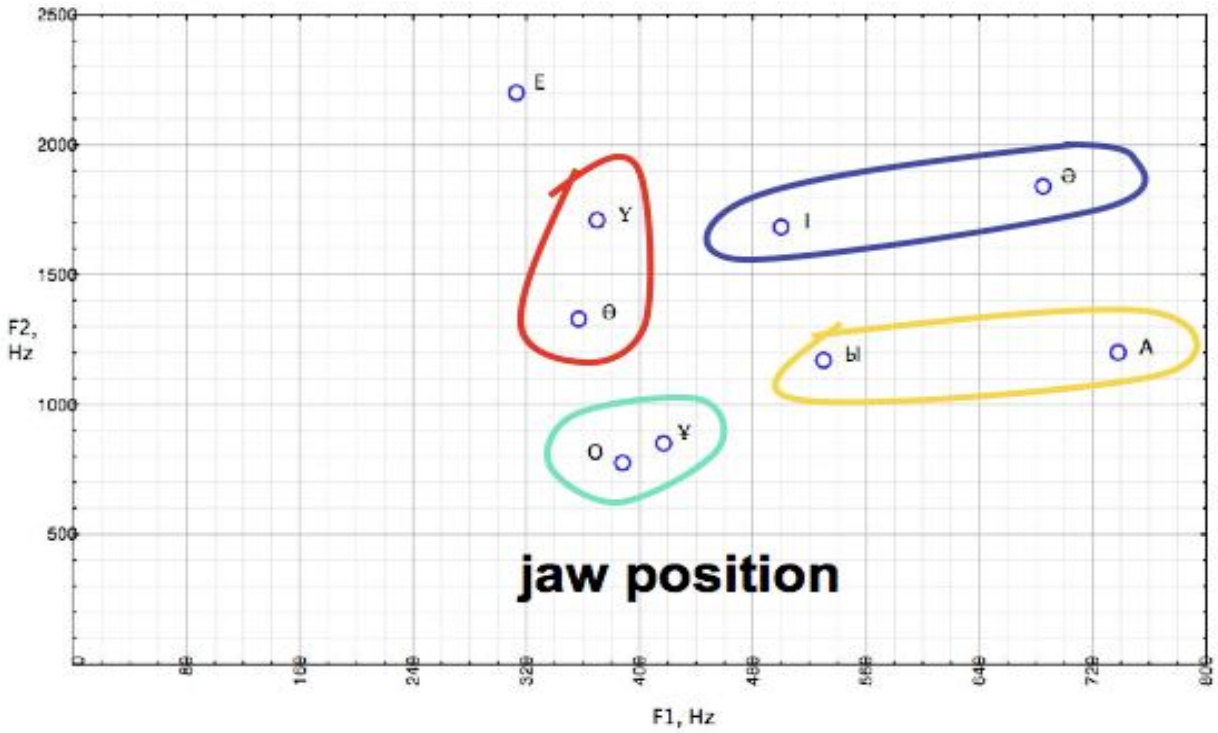


Рисунок 2.2.1.7. Классификация открытых и закрытых гласных по отношению к положению челюсти.

НИЦ ЭНЕРГИЙ ИНТЕЛЛЕКТ

Показанные на рисунке 2.2.1.7 пары {I - [ɪ], Ə - [æ]} и {Ы - [ɤ], А - [ɑ]} неокругленных гласных получены в результате операции пересечения мягких (переднеязычных) и твердых (заднеязычных) гласных с низкими гласными, а пары округленных гласных {Ө - [ɵ], Y- [ʏ]} и {О - [ɔ], Ɣ - [ʊ]} – пересечения мягких и твердых гласных с высокими гласными звуками.

Гласные в каждой паре {I - [ɪ], Ə - [æ]} и {Ы - [ɤ], А - [ɑ]} очень похожи в артикуляции, кроме положения челюсти, которое для первой гласной ниже, чем для второй. Следовательно, положение челюсти является тем признаком, который способен различать эти гласные в парах. Эффект влияния положения челюсти на положение языка можно наблюдать в одновременном уменьшении формант F1 и F2 в каждой паре, т.е. положение языка изменяется как в вертикальном, так и в горизонтальном направлениях. Что касается пар {Ө - [ɵ], Y- [ʏ]} и {О - [ɔ], Ɣ - [ʊ]} - округленных гласных, то положение челюсти для этих пар также влияет на вертикальное и горизонтальное положения языка, но в иной манере. Для них значения формантов F1 и F2 увеличиваются, а не уменьшаются, как это было в предыдущих парах.

Таким образом, эти факты оправдывают классификацию гласных А - [ɑ], Ə - [æ], О - [ɔ], Ө - [ɵ] как открытых, а гласные Ы - [ɤ], I - [ɪ], Ɣ - [ʊ], Y- [ʏ] – как закрытых по отношению к положению челюсти. Причем данный признак действительно необходим, потому что, как мы только что видели, положения языка и губ недостаточно, чтобы идентифицировать каждую гласную.

Особое внимание стоит уделить гласной [e]. Относительно первой и второй формант [e] имеет наиболее низкое значение F1 и наиболее высокое значение F2. Это означает, что звук [e] является высоким и переднеязычным гласным. Касательно положений губ и челюсти, то согласно традиционной классификации звук [e] неокругленный и открытый гласный, соответственно. При таком описании гласный [e] становится отличимым от гласного звука [æ], т.к. по вертикальному положению языка звук [æ] является низким

НИИ Искусственный интеллект

гласным. Поэтому можно сделать однозначный вывод, что для корректной и уникальной классификации казахских гласных должно быть привлечено вертикальное положение языка со значением *низкое* и *высокое*.

В результате проведенного анализа была предложена новая классификация гласных звуков казахского языка (Таблица 2.1.2.5), отличающаяся от традиционной классификации наличием нового артикуляционного признака «Вертикальное положение языка».

Акустический анализ гласных звуков казахского языка основан на следующих сингармонических тембрах:

– твердые нижние сингармонические гласные звуки;

– мягкие нижние сингармонические гласные звуки;

– твердые верхние сингармонические гласные звуки;

– мягкие верхние сингармонические гласные звуки;

– мягкий низкий сингармонический гласный звук.

Система сингармонических тембров гласных звуков казахского языка приведена в таблице 2.2.1.3

Таблица 2.2.1.3. Система сингармонических признаков гласных.

Гласные звуки	**Палатальный тембр**		**Лабиальный тембр**	
Наименование и обозначение в МФА (IPA)	Твердый	Мягкий	Низкий	Высокий
(а) - [ɑ]	+	–	+	–
(ә) - [æ]	–	+	+	–
(о) - [ɔ]	+	–	–	+
(ө) - [ɵ]	–	+	–	+
(ұ) - [ʊ]	+	–	–	+
(ү) - [ʏ]	–	+	–	+
(ы) - [ɤ]	+	–	+	–
(і) - [ɪ]	–	+	+	–

НИИ Искусственный интеллект

(e) - [e]	–	+	–	+

НИИ Искусственный интеллект

2.2.2. Акустический анализ согласных звуков казахского языка

Акустическая характеристика согласных осложняется тем, что к периодическим колебаниям (тону), характерным для гласных, добавляются непериодические колебания (шумы). Изучать последние очень трудно.

Для проведения акустического анализа казахских согласных звуков будем использовать следующие акустические признаки:

Признак *вокальность-невокальность* основан на оценке участия тона и шума, разделяет казахские согласные звуки на сонорные и шумные. К сонорным согласным звукам относятся (й) - [j], (ыл) - [l], (мы) - [m], (ны) - [n], (ыңы) - [ŋ], (ыр) - [r], (ў) - [w]. К шумным согласным звукам относятся: (бы) - [b], (вы) - [v], (ғы) - [ɣ], (гі) - [g], (ды) - [d], (кі) - [k], (қы) - [q], (жы) - [ʒ], (зы) - [z], (пы) - [p], (сы) - [s], (ты) - [t], (фы) - [f], (хы) - [h], (шы) - [ʃ].

Признак *звонкость-глухость* основан на оценке участия голоса, разделяет казахских согласные звуки на шумные звонкие и глухие. К шумным звонким согласным звукам относят все согласные, в которых голос преобладает над шумом: (бы) - [b], (вы) - [v], (ғы) - [ɣ], (гі) - [g], (ды) - [d], (жы) - [ʒ], (зы) - [z]. К глухим согласным относятся согласные, состоящие только из шума: (кі) - [k], (қы) - [q], (пы) - [p], (сы) - [s], (ты) - [t], (фы) - [f], (хы) - [h], (шы) - [ʃ].

Признак *назальность-неназальность* основан на оценке участия носового резонатора в образовании звука. К назальным звукам относятся носовые согласные: (мы) - [m], (ны) - [n], (ыңы) - [ŋ], а к неназальным звукам – все остальные согласные.

Признак *прерывность-непрерывность* показывает поведение звуковой волны в спектрограмме. Прерывные звуки на спектрограммах характеризуются резко оборванным краем волны, а у непрерывных звуков существуют переходные тоны. К прерывным звукам относятся смычные согласные (бы) - [b], (вы) - [v], (пы) - [p], (ты) - [t], а к непрерывным – щелевые согласные: (сы) - [s], (шы) - [ʃ].

Акустические признаки казахских согласных звуков приводятся в таблице 2.2.1.2.

НИИ Искусственный интеллект

Таблица 2.2.1.2. Акустические признаки казахских согласных звуков.

Название согласных звуков	Обозначение согласных звуков в МФА	Вокальные (Сонорные)	Невокальные (Шумные)		Назальность	Неназальность	Прерывность	Непрерывность
			Звонкие	Глухие				
(бы)	[b]	–	+	–	–	+	+	–
(вы)	[v]	–	+	–	–	+	+	–
(ғы)	[ɣ]	–	+	–	–	+	+	–
(гі)	[g]	–	+	–	–	+	–	+
(ды)	[d]	–	+	–	–	+	–	+
(кі)	[k]	–	–	+	–	+	–	+
(қы)	[q]	–	–	+	–	+	–	+
(жы)	[ʒ]	–	+	–	–	+	–	+
(зы)	[z]	–	+	–	–	+	–	+
(й)	[j]	+	–	–	–	+	–	+
(ыл)	[l]	+	–	–	–	+	–	+
(мы)	[m]	+	–	–	+	–	–	+
(ны)	[n]	+	–	–	+	–	–	+
(ыңы)	[ŋ]	+	–	–	+	–	–	+
(пы)	[p]	–	–	+	–	+	+	–
(ыр)	[r]	+	–	–	–	+	–	+
(сы)	[s]	–	–	+	–	+	–	+
(ты)	[t]	–	–	+	–	+	+	–
(ў)	[w]	+	–	–	–	+	–	+
(фы)	[f]	–	–	+	–	+	–	+
(хы)	[x]	–	–	+	–	+	–	+
(шы)	[ʃ]	–	–	+	–	+	–	+

В таблице 2.2.1.3 приведена классификация казахских согласных символами международного фонетического алфавита.

НИ "Искусственный интеллект"

Таблица 2.2.1.3. Классификация казахских согласных.

	Губно-губные	Губно-зуюные	Зубные	Альвеольные	Палатальвеольные	Палатальные	Велярные	Увулярные
Взрывные (эксплозивные)	p b		t d				k g	q
Назальные (носовые)	m		n				ŋ	
Фрикативные		f v		s z	ʃ ʒ		x ɣ	
Вибранты (дрожащие)			r					
Аппроксиманты (сонанты)	w					j		
Латеральный сонант			l					

Взрывные [b], [d] и [g] голосовые имеют большое время начала голоса - Voice Onset Time (VOT), которое составляет 140 -160 мс. Неголосовые взрывные [p], [t], [k] и [q] могут быть произнесены с небольшим придыханием и имеют от среднего до большого значения VOT, изменяясь в пределах 10-40 мс для звуков [p], [t] и 30-80 мс для звуков [k], [q].

Согласные звуки [f], [v] и [x] заимствованы из русского языка и используются в заимствованных словах. Например, [for'mula] – ‘формула’, [vek'tor] – ‘вектор’, [hrom] – ‘хром’. Кроме того, звук [v] добавляется в конец некоторых фамилий вместе с русскими суффиксами «ова» и «ева». Например, [ɑ'bɑjeva] – ‘Абай+ева’, [ɑ'sɑnɔva] – ‘Асан+ова’.

НИИ искусственный интеллект

В некоторых случаях в устной казахской речи фонемы [q] и [x] взаимно заменяются. Например, слово [ha'lyq] – ‘халық’ на [qa'lyq] – ‘қалық’, [hɑ'bɑr] – ‘хабар’ на [qɑ'bɑr] – ‘қабар’.

Теперь можно дать общую классификацию казахских звуков. На рисунке 2.2.1.2 показана общая классификация казахских звуков.

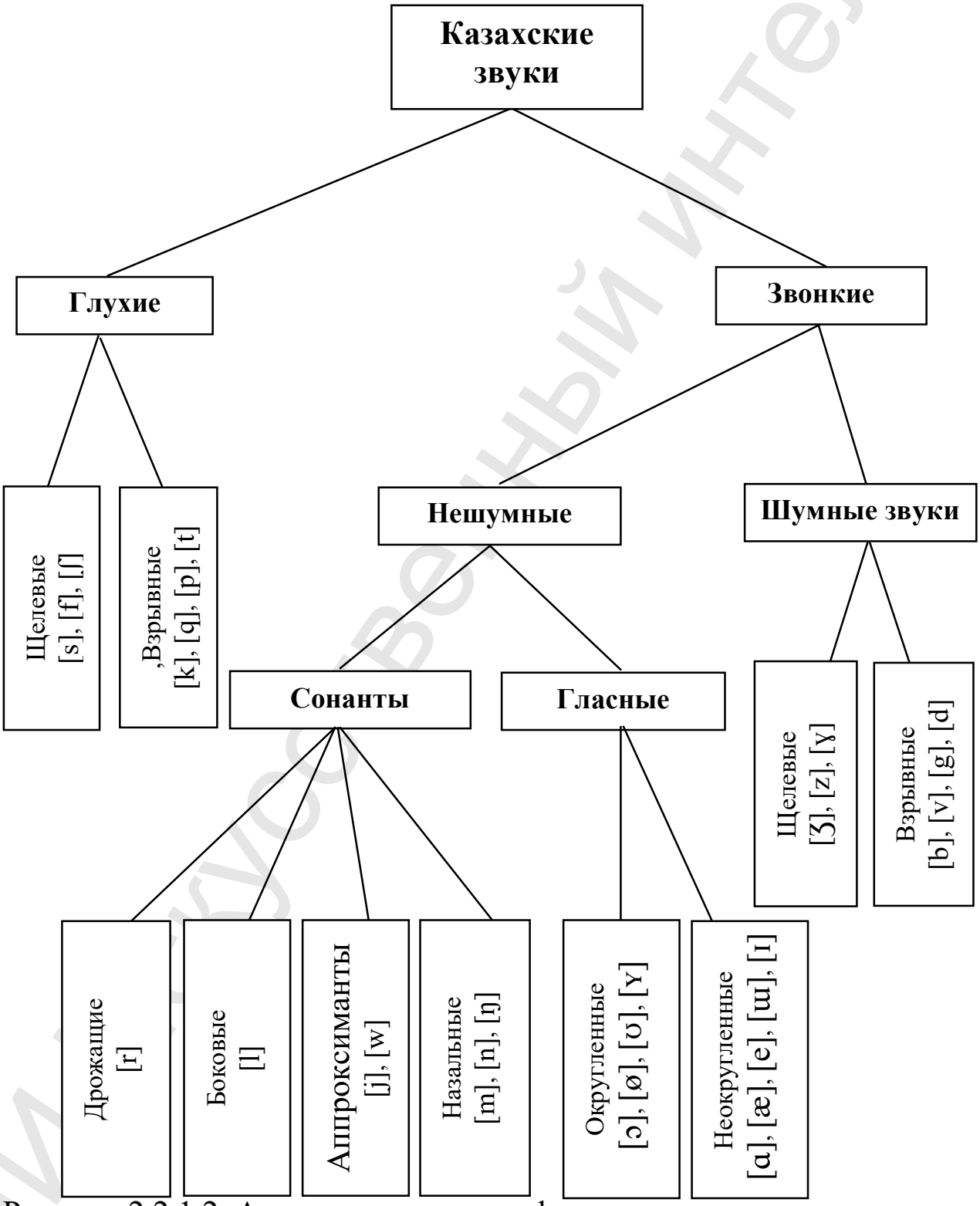


Рисунок 2.2.1.2. Акустическая классификация казахских звуков.

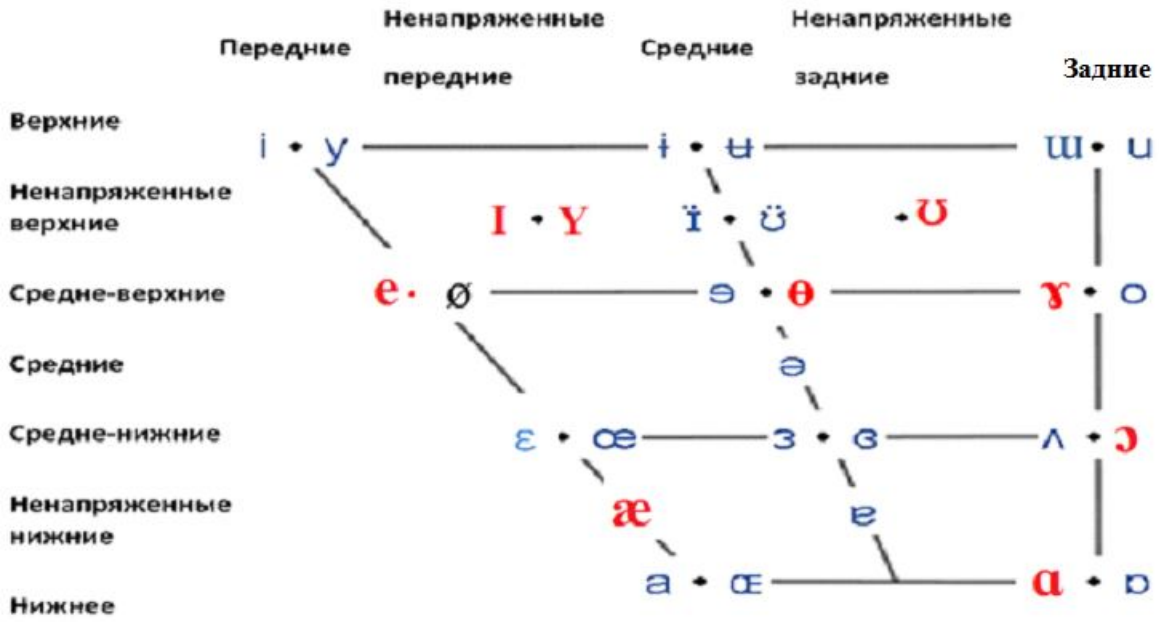
НИИ искусственный интеллект

2.2.3. Сравнение казахских звуков со звуками других языков

Большинство звуков казахского языка очень схожи по акустическим признакам с согласными звуками других естественных языков, но существуют отличительные характеристики, свойственные лишь казахскому языку.

Система гласных звуков естественных языков в МФА [31, 32] показана в таблице 2.2.3.1, в которой казахские гласные обозначены символами [ɑ], [æ], [e], [ɔ], [ɵ], [ʊ], [ʏ], [ɤ] и [ɪ].

Таблица 2.2.3.1. Гласные звуки естественных языков.



Вертикальная ось таблицы отражает подъём гласного. Гласные, произносящиеся с нижним положением языка, находятся внизу таблицы, а с поднятым языком − в верху таблицы. Отметим, что в МФА, подразумевая близость при подъеме языка к небу, термин «верхний» называется как «close», термин «нижние» – «open», аналогично признак «ненапряженный» означает «near» [31, 32].

Таким образом, горизонтальное положение гласного в таблице определяется его рядом. Гласные, для которых язык смещается вперёд (как [e]), находятся в левом столбце, а те, для которых язык смещается назад (как [ʌ]), - в правом.

В таблице, где гласные стоят парами, правый означает огублённый (лабиализованный) гласный (произносится с

НИИ Искусственного интеллект

округлёнными губами). Все остальные гласные − неогублённые (нелабиализованные).

Длительность казахских гласных не различима, но согласные могут быть удвоены за исключением велярного звука [ŋ] и увулярного звука [ɣ].

Согласные, предшествующие звуку [e], всегда палатальны (например, [bʲet] *бет* ‘лицо’). Однако увулярные согласные не могут предшествовать [e]. Фонема [e] в начале слова следует за [j], так же, как и фонемы [ɵ] и [ɔ] следуют за [w].

