

Масуд Махмудов

APFL= 'АНГЛАРИМ';
*
APFL= 'АНГЛАРЫ';
*
APFL= 'АНГЛАРЫ';
*
APFL= 'АНГЛАРЫ';
*
APFL= 'АНГЛЫ';
*
APFL= 'АНГР';
АНГ
L= 3;
APFL= 'АНГР';
AFFL= 'ЛАРЫМЫЗ';
К= 7;
APFL= 'ЛАРЫМЫЗ';
*
APFL= 'ЛАРЫМЫ';
*
APFL= 'ЛАРЫМЫ';
*
APFL= 'ЛАРЫМ';
*
APFL= 'ЛАР';
Л= 3;
= 'ЛАР';
AFFL= 'МММ3';
К= 4;
APFL= 'МММ3';
МММ3
L= 4;
AFFL= 'МММ3';
AFFL= 'ПОСЛЕДН.'
SLDOV= 'ПОСЛЕДН.', АФ



СЛОВО = ОСНОВА СРАВНИ

СИСТЕМА АВТОМАТИЧЕСКОЙ ПЕРЕРАБОТКИ ТЮРКСКОГО ТЕКСТА НА ЛЕКСИКО-МОРФОЛОГИЧЕСКОМ УРОВНЕ

АЗЕРБАЙЖАН ЕЛМƏР АКАДЕМИЯСЫ
НӘСИМИ АДЫНА ДИЛЧИЛИК ИНСТИТУТУ

АКАДЕМИЯ НАУК АЗЕРБАЙДЖАНА
ИНСТИТУТ ЯЗЫКОЗНАНИЯ им. НАСИМИ

МӘС'УД АХМЕД ОГЛУ МАХМУДОВ

ЛЕКСИК-МОРФОЛОГИ СƏВИЙЈƏДƏ
ТҮРК ДИЛlЕРИ МƏTİLLƏRİNİN
АВТОМАТИК İŞLƏNMƏSİ

БАКЫ - ЕЛМ - 1991

Масуд Ахмед оглы Махмудов

СИСТЕМА
АВТОМАТИЧЕСКОЙ ПЕРЕРАБОТКИ
ТЮРКСКОГО ТЕКСТА
НА ЛЕКСИКО-МОРФОЛОГИЧЕСКОМ
УРОВНЕ



БАКУ «ЭЛМ» 1991

ПРЕДИСЛОВИЕ

Спонсором издания является
Институт языкоznания им. Насими

Редактор к.филол.н. В.Я.ПИНС

Рецензенты: д-р филол.н., проф. А.А.АХУНДОВ
к. филол.н. К.А.ВЕЛИЕВА

Махмудов М.А.

Система автоматической переработки тюркского текста на лексико-
морфологическом уровне. - Баку: ЗЛМ, 1991. - 144 с.

ISBN 5-8066-0491-8

В работе построена лингвистическая модель системы автоматической
переработки тюркского текста, создан унифицированный алгоритм анализа тюркской словоформы, который может быть использован в качестве одного из основных модулей различных систем автоматической переработки текстов на тюркских языках, в том числе систем машинного перевода. Анализ на лексико-морфологическом уровне ориентирован на синтаксические процедуры, которые являются следующим шагом в процессе машинной обработки тюркских текстов. Особое внимание уделяется автоматическому распознаванию монимии корневых и аффиксальных морфем.

1404000000-236 Заказное
655(07)-91



Махмудов М.А., 1991

В настоящее время в связи с бурным развитием науки и техники резко возрастают объемы информационных потоков. Создание систем, осуществляющих оперативную переработку и анализ этих потоков - одна из актуальных и чрезвычайно важных проблем, стоящих перед специалистами и учеными, работающими в области кибернетики, информации и языкоznания.

Поиск, хранение и передача научно-технической информации, автоматическое составление рефератов и свертывание текстов с различными целями, машинный перевод /МП/, дешифровка текстов, а также построение обучаемых автоматов во многом зависят от разработки проблем искусственного интеллекта - направления, задачей которого является создание машин, имитирующих работу человеческого разума, в том числе его лингвистического поведения /91; 96; II2/.