Теперь мы рассмотрим особенности казахских гласных звуков (а) - [ɑ], (ә) - [æ], (е) - [e], (о) - [ɔ], (ө) - [ɵ], (ұ) - [ʊ], ү- [ʏ], (ы) - [ɤ], (і) - [ɪ], которые в действующем алфавите на основе кириллицы обозначаются буквами Аа, Әә, Ее, Оо, Өө, Ұұ, Үү, Ыы, Іі и сравним их с аналогичными звуками других естественных языков:

1) Звуки [ɑ], [e], [ɔ] и [ɤ] обозначаются русскими буквами Аа, Ее, Оо, но не все совпадают по звучанию с аналогичными русскими звуками.

2) Звук [ɑ] является неокругленным гласным нижнего подъема, как и русский (а), но в отличие от переднеязычного русского, является гласным заднего ряда. Подобный звук в IPA обозначается [ɑ]. Его можно услышать:

- в некоторых контекстах в русской речи, например, в слове «палка» [ˈpɑłkə];
- в английском языке, например, в слове «spa» [spɑː];
- во французском языке, например, в слове «pâte» [pɑt];
- в финском языке, например, в слове «kana» [ˈkɑnɑ];
- в немецком языке, например, в слове «Tag» [tɑːk], и др.

3) Звук [æ] является гласным переднего ряда нижнего подъема, и его аналоги можно встретить:

- в английском языке, например, в слове «cat» [kʰæt];
- в арабском языке, например, в слове «كتـــاب» [kiˈt̪æːb];
- в финском языке, например, в слове «mäki» [ˈmæki];
- в норвежском языке, например, в слове «lær» [læːɾ], и др.

НИИ Искусственный интеллект

4) Звук [e] является неокругленным гласным переднего ряда средне-верхнего подъема. Он может быть услышан:

- в русском языке перед мягкими согласными, например, в слове «шея» [ˈʂejə];
- во французском языке, например, в слове «beauté» [bote];
- в немецком языке, например, в слове «Seele» [ˈzeːlə];
- в итальянском языке, например, в слове «stelle» [ˈstelle];
- в польском языке, например, в слове «dzień» [dʑeɲ], и др.

5) Звук [ɔ] является округленным гласным заднего ряда, как русский (о), но русский звук имеет средне-верхний подъем, а казахский – средне-нижний. Его можно встретить:

- во французском языке, например, в слове «port» [pɔʀ];
- в немецком языке, например, в слове «voll» [fɔl];
- в итальянском языке, например, в слове «parola» [paˈɾɔːla];
- в польском языке, например, в слове «kot» [kɔt], и др.

6) Звук [ɵ] является округленным гласным среднего ряда средне-верхнего подъема. Подобный звук можно в IPA можно обозначить [ɵ] или [ɵ], и его можно встретить:

- в норвежском языке, например, в слове «søt» [sɵːt];
- в голландском языке, например, в слове «keuken» [kɵːkən];
- в монгольском языке, например, в слове «өгөх» [ɵgɵx];
- в русском языке, например, в слове «тётя» [ˈtʲɵtʲə], и др.

7) Звук (ɣ) является округленным гласным заднего ряда средне-верхнего подъема. Подобный звук в IPA обозначается [o], и его можно встретить:

- в шведском языке, например, в слове «åka» [ˈoːka];
- в исландском языке, например, в слове «bók» [bou̯k];
- в корейском языке, например, в слове «보수» [ˈpoːsu];
- в португальском языке, например, в слове «sou» [so];
- в чешском языке, например, в слове «oko» [ˈoko], и др.

8) Звук (ɣ) является округленным гласным переднего ряда верхнего подъема. Подобный звук можно в IPA можно обозначить [ʏ] или [ɵ], и его можно встретить:

НИИ Искусственный интеллект

- в датском языке, например, в слове «købe» [ˈkʰøːb̥ə];

- во французском языке, например, в слове «peu» [pø];

- в немецком языке, например, в слове «schützen» [ˈʃʏtsˑn]

- в шведском языке, например, в слове «ut» [ʏβ̞t];

- в армянском языке, например, в слове «qhɪɳ» [kʰʏɣ], и др.

9) Звук [ɤ] является гласным заднего ряда средне-верхнего подъема и, по сути, представляет собой нейтральный звук, встречающийся во многих языках мира в безударной позиции. Однако в казахском языке он может быть ударным. Под ударением этот гласный может повышаться в подъеме и сдвигаться назад в ряде. Это звук можно встретить:

- в русском языке в качестве редуцированного безударного гласного, например, в слове «тыква» [tɤˈkvʌ];

- в болгарском языке, например, в слове «зъб» [zɤp];

- в эстонском языке, например, в слове «kõrv» [kɤrv];

- в диалектах китайского языка, например, в слове «喝» [xɤ].

10) Звук [ɪ] является неокругленным гласным среднего ряда среднего или средне-верхнего подъема. Он во многом похож на нейтральный звук, однако его подъем несколько выше. Он может быть ударным. Его можно встретить:

- в русском языке после [t͡ʃ], например, в слове «солнце» [ˈsont͡ɕ];

- в голландском языке, например, в слове «ik» [ə̞k];

- в польском языке, например, в слове «mysz» [mə̞ʂ];

- в монгольском языке, например, в слове «үсрэ» [usəɾɞ̆].

В действующем казахском алфавите имеются три русские буквы Ии, Уу и Ээ, соответствующие обозначениям русских гласных звуков (и), (у) и (э), но звучания этих букв в казахском языке существенно разнятся от них: первые две обозначают дитфтонги (ый,ій) и (ұу,үу) соответственно. Например, в словах «тиын - тыйын», «тиін - тійін» и «су - сұу», «куə - күуə». Звучание буквы Ээ соответствует твердому аллофону гласного [e], имеющего более низкий подъем языка, чем [e]. Буква Ээ используется только в записи словах, заимствованных из русского языка, например, в

НИИ Искусственный интеллект

слове «экономика». Однако в речи носителей казахского языка первый звук в слове «экономика» звучит как звук [е] - неокругленный гласный переднего ряда средне-нижнего подъема. Поэтому русские гласные звуки (и), (у) и (э) не включаются в состав казахских звуков.

Все указанные казахские гласные звуки являются фонемами. Примеры, доказывающие, что все перечисленные гласные звуки являются фонемами, показаны в таблице 2.2.3.2.

Таблица 2.2.3.2. Примеры использования гласных фонем.

Фонема	**Орфографическая транскрипция**	**Фонетическая транскрипция**	**Русское значение**
(а) - [ɑ]	*сан*	[sɑn]	число
(ә) - [æ]	*сән*	[sæn]	красота
(е) - [е]	*сен*	[sen]	ты
(о) - [ɔ]	*қос*	[qɔs]	пара
(ө) - [ɵ]	*көс*	[kɵs]	коси
(ұ) - [ʊ]	*тұр*	[tʊr]	стой
(ү) - [ʏ]	*түр*	[tʏr]	вид
(ы) - [ɤ]	*қыр*	[qɤr]	грань
(і) - [ɪ]	*кір*	[kɪr]	грязь

Исконно казахские слова подчиняются сингармонизму гласных, т.е. слова, включая суффиксы и окончания, могут состоять либо из твердых звуков ([ɑ], [ɔ], [ɤ], [ʊ]), либо из мягких звуков ([е], [æ], [ɵ], [I], [ʏ]). Например, [ʒɑ'rɤqta] - ‘жарық+та’ - ‘при свете’, [bʏgIn'gʲe] - ‘бүгін+ге’ - ‘на сегодня’, здесь «та» и «ге» являются окончаниями. Исключением является звук [е], который может встречаться с обоими типами гласных. Например, [qɑ'ʒʲet] ‘қажет’ - ‘нужда’, [kɵ'kek] - ‘көкек’ - ‘кукушка’. В таких случаях последний гласный определяет тип гласных в добавляемых суффиксах. Например, [qɑʒʲet'tɪ] - ‘қажет+ті’ - ‘необходимый’. Это же правило относится и к заимствованным словам. Например, [bɑlʲet'kʲe] ‘балет+ке’ - ‘в балет’, [kItɑp'tɑ] ‘кітап+та’ - ‘в книге’.

НИИ Искусственный интеллект

Следует отметить, что в русском и английском языках отсутствуют признаки мягкости гласных фонем.

В казахском языке также имеется губной сингармонизм при использовании губных (округленных) гласных [ʏ], [ɵ], [ɔ] и [ʊ], но это не явно выражено и не отражено в орфографии. Например, слово [bʏ'gIn] – «сегодня» произносится как [bʏ'gʏn].

Теперь представим казахские согласные звуки в международном фонетическом алфавите.

Таблица 2.2.3.3 Казахские согласные звуки в международном фонетическом алфавите.

	Билабиальные		Лабиодентальные		Дентальные		Альвеольные		Палатоальвеолярные		Ретрофлексные		Палатальные		Велярные		Увулярные		Глоттальные	
Взрывные	**p**	**b**	p̪	b̪	t̪	d̪	**t**	**d**			ʈ	ɖ	c	ɟ	**k**	**g**	**q**	ɢ	ʔ	
Назальные		**m**		ɱ		n̪		**n**				ɳ		ɲ		**ŋ**		ɴ		
Фрикативные	ɸ	β	**f**	**v**	θ	ð	**s**	**z**	**ʃ**	**ʒ**	ʂ	ʐ	ç	ʝ	**x**	**ɣ**	χ	ʁ	h	ɦ
Аппроксиманты	**w**			ʋ				ɹ				ɻ		**j**		ɰ				
Одноударные		ⱱ		ⱱ				ɾ				ɽ						ɢ̆		
Вибранты		ʙ						**r**				•						ʀ		
Латеральные аппроксиманты								**l**				ɭ		ʎ		ʟ				
Латеральные фрикативы							ɬ	ɮ			ɭ̊˔		ʎ̝̊˔		l̥	ɬ				

НИМ Искусственный интеллект

Все указанные казахские согласные звуки являются фонемами, что подтверждается приведенными примерами в таблице 2.2.3.4.

Таблица 2.2.3.4. Примеры использования согласных звуков.

№	Фонема	Фонетическая транскрипция	Орфографическая транскрипция	Русское значение
1	p	[pɑj]	пай	доля
2	b	[bɑj]	бай	богатый
3	t	[tɑq]	тақ	трон
4	d	[dɑq]	дақ	ржавчина
5	k	[kʏl]	күл	пепел
6	g	[gʏl]	гүл	цветок
7	q	[qɑn]	қан	кровь
8	x	[xɑn]	хан	хан
9	ɣ	[ɣɑ'lɑm]	ғалам	вселенная
10	q	[qɑ'lɑm]	қалам	перо
11	m	[tɔm]	том	темнота
12	n	[tɔn]	тон	шуба
13	ŋ	[tɔŋ]	тоң	холод
14	f	[fæni]	фәни	бренный
15	p	[pæni]	пәни	тленный
16	s	[sɔr]	сор	бедствие
17	z	[zɔr]	зор	огромный
18	ʃ	[ʃɑq]	шақ	подходящий
19	ʒ	[ʒɑq]	жақ	сторона
20	w	[wɑq]	уақ	мелкий
21	j	[tɔj]	той	свадьба
22	r	[tɔr]	тор	клетка
23	l	[lɑq]	лақ	козленок

НИИ Искусственный интеллект

Дрожащий согласный звук [r] казахского языка очень похож на русский звук, но существенно отличается от английского аналога, в котором кончик язык немного оттянут назад и не вибрирует.

Таким образом, видим, что в казахском языке некоторые звуки отличаются по артикуляционным и акустическим признакам от соответствующих звуков других языков.

Ударение в казахских словах всегда падает на последний слог. В заимствованных словах ударение подчиняется этому правилу, если к слову добавлен суффикс, содержащий гласный звук. В противном случае, ударение зависит от языка, из которого это слово заимствовано.

Для сравнения в таблице 2.2.3.5 покажем орфографическую и фонетическую транскрипции заданного текста.

Таблица 2.2.3.5. Орфографическая и фонетическая транскрипция.

Орфографическая транскрипция	Фонетическая транскрипция
Қалың тонға оранған саяхатшыны көрген Солтүстік жел мен Күн кім мықты екендігіне сөз таластырады.	qɑl'ɤŋ tɔn'ɣɑ ɔrɑn'ɣɑn sɑjɑxɑtʃɤ'nɤ kʊr'gʲen sɔltʏs'tIk 'ʒʲel 'mʲen 'kʏn 'kIm mɤq'tɤ ekʲendIgI'ɲʲe 'sʊz tɑlɑstɤrɑ'dɤ.
Олар кім бірінші саяхатшыға өз тонын шешкізсе, сол екіншісінен мықты саналады деген шешімге келеді.	ɔl'ɑr 'kIm bIrIn'ʃI sɑjɑxɑtʃɤ'ɣɑ 'wʊz tɔ'nɤn ʃʲeʃkIz'sʲe, 'sɔl ekInʃIsI'ɲʲen mɤq'tɤ sɑnɑlɑ'dɤ dʲe'gʲen ʃʲeʃIm'gʲe kʲelʲe'dI.
Алдымен Солтүстік жел барынша үрлейді, бірақ не құрлым қатты үрлесе сол құрлым саяхатшы тонына орана береді.	ɑldɤ'mʲen sɔltʏs'tIk 'ʒʲel bɑrɤn'ʃɑ ʏrlʲej'dI, bI'rɑq 'ɲʲe qʊr'lɤm qɑt'tɤ ʏrlʲe'sʲe 'sɔl qʊr'lɤm sɑjɑxɑt'ʃɤ tɔnɤ'nɑ ɔrɑ'nɑ bʲerʲedI.
Соңында Солтүстік жел өз талпынуын тоқтатады.	sɔŋɤn'dɑ sɔltʏs'tIk 'ʒʲel 'wʊz tɑlpɤnʊ'wɤn tɔqtɑtɑ'dɤ.
Сосын Күн шығады, жылы шуағын шашады, сол кезде саяхатшы тонын дереу шешеді.	sɔ'sɤn 'kʏn ʃɤɣɑ'dɤ, ʒɤ'lɤ ʃʊwɑ'ɣɤn ʃɑʃɑ'dɤ, 'sɔl kʲez'dʲe sɑjɑxɑt'ʃɤ tɔ'nɤn dʲe'rʲew ʃʲeʃʲe'dI.

НИИ искусственный интеллект

Осылайша, Солтүстік жел екеуінің арасында Күннің мықты екенін мойындауға мәжбүр болады.	ɔsɤlɑj'ʃɑ, sɔltʏs'tIk 'ʒʲel ekʲewI'nIŋ ɑrɑsɤn'dɑ kʏn'nIŋ mɤq'tɤ ekʲe'nIn mɔjɤndɑw'ɣɑ mæʒ'bʏr bɔlɑ'dɤ.

3. КАЗАХСКИЙ АЛФАВИТ

3.1. Перевод казахского алфавита на латиницу

3.1.1. Латинский алфавит и его разновидности

Латинский алфавит (латиница) произошел от западного варианта греческого алфавита в VII веке до н. э. и впоследствии распространился по всему миру. Первоначально латинский алфавит включал только 21 прописную (заглавные) букву: A, B, C, D, E, F, Z, H, I, K, L, M, N, O, P, Q, R, S, T, V, X. Затем в I веке до н.э. в него были добавлены буквы Y и Z для записи слов, заимствованных из греческого языка и, в итоге, получился классический латинский алфавит из 23 букв. В этом алфавите на рубеже IX века появились строчные (маленькие) буквы, а в XI веке в письме германских языков стали использовать диграф VV. Затем в XVI веке в него ввели 2 новые буквы J и U для дифференциации слоговых и неслоговых вариантов букв I и V (I/J и U/V). Примерно в то же время в северной Европе диграф VV стал считаться отдельной буквой W и тем самым в латинском алфавите количество букв стало равно 26: Aa, Bb, Cc, Dd, Ee, Ff, Gg, Hh, Ii, Jj, Kk, Ll, Mm, Nn, Oo, P, Qq, Rr, Ss, Tt, Uu, Vv, Ww, Xx, Yy, Zz. Этот 26-буквенный латинский алфавит зафиксирован Международной организацией по стандартизации– International Organization for Standardization (ISO) как «*базовый латинский алфавит*». Он совпадает с современным английским алфавитом и является основой письменности большинства романских, германских и других языков [34].

В настоящее время базовый латинский алфавит де-факто стал «алфавитом международного общения», так как граждане любой страны знакомятся с ним еще со школы на уроках математики, химии и иностранного языка.

Неоднократно выдвигалась идея перевода письменностей всех языков на базовый латинский алфавит. Однако в некоторых языках количество звуков больше, чем количество букв в базовом латинском алфавите, и в зависимости от их фонетических систем некоторые буквы должны иметь другие значения. Поэтому для таких языков приходится адаптировать базовый латинский алфавит [35].

В процессе адаптации базового латинского алфавита к фонетическим системам некоторых языков возникла проблема обозначения звуков, для которых отсутствовали буквы в этом алфавите. Для ее решения можно применить следующие методы адаптации алфавита:

1. Использование цепочек букв (диграфов, триграфов, тетраграфов) для написания одного звука: например, для обозначения звуков (ш) - [ʃ] и (ч) - [tʃ] в английском языке используются диграфы ***sh*** и ***ch***, а в немецком языке − триграф ***sch*** и тетраграф ***tsch,*** соответственно.

2. Использование положительной диакритики: например, во многих языках использовали буквы типа ***á***, ***ä***, ***ğ,*** полученные путем добавления точек, птичек и других диакритических знаков существующим буквам алфавита;

3. Использование отрицательной диакритики: например в турецком языке звук (ы) - [ɤ] обозначается буквой ***ı***, полученной путем удаления точки от буквы базового латинского алфавита ***i***;

4. Введение новых букв, модифицируя имеющиеся в алфавите буквы: например, модифицируя имеющиеся в алфавите буквы **c** и **s** для обозначения звуков (си) - [si:] и (эс) - [es] в английском языке получим буквы **ç** и **ş** для обозначения звуков (че) - [tʃ] и (ше) - [ʃ] в турецком языке, соответственно.

5. Введение новых букв за счет заимственных букв из других алфавитов: например, в некоторых языках буквы ***ŋ*** и ***ɔ*** заимствованы из греческого алфавита, а буква ***þ*** − из рунической азбуки;

НИИ Искусственный интеллект

6. Использование лигатуры, полученной из сращения двух букв: например, вышеуказанное образование W из двойной VV в поздней латыни, английском, немецком и польском языках или ***æ*** из ***a*** и ***e*** в датском, норвежском и исландском языках.

7. Использование нескольких букв и диакритического знака для обозначения одного звука, например, в бретонском языке звук «*х*» обозначается диграфом ***c*** и ***h*** с апострофом ***'***, т.е. цепочкой ***c'h***.

8. Изменение значений отдельных букв в базовом латинском алфавите: например, буква ***x*** в английском языке обозначает звук (экс) - [eks], а в португальском языке обозначает звук (шиш) - [ʃiʃ].

Замечание: Диграфы, триграфы, тетраграфы по существу являются орфографическими правилами того или иного языка, в котором они используются.

В настоящее время термин «латинский алфавит» подразумевает алфавит любого языка, использующего адаптированный вариант латинского алфавита, который может не включать отдельные буквы базового латинского алфавита или содержать дополнительные буквы. Кроме того, в латинизированных алфавитах взаимное расположение букв может быть разным.

Для всех языков с нелатинской письменностью существуют также системы записи латиницей (романизации). Если иностранец даже не знает правильного чтения, ему гораздо легче иметь дело со знакомыми латинскими буквами, чем с китайскими иероглифами. В ряде стран вспомогательное письмо латиницей стандартизировано, и дети изучают его в школе (в Японии, Китае).

Латинский алфавит используется во всем мире для латинизации языков, которые пользуются нелатинскими алфавитами, с целью упрощения коммуникации. Большинство языков с нелатинскими алфавитами имеют официальные правила *транслитерации* на базе латиницы, например для славянских языков с алфавитами на основе кириллицы [36-38]. Иногда такие системы транслитерации называют романизированными, т.е. производными от романских, а следовательно, латинских языков.

НИ интеллект

3.1.2. Опыт перевода алфавита тюркских языков на латиницу

Используя некоторые сочетания вышеуказанных методов можно построить новый алфавит казахского языка на основе латинской графики. Для этого мы должны определить критерии и задачи перевода казахской письменности на новый алфавит, а также разработать правила избавления от ненужных казахскому языку русских звуков, внесенных в 1940 году, и разработать алгоритм замены кириллических букв на латинские буквы. При этом мы должны учесть опыт латинизации письменности в независимых тюркских странах, а именно опыт перевода алфавита на латиницу турецкого, туркменского, узбекского и азербайжанского языков.

Известно, что наиболее ранним из известных примеров тюркской письменности является древнетюркское руническое письмо, называемое орхоно-енисейской письменностью [3, 10]. В настоящее время тюркские языки используют множество систем письма: арабскую, греческую, кириллическую, латинскую и другие.

В Турции на протяжении долгого времени, вплоть до XX века, использовался турецкий вариант арабского алфавита, который не совсем подходил для записи исконно турецких слов. Арабский язык имеет много согласных, турецкий же богат на гласные: тогдашний алфавит не мог в полной мере передать звуковой строй речи, поэтому и возникла идея реформы. Современный турецкий алфавит на основе латинской графики был принят в 1928 году в рамках реформы турецкого языка, начатой Ататюрком. Турция с целью идеологического обоснования национальной идеи турецкого государства перешла с арабского на латинский алфавит. Символы турецкого алфавита были заимствованы из различных алфавитов языков Европейских стран [39].

Современный турецкий алфавит на основе латинской графики адаптирован к фонетике турецкого языка и содержит 29 букв *Aa*, *Bb*, *Cc*, Çç, *Dd*, *Ee*, *Ff*, *Gg*, Ğğ, *Hh*, Iı, İi, *Jj*, *Kk*, *Ll*, *Mm*, *Nn*, *Oo*, Öö, P, *Rr*, *Ss*, Şş, *Tt*, *Uu*, Üü, *Vv*, *Yy*, Z z, из них 6 букв Çç, Ğğ, İı, Öö, Şş, Üü

НИИ искусственный интеллект

не входят в базовый латинский алфавит и 3 буквы базового латинского алфавита *Qq*, *Xx*,*Ww* не используются (Таблица 4.3.1).

В начале 1990-х годов в целях упрощения чтения Туркменской орфографии, туркменская письменность была переведена на латинскую графическую основу. Туркменский алфавит на основе латинской графики принимался трижды [40]:

– вариант 1992 года содержал 26 букв *Aa*, *Bb*, *Cc*, *Dd*, *Ee*, *Ff*, *Gg*, *Hh*, *Ii*, *Jj*, *Kk*, *Ll*, *Mm*, *Nn*, *Oo*, *Qq*, P, *Rr*, *Ss*, *Tt*, *Uu*, *Vv*, *Ww*, *Xx*, *Yy*, *Zz*, из них 3 буквы *Qq*, *Vv*, *Xx* использованы нетрадиционно для замены кириллических букв Өө, Үү, Ыы соответственно, а 5 диграфов Ea ea, Jh jh, Ng ng, Sh sh, Ts ts – для обозначения специфических звуков туркменского языка (Таблица 4.3.2);

– вариант 1993 года содержал 30 букв *Aa*, *Bb*, Çç, *Dd*, *Ee*, Ää, *Ff*, *Gg*, *Hh*, *Ii*, *Jj*, £ƒ, *Kk*, *Ll*, *Mm*, *Nn*, Ññ, *Oo*, Öö, P, *Rr*, *Ss*, $¢, *Tt*, *Uu*, Üü, *Ww*, *Yy*, ¥ÿ, *Zz*, из них 3 знака валют £ƒ, $¢, ¥ÿ и 5 диакритических букв Çç, Ää, Ññ, Öö, Üü предложены для обозначения специфических звуков туркменского языка (Таблица 4.3.2);

– вариант 1995 года содержит 30 букв *Aa*, *Bb*, Çç, *Dd*, *Ee*, Ää, *Ff*, *Gg*, *Hh*, *Ii*, *Jj*, Žž, *Kk*, *Ll*, *Mm*, *Nn*, Ňň, *Oo*, Öö, P, *Rr*, *Ss*, Şş, *Tt*, *Uu*, Üü, *Ww*,*Yy*,Ýý, *Zz*, из них 8 букв Çç, Ää, Žž, Ňň, Öö, Şş, Üü, Ýý используются для обозначения специфических туркменских звуков и не входят в базовый латинский алфавит, а 4 буквы базового латинского алфавита *Cc*, *Qq*, *Vv*, *Xx* не используются (Таблица 4.3.2).

Последний вариант с 2000 года стал единственным допустимым во всех официальных сферах Туркмении.

Перевод туркменской письменности на латинскую графику был осуществлён без учёта реальных возможностей страны: в течение 6 лет после реформы языка отсутствовали учебники на латинице, что отрицательно сказалось на качестве образования [41].

Узбекский алфавит на основе латиницы в 1990-х годах был принят дважды [42] :

НИИ Искусственный интеллект

– вариант 1993 года содержит 31 букву (*Aa*, *Bb*, *Cc*, *Dd*, *Ee*, *Ff*, *Gg*, *Hh*, *Ii*, *Jj*, *Kk*, *Ll*, *Mm*, *Nn*, *Oo*, P, *Qq*, *Rr*, *Ss*, *Tt*, *Uu*, *Vv*, *Xx*, *Yy*, *Zz*, Çç, Ğğ, Ɉɉ, Ññ, Öö, Şş), из них 6 букв Çç, Ğğ, Ɉɉ, Ññ, Öö, Şş не входят в базовый латинский алфавит и используются для обозначения специфических звуков узбекского языка;

– вариант 1995 года содержит 24 буквы базового латинского алфавита и 1 знак ′ - аксант эгю, при этом 2 буквы базового латинского алфавита *Cc* и *Ww* не используются и ранее используемые 5 диакритических букв Çç, Ññ, Şş, Ğğ, Öö заменены на 5 диграфов Ch ch, Ng ng, Sh sh, G′ g′, O′ o′, а специфическая буква Ɉɉ замена на букву *Jj* (Таблица 4.3.3).

Следует учесть, что в узбекском языке диграфы Ch ch, Ng ng, Sh sh, G′ g′, *O′o′* не надо считать буквами алфавита, они являются орфографическими правилами языка.

Перевод азербайджанской письменности на латинский алфавит начали в 1992 году и завершили в 2001 году. В новом алфавите имеются 32 буквы *Aa*, *Bb*, *Cc*, Çç, *Dd*, *Ee*, Əə, *Ff*, *Gg*, Ğğ, *Hh*, *Xx*, Iı, İi, *Jj*, *Kk*, *Qq*, *Ll*, *Mm*, *Nn*, *Oo*, Öö, P, *Rr*, *Ss*, Şş, *Tt*, *Uu*, Üü, *Vv*, *Yy*, *Zz*, из них 8 букв Çç, Əə, Ğğ, Iı, İi , Öö, Şş, Üü не входят в базовый латинский алфавит и 2 буквы базового латинского алфавита *Cc* и *Ww* не используются (Таблица 4.3.4).

Следует отметить, что задолго до перехода на латиницу из состава кириллического алфавита азербайджанского языка в 1947 году была исключена буква Цц, а в 1958 году эта же участь постигла и буквы Ээ, Юю и Яя, которые использовались для записи заимствованных из русского языка слов. В первоначальном проекте азербайджанского алфавита на основе латинской графики вместо буквы Əə была использована буква Ää [43].

Выводы:

1. В указанных алфавитах тюркских языков, кроме узбекского, некоторые буквы не входят в состав базового латинского алфавита, и порядок следования букв в них не соответствует порядку их следования в базовом латинском алфавите, что потребует **немалых затрат на разработку программ для сортировки и поиска**

НИИ Искусственный интеллект

данных при каждой смене поколений компьютеров и программного обеспечения, так как для базового латинского алфавита они встраиваются автоматически.

2. В в узбекском алфавите в качестве второй пары диграфа используется знак ′ - аксант эгю (апостроф), который в языках программирования применяется как специальный знак для выделения в тексте отдельных его фрагментов. В случае его использования при компьютерной обработке узбекских текстов звуки [ɔ] и [ɣ], обоначаемые как *O′o′* и *G′g′* соответственно, **могут быть не распознанными**. В принципе, знак ′ - аксант эгю можно легко заменить на знак ` – *аксант грав* (см. Рисунок 3.5.2).

НИИ Искусственный интеллект

Таблица 3.3.1. Турецкий алфавит.

№	Латиница	МФА	Название на кириллице	Характеристика звука
1	A a	[a]	а	Звучит более открыто, чем русский (а)
2	B b	[b]	бе	Звучит менее звонко, чем русский (б)
3	C c	[dʒ]	дже	Звучит сочетанием русских (д) и (ж)
4	Ç ç	[tʃ]	че	
5	D d	[d]	де	
6	E e	[e]	э, е	Звучит как /æ/ перед ı, **m**, **n**, **r** как /э/ в конце слово и как /е/ в других случаях
7	F f	[f]	фе	
8	G g	[g, ɟ]	ге	
9	Ğ ğ	[ɣ]	ғе	Удлиняет предыдущую согласную, придавая ей горловое звучание
10	H h	[h]	хе	
11	I ı	[ɤ]	ы	Более задний, чем русский (ы)
12	İ i	[I]	и	
13	J j	[ʒ]	жэ	Встречается в заимствованных словах
14	K k	[k, c]	ке	В конце слога или слова смягчается
15	L l	[l, ɫ]	ле	
16	M m	[m]	ме	
17	N n	[n]	не	
18	O o	[o]	о	
19	Ö ö	[ɵ]	ө	Звучит, как слияние русских (о) и (ё)
20	P p	[p]	пе	Характерно придыхание в начале слова
21	R r	[ɾ]	ре	Звучит в конце слова как (ж/ш)
22	S s	[s]	се	
23	Ş ş	[ʃ]	ше	
24	T t	[t]	те	Придыхание в начале слова
25	U u	[u]	у	
26	Ü ü	[y]	ү	Звучит, как слияние русских (у) и (ю)
27	V v	[v]	ве	
28	Y y	[j]	йе	
29	Z z	[z]	зе	

НИИ Искусственный интеллект

Таблица 3.3.2. Туркменский алфавит.

№	Латиница	МФА	Кириллица	Название на кириллице
1	A a	[a]	Аа	а
2	B b	[b]	Бб	бэ
3	Çç	[ʧ]	Чч	че
4	D d	[d]	Дд	де
5	E e	[e]	Ее	е
6	Ä ä	[æ]	Әә	ае
7	F f	[ɸ]	Фф	фе
8	G g	[g~ʁ]	Гг	ге
9	H h	[h~x]	Хх	хе
10	I i	[i]	Ии	и
11	J j	[dʒ]	Жж	дже
12	Ž ž	[ʒ]	Жж	же
13	K k	[k~q]	Кк	ka
14	L l	[l]	Лл	эль
15	M m	[m]	Мм	эм
16	N n	[n]	Нн	эн
17	Ň ň	[ň]	Ңң	энг
18	O o	[o]	Оо	о
19	Ö ö	[ɵ]	Өө	ое
20	P p	[p]	Пп	ре
21	R r	[r]	Рр	эг
22	S s	[θ]	Сс	эs
23	Ş ş	[ş]	Шш	ше
24	T t	[t]	Тт	te
25	U u	[u]	Уу	у
26	Ü ü	[y]	Үү	уе
27	W w	[β]	Вв	ве
28	*Yy*	[ɤ]	Ыы	ы
29	Ý ý	[j]	Йй	йе
30	Z z	[ð]	Зз	зе

НИИ искусственный интеллект

Таблица 3.3.3. Узбекский алфавит.

№	Латиница	МФА	Кириллица	Название на кириллице
1	A a	[a, æ]	Аа	а
2	B b	[b]	Бб	бе
3	D d	[d]	Дд	де
4	E e	[e]	Ее	е
5	F f	[f]	Фф	фе
6	G g	[g]	Гг	ге
7	H h	[h]	Ҳҳ	хе
8	I i	[I]	Ии	і
9	J j	[ʒ]	Жж	дже
10	K k	[k]	Кк	ке
11	L l	[l]	Лл	ле
12	M m	[m]	Мм	мә
13	N n	[n]	Нн	нә
14	O o	[ɑ]	Оо	о
15	P p	[p]	Пп	пе
16	Q q	[q]	Ққ	қы
17	R r	[r]	Рр	ры
18	S s	[s]	Сс	сы
19	T t	[t]	Тт	ты
20	U u	[u]	Уу	у
21	V v	[v, w]	Вв	ве
22	**X x**	**[x]**	**Хх**	**хе**
23	Y y	[j]	Йй	й
24	Z z	[z]	Зз	зе
25	O′ o′	[ɔ]	Ўў	ө
26	G′ g′	[ɣ]	Ғғ	ғы
27	Ng ng	[ŋ]	Нг нг	ңы
28	Sh sh	[ʃ]	Шш	ше
29	Ch ch	[ʧ]	Чч	че
30	′	[ʔ]	Ъ ъ	Апостроф

НИИ Искусственный интеллект

Таблица 3.3.4. Азербайджанский алфавит.

№	Латиница	МФА	Кириллица	Название на кириллице
1	A a	[ɑ:]	Аа	А
2	B b	[b]	Бб	Бе
3	C c	[dʒ]	Сс	Се
4	Ç ç	[tʃ]	Чч	Че
5	D d	[d]	Дд	Ле
6	E e	[ɛ]	Ее	Е
7	Ə ə	[æ]	Әә	Ә
8	F f	[f]	Фф	Фе
9	G g	[g’]	Кк	Ге
10	Ğ ğ	[ɣ]	Ғғ	Ғе
11	H h	[h]	Һh	Ха
12	**X x**	**[x]**	**Хх**	**Хе**
13	I ı	[ɤ]	Ыы	Ы
14	İ i	[I]	Ии	И
15	J j	[Ʒ]	Жж	Дже
16	K k	[k]	Кк	Ке
17	Q q	[g]	Гг	Ку
18	L l	[l]	Лл	Эль
19	M m	[m]	Мм	Эм
20	N n	[n]	Нн	Эн
21	O o	[ɔ]	Оо	О
22	Ö ö	[œ]	Өө	Ө
23	P p	[p]	Пп	Пе
24	R r	[r]	Рр	Ре
25	S s	[s]	Сс	Се
26	Ş ş	[ʃ]	Шш	Ше
27	T t	[t]	Тт	Те
28	U u	[u]	Уу	У
29	Ü ü	[y]	Үү	Ү
30	V v	[v]	Вв	Ве
31	Y y	[j]	*Jj*	Йе
32	Z z	[z]	Зз	Зе

НИИ Государственный интеллект

3.1.3. Критерий и задачи перевода казахской письменности на латинский алфавит

Прежде чем осуществить перевод казахской письменности на новый алфавит, следует обратить внимание на следующие важные тенденции и обстоятельства:

1. В эпоху цифровизации (информатизации) общества во всем мире с помощью информационных коммуникационных технологий (ИКТ) создаются и обрабатываются цифровые (информационные, образовательные, нормативные и другие) ресурсы массового потребления очень большого объема - Big Data, которые требуют прямо пропорциональных по объему интеллектуальных, трудовых и финансовых затрат.
2. Все производимые в мире компьютеры изначально поддерживают и будут поддерживать 26-ти буквенный базовый латинский алфавит, при этом адаптация этих компьютеров к национальным алфавитам, имеющим буквы, не входящие в базовый латинский алфавит, требует разработки и применения специфических шрифтов, драйверов, программ сортировки и поиска информации, которые ориентированы на эти алфавиты и требуют немалых дополнительных затрат.
3. С учетом тенденции быстрого роста объема цифровых ресурсов и расширения сферы применения ИКТ мировой общественностью предлагается использовать в глобальном информационном пространстве единый алфавит, совпадающий с базовым латинским алфавитом, в целях упрощения и снижения затрат для создания и обработки этих ресурсов, повсеместно создаются стандарты для транслитерации букв национальных алфавитов на буквы базового латинского алфавита [36-38].
4. Государственный статус казахского языка обязывает к тому, чтобы в нашей стране цифровые ресурсы создавались и обрабатывались на казахском языке.
5. Существующие информционные ресурсы на казахском языке, находящиеся в различных фондах на бумажных носителях в алфавите на основе арабицы, латиницы-жаналип и кириллицы, должны оцифровываться после их отбора специальной комиссией.
6. Существующие цифровые информационные ресурсы в алфавите на основе арабицы, латиницы-жаналип и кириллицы

должны конвертироваться на базовый латинский алфавит.

7. Перевод казахской письменности на базовый латинский алфавит значительно упростит изучение и улучшит развитие казахского языка, так как этот алфавит знают все грамотные люди во всем мире, начиная со школьных лет.

Учитывая опыт адаптации базового латинского к другим языкам и перспективы создания и развития информационных ресурсов и технологий на казахском языке, предлагается следующий **критерий перевода казахской письменности на новый алфавит** [44]:

1. Новый алфавит должен создаваться только на основе звуковой системы казахского языка, которая содержит **31** фонему, из них **9** гласных: (а) - [ɑ], (ә) - [æ], (е) - [e], (о) - [ɔ], (ө) - [ɵ], (ұ) - [ʊ, u], (ү) - [ʏ], (ы) - [ɤ], (і) - [ɪ, i] и **22** согласные: (бы) - [b], (вы) - [v], (ғы) - [ɣ], (гі) - [g], (ды) - [d], (кі) - [k], (қы) - [q], (жы) - [ʒ], (зы) - [z], (й) - [j], (ыл) - [l], (мы) - [m], (ны) - [n], (ың) - [ŋ], (пы) - [p], (ыр) - [r], (сы) - [s], (ты) - [t], (ў) - [w], (фы) - [f], (хы) - [h], (шы) - [ʃ].

Это позволит построить строгие орфоэпические и орфографические нормы в соответствии с внутренними законами казахского языка. В новом алфавите будут отсутствовать буквы, обозначающие звуки русского языка, которые противоречат фонетике казахского языка.

2. Новый алфавит должен образовываться только из букв базового латинского алфавита.

Это позволит людям облегчить письменное общение с друг другом, создание и обрабатку цифровых информационных ресурсов на казахском языке с помощью электронных устройств (компьютеров, телефонных аппаратов и др.) без разработки и установки дополнительных шрифтов и драйверов, а также программ для сортировки и поиска данных, так как они для базового латинского алфавита будут автоматически встроены во всех производимых в мире электронных устройствах.

НИИ искусственный интеллект

3. Новый алфавит может менять звуковое значение некоторых букв базового латинского алфавита для обозначения звуков казахского языка.

Это позволит в новом алфавите для некоторых букв определить новые звуковые значения из звуковой системы казахского языка. Но при их подборе необходимо учитывать близость звучания казахского звука со звучанием звука-оригинала, обозначение которого включено в базовый латинский алфавит. Наличие близости облегчит подготовку орфоэпических и орфографических норм казахского языка на основе нового алфавита, ускорит его изучение и использование на практике.

4. Новый алфавит при отсутствии адекватного однобуквенного обозначения отдельных казахских звуков должен позволить использовать диграфы для их обозначения.

Это потребует, чтобы в качестве диграфов были использованы цепочки, во-первых, из двух букв, не встречающихся рядом в казахских простых словах, или, во-вторых, из буквы и специального знака, не включенного в алфавит языков программирования. Так как, во-первых, если буквы, включенные в конкретный диграф, будут встречаться рядом в казахском простом слове, то звук, который обозначен этим диграфом, не будет распознаваться, во-вторых, если специальный знак, используемый в диграфе включен в алфавит языка программирования, с помощью которого обрабатывается казахский текст, то звук, в обозначении которого участвует этот специальный знак, не будет распознаваться (см. 3.3, выводы п.2, узбекский язык).

5. Новый алфавит должен иметь порядок взаиморасположения букв, полностью совпадающий с порядком взаиморасположения букв в базовом латинском алфавите.

Это позволит при создании или применении информационных технологий и ресурсов на казахском языке без дополнительных затрат использовать программы для сортировки и поиска данных, так как для базового латинского алфавита программы для сортировки и поиска данных автоматически встроены в состав всех

НИИ Искусственный интеллект

системных и прикладных программ. Если новый алфавит имеет порядок взаиморасположения букв, не совпадающий с порядком взаиморасположения букв в базовом латинском алфавите, то необходимо заново разработать и установить программы для сортировки и поиска данных, что требует немалых затрат.

Теперь учитывая, что в базовом латинском алфавите имеются **26** букв, а в казахском языке - **31** фонема, для определения нового алфавита казахского языка мы должны решить следующие задачи:

1) определить звуки, обозначаемые только одной буквой;
2) определить звуки, обозначаемые диграфами;
3) определить значения диграфов;
4) обновить орфоэпические и орфографические нормы, а также морфологические и синтаксические правила казахского языка;
5) разработать технологии для их компьютерной обработки и подготовить интеллектуальные сетевые многоуровневые учебные пособия и учебники с речевыми интерфейсами, позволяющие обучаться казахскому языку через любые мобильные устройства.