Известно, что задача полного автоматического анализа текста пока еще находится в начальной стадии ее решения, однако научно-технический прогресс и развитие культурных, экономических связей в обществе требуют скорейшего выполнения социального заказа на построение систем лингвистического обеспечения АСУ и ИПС, семантического МП, а также на разработку лингвистических аспектов искусственного интеллекта /25; 40; 62; 75; 88; 87; 107/.

Смысловой и грамматический анализ с помощью ЭВМ является центральным модулем всей автоматической системы переработки текста /АПТ/. Вместе с тем, при разработке профилей АПТ и МП в настоящее время речь чаще всего идет о пословно-пооборотной переработке текста. Однако, остановиться на данном этапе и отказаться от разработки более полного и глубокого анализа текста было бы ошибочно /18; 21/.

В частности, если говорить о тюркских и вообще агглютинатив-

ных языках, то задача автоматического анализа должна здесь решаться иначе, чем это делается относительно флексивных и флексивноаналитических языков.

В индоевропейских языках любая словоформа /с/ф/ употребляется с одним или двумя аффиксами и приведение этих с/ф к словарному виду не представляет особых трудностей. Что же касается тюркских языков, то здесь словоформы могут содержать значительное число морфологических показателей, причем от каждой основы можно образовать тысячи форм /51; 75/. Исследования на материале тюркских языков показывают, что "для обеспечения 70 процентного покрытия словоформами наугад взятого текста достаточный объем выборки в агглютинирующем языке должен почти в 5 раз превосходить объем выборки флексивно-аналитического языка и более чем в два раза быть больше, чем выборка флексивно-синтетического языка" /12,87/. Поэтому для АПТ на тюркских, как, впрочем, и на всех агглютинативных языках, приходится строить автоматический словарь /АС/, включая машинные или традиционные основы /канонические формы/, к которым с помощью специального алгоритма должны приводиться текстовые словоформы. Вопрос этот чрезвычайно важен, поскольку применяющиеся в действующих системах МП алгоритмы анализа, оперирующие со словоформами, сводят до минимума морфологический анализ с/ф. Решение этой проблемы особенно актуально для агглютинирующих тюркских языков, для которых в силу особой специфики структуры с/ф, пословно-пооборотный МП малоэффективен.

Исходя из приведенных выше соображений целью настоящей работы является разработка принципов автоматического морфологического анализа тюркской словоформы и построение математической модели морфологии тюркского текста с последующей реализацией ее в виде воспроизводящей инженерно-лингвистической модели /ВИЛМ/

на материале азербайджанского языка. Созданный на базе этой модели машинный алгоритм автоматического морфологического анализа /AMA/ азербайджанского текста, в свою очередь, может быть использован в качестве одного из основных модулей различных систем АПТ - машинного перевода, индексирования, аннотирования, реферирования и т.д.

ГЛАВА I

МАШИННАЯ МОРФОЛОГИЯ ТУРСКИХ ЯЗЫКОВ

I. Место морфологического анализа в системе АЛТ.

Автоматический морфологический анализ /AMA/ является одним из наиболее важных и необходимых этапов лингвистического анализа при АЛТ. В результате АМА при обработке отдельно взятых с/Ф обеспечивается получение первичных грамматических характеристик необходимых для последующих этапов анализа. Формализация морфологического анализа дает возможность автоматизировать процессы сбора, накопления, поиска, хранения и передачи информации в автоматизированных информационных системах /16; 17/.

По существу, АМА - это переход от формы /графической/ к смыслу. Однако, с помощью одного морфологического анализа полный переход от графической формы к смыслу невозможен. Омонимия /омография/, порядок слов и т.д. "регулируются" с помощью синтаксических-семантических блоков /рис. I/. Эти блоки функционируют после морфологического анализа. Слова, представляющие трудность для синтаксического анализа /омонимичные, многозначные слова, которые не зафиксированы в АС/, проверяются при помощи семантического блока. Семантический блок в свою очередь связан с морфологическим блоком. В блоке морфологического анализа "трудноанализируемые" слова при помощи семантического блока пересматриваются и потом снова возвращаются в синтаксический блок. В зависимости от специфики отдельных языков при расширении и совершенствовании алгоритмов автоматического лингвистического анализа /АЛ/ морфологический, синтаксический и семантический блоки при необходимости подвергаются изменениям, дополняются^x /18; 79; 86/.