Решения этих задач будут основываться на:

– результатах исследования текстового корпуса казахского языка, содержащего 100 миллионов букв действующего алфавита на основе кириллицы [45], с целью определения частоты употребления букв и сочетания букв в казахском тексте (см. Таблица 3.3.1);

– математической модели и геометрической модели системы гласных казахских звуков (см. параграф 2.1.2, формула 2.2, рисунок 2.1.2.3 и рисунок 2.1.2.4), отражающие их классификацию;

– осцилограммы гласных звуков (см. параграф 2.1.2, рисунок 2.1.2.2), которая позволила установить между мягкими и твердыми гласными звуками казахского языка следующие отношения:

(*ə*) = (*a*)+(*e*) или [æ] = [ɑ] +[e]
(*ө*) = (*i*)+(*e*) или [ө] = [ɔ] +[e]
(*ɣ*) = (*ɣ*)+(*e*) или [ʏ] = [ʊ] +[e]
(*i*) = (*ы*)+(*e*) или [ɪ] = [ɤ] +[e]

Эти отношения показывают, что для обозначения мягких гласных можно использовать диграфы. Такой способ используется

в других языках. Например, в английском языке звук (***ш***) обозначается диграфом ***sh***, полученного с помощью сочетания букв ***s*** и ***h***, а в узбекском языке звук (***ə***) представлен диграфом ***a'***.

Результаты исследования текстового корпуса казахского языка были использованы для определения того, какой звук будет обозначаться одной буквой и какой звук будет представлен диграфом, чтобы при переводе казахской письменности на латиницу не сильно удлинять запись казахских слов. Например, буквы в действующем алфавите на основе кириллицы, обозначающие звуки (***ə***) - [æ], (***ө***) - [ɵ] и (***ұ***) **-** [ʊ] имеют низкие частоты употребления, так как они встречаются только в корневых словах и не применяются в суффиксах и окончаниях, т.е. не будут появляться в производных словах и словоформах. Так, если в новом алфавите обозначим эти звуки диграфами, то запись слова с их участием не будет сильно удлиняться. Однако частота употребления буквы, обозначающей звука (***і***) - [ɪ] является самой высокой, кроме того, этот звук входит в состав суффиксов и окончаний и может встречаться несколько раз в одном слове (например, в слове '*біліктіліктің*' встречается 5 раз). Поэтому, несмотря на то, что этот звук может быть получен с помощью выражения (***ы***)+(***е***) или [ɤ]+[e], в новом алфавите его следует обозначать одной латинской буквой ***Ii***, чтобы длина записи казахского слова была короче.

Результаты исследования этого казахского текстового корпуса также помогли разработать правила избавления казахского текста на действующем на основе кириллицы алфавите от букв Ёё, Ии, Ээ, Юю, Яя, Цц, Чч, Щщ, Ьь, Ъъ, звуковые значения которых не будут входить в звуковую систему казахского языка. На основании этих правил был создан переходной орфографический словарь казахского языка на кириллице без этих букв [228, 229].

Этот словарь будет служить исходной базой данных для автоматического перевода казахской письменности на латиницу, который будет состоять из двух этапов: первый этап - перевод исходного текста на промежуточный текст без букв Ёё, Ии, Ээ,

НИИ ИСУ ЛКТ

Юю, Яя, Цц, Чч, Щщ, Ьь, Ъъ; второй этап – перевод промежуточного текста на латиницу. Реализуя эти этапы, мы разработали и внедрили конвертор для автоматического перевода любого казахского текста с кириллицы на латиницу: http//alphabet.kz/convertor.

Таблица 3.3.1. Частотность букв в корпусе казахского языка на кириллице, %

№	Буква	Литература	Публицистика	Официал	Наука	Разговор	Всего
1	**А**	12,85	12,89	12,18	11,74	12,65	12,52
2	**Ә**	0,71	0,81	0,98	0,87	0,73	0,88
3	**Б**	2,87	2,42	2,14	2,21	2,69	2,31
4	**В**	0,15	0,29	0,25	0,33	0,22	0,26
5	**Г**	1,13	1,16	1,24	1,29	1,19	1,20
6	**Ғ**	1,74	1,64	1,42	1,51	1,59	1,54
7	**Д**	4,59	4,45	4,04	4,50	4,60	4,27
8	**Е**	8,00	7,45	8,10	7,98	8,11	7,82
9	**Ё**	0,00	0,00	0,00	0,00	0,00	0,00
10	**Ж**	1,84	1,67	1,94	1,57	1,64	1,81
11	**З**	1,58	1,62	1,32	1,42	1,74	1,47
12	**И**	0,79	1,55	1,44	1,92	1,25	1,45
13	**Й**	2,28	1,41	0,98	1,33	1,96	1,27
14	**К**	2,88	2,72	3,23	2,95	2,90	2,98
15	**Қ**	3,52	3,33	3,36	3,24	3,13	3,35
16	**Л**	4,52	5,16	5,41	5,13	4,62	5,23
17	**М**	2,90	3,22	3,69	3,19	3,39	3,42
18	**Н**	6,90	6,85	6,57	6,75	6,84	6,72
19	**Ң**	1,68	1,58	1,50	1,53	1,56	1,55
20	**О**	2,58	2,56	2,30	2,82	2,72	2,44
21	**Ө**	1,11	0,89	0,86	0,83	0,96	0,89
22	**П**	2,42	1,52	1,12	1,48	2,02	1,40
23	**Р**	5,10	5,54	5,78	5,77	5,35	5,62
24	**С**	4,03	4,29	4,31	3,99	4,14	4,27
25	**Т**	5,19	6,01	6,52	6,31	5,54	6,19
26	**У**	1,04	0,9	0,73	0,83	0,86	0,83
27	**Ұ**	1,18	1,71	2,78	1,97	1,42	2,18
28	**Ү**	1,08	0,73	0,59	0,71	0,89	0,69
29	**Ф**	0,04	0,12	0,13	0,20	0,17	0,12
30	**Х**	0,18	0,28	0,18	0,25	0,20	0,23
31	**Һ**	0,02	0,02	0,00	0,01	0,01	0,01
32	**Ц**	0,03	0,12	0,19	0,23	0,08	0,15
33	**Ч**	0,04	0,03	0,03	0,05	0,05	0,03
34	**Ш**	1,44	1,23	1,24	1,14	1,35	1,25
35	**Щ**	0,01	0,00	0,00	0,01	0,01	0,01
36	**Ъ**	0,01	0,01	0,03	0,02	0,00	0,02
37	**Ы**	7,36	7,82	7,89	7,65	7,23	7,81

НИИ ИЯ ЕНУ им. Л.Н. Гумилева

38	**І**	5,84	5,41	4,90	5,41	5,64	5,20
39	**Ь**	0,06	0,07	0,07	0,11	0,11	0,07
40	**Э**	0,02	0,08	0,08	0,18	0,06	0,08
41	**Ю**	0,03	0,04	0,14	0,06	0,05	0,09
42	**Я**	0,26	0,40	0,34	0,51	0,33	0,37

НИИ Искусственный интеллект

3.1.4. Правила избавления казахского текста от инородных букв

Осуществление перевода письменности любого языка с одной графики на другую требует комплекса работ, связанных с уточнением его звуковой системы, с изменением его орфоэпических, орфографических, морфологических и синтаксических правил, с разработкой компьютерных программ прямого и обратного преобразования текстов в исходном алфавите на тексты в новом алфавите и наоборот, а также с созданием соответствующих словарей и подготовки многоуровневых учебных изданий [181-222].

Во введении сказано, что в реформе по переводу казахской письменности на латиницу будет участвовать **31** фонема, из них **9** гласных фонем, **22** согласные фонемы. В действующем на основе кириллицы алфавите имеются **42** буквы, из которых **10** букв (Ёё, Ии, Ээ, Юю, Яя, Цц, Чч, Щщ. Ьь, Ъъ) используются для обозначения звуков русского языка и, в основном, встречаются в заимствованных словах. Правописание таких слов противоречит грамматическим нормам казахского языка. Так, например, в русском языке мягкость или твердость произношения звука определяется наличием в нем знака Ьь или Ъъ, который ставится после согласной буквы. В казахском языке мягкость и твердость произношения слова зависит от наличия в нем мягких гласных Әә, Ее, Өө, Үү, Іі или твердых гласных Аа, Оо, Ұұ, Ыы и действует закон сингармонизма (к мягкому слогу добавляются мягкие аффиксы, а к твердому - твердые).

Поэтому, чтобы автоматизировать преобразования казахского текста в действующем алфавите на основе кириллицы в текст в новом алфавите на основе латиницы, необходимо построить следующие группы правил обработки казахских слов [229]:

– правила избавления в словах от букв Ёё, Ии, Ээ, Юю, Яя, Цц, Чч, Щщ, Ьь, Ъъ;

– правила обработки буквы Уу в словах;

НИИ искусственный интеллект

– правила обработки букв Аа, Әә, Оо, Өө, Ұұ, Үү, Ы,ы, Іі в словах.

I. Правила исключения буквы Ьь

I.1. Если буква **Ьь** встречается непосредственно перед буквой Ее, то она заменяется на цепочку «ій» (например, Вьетнам = Війетнам, премьер = премійер, фельетон = фелійетон).

I.2. Если буква **Ьь** встречается непостредственно перед буквой Юю, то цепочка «ью» заменяется на «үй» (например, компьютер = компүйтер, тьютор = түйтор).

I.3. Если буква **Ьь** встречается между двумя согласными, то она удаляется (например, альфа = алфа, сальдо = салдо, тальго = талго, Харьков = Харков).

I.4. Если буква **Ьь** встречается в конце слова, то она удаляется (например, бронь = брон, модель = модел, модуль = модул, Тверь = Твер, Мальта = Малта, Польша = Полша).

II. Правило исключения буквы Ъъ.

II.1. Буква **Ъъ** заменяется на цепочку «ый» везде, где она встречается (например, съезд=сыйезд, подъезд=подыйезд).

III. Правила исключения буквы Ёё.

III.1. Если буква **Ёё** встречается в начале слово, то она заменяется на цепочку «Йо йо» (например, ёлка = йолка, ёрник =йорник).

III.2. Если буква **Ёё** встречается после буквы согласного звука, то она заменяется на букву Өө (например, актёр = актөр, монтёр = монтөр, сапёр = сапөр, шахтёр = шахтөр, шофёр = шофөр).

IV. Правило исключения букв Ээ.

IV.1. Буква **Ээ** заменяется на букву Ее везде, где она встречается (например, эвенк = евенк, экскаватор = екскаватор, элеватор = елеватор, экология = екология, экономика = економика, экспорт = експорт, электр = електр, этюд = етюд, этаж = етаж, эскимос = ескимос, эскалатор = ескалатор, эфир = ефир).

НИИ искусственный интеллект

V. Правила исключения буквы Юю.

V.1.Если буква **Юю** встречается в начале слова перед твердым слогом, то она заменяется на цепочку «йұ» (например, юань = йұань, Югославия = Йұгославия, Юра = Йұра, Юрмала = Йұрмала, юкагир = йұкагир, юла = йұла, юмор = йұмор, юнга = йұнга, юскас = йұскас).

V.2.Если буква **Юю** встречается после буквы твердого гласного, то она заменяется на цепочку «Йұ йұ» (например, аю = айұ, ою= ойұ, ұю = ұйұ, баю = байұ, қою= қойұ, сою= сойұ, құю = құйұ, қыю = қыйұу, сыю = сыйұ, тыю = тыйұ).

V.3.Если буква **Юю** встречается в начале слова перед мягким слогом, то она заменяется на цепочку «йү» (например, юнкер = йүнкер, Юпитер = Йүпитер, юриспруденция = йүриспруденция).

V.4.Если буква **Юю** встречается после мягкого гласного, то она заменяется на цепочку «йү» (например, кею = кейү, есею = есейү, ию = ийү, күю = күйү, сүю = сүйү, тию = тийү, түю = түйү, үлкею = ұлкейү, үю = үйү).

V.5.Если буква **Юю** встречается в закрытом слоге между двумя согласными, то она заменяется на букву «ү» (например, бюллетень = бүллетень, карбюратор = карбүратор, люфт = лүфт, пюпитр = пүпитер, увертюра = увертүра, Цюрих = Цүрих).

VI. Правила исключения буквы Яя.

VI.1. Если буква **Яя** встречается в начале слова перед согласной буквой в твердом слоге, то она заменяется на цепочку «йа» (например, Ява = Йава, ядро=йадро, Ямал = Йамал, япырай = йапырай, яранга = йаранга, ясақ = йасақ, яшма = йашма).

VI.2. Если буква **Яя** встречается в слове после твердого гласного или твердого слога, то она заменяется на цепочку «йа» (например, ая = айа, саян = сайан, оян = ойан, қоян = қойан, туя = туйа, тұяқ = тұйақ).

VI.3. Если буква **Яя** встречается в начале слова перед согласной

НИИ Искусственный интеллект

буквой в мягком слоге, то она заменяется на цепочку «йә» (например, ягер = йәгер, яки = йәки, якорь = йакорь, Ямайка = Йамайка, яранга = йаранга, Ясави = Йасави, ясли = йәсли).

VI.4. Если буква **Яя** встречается в слове после мягкого гласного, то она заменяется на цепочку «йә» (например, фея = фейә, секция = секцийә, сессия = сессийә).

VI.5. Если буква **Яя** встречается в закрытом слоге между двумя согласными, то она заменяется на букву «ә» (например, вентилятор = вентиләтор, отряд = отрәд, снаряд = снарәд, аккумулятор = аккумуләтор, ассимилятор = ассимиләтор, регулятор = регуләтор).

VII. Правила исключения буквы Ии.

VII.1. Если буква **Ии** встречается в начале заимствованного слова (термина) перед буквами согласных Бб, Вв, Ғғ, Гг, Дд, Жж, Зз, Кк, Ққ, Лл, Мм, Нн, Ңң, Пп, Рр, *Сс*, Тт, Фф, Хх, Шш, то она заменяется на букву Іі (например, иберия = іберия, иврит = іврiт, игумен = ігумен, идальго = ідалго, идеа = ідеа, идиома = ідиеома, идиот = ідиот, известь = ізвесть, икон = ікон, иллюстратор = іллюстратор, имитатор = імитатор, инвариант = інвариант, инфекция = інфекция, институт = інститут, ипотека = іпотека, иприт = іприт, Иран =Іран, испан = іспан, итерация = ітерация, ихтизавр = іхтизавр, ихтиолог = іхтиолог, ишемия = ішемия).

VII.2. Если буква **Ии** встречается в заимствованных словах в закрытом слоге после букв согласных, кроме букв Ғғ и Ққ, то она заменяется на букву «і» (например, бит = біт, вирус = вірус, гид = гід, диктор= діктор, гир = гір, Алжир = Алжір, инжинер = инжінер, кризис = крізіс, кит = кіт, лира = ліра, Памир = Памір, теннис = тенніс, папирус = папірус, тарих = таріх, университет = університет, эликсир = эліксір, тир = тір, институт = інститут, тахир = тахір, эфир = эфір, кашир = кашір).

VII.3. Буква **Ии** заменяется на цепочку «ый» в словах «ми, мия, биязы, сия», т.е. ми = мый, мия = мыя, биязы = быйязы, сия = сыйа.

VII.4. Буква **Ии** заменяется на цепочку «ій» в следующих

НИИ Искусственный интеллект

однослоговых словах: би = бій, ки = кій, ли = лій, пи = пій, си = сій, фи = фій, хи = хій, ти = тій, ши = шій. Здесь у некоторых слов в казахском языке нет значения, они используются для связки.

VII.5. Если буква **Ии** встречается после буквы «Ғғ» или «Ққ», то она заменяется на цепочку «ый» (например, ғибрат = ғыйбрат, ғибадат = ғыйбадат, ғиззат = ғыйззат, ғимарат = ғыймарат, қима = қыйма, қисық = қыйсық, қиық = қыйық, қию = қыйю, қия = қыйя, қияр=қыйяр).

VII.6. Если буква **Ии** встречается в начале слова перед буквами гласных Аа, Әә, Ее, Оо, Өө, Уу, Ұұ, Үү, Ыы, Іі, то она заменяется на букву Йй (например, иайдо = йайдо, иә = йә, ие = йе, иегер = йегер, иек = йек, иезуит = йезуит, иодофермен = йодофермен, ион = йон, иордан = йордан, иудей= йудей, иудаизм = йудаизм, ыығы = йыығы, ыық = йық, иі = йі, иігу = йігу, иіл = йіл, июль = йюль).

VII.7. Если буква **Ии** встречается в слове перед мягким слогом, то она заменяется на цепочку «ій» (например, биік = бійік, игі = ійгі, ине = ійне, иін = ійін, жиде = жійде).

VII.8. Если буква **Ии** встречается в слове перед твердым слогом, то она заменяется на цепочку «ый» (например, бипаз = быйпаз, биыл = быйыл, сиыр = сыйыр, иық = ыйық, қитар = қыйтар).

VII.9. Если буква **Ии** встречается в последнем открытом слоге слова, следующего за мягким слогом слова, после букв Бб, Вв, Гг, Дд, Жж, Зз, Кк, Лл, Мм, Нн, Пп, Рр, *Сс*, Тт, Хх, Шш, то заменяется на цепочку «ій» (например, әліпби = әліпбій, әлди = әлдій, визави = визавій, яки = якій, пәни = пәній, ресми = ресмій).

VII.10. Если буква **Ии** встречается в последнем открытом слоге, следующего за твердым слогом слова, после букв Бб, Вв, Ғғ, Дд, Жж, Зз, Ққ , Лл, Мм, Нн, Пп, Рр, *Сс*, Тт, Хх, Шш, то заменяется на цепочку «ый» » (например құдағи = құдағый, ылғи = ылғый, ылди = ылдый, қияли = қыйалый, адами = адамый, ғалами = ғаламый, ғылыми = ғылымый, тарази = таразый, тарихи = тарихый, ылғи = ылғый, шұлғи = шұлғый, зымзия =зымзыйя, сұрқия = сұрқыйя).

НИИ интеллектуальной собственности

VIII. Правила обработки буквы Уу.

IX. В действующем казахском алфавите на основе кириллицы имеется буква **Уу**, которая обозначает гласный звук русского языка [u:]. В звуковой системе казахского языка аналогичного гласного нет, но имеется похожий согласный звук (ў) - [w], который в зависимости от твердости и мягкости контекста может произноситься (озвучиваться) как сочетание звуков (ұў) или (үў).

Следует отметить, что в казахской письменности в действующем алфавите на основе кириллицы имеются противоречия относительно использования буквы **Уу** (см. ВВЕДЕНИЕ, стр. 19-27). Для устранения этих противоречий в записи казахских слов, используя знак «#» в промежуточных заменах, предлагается использовать следующие правила замены:

VIII.1. Если слово состоит только из одной буквы **Уу**, то она заменяется на цепочку «ұ#» (например, у = ұ#)

VIII.2. Если буква **Уу** встречается в слове после буквы твердого или мягкого гласного, то она заменяется на знак «#» (например, ау = а#, әу = ә#, еу = е#, оу = о#, аулау = а#ла#, әуелеу = ә#еле#, баулау = ба#ла#, бұғаулау = бұға#ла#, елеулеу = еле#ле#, жалаулау = жала#ла#, лаулау = ла#ла#, кәуе = кә#ә, қараулау = қара#ла#, сабаулау = саба#ла#, сауалдау = са#алда#, сауыт = са#ыт, сәулелеу = сә#леле#, таулау = та#ла#, шоу = шо#).

VIII.3. Если буква **Уу** встречается в начале заимствованного слова, то она заменяется на букву «ұ» (например, Уганда = Ұганда, университет = ұниверситет, универсам = ұниверсам, унтер = ұнтер, уран = ұран, уролог = ұролог, Уругвай = Ұругвай, утелита = ұтелита, утопия = ұтопия, Уфа = Ұфа).

VIII.4. Если буква **Уу** встречается в слове перед буквой твердого гласного, то она заменяется на цепочку «ұ#» (например, уа = ұ#а, уайым = ұ#айым, уақ = ұ#ақ, уыз = ұ#ыз, уық = ұ#ық, уыт = ұ#ыт, буаз = бұ#аз, буын = бұ#ын, дуа = дұ#а, қуат = құ#ат, қуыршақ = құ#ыршақ, суат = сұ#ат, суық = сұуық, уыс = ұуыс, туар

НИИ Искусственный интеллект

= тұуар, туыс = тұ#ыс, туынды = тұ#ынды, шуақ = шұуақ, шуыл = шұуыл).

VIII.5. Если буква **Уу** встречается в слове перед твердым слогом, то она заменяется на букву «ұ» (например, вулкан = вұлқан, убалдырған = ұбалдырған, удмурт = ұдмурт, удай = ұдай, уқалау = ұқалау, улану = ұлану, улы = ұлы, умажда = ұмажда, усасыр = ұсасыр, ушығу = ұшығу, тулпар = тұлпар, туған = тұған, футбол = фұтбол, формула = формұла).

VIII.6. Если буква **Уу** встречается в слове перед мягким слогом, то она заменяется на букву «ү» (например, угия = үгия, укекіре = үкекіре, улеу = үлеу, букет = бүкет, сурет = сүрет, купе = күпе, кетчуп = кетчүп).

VIII.7. Если буква **Уу** встречается в конце слова после согласной буквы твердого слога, то она заменяется на цепочку «ұ#» (например, абу = абұ#, ағу = ағұ#, азу = азұ#, алу = алұ#, аққу = аққұ#, алжу = алжұ#, асу = асұ#, беру = берү#, бұйырту = бұйыртұ#, жабу = жабұ#, жору = жорұ#, даму = дамұ#, қабысу = қабысұ#, қайыру = қайырұ#, қону = қонұ#, ру = рұ#, су = сұ#, табысу = табысұ#, тасу = тасұ#, тарту = тартұ#, ту = тұ#, шашу = шашұ#, шолу = шолұ#).

VIII.8. Если буква **Уу** встречается в конце слова после согласной буквы мягкого слога, то она заменяется на цепочку «ү#» (например, елу = елү#, кебу = кебү#, келісу = келісү#, кешігу = кешігү#, езу = езү#, ілу = ілү#, кему = кемү#, сену = сенү#, кіру = кірү#, кірісу = кірісү#, кету = кетү#, сену = сенү#, шешу = шешү#).

VIII.9. В промежуточном тексте символ «#» заменяется на **Уу** везде, где он встречается.

X. Правила исключения букв Цц.

IX.1. Если буква **Цц** встречается в начале слова или после буквы согласного звука, то она заменяется на «с» (например, акция = аксия, абстракция = абстраксия, цирк = сирк, цемент = семент, цифр = сифр, дирекция = дирексия, индукция = индуксия, концепция = концепсия, коференция = конференсия, реакция =

НИИ Искусственный интеллект

реаксия, редукция = редуксия, резиденция = резиденсия, селекция = селексия, транскрипция = транскрипсия, функция = функсия).

IX.2. Если буква **Цц** встречается после буквы гласного звука, то она заменяется на «тс» (например, артикуляция = артикулятсия, ассимиляция = ассимилятсия, бифуркация = бифуркатсия, девальвация = девальватсия, декларация = декларатсия, делегация = делегатсия, делимитация = делемитатсия, дефиниция = дефинитсия, диссертация = диссертатсия, инновация = инноватсия, интуиция = интуитсия, капитуляция = капитулятсия, коалиция = коалитсия, композиция = композитсия, конституция = конститутсия, манипуляция = манипулятсия, рация = ратсия, ротация = ротатсия, революция = револютсия, трансляция = транслятсия).

XI. Правила исключения букв Чч.

X.1. Буква **Чч** заменяется на букву Шш везде, где она встречается (например, чай = шай, чад = шад, чартер = шартер, чип = шип, чемпион = шемпион).

XII.Правила исключения букв Щщ.

XI.1. Если буква **Щщ** встречается между двумя буквами гласных звуков, то она заменяется на цепочку «шш» (например, ащы = ашшы, тұщы = тұшшы, Клещев = Клешшев), иначе она заменяется на букву «ш» (например, Шедрин = Шедрин, Щелоков = Шелоков, Щербаков = Шербаков, плащ = плаш).

XIII. Правило очищения от лишних букв Йй.

Если в тексте от предыдущих замен осталась цепочка «йй» , то из неё удаляется одна буква «й» (например, гвардия = гвардыйййа = гвардыйа, Испания = Ыйспаныйййа = Ыйспаныйа, комиссия = комыйссыйййа = комыйссыйа, Ливия = Лійвійййə = Лійвійə, морфология = морфологыйййа = морфологыйа, орфография = орфографыйййа = орфографыйа, Сирия = Сійрійййə = Сійрійə, сессия = сессійййə = сессійə, Сицилия = Сиційлійййə = Сиційлійə, экспансия = экспансыйййа = экспансыйа).

НИИ искусственный интеллект

XIV. Правила обработки букв Нн и Ңң

XIII.1. Если в записи слова после буквы сонорного носового согласного **Нн** встречается буква **Гг**, то в соответствии с нормой произношения в казахском языке вместо бувы Нн запишется буква **Ңң (**например, күнгі = күңгі, түнгі = түңгі, жүнге = жүңге, сенген = сеңген, тенге = теңге).

НИИ Искусственный интеллект

3.1.5. Правила перевода казахской письменности на латиницу

Для представления 31 казахской фонемы в 26 буквенном базовом латинском алфавите необходимо применять некоторые методы его адаптации к другим языкам (см. 3.1.1).

Для обозначения казахских фонем буквами базового латинского алфавита используется схожесть их обозначения в МФА [31, 32] и предлагается 3 группы правил:

1. Один звук и одна буква:

1.1. **5** казахских гласных фонем (а) - [ɑ], (е) - [e], (о) - (o), (ұ) - [ʊ] и (ы) - [ɤ] будут обозначаться **5** латинскими буквами *Aa*, *Ee*, *Oo*, *Uu* и *Yy*, значениями которых являются созвучные с ними звуки, имеющие в МФА одинаковые обозначения с обозначениями указанных казахских фонем, соответственно;

1.2. **18** казахских согласных фонем (бы) - [b], (вы) - [v], (гі) - [g], (ды) - [d], (фы) - [f], (й) - [j], (кі) - [k], (қы) - [q], (ыл) - [l], (мы) - [m], (ны) - [n], (пы) - [p], (ыр) - [r], (сы) - [*s*], (ты) - [t], (ў) - [w], (хы) - [x] и (зы) - [z] будут обозначаться **18** латинскими буквами *Bb*, *Vv, Gg*, *Dd*, *Ff*, *Jj, Kk*, *Qq, Ll*, *Mm*, *Nn*, *P*, *Rr*, *Ss*, *Tt*, *Ww, Xx* и *Zz*, значениями которых являются созвучные с ними звуки, имеющие в МФА одинаковые обозначения с обозначениями указанных казахских фонем, соответственно;

1.3. Согласная фонема (цы) - [tc], заимствованная из русского языка, будет обозначаться латинской буквой *Cc.* Согласный звук (цы) - [tc] является глухим и твердым. При артикуляции вначале спинка языка круто выгнута и касается бугорков за верхними зубами, кончик языка упирается в нижние зубы, затем спинка языка опускается до положения, занимаемого при произнесении созвучного звука (сы) - [*s*], а кончик языка остается на месте, голосовые связки отдыхают и горло не дрожит (нет голоса).

НИИ искусственный интеллект

2. Один звук и диграф:

2.1. Казахские **3** гласные мягкие фонемы (*ә*) - [æ], (*ө*) - [ɵ], (ү) - [ʏ] будут обозначаться диграфами, в которых первыми парами являются латинские буквы *Aa*, *Oo* и *Uu*, обозначающие близкие по звучанию твердые гласные (а) - [ɑ], (о) – [о] и (ұ) - [ʊ], соответственно, а второй парой каждого диграфа будет специальный символ ` – *аксант грав, который входит в состав алфавита* для первого варианта алфавита (как это принято в узбекском алфавите) или латинская буква *Ee* для второго варианта алфавита (как это принято в «компьютерном» немецком алфавите).

2.2. Казахские согласные фонемы (*ғы*) - [ɣ], (*шы*) - [ʃ], (*жы*) - [ʒ] и замствованная согласная фонема (*чы*) - [tʃ] будут обозначаться диграфами *Gh gh, Sh sh, Zh zh* и *Ch ch*, соответственно. Первыми парами диграфов является латинские буквы, которые были использованы для обозначения созвучных согласных фонем (*гі*) - [g], (*сы*) - [s], (*зы*) - [z] и (*цы*) - [tc], соответственно. Такие диграфы в новом алфавите не будут порождать орфографические противоречия в казахской письменности.

2.3. Казахская согласная сонорная фонема (ыңҗ) - [ŋ] будет обозначаться диграфом. Первой парой диграфа будет латинская буква *Nn*, обозначающая фонему (ны) - [n], которая по своему значению близка к фонеме (ыңҗ) - [ŋ]. Второй парой диграфа будет латинская буква *Gg,* обозначающая звук (гі) - [g] (как это принято в английском и узбекском языках). В записи в новом алфавите диграф «*ng*» должен содержаться только в одном слоге, так как в казахском слове никогда перед звуком (ыңҗ) - [ŋ] не встречается согласный звук и нет слова, которое начинается со звука (ыңҗ) - [ŋ].

Однако некоторые, разные по произношению и смыслу слова в предлагаемом алфавите будут писаться идентично. Например, разные слова «*күңі*» и «*күнгі*» в предлагаемом алфавите будут писаться идентично:

1-вариант: «*күңі = ku`ngi*» және «*күнгі = ku`ngi*».

2-вариант: «*күңі = kuengi*» және «*күнгі = kuengi*».

НИИ Искусственный интеллект

Для того, чтобы не допустить подобные случаи, необходимо принять следующее правило.

Если при записи в латинском алфавите казахского слова после буквы *Nn*, обозначающей согласный звук (ны) - [n], необходимо писать букву *Gg,* обозначающей согласный звук (гі) - [g] (такой случай возможен только в исконно казахских словах), то вместо буквы *Nn* надо писать диграф «*ng*», который обозначает согласный звук (ыңг) - [ŋ]. Это будет соответствовать закономерности произношения в казахском языке и в его письменности не порождает никакие противоречия (один из видов этого правила показан в параграфе 3.1.4, в правиле XIII для обработки казахских текстов на кириллице).

После применения этого правила вышеуказанный пример перепишется как:

1-вариант: «*кү**ңі** = ku`ngi*» және «*кү**н**гі = кү**ң**гі = ku`nggi*».

2-вариант: «*кү**ңі** = kuengi*» және «*кү**н**гі = кү**ң**гі =kuenggi*».

Замечание: Диграфы являются орфографическими правилами и не считаются элементами алфавита.

3. Одна буква и два звука.

В современном казахском языке имеется заимствованный из русского языка гласный звук (и) - [i], который образуется при закрытой челюсти, неокругленных губах, высоком положении языка и продвинутом вперед языке. Следовательно, звук (и) - [i] по положению челюсти является закрытым, по положению губ является неокругленным, по вертикальному положению языка является верхним, а по горизонтальному положению - переднеязычным гласным звуком. Артикуляционные признаки звука [и]: *закрытый, неокругленный, верхний, переднеязычный* [18, 19].

В действующем алфавите есть кириллическая буква Ии. Во-первых, она встречается в заимствованных словах и на практике носителями казахского языка произносится, как исконно казахский звук (і), например, *интернет = (інтернет)*, *инвестор = (інвестор)*,

НИИ Искусственный интеллект

институт = (інстітұт). Во-вторых, в исконно казахских словах во всех случаях произносится, как сочетание звуков (*ій*) или (*ый*) в зависимости от мягкости или твердости контекста, например, *киік* – (*кійік*), *қиық* – (*қыйык*), *тиін – (тійін), тиын – (тыйын)*, *ғылыми – (ғылымый), ресми – (ресмій), ылғи – (ылғый),* ылди = ылдый, қияли = қыйалый, адами = адамый, ғылыми = ғылымый, тарихи = тарихый.

Поэтому в первом случае русский гласный звук (и) нужно принять как аллофон казахской фонемы (і). Во втором случае просто привести в порядок орфографические правила, и тогда не будет никаких орфографических противоречий, указанных во Введении настоящей работы (см. стр. 24-25). Тогда латинская буква Іі будет обозначать как фонему (і), так и её аллофон (и), которые в МФА будут обозначаться символами [ɪ] и [i], соответственно, как это принято во многих алфавитах других языков (например, в английском языке: bit [bɪt] и bite [baɪt]), то есть гласный аллофон (и) не будет обозначаться отдельной графемой, он будет обозначаться той же латинской буквой *Ii*, которой обозначается казахская гласная фонема (і).

По аналогии со звуком (и), в современном казахском языке имеется заимствованный из русского языка гласный звук (у) - [u], который образуется при закрытой челюсти, округленных губах, высоком положении языка и отодвинутом назад языке. Следовательно, звук (у) - [u] по положению челюсти является закрытым, по положению губ является округленным, по вертикальному положению языка является верхним, а по горизонтальному положению - заднеязычным гласным звуком. Артикуляционные признаки звука (у): *закрытый, округленный, верхний, заднеязычный* [18, 19].

Также в действующем алфавите есть кириллическая буква Уу. Во-первых, она встречается в заимствованных словах и на практике носителями казахского языка произносится, как исконно казахский звук (ұ), например, *универсам* = (*ұниверсам*), *университет* = (*ұниверситет), формула = (формұла)*. Во-вторых, она в исконно казахских словах во всех случаях произносится как сочетание

НИИ Искусственный интеллект

звуков (*үў*) или (*ұў*) в зависимости от мягкости или твердости контекста, например, ағу = (ағұў), азу = (азұў), алу = (алұў), аққу = (аққұў), алжу = (алжұў), асу = (асұў), бұйырту = (бұйыртұў), жабу = (жабұў), жору = (жорұў), даму = (дамұў), қабысу = (қабысұў), қайыру = (қайырұў), қону = (қонұў), сурет = (сүўрет), келісу = (келісүў), кешігу = (кешігүў), езу = (езүў), кему = (кемүў), сену = (сенүў), кіру = (кірүў), кірісу = (кірісүў), кету = (кетүў), сену = (сенүў), шешу = (шешүў), ушығу = (ұўшығұў), тулпар = (тұлпар), туған = (тұған), табысу = (табысұў), тасу = (тасұў), тарту = (тартұў), ту = (тұў), шашу = (шашұў), шолу = (шолұў).

Поэтому в первом случае, русский гласный звук (у) нужно принять как аллофон казахской фонемы (ұ). Во втором случае, просто привести в порядок орфографические правила и тогда не будет никаких орфографических противоречий, указанных во Введении настоящей работы (см. стр. 24-25). Тогда латинская буква *Uu* будет обозначать как фонему (ұ), так и её аллофон (у), которые в МФА будут обозначаться символами [ʊ] и [u], соответственно, как это принято в алфавитах многих других языков (например, в английском языке: put [pʊt], but [bʌt] и tube [tu:b]), то есть гласный аллофон (у) не будет обозначаться отдельной графемой, он будет обозначаться той же латинской буквой *Uu*, которой обозначается казахская гласная фонема (ұ).

Вышеуказанные аллофоны (и) и (у) гласных казахских фонем (і) и (ұ) позволит написание многих заимствованных и подлежащих к заимствованию непереводимых на казахский язык международных терминов в новом алфавите приблизить к их оригинальному написанию, сохраняя казахское произношение, например*, internet= (інтернет), investor= (інвестор), institute = (інстітұт), universitet = (ұніверсітет), universam = (ұніверсам), formula = (формұла).*

Итак, с **26** буквами базового латинского алфавита удалось обозначить **31** фонему казахского языка, из них **26** фонем обозначены отдельными буквами, а **5** фонем обозначены диграфами. Для всех фонем определены обозначения в МФА, в том числе и для аллафонов:

НИИ Исследовательский институт

(i) - [ɪ, i] и (ұ) - [ʊ, u].

В результате предложены 2 варианта алфавита казахского языка на основе латинской графики с указанием соответствующих обозначений в МФА: в первом варианте алфавита содержатся только 26 букв базового латинского алфавита, а во втором варианте, помимо этих букв, содержится и знак ` – *аксант грав,* который будет использоваться, как вторая пара диграфа и не будет порождать проблему, как в узбекском алфавите, связанную со знаком ′ – *аксант эгю* (см. параграф 3.1.2).

Предложенный алфавит обладает следующими свойствами:

1) алфавит не порождает противоречий в казахской орфоэпии и орфографии, так как он основан на системе казахских звуков и на системе знаков в МФА;

2) алфавит при создании и обработке информационных ресурсов на казахском языке не нуждается в программных средствах, которые требуют дополнительные затраты таких, как:

– шрифтов и драйверов, так как он состоит только из букв английского алфавита, размещенного на клавиатуре всех типов компьютеров с многочисленными шрифтами и драйверами для всех типов устройств ввода и вывода данных;

– программ сортировок и поиска информации, так как порядок следования букв в нём совпадает с порядком следования букв в английском алфавите;

3) алфавит позволяет сблизить написание непереводимых на казахский язык международных заимствованных терминов - слов с их написанием в оригинале без изменения звуковой системы казахского языка;

4) алфавит позволяет легко и быстро набрать казахский текст с помощью клавиатуры любого компьютера без перехода на другие вкладки.

В Таблице 3.5.1 показан новый казахский алфавит на латинице, а в Таблице 3.5.2 описываются орфографические правила для диграфов, применяемых для обозначения некоторых звуков казахского языка.

НИИ Искусственный интеллект

НИИ Искусственный интеллект

Таблица 3.5.1. Казахский алфавит на латинице

№	Латиница	Название звука	МФА	Кириллица	Пояснение
1	A a	(а)	[ɑ]	А а	
2	B b	(бы)	[b]	Б б	
3	C c	(цы)	[tc]	Ц ц	Заимствованный звук
4	D d	(ды)	[d]	Д д	
5	E e	(е)	[e]	Е е	
6	F f	(фы)	[f]	Ф ф	Заимствованный звук
7	G g	(гі)	[g]	Г г	
8	H h				Вторая пара диграфа
9	I i	(і), (и)	[ɪ, i]	І і, Ии	(і) – фонема, (и) – аллофон
10	J j	(й)	[j]	Й й	
11	K k	(кі)	[k]	К к	
12	L l	(ыл)	[l]	Л л	
13	M m	(мы)	[m]	М м	
14	N n	(ны)	[n]	Н н	
15	O o	(о)	[ɔ]	О о	
16	P p	(пы)	[p]	П п	
17	Q q	(қы)	[q]	Қ қ	
18	R r	(ры)	[r]	Р р	
19	S s	(сы)	[s]	С с	
20	T t	(ты)	[t]	Т т	
21	U u	(ұ), (у)	[ʊ, u]	Ұ ұ, Уу	(ұ) – фонема, (у) – аллофон
22	V v	(вы)	[v]	В в	
23	W w	(ў)	[w]		Согласная фонема
24	X x	(хы), (хі)	[x, h]	Хх, Hh	(хы) – фонема, (хі) – аллофон
25	Y y	(ы)	[ɤ]	(ы)	
26	Z z	(зы)	[z]	(з)	
27	`	аксант грав			Вторая пара диграфа

НИИ ИСПУ собственный интеллект

Таблица 3.5.2. Орфографические правила для диграфов

№	Название звуков	МФА	Буквы кириллицы	Диграфы на латинице	Пример	Транскрипция
1	(ə)	[æ]	Ә ə	Ae ae	saebiz	[sæbɪz]
				A` a`	sa`biz	[sæbɪz]
2	(ө)	[ɵ]	Ө ө	Oe oe	oemir	[ɵmir]
				O` o`	o`mir	[ɵmir]
3	(ү)	[ʏ]	Ү ү	Ue ue	uemit	[ʏmɪt]
				U` u`	u`mit	[ʏmɪt]
4	(ың)	[ŋ]	Ң ң	Ng ng	ku`nggi	[kʏŋgi]
5	(чы)	[tʃ]	Ч ч	Ch ch	chip	[tʃɪp]
6	(ғы)	[ɣ]	Ғ ғ	Gh gh	ghalym	[ɣɑlɤm]
7	(шы)	[ʃ]	Ш ш	Sh sh	shar	[ʃɑr]
8	(жы)	[ʒ]	Ж ж	Zh zh	zhanga	[ʒɑŋɑ]

Здесь знак ` – *аксант грав* служит в качестве второй пары диграфов A` a`, O` o` и U` u` для представления мягких гласных казахских звуков (*ə*) - [æ], (*ө*) - [ɵ] и (ү) - [ʏ], соответственно.

Напомним, что в узбекском языке аналогичную роль играет знак ′ – *аксант эгю,* используемый в качестве второй пары диграфов O′ o′ и G′ g′ для представления звуков (о) - [ɔ] и (ғе) - [ɣ] (см. Параграф 3.1.2).

На рисунках 3.5.1 и 3.5.2 в желтой рамке показано место расположения знака ` - *аксант грав* и в красной рамке – место расположенния знака ′ - *аксант эгю* на клавиатурах компьютеров типа IBM и типа Apple, соответственно. Важно знать это, только при компьютерной обработке текста, поскольку коды этих знаков разные, а в ручной записи не важно какой из них использовать.