^x Одной из первых посвященных морфологическому анализу при машинном переводе с татарского языка на русский в тюркологии

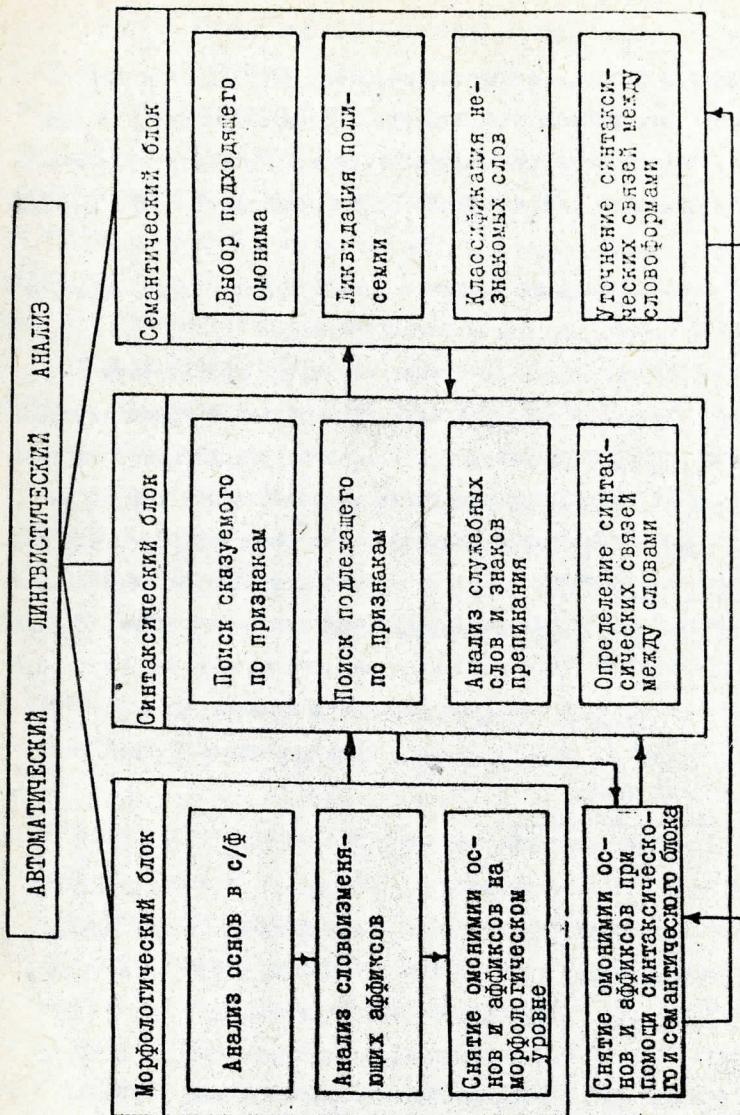


Рис. I. Общая схема автоматического лингвистического анализа

Структура алгоритма морфологического анализа не является существенной для последующих этапов. Задача заключается в том, чтобы получить в результате морфологического анализа грамматические характеристики с/ф, необходимые для синтаксического анализа /45; 83/.

2. О необходимости создания машинной морфологии тюркских языков с АПТ.

На современном этапе развития инженерной лингвистики нецелесообразно вводить в машину только те сведения о грамматическом строе языка, которые необходимы для решения ограниченной частной задачи. Гораздо эффективнее исходить из более общей стратегии предварительного построения уже целостной действующей модели языкового механизма с той или иной степенью детальности. Такой подход позволяет решение исходной задачи свести к одному из частных аспектов приложений этого машинного аналога естественного языка, а построенная модель становится гарантией практической разумности целого ряда не только заранее предусмотренных, но и не предусмотренных задач.

продолжение сноски со стр8

считается работа Н.А.Халитовой, Р.А.Закировой, Р.У.Гимадутдиновой. Авторами выделяются следующие этапы анализа: Общие правила /А/, Список информации /Б/, Список информации с суффиксами /В/. Нестандартная запись /Г/. Омонимичные случаи /Д/.