НИИ искусственного интеллекта

1. Клавиатура офисная.

2. Клавиатура ASUS.

3. Клавиатура Тачпад.

4. Клавиатура Asus B43

Рисунок 3.5.1. Клавиатура для компьютеров типа IBM

1. Стандартная клавиатура Apple.

2. Клавиатура Apple Wireless.

3. Клавиатура iMac.

4. Клавиатура Apple Wireless MC184.

Рисунок 3.5.2. Клавиатура для компьютеров типа Apple

ПРИЛОЖЕНИЕ. ГИМН РЕСПУБЛИКИ КАЗАХСТАН

Действующий алфавит	1- вариант нового алфавита	2- вариант нового алфавита
Алтын күн аспаны, Алтын дән даласы, Ерліктің дастаны, Еліме қарашы! Ежелден ер деген, Даңқымыз шықты ғой. Намысын бермеген, Қазағым мықты ғой!	Altyn kuen aspany, Altyn daen dalasy, Erlikting dastany, Elime qarashy! Ezhelden er degen, Dangqymyz shyqty ghoj. Namysyn bermegen, Qazaghym myqty ghoj!	Altyn ku`n aspany, Altyn da`n dalasy, Erlikting dastany, Elime qarashy! Ezhelden er degen, Dangqymyz shyqty ghoj. Namysyn bermegen, Qazaghym myqty ghoj!
Қайырмасы: Менің елім, менің елім, Гүлің болып егілемін, Жырың болып төгілемін, елім! Туған жерім менің – Қазақстаным!	**Qajyrmasy:** Mening elim, mening elim, Gueling bolyp egilemin, Zhyryng bolyp toegilemin, elim! Tughan zherim mening – Qazaqstanym!	**Qajyrmasy:** Mening elim, mening elim, Gu`ling bolyp egilemin, Zhyryng bolyp to`gilemin, elim! Tughan zherim mening – Qazaqstanym!
Ұрпаққа жол ашқан, Кең байтақ жерім бар. Бірлігі жарасқан, Тәуелсіз елім бар. Қарсы алған уақытты, Мәңгілік досындай. Біздің ел бақытты, Біздің ел осындай!	Urpaqqa zhol ashqan, Keng bajtaq zherim bar, Birligi zharasqan, Taewelsiz elim bar Qarsy alghan waqytty, Maenggilik dosyndaj, Bizding el baqytty, Bizding el osyndaj!	Urpaqqa zhol ashqan, Keng bajtaq zherim bar, Birligi zharasqan, Ta`welsiz elim bar, Qarsy alghan waqytty, Ma`nggilik dosyndaj, Bizding el baqytty, Bizding el osyndaj!
Қазақ елі – мәңгілік ел	**Qazaq eli – maenggilik el!**	**Qazaq eli – ma`nggilik el!**

НИИ Искусственный интеллект

СПИСОК ЛИТЕРАТУРЫ

1. Казахский язык // Большая советская энциклопедия : в 30 т. / 3-е изд. – М.: Советская энциклопедия, 1969–1978.
2. Баскаков Н.А. К вопросу о классификации тюркских языков. Известия Академии наук СССР. Отделение литературы и языка. Т. XI, вып. 2. – М., 1952. - С.121–134.
3. Аманжолов А.С. История и теория древнетюркского письма. - Алматы: Мектеп, 2003. - 368 С.
4. https://ru.wikipedia.org/wiki/Казахский_язык.
5. https://ru.wikipedia.org/wiki/Казахи.
6. Конституция Республики Казахстан от 30.05.1995 с изменениями и дополнениями 07.10.1998, 21.05.2007, 02.02.2011.
7. http://www.akorda.kz/ru/addresses/addresses_of_president/posla.nie-prezidenta-respubliki-kazahstan-nnazarbaeva-narodu-kazahstana-14-dekabrya2012-g.
8. https://egemen.kz/article/nursultan-nazarbaev-bolashaqqa-baghdar-rukhani-zhanhghyru.
9. https://ru.wikipedia.org/wiki/Древнетюркское_руническое_письмо.
10. https://ru.wikipedia.org/wiki/Иссыкское_письмо.
11. http://www.niac.gov.kz/kz/expertiza/item/138-dzhadidizm-i-reforma-musulmanskogo-obrazovaniya.
12. Байтурсынов А. Тіл тағылымы. - Алматы, 1992. – 448 б.
13. https://ru.wikipedia.org/wiki/Казахская_письменность
14. Қазақ грамматикасы. Фонетика, сөзжасам, морфология, синтаксис. – Астана, 2002.
15. Мырзабеков С. Қазіргі қазақ тілінің фонетикасы. – Алматы: Дәуір-Кітап, 2013. – 219 с.
16. Калиева Б.А. Қазақ тілінің фонетикасы. Оқу құралы. – Алматы: Эверо, 2014. – 162 с.
17. http://onstudy.kz/dausty_dybys.
18. Зиндер Л.Р. Общая фонетика. - Л., 1960; 2 изд., М., 1979.
19. https://ru.wikipedia.org/wiki/Фонетика.
20. https://traditio.wiki/Органы_речи_и_общение.

НИИ Искусственный интеллект

21. https://ru.wikipedia.org/ wiki/Принцип_Дирихле

22. Шәріпбаев А.А., Төреханов С.И. Қазақ тіліндегі дауысты звуковтар???? жүйесінің математикалық моделі. Л.Н.Гумилев атындағы ЕҰУ Хабаршысы, №3-4. - Астана, 2002. - Б.21-27.

23. Рабинер Л., Гоулд Б. Теория и применение цифровой обработки сигналов. – М.: Издательство «Мир», 1975. – 839 с.

24. https://ru.wikipedia.org/wiki/Булева_алгебра.

25. Z. Yessenbayev, M. Karabalayeva and A. Sharipbayev, "Formant Analysis and Mathematical Model of Kazakh Vowels," *2012* UKSim 14th International Conference on Computer Modelling and Simulation, Cambridge, 2012. - P.427-431.

26. Olive J.P., Greenwood A. and Coleman J., “Acoustics of American English speech. A dynamic aProach,” Springer, 1993. – 396 p.

27. Hayward K., “Experimental Phonetics,” Harlow, UK: Pearson, 2000. – 149 p.

28. Есенбаев Ж.А. Распознавание казахской речи по определенной словарной базе в условиях шумов. Дис.доктора PhD : *6D060200*-Информатика. - Астана: ЕНУ им. Л.Н.Гумилева, 2014.

29. Фант Г. Акустическая теория речеобразования. - М., 1964.

30. Praat, URL: http://www.fon.hum.uva.nl/praat/.

31. International Phonetic Association. Handbook of the International Phonetic Association: A Guide to the Use of the International Phonetic Alphabet. Cambridge, U.K.: Cambridge University Press, 2003.

32. https://en.wikipedia.org/wiki/International_Phonetic_Alphabet

33. Лайонз Дж. Структура языка // Введение в лингвистику = Introduction to Theoretical Linguistics / Пер. с англ. Под ред. и с предисл. В. А. Звегинцева. — М.: Прогресс, 1978.

34. https://ru.wikipedia.org/wiki/Латинский_алфавит.

35. https://ru.wikipedia.org/wiki/Алфавиты_на_основе_ латинскогого.

36. ГОСТ 16876-71 Правила транслитерации букв кирилловского алфавита буквами латинского алфавита.

НИИ искусственный интеллект

37. https://ru.wikipedia.org/wiki/ISO_9.

38. https://ru.wikipedia.org/wiki/Транслитерация_русского_алфавита_латиницей.

39. https://ru.wikipedia.org/wiki/ Турецкая_письменность.

40. https://traditio.wiki/Туркменская_письменность.

41. Образование в Туркменистане. Доклад. Туркменская Инициатива по Правам Человека. 4 декабря 2012.

42. https://ru.wikipedia.org/wiki/Узбекская_письменность

43. https://ru.wikipedia.org/wiki/Азербайджанская_письменность.

44. Шарипбай А., Қазақ жазуын латын әліпбиіне көшіру жобасы. Л.Н.Гумилев атындағы ЕҰУ, Астана, 2013. –29 б.

45. http://kazcorpus.kz/klcweb/en/.

Список публикаций научной школы

Кодировка и распознавание казахских букв, драйвера

46. Шарипбаев А., Алпысбаев Т., Туреханов С. СТ РК 1048-2002. Информационная технология. 8-битовая кодовая таблица казахского алфавита. Комитет по стандартизации, метрологии и сертификации Министерства индустрии и торговли РК. - Астана, 2002. -12 С.

47. Шәріпбаев А.А., Тренкеншу В. Драйвера, шрифты и конвертеры казахских букв. Труды I-ой Международной научно-практической конференции «Информатизация общества». - Астана, ЕНУ им. Л.Н.Гумилева, 2004. - С.174-178.

48. Шарипбаев А.А., Тренкеншу В.П. Нейросетевой метод оптического распознавания символов. Вестник ЕНУ им. Л.Н.Гумилева, №1-2, 2003. - С.63-70.

49. Шарипбаев А.А. СТ РК 2045-2010. Аппараты телефонные. Технические требования (Требования к кодировке казахских букв на основе кириллицы). Комитет технического регулирования и метрологии Министерства индустрии и новых технологий РК. - Астана, 2010. - 34 С.

НИИ Искусственный интеллект

50. Шарипбаев А. СТ РК 2046-2010. Абонентские станции сотовой связи.Технические требования (Требования к кодировке букв казахского алфавита). Комитет технического регулирования и метрологии Министерства индустрии и новых технологий РК. - Астана, 2010. - 38 С.

Терминологические, толковые и орфографические словари

51. Шарипбаев А.А., Бурибаев Б., Алдажаров К. и др. Казахско-русский, русско-казахский словарь терминов. Информатика и вычислительная техника. – Алматы: Рауан, 1999. - 240 с.

52. Шәріпбаев А.А. Қазақ тілінде термин жасау принциптері мен шарттары. МАҚКМ, Тілдерді дамыту департаменті. - Астана, 2001.

53. Шарипбаев А.А., Рахлетова А., Искакова А.С. Толковый словарь на английском, русском и казахском языках по информатике. – Алматы: Сөздік-Словарь, 2002. - 175 с.

54. Шарипбаев А.А., Абдиев Қ., Бурибаев Б. и др. Научно-отраслевой толковый словарь. Информатика и вычилительная техника. – Алматы: Мектеп, 2002. - 457 с.

55. Шәріпбаев А.А., Сулейменов Б.А., Әубакир Д.А. СТ РК 34.014. Информационная технология. Комплекс стандартов на автоматизированные системы. Автоматизированные системы. Термины и определения на английском, русском и казахском языках. КСМС МИТ РК. - Астана, 2002. - 48 с.

56. Шарипбай А.А., Буркитбаев М.Т. СТ РК 1091. Единая система программной документации. Термины и определения на английском, русском и казахском языках. КСМС МИТ РК. - Астана, 2002. - 114 с.

57. Шарипбаев А.А., Адамов А.А. Об электронном словаре нефтяных морских терминов на русском и казахском языках. Труды 1-ой Международной научно-практической конференции «Информатизация общества». – Астана: ЕНУ им. Л.Н.Гумилева, 2004. - С.183-184.

НИИ Интеллектуальный интеллект

58. Шарипбаев А.А., Тренкеншу В. Автоматизация создания и редактирования многоотраслевого казахско-русского электронного словаря. Труды 1-ой Международной научно-практической конференции «Информатизация общества». Астана: ЕНУ им. Л.Н.Гумилева, 2004. - С.173-178.

59. Шарипбай А.А., Адамов А.А., Боранбаев С.Н., Дюзелбаев С. Казахско-русский, русско-казахский терминологический словарь «Информационные и коммуникационные технологии». – Астана: Алтын кітап, 2013. - 490 с.

60. Шарипбай А.А., Адамов А.А. и др. Казахско-русский, русско-казахский терминологический словарь «Информатика и вычислительная техника». 10000 терминов. (Одобрен Республиканской терминологической комиссией при Правительстве РК). – Алматы: КАЗақпарат, 2014. - 452 с.

61. А. Шәріпбай. Қазақ жазуын латын әліпбиіне көшіруге барысында қажетті орфографиялық сөздік: Сөздік. - Астана: Л.Н.Гумилев атындағы ЕҰУ, 2014. - 87 б.

Фонетика

62. Шәріпбаев А.А., Төреханов С.І. Фонетикалық жүйеміз неге ала-құла? “Егеменді Қазақстан”, № 209 (22813). 19.09.2001.

63. Шәріпбаев А.А., Төреханов С. Қазақ тілінің қазіргі фонетикасы жайында. Л.Н.Гумилев атындағы ЕҰУ Хабаршысы, №3-4. - Астана, 2001. - Б.43-48.

64. Шәріпбаев А.А., Тұрғанбекова. Қазақ тілі төл дыбыстары негізіндегі жалғаулар ережелерін формалдау. Л.Н.Гумилев атындағы ЕҰУ Хабаршысы. №8 - Астана, 2002. - С.61-66.

65. Шәріпбаев А.А., Төреханов С. Қазақ тілі дыбыстар тіркесін талдауды автоматтандыру. Труды І-ой Международной научно-практическрй конференции «Информатизация общества». – Астана: ЕНУ им. Л.Н.Гумилева, 2004. - С.171-172.

66. Шарипбаев А.А., Бекманова Г.Т. Формализация фонологических закономерностей казахского языка для систем

НИИ интеллектуальной собственности

автоматического распознавания речи // Вестник Национальной академии наук РК. №6. - Алматы, 2008. - С.42-46.

67. Бекманова Г.Т. Транскриптор казахских слов в транскрипционную запись для систем обработки естественного языка. // Вестник науки. Костанайский социально-технический университет имени З.Алдамжар. - № 2. - Костанай, 2009. - С.33-42.

68. Sharipbayev A., Bekmanova G. Some questions of the automatic transcription of Kazakh language. The module of transcription of kazakh speech recognition system. Proceedings of II International Scientific and Practical Conf. «Informatization of Society". - Astana, 2010. - P.4-9.

69. Sharipbayev A., Yessenbayev Z., Karabalayeva M. The mathematical description of Kazakh phonetics. UKSim2012: UKSim 14th International Conference on Computer Modelling and Simlulation. – Cambridge, United Kingdom, 2012. - P.427-431.

70. Zh.Yessenbayev, M.Karabalayeva, A.Sharipbay. Formant Analysis and Mathematical Model of Kazakh Vowels // Computer Modelling and Simulation (UKSim), 2012 UKSim 14th International Conference on. – 2012. - P.427-431 (индексируется в Scopus).

71. Sharipbayev A., Bekmanova G., Karabalayeva M. Mathematical models of designing automatic transcription. The 4th Congress of the Turkic World Mathematical Society. - Baku, Azerbaijan, 2011. - P.68-73.

72. Bekmanova G. T., Sharipbaeyv A.A., Shelepov V.U. Formalization of phonological rules of the Kazakh language for system automatic speech recognition // Вестник Евразийского национального университета им. Л.Н.Гумилева. – Специальный выпуск. - Астана, 2012. - С.51-59.

73. Шарипбай А., Есенбаев Ж., Карабалаева М. Математическое моделирование. Геометрия казахских гласных// Доклады Национальной академии наук РК. - №1. - Алматы, 2012. - С.24-32.

74. Шарипбай А.А., Бурибаева А.К., Бекманова Г.Т., Калиев А.К. Гласные звуки казахского языка и методы их выделения в

НИИ искусственный интеллект

начале слова// Доклады Национальной академии наук РК. - №6. - Алматы, 2012. - С.10-16.

75. Шәріпбай А.Ә., Бөрібаева Ә.К. Қазақ тілі звуковтарын фонетикалық және фонологиялық талдау // Л.Н.Гумилев атындағы ЕҰУ-нің Хабаршысы. - №6 (97). - Астана: 2013. - Б.75-90.

76. Ә.К. Бөрібаева, А.Ә. Шәріпбай. Қазақ тіліндегі дифондар үндесімі // «Ғылым және білімде ақпараттық-коммуникациялық технологияларды қолдану» халықаралық ғылыми-тәжірибелік конференция материалдары. 22 қараша 2013 жыл. - Б.45-49.

77. Шарипбай А., Есенбаев Ж.А., Карабалаева М., Шамаева Ф. К вопросу об уточнении фонетического строя казахского языка посредством акустического анализа звуков// «Түркі тілдерін компьютерлік өңдеу» халықаралық конференция еңбектері. – Астана: Л.Н.Гумилев атындағы ЕҰУ баспасы, 2013. - Б.105-111.

78. A.Sharipbay, A.Buribayeva. The advantage of interphoneme processing at diphone recognition of Kazakh words // Proceedings of the II International Conference on Computer processing of Turkic Languages (Turklang 2014). - Istanbul: Istanbul Technical University. - P.75-78.

79. Makazhanov, A., Yessenbayev, Zh., Sabyrgaliyev, I., Sharafudinov, A., and Makhambetov O., On Certain Aspects of Kazakh Part-of-Speech Tagging // In the Proc. of 2014 IEEE 8th International Conference on Application of Information and Communication Technologies (AICT2014) – 2014. – P. 240-243.

Морфология, морфологические анализаторы и синтезаторы

80. Шәріпбаев А.А., Тұрғанбекова Ж. Қазақ морфемасын синтездеу мен талдау// Труды Международной научной конференции «Развитие информационных технологий». – Алматы: КазНУ им. Аль-Фараби, 2003. - С.86-88.

81. Шәріпбаев А.А., Тұрғанбекова Ж. Қазақ тілі жалғау ержелерінің алгебрасы// ҚР Ұлттық ғылым академиясының Баяндамасы. - №2. – Алматы, 2005. - Б.59-65.

НИМ искусственный интеллект

82. Шәріпбаев А.А., Тұрғанбекова Ж. Қазақ тілі морфологиялық ережелерін талдау мен тудыруды автоматтандыру// Труды I-ой международной научно-практич. конференции «Информатизация общества». – Астана: ЕНУ им. Л.Н.Гумилева, 2004. - С.175-176.

83. Шарипбаев А.А., Бекманова Г.Т. Формализация морфологических правил казахского языка с помощью семантической нейронной сети// Доклады НАН РК. - №4. - Алматы, 2009. - С.11-16.

84. Бекманова Г.Т. Некоторые подходы к проблемам автоматического словоизменения и морфологического анализа в казахском языке // Вестник Восточно-Казахстанского государственного технического университета им. Д. Серикбаева. - № 4. - Усть-Каменогорск, 2009. - С.192-197.

85. Шәріпбаев А.А., Бөрібаева А. Қазақ тілінің морфологиялық ережелерін схемотехникалық моделдеу// Л.Н.Гумилев атындағы ЕҰУ Хабаршысы. - № 2. - Астана, 2010. - С.73-79.

86. Шәріпбаев А.А., Бөрібаева А. Verilog тілінде қазақ сөздерінің морфологиялық синтезін автоматтандыратын программа құру// Л.Н.Гумилев атындағы ЕҰУ Хабаршысы. - № 4. - Астана, 2010. - С.182-187.

87. Шәріпбаев А.А., Бөрібаева А. Қазақ сөздерінің морфологиялық синтезін Quartus II ортасында аппараты жүзеге асыру// Л.Н.Гумилев атындағы ЕҰУ Хабаршысы. - № 4. - Астана, 2010. - С.197-203.

88. Ергеш Б.Ж., Бекманова Г.Т. Система автоматического словоизменения казахских слов: модуль генерации словоформ местоимений // Труды Международной научной конференции студентов, магистрантов и молодых ученых. -Астана, 2010. -С.7-9.

89. Шәріпбаев А. Ә., Ергеш Б.Ж. Формалдық грамматикалар негізінде сын есімдердің сөзжасам ережелерін жазу // «Қоғамды ақпараттандыру» II Халықаралық ғылыми-практикалық конференция еңбектері. - Астана, 2010. - Б.42-47.

НИИ ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ

90. Бурибаева А.К. Қазақ тілінің морфологиялық ережелерін шешімдер қабылдау жүйесіне салу// Труды VII Казахстанско-Российской научно-практической конференции «Математическое моделирование научно-технических и экологических проблем в нефтегазодобывающей промышленности». -Алматы, 2010. -С.67-72.

91. Бурибаева А.К. Қазақ тілі жалғауларына байланысты шешімдер қабылдау жүйесін құру // Қазақстан Республикасы тәуелсіздігінің 20 жылдығына арналған II Халықаралық «Ақпараттық - инновациялық технологиялар: ғылым, білім және бизнестің интеграциялануы» ғылыми-тәжірибелік конференциясы. - Алматы, 2011. - Б.81-86.