Программа для выборки информации к словоформам и их анализа составлялась для небольшого математического отрывка /120с/ф/. Поиск основы в словаре осуществляется побуквенным сравнением текстовой единицы с основами словаря. Каждой основе и аффиксу приписывалась соответствующая информация. Авторами составлена программа для поиска аффиксов по специальной таблице /110/.

II

Единицы и категории "машинной морфологии" не обязательно могут быть тождественны единицам и категориям, предоставленным в обычных грамматических описаниях. При таком широком "машинноморфологическом" подходе к языку, как к объекту кибернетического моделирования подобные расхождения сводятся к минимуму, а если они и обнаруживаются, то выявление их причин бывает полезным не только для разработчиков модели, но и соответствующего раздела "чистого" теоретического и прикладного методического языкознания.

При всей прикладной ориентации данного исследования оно несомненно, представляет и общетеоретический интерес как для тюркологии, так и для типологических разделов языкознания.

В данном исследовании мы опираемся на выработанную целостную лингвистическую концепцию системно-структурного семиологического поэтапно-уровневого и многосвязно-межуровневого моделирования единиц текста, составляющих диалектическое единство речи и языка.

При этом необходимо выяснить следующие вопросы:

- возможности и границы формального анализа тюркской словоформы;
- возможности построения полного цикла лингвистического моделирования.

Для решения первого вопроса принят гипотеза, согласно которой тюркская словоформа представляет собой такую лингвистическую единицу, которая может быть в подавляющем большинстве случаев формально проанализирована с точки зрения ее разделения на основу и аффиксы и с которой может быть снята конверсионная омонимия без обращения к окружающему контексту. Иными словами, в работе выдвигается гипотеза об относительной автограмматичности и автосемантичности тюркской словоформы.

Опыт промышленной эксплуатации систем лексического АП /6; 18; 26; 27; 35; 41; 61; 117/ автоматического рефериования /5/ и других систем АП, выдающих пословно-оборотную переработку текста без учета грамматики, показывает, что область применения этих систем ограничивается изолирующими языками типа китайского, а также флексивно-аналитическими языками типа английского. Что же касается языков с развитой флексией и агглютинативной аффиксацией /в частности тюркских языков/, то здесь лексико-фразологическая АП оказывается бесперспективной, поскольку одна глагольная или именная основа может порождать в этих языках сотни и даже тысячи с/ф. Поэтому система АП указанных языков должна обязательно включать блок морфологического анализа /и синтеза, если речь идет о двухязычной ситуации/ /83, 272-290/ *.

* Так, максимальное число словоформ, которые могут быть образованы от одной глагольной основы, для татарского языка составляет 17947, для турецкого 11690, для узбекского 1392 /51,17/. Даже при длине цепочки аффиксов, ограниченной пятью формантами, соответствующий показатель для азербайджанского языка превышает 8000. Максимальное же число аффиксов в составе азербайджанской глагольной словоформы доходит до десяти /75,104/. Например, азербайджанская словоформа *мектеблеримиздекилдердендирилрем*? состоит из 10 морфем и эта сложная композиция в целом переводится как "эти и те, которые есть у нас в школе?".

В казахском языке только у существительных количество формообразующих аффиксов достигает 500, а многие казахские глаголы имеют до 1000 различных форм /12,39/. Здесь привлекает внимание еще один важный факт: если английские научно-технические подызыки используют не более 18 тыс. разных словоформ, во флексивных язы-

Опыт моделирования лексики и морфологии некоторых индоевропейских языков /65; 101; 102; 108; 104; 105/ убедительно показал, что в связи с инженерно-лингвистическим моделированием и АП необходимо создание машинных морфологий, "опирающихся на нетрадиционное видение так называемых машинных флексий и создание словаря машинных основ /квазиоснов/"/ /42,245; 58/.

В нашем понимании машинная морфология тюркских языков /ММЯ/ - это морфология особого рода, позволяющая проведение полного морфологического анализа и получение необходимых грамматических характеристик для последующих этапов лингвистического анализа по чисто формальным признакам при АП с различными целями. При этом следует выявить и учсть целый ряд специфических особенностей словообразования и словоизменения тюркских языков, которые обычно в явной форме не фиксируются в грамматиках /50; 55; 97; 100/.