92. Ергеш Б.Ж. Қазақ тілінде жұрнақтарды жалғауды жақшалық жазба арқылы жүзеге асыру // Математика, механика және информатика пәндерінен III Республикалық студенттік ғылыми-практикалық конференциясының баяндамалар жинағы. - Астана, 2011. - Б.232-234.

93. Bekmanova G.N., Sharipbayev A.A, Yergesh B., Mukanova A.S. Generation of kazakh words// Вестник ЕНУ им. Л.Н.Гумилева. – Специальный выпуск. - Астана, 2012. - С.27-35.

94. Bekmanova G.N., Sharipbayev A.A, Buribayeva A. Formalization of morphological rules of inflection in the kazakh language // Вестник ЕНУ им. Л.Н.Гумилева. – Специальный выпуск. - Астана, 2012. - С.18-26.

95. A.Sharipbay, G.Bekmanova, A.Mukanova, B.Yergesh. Automatisation and transformation of verbs in Kazakh language // International scientific-theoretical conference. -Almaty, 2012. -P.14-18.

96. Шарипбай А.А., Бекманова Г.Т., Ергеш Б., Бурибаева А.К., Карабалаева М.Х. Интеллектуальный морфологический анализатор, основанный на семантических сетях // II Международная научно-техническая конференция «Открытые семантические технологии проектирования интеллектуальных систем». – Минск: БГУИР, Белоруссия, 2012. - С.397-400.

97. Бурибаева А.К., Шарипбаев А.А., Бекманова Г.Т., Карабалаева М.Х., Ергеш Б.Ж. Аппаратная реализация синтеза

НИИ Искусственный интеллект

словоформ казахского языка с помощью ассоциативной памяти// Вестник ЕНУ им. Л.Н.Гумилева. – Специальный выпуск. - Астана, 2012. - С.180-183.

98. Шәріпбай А.Ә., Бекманова Г.Т., Муканова А.С., Ергеш Б.Ж. Қазақ тілінде етістіктің жаңа сөзжасамдарын тудыруды автоматтандыру // А.Қ. Жұбановтың 75 жылдық мерейтойына орай «Қазіргі қазақ тіл білімі: қолданбалы лингвистиканың өзекті мәселелері» халықаралық ғылыми-теориялық конференция еңбектері. - Алматы, 2012. - Б.114-118.

99. Шарипбаев А.А., Бекманова Г., Муканова А., Ергеш Б. Формальная модель морфологических правил казахского языка//Доклады Науиональной академии наук РК. - Алматы:, 2012. - С.16-22.

100.Бекманова Г.Т., Жеткенбай Л. Қазақ тілінің күрделі сөздерін формалдау негізінде жасау // Труды 1-ой международной конференции «Компьютерная обработка тюркских языков». - Астана, 2013. - С.237-243.

101.Шарипбай А., Бекманова Г.Т., Муканова А.С., Ергеш Б.Ж. Онтологическая модель представления морфологических правил казахского языка в виде семантических гиперграфов// Материалы III международной научно-технической конференции «Открытые семантические технологии проектирования интеллектуальных систем». - Минск, 2013. - С.337-340.

102.Шарипбай А., Муканова А.С., Ергеш Б.Ж., Разахова Б.Ш. Морфологиялық ережелерді онтологиялық моделдеу// «Түркі тілдерін компьютерлік өңдеу» атты I халықаралық конференция еңбектері. - Астана, 2013. - Б.188-194.

103.B.Yergesh, A.Mukanova, G.Bekmanova, A.Sharipbay, B.Razakhova. Formal models of nouns in the Kazakh language. Leonardo Electronic Journal of Practices and Technologies; Issue 25, 2014 (13). - P. 264-273.

104.Ергеш Б. Ж., Муканова А.С., Шарипбай А.А., Бекманова Г.Т. Формальная модель именных частей речи казахского языка//

НИИ Искусственный интеллект

Труды XIV национальной конференции по искусственному интеллекту с международным участием. - Казань, 2014. - С.101-107.

105.А. Шәріпбай, А.Муканова, Г.Бекманова. Қазақ тілі сын есімінің формалды моделі// Л.Н.Гумилев атындағы ЕҰУ-нің Хабаршысы. - №??? – Астана, 2014. - С.114-119.

106.L.Zhetkenbay, A.A.Sharipbay, G.T.Bekmanova. Semantic Neural Network Model of the Morphological Rules of the Kazakh and Turkish Noun for NLP System. Proceeding of the 2nd International conference on Computer, Control and Communication Technologies (CCCT'15), 2015.– Turkey, Antalya, P.95-100. ISBN No.:978-93-84422-46-2.

107.B.Yergesh, A.Sharipbay, G.Bekmanova, Lipnitskii S. Sentiment analysis of Kazakh phrases based on morphological rules // Journal Of Kyrgyz State Technical University named after I.Razzakov, Theoretical And APlied Scientific Technical Journal. - № 2 (38). - Bishkek, Kyrgyzstan, 2016. - P.39-42. ISSN 1694-5557.

108.Yergesh Banu, Sharipbay Altynbek, Bekmanova Gulmira.The role of Adjectives in the Sentiment Analisys (2016) In Proceedings of the 14th International Conference on Computational and Cognitive Linguistics TEL'2016 "Semantic Models and Technologies". - Kazan, Tatarstan, Russia. - P.85-89. In Russian.

109.Sharipbay A., Bekmanova G., Zetkenbay L., Kamanur U. Ontological modeling of morphological rules for the adjectives in Kazakh and Turkish languages// Journal of Theoretical and APlied Information Technology. - Vol. 91. N.2. – 2016. - P.257-263. ISSN: 1992-8645.

110.Шарипбай А., Зулхажав А., Разахова Б.Ш. Морфологический анализатор казахского языка на основе онтологического моделирования морфологических правил// Материалы международной конференции по компьютерной и когнитивной лингвистике TEL. – Казань, 2016. - С.248-252.

Синтаксис, синтаксические анализаторы и синтезаторы

НИИ Искусственный интеллект

111.Шәріпбаев А.А., Уталина Б.Ш. Қазақ тілінің синтаксистік ережелерін синтездеуді автоматтандыру// Труды II Международной конференции «New Trends in the Computer Science Master's Curriculum». - Алматы, 2004. - С.347-350.

112.Шәріпбаев А.А., Уталина Б. Қазақ тілінің сөз тіркестерін талдау ережелерін автоматтандыру// Материалы I-ой международной научно-практ. конференции «Информатизация общества». – Астана: ЕНУ им. Л.Н.Гумилева, 2004. - С.178-179.

113.Шәріпбаев А.А., Уталина Б. Қазақ тілінің сөйлемдерді талдау ережелерін автоматтандыру// Труды I-ой международной научно-практ. .конференции «Информатизация общества». – Астана: ЕНУ им. Л.Н.Гумилева, 2004. - С.179-180.

114.Шәріпбаев А.А., Уталина Б.Ш. Қазақ тілі сөйлемдерін тудыру процесін математикалық моделдеу// Материалы III-ей международной конференции «Математическое моделирование, информационные технологии в образовании и науке». – Алматы: КазНПУ им.Абая, 2005. - Том 1. - С.221-226.

115.Шәріпбаев А.А., Уталина Б.Ш. Қазақ тіліндегі сөйлемдерді талдау ережелерін автоматтандыру// Қазақстан Республикасы Ұлттық ғылым академиясының Хабарлары. - №2(150). - Алматы, 2005. - Б.16-20.

116.Шәріпбаев А.А., Уталина Б.Ш. Қазақ тілі сөйлемдерін тудыру процессін математикалық моделдеу// Материалы III-ей международной конференции «Математическое моделирование, информационные технологии в образовании и науке».- Алматы: КазНПУ им.Абая, 2005. - С.221-225.

117.Шәріпбаев А.А., Уталина Б.Ш. Қазақ тіліндегі сөз тіркестерін талдау ережелерін автоматтандыру// Л.Н.Гумилев атындағы ЕҰУ Хабаршысы. - №5. – Астана, 2005. - Б.36-41.

118.Шәріпбаев А.А., Уталина Б.Ш. Контексті бос грамматика арқылы қазақ тілі сөйлемдер жиынының анықталуы// ҚР Ұлттық ғылым академиясының Баяндамасы. - № 5. - Алматы, 2005. - Б.123-128.

НИИ искусственный интеллект

119.Шәріпбаев А.А., Разакова Б.Ш. Қазақ тіліндегі сөйлемдерді танудың математикалық әдісі// ҚР Ұлттық ғылым академиясының Хабаршысы. - №2. - Алматы, 2008. - Б.16-20.

120.Шәріпбаев А.А., Разакова Б.Ш. Қазақ тіліндегі сөйлемдерді танудың математикалық әдісі// Международная научно-практическая конф. «Информационно-инновационные технологии: Интеграция науки, образования и бизнеса».. – Алматы: КазНТУ им. К.И. Сатпаева, 2008. - С.220-225.

121.Sharipbayev A., Razakhova B. Mathematical models of syntactical rules of the Kazakh Language subject to semantics of parts of the sentence// The 4th Congress of the Turkic World Mathematical Society. - Baku, Azerbaijan, 1-3 July 2011. - P 463.

122.Ергеш Б.Ж., Муканова А.С., Разахова Б.Ш. Қазақ тіліндегі жай сөйлемдердің онтологиялық моделі// «Түркі тілдерін компьютерлік өңдеу» I халықаралық конференция еңбектері. - Астана, 2013. - Б. 186-188.

123.Шарипбай А., Ергеш Б.Ж., Разахова Б.Ш., Муканова А.С. Қазақ тіліндегі жай сөйлемдердің онтологиялық моделі// «Түркі тілдерін компьютерлік өңдеу» атты I халықаралық конференция еңбектері. - Астана, 2013. - Б.188-194.

124.Sharipbay A. Problems and prospects of computer processing of the Kazakh language // Proceedings of the I International Conference on Computer processing of Turkic Languages. - Astana: L.N.Gumilyov Eurasain National University, 2013. - P.13-17.

125.Шәріпбай А., Голенков В., Разахова Б., Кабенов Д. Онтология негізінде мәтінді табиғи тілде талдау// Л.Н.Гумилев атындағы ЕҰУ-нің Хабаршысы. - №2(99). –Астана, 2014. -Б.94-100.

Семантика

126.Бекманова Г.Т., Шарипбаев А.А. Построение логической семантики слов казахского языка // Труды всероссийской конференции с международным участием. Знания-Онтологии-Теории. – Новосибирск, 2009. - Том 2. - С.246-249.

НИИ Искусственный интеллект

127.Шәріпбаев А.А., Бөрібаева А., Қазақ тілі жалғауларына байланысты шешімдер қабылдауды схемотехникалық моделдеу// Труды Международной научно-практической конференции «Актуальные проблемы математики, информатики, механики и теории управления». - Алматы: РГП «Институт математики, информатики и механики» КН МОН РК, 2009. - С.44-49.

128.Sharipbayev A., Bekmanova G. The synthesis of world forms of Turkic languages using semantic neural network. Modern problems of aPlied mathematics and information technologies - al Khorezmiy 2009, Tashkent, Transaction of the conference, Volume № 2. - P.148-152.

129.Шарипбай А., Омарбекова А.С. Семантическая модель электронного учебного издания// Вестник ЕНУ им. Л. Гумилева. Спецвыпуск. – Астана, 2012. - С.432-435.

130.Бекманова Г.Т., Жеткенбай Л. Қазақ күрделі сөздерін түрлендірудің семантикалық моделі // Труды 1-ой международной конференции «Компьютерная обработка тюркских языков». - Астана, 2013. - С.243-248.

131.Муканова А.С. Семантикалық желіні қазақ тілі сөздерін түрлендіруде қолдану // Математика, механика және информатика пәндерінен III - Республикалық студенттік ғылыми-практикалық конференциясы: баяндамалар жинағы. - Астана, 2011. - Б. 397-398.

132.Бекманова Г.Т., Муканова А.С., Ергеш Б. Математические модели морфологических правил казахского языка с учетом семантики слов// Труды международной конференции «Актуальные проблемы современной математики, информатики и механики». – Алматы: КазНУ, 2011. - С.307-308.

133.Бекманова Г.Т., Муканова А.С., Ергеш Б. Қазақ тіліндегі жаңа етістік формаларының семантикалық нейрондық желі арқылы түрленуі // Труды Международной конференции «Актуальные проблемы современной математики, информатики и механики». – Алматы: КазНУ, 2011. - С.308-309.

134.Ергеш Б.Ж., Муканова А.С. Қазақ тіліндегі зат есімдерді зерделі талдауды автоматтандыру// Труды Международной

НИИ Искусственный интеллект

конференции «Актуальные проблемы современной математики, информатики и механики». – Алматы: КазНУ, 2011. - С.311-315.

135.Бекманова Г.Т., Махимов А. К. Графематический анализ казахского неструктурированного текста // Труды III международной научно-практической конференции «Информатизация общества». - Астана, 2012. - С.509-511.

136.Бекманова Г.Т., Махимов А.К., Калиев А. Транслятор казахского неструктурированного текста в морфологическую скобочную запись // Труды III Международной научно-практической конференции «Информатизация общества». - Астана, 2012. - С.511-514.

137.Yergesh B., Sharipbay A., Bekmanova G., Buribayeva A., Mukanova A.. Semantic retrieval of information in the Kazakh language in e-libraries // Journal of International Scientific Publications: Educational Alternatives, Volume 10. Part 1, ISSN:1313-2571, P.108-115, Published at: http:/www.scientific publication.net.

138.Sharipbay A., Yergesh B., Mukanova A., Bekmanova G., Razakhova B. Semantic Hyper-graph Based Representation of Nouns in the Kazakh Language// Computacion y Sistemas. - Vol. 18. - №. 3. – 2014. - P.627–635 ISSN 1405-5546 DOI (Scopus).

139.Sharipbay A., Mukanova A. Smart question answering system for e-government web-portal //Proceedings of the International Scientific-Practical Conf. «Smart government: Science &Technology». - Astana, 2014. - P.20-24.

140.Sharipbay A., Karimzhan B., Razakhova B. "Smart System for Analysis and Synthesis of the Kazakh Sentences". International Scientific-Practical Conference "Smart Goverment: Science and Technology". Proceedings. October 7-8, 2014. - P.71-74.

141.Sharipbay A., Bekmanova G., Buribayeva A., Yergesh B., Mukanova A. Semantic neural network model of morphological rules of the agglutinative languages // 6th International Conference on Soft Computing and Intelligent Systems, and 13th International Symposium on Advanced Intelligence Systems. - Kobe, Japan, 2012. art. no. 6505337. - P.1094-1099.

НИИ Искусственный интеллект

142.Sharipbay A., Zhetkenbay L., Bekmanova G.T., Kamanur U. The Semantical, Ontological Models and Formalization Rules Kazakh Compound Words. – TurkLang, 2014.

143.Sharipbay A., Bekmanova G., Mukanova A., Yergesh B. Ontological representation of the kazakh language noun // Proc. 2nd international scientific and practical conference “Intellectual information and communication technologies as a tool for realization of the third industrial revolution within the framework of the Kazakhstan 2050 strategy”. - Astana, 2014. - P.242-246.

144.Sharipbay A., Mukanova A., Yergesh B., Bekmanova G. Formal Model of Adjective In the Kazakh Language // Proceedings of the II International Conference on Computer processing of Turkic Languages. – Istanbul: Technical University, 2014. - P.57-61.

145.Sharipbay A., Bekmanova G., Yergesh B., Mukanova A. Synchronized Linear Tree for Morphological Analysis and Generation of the Kazakh Language // Proceedings of the II International Conference on Computer processing of Turkic Languages. – Istanbul: Technical University, 2014. - P.113-117.

146.Yergesh B., Mukanova A., Bekmanova G., Sharipbay A., Razakhova B. Semantic hyper-graph based representation of nouns in the Kazakh language Computacion y Sistemas. - Volume 18. - Issue 3, 1. - 2014. - P.627-635.

147.Шарипбай А., Головинов М.В., Шелепов В.Ю. Технология создания казахско-английского электронного переводчика с голосовым вводом слов для мобильных устройств// Вестник Евразийского национального университета им. Л.Н.Гумилева. – № 4. - Астана, 2014. - С.128-133.

148.Sharipbay A., Zhetkenbay L., Bekmanova G., Kamanur U. The ontological model of noun for Kazakh-turkish machine translation system// Proceedings of the international conference “Turkic languages processing». TurkLang 2015, September 17-19, 2015. - Kazan, Tatarstan, Russia, ISBN 978-59690-0262-3. - P.15.

149.Barlybayev A., Sabyrov T., Sharipbay A., Omarbekova A. Data Base Processing Programs with Using Extended Base Semantic

НИИ Искусственный интеллект

Hypergraph. In: Rocha Á., Correia A., Adeli H., Reis L., Costanzo S. (eds) Recent Advances in Information Systems and Technologies. WorldCIST 2017. Advances in Intelligent Systems and Computing, vol 569. Springer, Cham.

Речевые технологии

150.Шарипбаев А.А., Карабалаева М.Х. Дикторозависимая система распознавания казахской речи: модель, реализация, тестирование// Доклады НАН РК. - №5. - Алматы, 2007. - С.10-17.

151.Шарипбаев А.А., Карабалаева М.Х. Реализация системы автоматического распознавания казахской речи// ҚР Ұлттық академиясының Хабарлары. - №6. - Алматы, 2007. - С.23 -30.

152.Шарипбаев А.А., Бекманова Г.Т. Математические методы распознавания сочетания звуков казахского языка// Труды Международной научно-практической конф. «Информационно-инновационные технологии: Интеграция науки, образования и бизнеса». – Алматы: КазНТУ им. К.И. Сатпаева, 2008. - С.220-225.

153.Бекманова Г.Т., Шарипбаев А.А. К вопросу распознавания казахской речи // Современные техника и технологии: сборник трудов 14 Международной научно-практической конференции. - Томск, 2008. - С.240-242.

154.Бекманова Г.Т., Шарипбаев А.А. Математические методы распознавания сочетания звуков казахского языка // Информационно-инновационные технологии: интеграция науки, образования и бизнеса: труды международной научно-практической конференции. - Алматы, 2008. - С.216-220.

155.Бекманова Г.Т. Транскриптор казахских слов для распознавания речи // Вестник Национальной академии наук Республики Казахстан. - № 6. – Алматы, 2009. - С.12-17.

156.Шарипбаев А.А., Карабалаева М. Алгоритмы пофонемного распознавания казахской речи в амплитудно-временном пространстве// Труды I Всероссийской конференции "Знания -

НИИ Искусственный интеллект

онтологии - теории" с международным участием. – Новосибирск, 2009. - С.168-171.

157.Бекманова Г.И., Ниценко А.В., Шелепов В.Ю., Шарипбаев А.А. О некоторых вопросах, связанных с распознаванием казахской речи// Вестник ЕНУ им.Л.Н.Гумилева. - № 6- Астана, 2009. - С.172-178.

158.Бекманова Г.И., Ниценко А.В., Шелепов В.Ю., Шарипбаев А.А. Алгоритмы распознавания казахского слова как целого// Вестник ЕНУ им.Л.Н.Гумилева. - № 2. - Астана, 2010. - С.45-52.

159.Шарипбаев А.А.,Тренкеншу В.П. Распознавание речи на нейронных сетях. Классификация изолированных слов многослойным персептроном// Вестник КазНУ им. Аль-Фараби. - № 4(67). - Алматы, 2010. - С.294-299.

160.Бекманова Г.Т., Ниценко А.В., Шарипбаев А.А., Шелепов В.Ю. Некоторые аспекты распознавания казахской речи // Информатизация общества: труды 2-ой международной научно-практической конференции. - Астана, 2010. - С.470-475.

161.Бөрібаева Ә.К. Қазақша сөйлеу синтезін аппаратты жүзеге асыру // Қоғамды ақпараттандыру» II-Халықаралық ғылыми-практикалық конференция еңбектері. - Астана Л.Н.: Гумилев атындағы ЕҰУ, 2010. - С.557-558.

162.Шарипбаев А.А., Бурибаева А., Кокушев Т. Синтез речи на основе дифонной базы казахского языка// Труды II-ой Международной конференции «Информационные технологии». – Алматы: КазНТУ им. К.Сатпаева, 2011. - С.255-257.

163.Бөрібаева А.К. Қазақша сөйлеу синтезіне арналған дифондар классификациясы// ҚР тәуелсіздігінің 20 жылдығына арналған II Халықаралық «Ақпараттық - инновациялық технологиялар: ғылым, білім және бизнестің интеграциялануы» ғылыми-тәжірибелік конференциясы. - Алматы, 2011. - С.21-25.