Дело в том, что процессы анализа с/ф человеком и машиной коренным образом отличаются друг от друга. Если человек при этом опирается, помимо формальных признаков, также на другие лингвистические и экстралингвистические знания, то ЭМ такой возможностью не располагает *. Информация, предназначенная для ввода ЭМ, продолжение сноски со стр. 12.

как их число достигает 20-25 тыс., а в агглютинативных, например, в казахском, число разных с/ф увеличивается до 60 тыс. /52,13/. Частотный словарь языка романа М.Ауэзова "Абай жолы" охватывает 466 тыс. словоупотреблений, в которых имеется 61 тыс. различных с/ф. Аналогичную картину дают и другие тюркские языки /14/.

* Известно, что "машина не воспринимает", а опознает слово, так как она имеет дело с его кодом, а не с "образом", поэтому процесс опознания слова сводится к выполнению строго определенных последовательных операций, приводящих к искомому коду /46/.

должна быть предельно эксплицитной.

Создание ММ в связи с АЛТ необходимо для исследования логики построения тюркской с/ф и создания ее структурно-функциональной модели, а также проверки с помощью лингвистического автомата /ЛА/. Без ММ и морфологического блока этой модели невозможна автоматическая переработка тюркского текста, поскольку, как уже отмечалось выше, каждая основа в этих языках образует несколько тыс. с/ф, которые практически нельзя задать в словаре, как это делается для аналитических языков /77; 97/.

Вообще, при морфологическом анализе считается целесообразным использование машинных основ и аффиксов в сочетании с системой алгоритмов АМА /43; 44; 72; 78/.

Прежде чем перейти к описанию ММЯ необходимо дать формальное рабочее определение частей речи и грамматических категорий в машинной морфологии, что имеет существенное значение при построении воспроизводящей инженерно-лингвистической модели ММЯ.

3. Формальная классификация частей речи и грамматических категорий в машинной морфологии.

В тюркских языках части речи традиционно классифицируются в основном в зависимости от смыслового содержания слова /лексического значения/, принадлежности к той или иной системе словообразования или словоизменения /морфологических особенностей/, а также синтаксических функций /3; 9; 10; II; 49; 113; ср. также 29/.

Исходя из этих признаков в тюркских языках первичные и вторичные морфемы делятся на следующие три группы /II/:

1/ корневые морфемы, выражающие имя, т.е. название предмета или статического признака предмета,

2/ корневые морфемы, выражающие глагол, т.е. обозначающие действие - динамический признак;

3/ корневые служебные морфемы, выражающие грамматические отношения слов в составе словосочетаний и предложений.

К первой группе относятся имена существительные, имена прилагательные, наречия, числительные, междометия и мимесы, местоимения; ко второй группе - глаголы и к третьей - частицы, последлоги, союзы, модальные слова.

Первые две группы относятся к самостоятельным, или знаменательным, частям речи, а третья группа относится к несамостоятельным, или служебным частям речи. Всю совокупность корневых морфем можно подразделить также на два больших класса:

1/ самостоятельные, или знаменательные,

2/ несамостоятельные, или служебные.

Кроме того, в тюркских языках существуют группы слов, которые на морфологическом уровне, вне контекста, нельзя отнести к определенной части речи /омонимичные слова/.

Слова, относящиеся к самостоятельным частям речи, имеют особые, отличные друг от друга системы аффиксального словообразования и словоизменения /96; 116/. Однако, формально некоторые фиксы неразличимы /сиграфы/. Омонимичные слова могут иметь фиксы, присоединяемые к корневым морфемам как первого, так и второго типа. Служебные части речи вообще не присоединяют аффиксальных элементов.

Следует отметить, что в тюркских языках нет строгой дифференциации определенных морфологических признаков частей речи, поэтому не существует единого четкого критерия, лежащего в основе классификации частей речи. Кроме того, в тюркских языках имеются слова, которые в предложении могут выполнять разные синтаксические функции. Так, азербайджанское слово даш "камень" традиционно в лексико-грамматическом плане квалифицируется как имя суще-

существительное, но в сочетании /в препозиции/ с другим существительным оно выполняет синтаксическую функцию прилагательного; ср.: даш дивар "каменная стена" *. Слово јахшы "хорошиЙ", будучи прилагательным, тем не менее в предложении выступает и в функции наречия. Некоторые неопределенные числительные также имеют различные синтаксические функции: чох "много", аз "мало" и др.