164.Бекманова Г.Т. Математические методы распознавания речи на основе методов динамического программирования // Информационно-иновационные технологии: Интеграция науки, образования и бизнеса: труды II международной научно-

НИИ Искусственный интеллект

практической конференции, посвященной 20-летию Независимости Республики Казахстан. - Том 2. - Алматы, 2011. - С.316-317.

165.Бекманова Г.Т. Математические методы распознавания речи на основе скрытых Марковских моделей // Информационно-иновационные технологии: Интеграция науки, образования и бизнеса: труды II международной научно-практической конференции, посвященной 20-летию Независимости Республики Казахстан. –Том 2. - Алматы, 2011. - С.317-319.

166.Жантолин Д. Б., Бекманова Г.Т. Обзор некоторых алгоритмов определения начала и конца и речи // Информатизация общества: труды 3-ой международной научно-практической конференции. – Астана, 2012. - С.514-516.

167.Buribayeva A.K., Barlybayev A.B., Sharipbayev A.A., Yergesh B.Zh. Speech user interfaces and contents for e-learning // Journal of International Scientific Publications: Educational Alternatives. - Volume 10, part 4. - P.244-251. ISSN: 1313–2571, published at: http://www.scientific-publications.netм

168.Sharipbay A., Barlybayev A., Buribayeva A., Yergesh B. Speech user interfaces and contents for e-learning. Journal of International Scientific Publications: Educational Alternatives. - Volume 10, Part 4. ISSN:1313-2571. - Bulgaria, 2012. - P.244-251.

169.Бурибаева А.К. Речевые технологии для e-learning// Л.Н.Гумилев 100 жылдығына арналған Қоғамды ақпараттандыру III Халықаралық ғылыми-практикалық конференция еңбектері. - Астана, 2012. - Б. 516-519.

170.Шелепов В.Ю., Ниценко А.В., Карабалаева М.Х., Бурибаева А.К. О распознавании речи на основе межфонемных переходов// Вестник Евразийского национального университета им. Л.Н.Гумилева. – Специальный выпуск. - Астана, 2012. - С.24-29.

171. Шарипбаев А.А., Бурибаева А.К., Калиев А.К. Гласные звуки казахского языка и методы их выделения в начале слова //Доклады Национальной Академии наук РК. - №6.- 2012. - С.10-16.

НИИ Искусственный интеллект

172.Бурибаева А.К., Дорохина Г.В., Ниценко А.В., Шелепов В.Ю. Сегментация и дифонное распознавание речевых сигналов// Труды СПИИРАН. - Вып. 31. – 2013. - С.20–42.

173.Buribayeva A.K., Yergesh B.Zh., Mukanova A.S. Speech user interfaces and contents for e-learning// AWERProcedia Advences in APlied Sciences. - Vol 1 (2013). - P.594-601.

174.Бурибаева А.К. Распознавание казахских слов на основе дифонной базы// Труды I-ой Международной конференции "Компьютерная обработка тюркских языков". - Астана, 2013. - С.30-39.

175.Бурибаева А.К., Шарипбаев А.А. Метод выделения гласных звуков казахского языка в начале слова// Материалы конференции по компьютерной и когнитивной лингвистике TEL'2014 «Языковая семантика: модели и технологии». - Казань, 2014. - С.55-59.

176.Шарипбай А., Головинов М.В., Шелепов В.Ю., Бекманова Г.Т. Реализация алгоритма звукового распознавания казахской речи мобильными устройствами// Вестник Евразийского национального университета им. Л.Н.Гумилева. – № 3. - Астана, 2014. - С.86-94.

177.Buribayeva A., Sharipbay A. Kazakh Vowel Recognition at the Beginning of Words // Mediterranean Journal of Social Sciences ISSN 2039-9340(Print) ISSN 2039-2117(Online), Vol 6. No 2 S4. - P 121-127.

178.Buribayeva A., Sharipbay A., Bekmanova G. Kazakh words recognition based on diphone database // Proceedings of the III-th International Conference on Turkic Language Processing (TurkLang). - Kazan, Tatarstan, 2015. - P.454-466.

179.Buribayeva A., Sharipbay A. Method of vowel recognition at the beginning of words//The International Arab Journal of Information Technology», vol. 2, 2015 (IAJIT), импакт-фактор - 0,39.

Электронное обучения и интпеллектуальная оценки знаний

НИИ Искусственный интеллект

180.Шарипбай А.А., Мамырулы Т. Қазақ тілін автоматты окытуға арналган формалды сипаттамасы// Вестник МН-АН РК. - № 5. - Алматы, 1997. - С.47-51.

181.Шарипбай А.А., Дуйсенбаев Д. Қазақ тілін окытуды автоматтандыру// Материалы РК КО ПП. – Алматы: КазГУ, 1998.

182.Шарипбай А.А., Дузельбаева А.С. Электронный учебник по делопроизводству на казахском языке// Труды международной конференции «Развитие информационных технологий в высшей школе». КазНУ им. Аль-Фараби, 2003. - С.176-180.

183.Шәріпбай А.Ә., Омарбекова А. Қазақ тілінде Іс жүргізудің мультимедиялық оқыту программасы// Материалы 1 международной научно-практической конференции «Информатизация общества». - Астана: ЕНУ им. Л.Н.Гумилева, 2004. - С.81-83.

184.ШәріпбайА., Омарбекова А., Туреханов С. Қазақ тілінің мультимедиялық оқыту программасы// Материалы I международной научно-практической конференции «Информатизация общества». – Астана: ЕНУ им. Л.Н.Гумилева, 2004. - С.98-101.

185.Шарипбай А.А., Омарбекова А.С. и др. СТ РК 34.017. Информационная технология. Электронное издание. Электронное учебное издание. - Астана, 2005. -31С.

186.Шарипбай А.А., Туреханов С. Қазақ тілі грамматикасын автоматты оқыту мәселері// Труды международного форума “Информатизация системы образования”. - Алматы, 2002. - С.5-9.

187.Шарипбай А. Построение онтологической модели генерации интеллектуальных электронных учебных изданий// Труды III международной научно-практической конференции «Информатизация общества». - Астана: ЕНУ им. Л.Н.Гумилева, 2012. - С.580-584.

188.Шарипбай А., Омарбекова А., Сундетова А., Исмаганбетов Б. Автоматизация создания электронного обучения// Вестник ЕНУ им. Л. Гумилева. Спецвыпуск. – Астана, 2012. - С.36-41.

НИИ Искусственный интеллект

189.Шарипбай А., Омарбекова А.С., Исмагамбетов Б., Сундетова А. Интеллектуальные обучающие системы// Вестник ЕНУ им. Л. Гумилева. Спецвыпуск. – Астана, 2012. - С.354-356.

190.Шарипбай А., Омарбекова А.С., Кинтонова А.Ж. Онтологическое моделирование в автоматизированных обучающих системах// Труды III международной научно-практической конференции «Информатизация общества». - Астана: ЕНУ им. Л.Н.Гумилева, 2012. - С.533-535.

191.Шарипбай А., Омарбекова А.С., Кинтонова А.Ж., Барлыбаев А. Инструментальные средства разработки электронных учебников// Вестник Национальной Академии наук РК. - №6. - Астана, 2012. - С.3-11.

192.Buribayeva A., Barlybayev A., Sharipbay A., Yergesh B. Speech user interfaces and contents for e-learning// Journal of International Scientific Publications: Educational Alternatives. - Volume 10, part 4. - P.244-251. ISSN: 1313–2571, published at: http://www.scientific-publications.net.

193.Sharipbay A., Seifullina A., Omarbekova A. Building automation ontological generation of intellectual electronic textbooks// The 6th International Conference on Soft Computing and Intelligent Systems. - Kobe, Japan, 2012. - P.17-19.

194.Sharipbay A., Omarbekova A., Seifullina A. Automatization create electronic learning// Journal of International Scientific Publications: Education Alternatives. - Volume 10, Part 1, ISSN 1313-2571. - Bulgaria, 2012. - P.242-250.

195.Sharipbay A., Omarbekova A., Zakirova A. Frame concept for generator of electronic educational Publications// The 14th International Symposium on Advanced Intelligent Systems (ISIS2013). - Daeieon, Korea, 2013. - P.259-264.

196.Sharipbay A., Omarbekova A., Seifullina A. Creating Intelligent Eletronic Textbooks based on Knowledge Management// World Congress on Enginering2013. - London, U.K., 2013. - C.351-357.

НИИ Искусственный интеллект

197.Sharipbay A., Barlybaev A., Omarbekova A., Akimbekova G.M. Easy Definition of the Situation About the Learning Process of Students, Pensee Journal, 75 (10). - France, 2013. - P.150-155.

198.Sharipbay A., Omarbekova A., Seifullina A. The Creation Intelligent E-learning System by Higher Education in Kazakhstan// The 14th International Symposium on Advanced Intelligent Systems (ISIS2013). November 13 – 16, 2013/Daeieon, Korea. - P.115-121.

199.Sharipbay A., Omarbekova A., Zakirova A., Barlybayev A. Generation of intelligent e-learning editions on the basis of ontologic model// Journal of International Scientific Publications: Education Alternatives, Volume 11, Part 1. - Bulgaria, 2013. - P.47-53.

200.Sharipbay A., Omarbekova A. Generator of electronic educational publications// 4th International Academic Research Conference on Business, Education, Nature and Technology. - Seattle, USA, 2013. - P.241-248.

201.Sharipbay A., Barlybayev A., Omarbekova A. Semantic Model of an Intellectual e-university for the Environmental Education// 1st Global Conference on Environmental Studies: AWERProcedia dvances in APlied Sciences. - V.1. -Turkey, -Antalya, 2013. - P.602-608.

202.Sharipbay A., Barlybayev A., Omarbekova A. Semantic Model of an Intellectual e-university for the Environmental Education// AWERProcedia Advances in APlied Sciences Proceedings of 1st Global Conference on Environmental Studies (CENVISU2013), Vol 1. ISSN: 2301-2706. – Turkey, Antalya, 2013. - P.602-608.

203.Шарипбай А., Омарбекова А.С., Кинтонова А.Ж., Ниязова Р.С., Барлыбаев А. Генерация интеллектуальных электронных учебных изданий на основе онтологической модели// Материалы международной научно-технической конференции «Открытые семантические технологии проектирования интеллектуальных систем». - Минск: БГУИР, 2013. - С.551-554.

204.Sharipbay A., Omarbekova A., Seifullina A., Zakirova A. Creating Intelligent Electronic Textbooks based on Knowledge Management// The World Congress on Engineering 2014 (WCE 2014). Volume 1. – 2014. - P.224-227.

НИИ Искусственный интеллект

205.Sharipbay A., Omarbekova A., Nurgazinova G. Creating Intelligent Electronic Textbooks in Kazakhstan// 6th International Conference on Education and New Learning Technologies. - Barcelona, Spain, 2014. - P.2926-2933.

206.Sharipbay A., Omarbekova A., Zakirova A. Frame concept for generator of electronic educational Publications. Advances in Intelligent Systems and Computing. Soft Computing in Machine Learning. Volume 273, 2014. - P 1120. (Springer).

207.Sharipbay A., Turmaganbetova Sh., Omarbekova A. The creating ontological model of the tax systems in Kazakhstan. Advances in Intelligent Systems and Computing. Soft Computing in Intelligent Control.Volume 272, 2014. - P 63-67. (Springer)

208.Sharipbay A., Barlybayev A., Sabyrov T. Measure the Usability of Graphical User Interface. New Advances in Information Systems and Technologies New Advances in Information Systems and Technologies, Advances in Intelligent Systems and Computing. 2016. – Vol.444, ISSN: 2194-5357. P.1037-1045. (Scopus).

209.Sharipbay A., Barlybayev A., Sabyrov T., Akimbekova E. Software for automation of activity of the personnel in the smart-universit. Proceedings of EDULEARN16 Conference. - Barcelona, 2016. - ISBN: 978-84-608-8860-4. - P.3731- 3735. (Thomson Reuters Web of Science).

210.Kuzenbayev B.A., Sharipbay A.A., Melnikov A.V., Niyazova R.S. Intellectual management of the educational process of higher educational institutions on the basis of ontological modeling// Journal of Theoretical and APlied Information Technology. 2016. Vol.89, Issue 2. - P.404-408. (Scopus).

211.Кузенбаев Б.А., Шарипбай А.А., Ниязова Р.С. Экспертные методы оценки, качества и надежности автоматизированной системы управления образовательным процессом// Труды международной конференции «Открытые семантические технологии проектирования интеллектуальных систем». - Минск, 2016. - С.527-530.

НИИ ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ

212.Шарипбай А.А., Барлыбаев А.Б., Сабыров Т.С. Представление и обработка знаний, хранящихся в многомерных данных посредством расширенного базового семантического гиперграфа// Труды международной конференции «Открытые семантические технологии проектирования интеллектуальных систем». - Минск, 2016. - С.197-202.

213.Nurgazinova G., Golenkov V., Sharipbay A., Omarbekova A., Barlybayev A. On creation of the intelligent help system// Journal of Theoretical and APlied Information Technology. 2016. -Vol.86, Issue 1. - P.26-33. (Scopus).

214.Nurgazinova G., Golenkov V., Sharipbay A., Omarbekova A., Barlybayev A. Designing of Intelligent Reference System for Algebra Based on the Knowledge Database// international conference on control, automation and artificial intelligence. 2015.- ISBN:978-1-60595-278-9. - P.230-235. (Thomson Reuters Web of Science).

215.Шарипбай А. Моделирование оценки знаний испытуемого// Труды международной научной конференции «Информатика и прикладная математика», посвященной 25-летию Независимости Республики Казахстан и 25-летию Института информационных и вычислительных технологий. - Алматы, 2016. - С.356-361.

216.Шарипбай А.А., Барлыбаев А., Сабыров Т., Улюкова Г.Б. Программа обработки базы знаний для интеллектуальных систем// Вестник ЕНУ имени Л.Н. Гумилева. № 4. –Астана, 2016. -С.51-60.

217.Barlybayev A., Sharipbay A., Sabyrov T. Student’s performance evaluation by fuzzy logic. Procedia Computer Science. 2016.- ISSN: 1877-0509. - P.98-105. (Scopus).

218.Sharipbay A., Barlybayev A. An Intelligent System for Learning, Controlling and Assessment Knowledge. INFORMATION, 18 (5(A)). ISSN: 134.3-4500. – Japan, 2015. - P.1817-1828.

219.Sharipbay A., Barlybayev A. Formal Models of the Intelligent Electronic University. International Journal of Information Technology & Computer Science (IJITCS). Volume 21: Issue No.: 1 / August 2015. ISSN : 2091-1610.

НИИ Искусственный интеллект

220.Шарипбай А., Омарбекова А.С. Генератор электронных учебных изданий на латинице// Материалы международной научно-практической конференции «Научные исследования и их практическое применение». - Одесса, 2013. - С.101-109.

221.Шарипбай А., Омарбекова А.С. Технология создания электронных учебных изданий на латинице// Труды I международной конференции "Компьютерная обработка тюркских языков". – Астана, 2013. - С.336-342.

Латинизация казахской письменности

222.Шәріпбаев А.А. Латынсыз компьютерсіз көзіміз ашылмайды// Егеменді Қазақстан”, 27.12.2000.

223.Sharipbay A., Anuarbekov K. Conversion of kazakh writing into latin alphabet project // The Proceedings of the I International Conference on Computer processing of Turkic Languages. – Astana: L.N.Gumilyov Eurasain National University, 2013. - P.121-140.

224.Шарипбай А. Қазақ жазуын латын әліпбиіне көшіру жобасы // «Латын әліпбиіне көшу: тәжірибе және ықпалдастық» атты халықаралық ғылыми-тәжірибелік конференциясы. – Астана, 7-маусым, 2013. - Б.27-34.

225.Шарипбай А. Қазақ жазуын латын әліпбиіне көшіруге ұсыныстар // Қазақ әліпбиін латын қарпіне көшірудің өзекті мәселелері: Республикалық ғылыми-практикалық конференция материалдары. – Астана, 2013. - Б.48-62.

226.Шарипбай А. Қазақ жазуын латын әліпбиіне көшірудің ғылыми, технологиялық және методикалық негіздері // Қазақстан 2050: стратегиясы және латын әліпбиіне көшу үрдісі. Халықаралық ғылыми-практикалық симпозиум. – Алматы, 2013. - Б. 37-43.

227.Шарипбай А. Қазақ тіліне кірме терминдерді латын графикасымен таңбалау алдындағы өтпелі ережелер // ҚР Мәдениет және ақпарат министрлігі Тіл комитеті өткізген «Кірме терминдерді латын графикасымен таңбалау мәселелері» атты республикалық дөңгелек үстелі. – Астана: Л.Н.Гумилев атындағы ЕҰУ, 2013. - С.23-29.

НИИ УрГУ Государственный интеллект

228.Шарипбай А., Омарбекова А.С. Конвертация текста на казахском языке с кириллицы на латиницу // Труды I международной конференции "Компьютерная обработка тюркских языков". – Астана: ЕНУ им.Л.Н.Гумилева, 2013. - С.115-121.

229.Шарипбай А. Қазақ тілі жазуын латын графикасына көшіру жобалары // Л.Н. Гумилев атындағы ЕҰУ Хабаршысы. - № 3-4. – Астана, 2013. - Б. 58-64.

230.Шәріпбай А. Қазақ жазуын латын әліпбиіне ауыстыру негіздемесі. – Астана: Л.Н.Гумилев атындағы ЕҰУ, 2014. – 43 Б.

231.Шәріпбай А. Қазақ жазуын латын әліпбиіне көшіруге барысында қажетті орфографиялық сөздік: Сөздік. - Астана: Л.Н.Гумилев атындағы ЕҰУ, 2014. – 87 Б.

Диссертации

232.Омарбеков А.С. Разработка технологии автоматизации создания электронных учебных изданий. Дис.канд. техн. наук: по специальности 05.13.11 - Математическое и программное обеспечение вычислительных машин и сетей / Астана: ЕНУ им. Л. Н. Гумилева, 2007.

233.Разакова Б.Ш. Разработка методов синтаксического анализа словосочетаний и предложений казахского языка. Дис.канд. техн. наук: по специальности 05.13.11 - Математическое и программное обеспечение вычислительных машин и сетей /Астана: ЕНУ им. Л. Н. Гумилева, 2008.

234.Бекманова Г.Т. Разработка методов звукового распознавания слов на основе их морфологического анализа и синтеза. Дис.канд. техн. наук: по специальности 05.13.11 - Математическое и программное обеспечение вычислительных машин и сетей / Астана : ЕНУ им. Л. Н. Гумилева, 2010.

235.Карабалаева М.Х. Математическое и программное обеспечение пофонемного распознавания казахской речи на основе сегментации речевого сигнала: Дис.кан. тех. наук: по специальности 05.13.11 - Математическое и программное

НИИ Искусственный интеллект

обеспечение вычислительных машин, систем и сетей / Астана, ЕНУ им. Л.Н.Гумилева, 2010.

236.Бекманова Г.Т. Разработка методов звукового распознавания слов на основе их морфологического анализа и синтеза. Дис.доктора PhD : по специальности 6D060200-Информатика / Астана: ЕНУ им. Л.Н.Гумилева, 2010.

237.Есенбаев Ж.А. Распознавание казахской речи по определенной словарной базе в условиях шумов. Дис.доктора PhD : 6D060200-Информатика / Астана: ЕНУ им. Л.Н.Гумилева, 2014.

238.Бурибаева А.К. Распознавание казахских слов на основе дифонной модели Дис.доктора PhD : по специальности 6D060200-Информатика / Астана : ЕНУ им. Л.Н.Гумилева, 2015.

239.Муканова А.С. Модели и методы семантической обработки знаний на казахском языке. Дис.доктора PhD : 6D060200-Информатика / Астана : ЕНУ им. Л.Н.Гумилева, 2015.

240.Кабенов Д.И. Модели и методы интеллектуальной оценки знаний обучающегося. Дис.доктора PhD : 6D060200-Информатика / Астана : ЕНУ им. Л.Н.Гумилева, 2015.

НИИ Искусственный интеллект

ШАРИПБАЙ Алтынбек Амирович

ПРОБЛЕМЫ ПЕРЕВОДА КАЗАХСКОЙ ПИСЬМЕННОСТИ НА ЛАТИНСКИЙ АЛФАВИТ

Монография

Тираж 100 экз.

Формат 60х84 1/16. Усл. 8,6 п.л.

Бумага офсетная. Шрифт «Times New Roman»

Заказ № 15602.

Отпечатано в типографии ЕНУ имени Л.Н.Гумилева

РК, Астана, 010008, ул. Кажымукана, 13/1

НИИ Искусственный интеллект