Вместе с тем большинство тюркологов по вопросу классификации частей речи в тюркских языках высказывает сходную точку зрения /II, 29; 42; 49; 76-77; 95, 35; II8; II9; I20; I25/. Так, в зависимости от смыслового содержания слова, принадлежности его к той или иной системе словообразования или изменения /морфологических особенностей/, а также синтаксических функций проф. М. Гусейнзаде выделяет четыре группы для частей речи азербайджанского языка /II8, I8/: 1/ самостоятельные /знаменательные/, 2/ звукообразно-подражательные /кальки/, 3/ междометия, 4/ служебные.

Кроме этого, в азербайджанском языке имеются модальные слова, которые образуют отдельную группу /рис. 2/.

Известно, что в машинной морфологии выделение тех или иных грамматических классов и категорий, частей речи опирается на строго формализованные принципы. Поэтому эта грамматика может не совпадать с традиционной грамматикой, хотя в обязательном порядке должна учитывать достижения общей грамматической теории и грамматическое описание конкретного языка и группы языков.

Оставляя в стороне спорные вопросы и опираясь на известные классификации частей речи современного азербайджанского языка

* Аналогичные сложности возникают и в некоторых индо-европейских языках; ср.: конверсионную адъективизацию английского существительного *stone* "камень", в словосочетании *stone wall* "каменная стена".

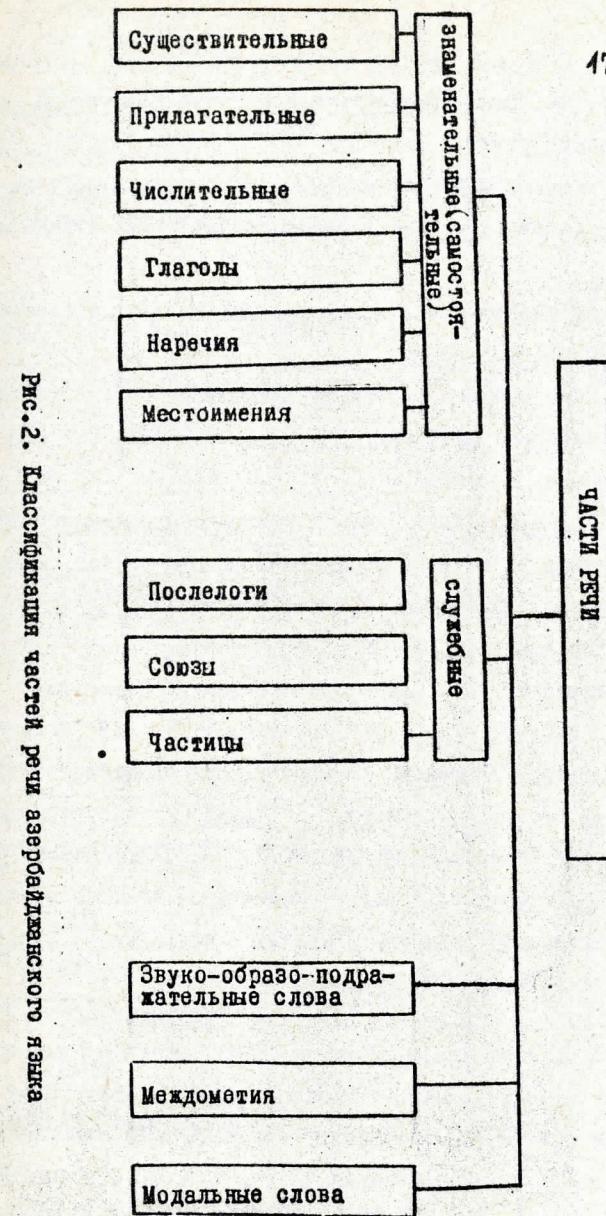


Рис.2. Классификация частей речи азербайджанского языка

/8 34 36;72;115/, мы принимаем следующую рабочую формальную классификацию частей речи:

Части речи в нашей машинной морфологии подразделяются на две группы по признаку отношения к словоизменительным аффиксам /рис. 8/:

I/ Изменяемые /части речи, в которых входят с/ф, содержащие словоизменительные аффиксы/;

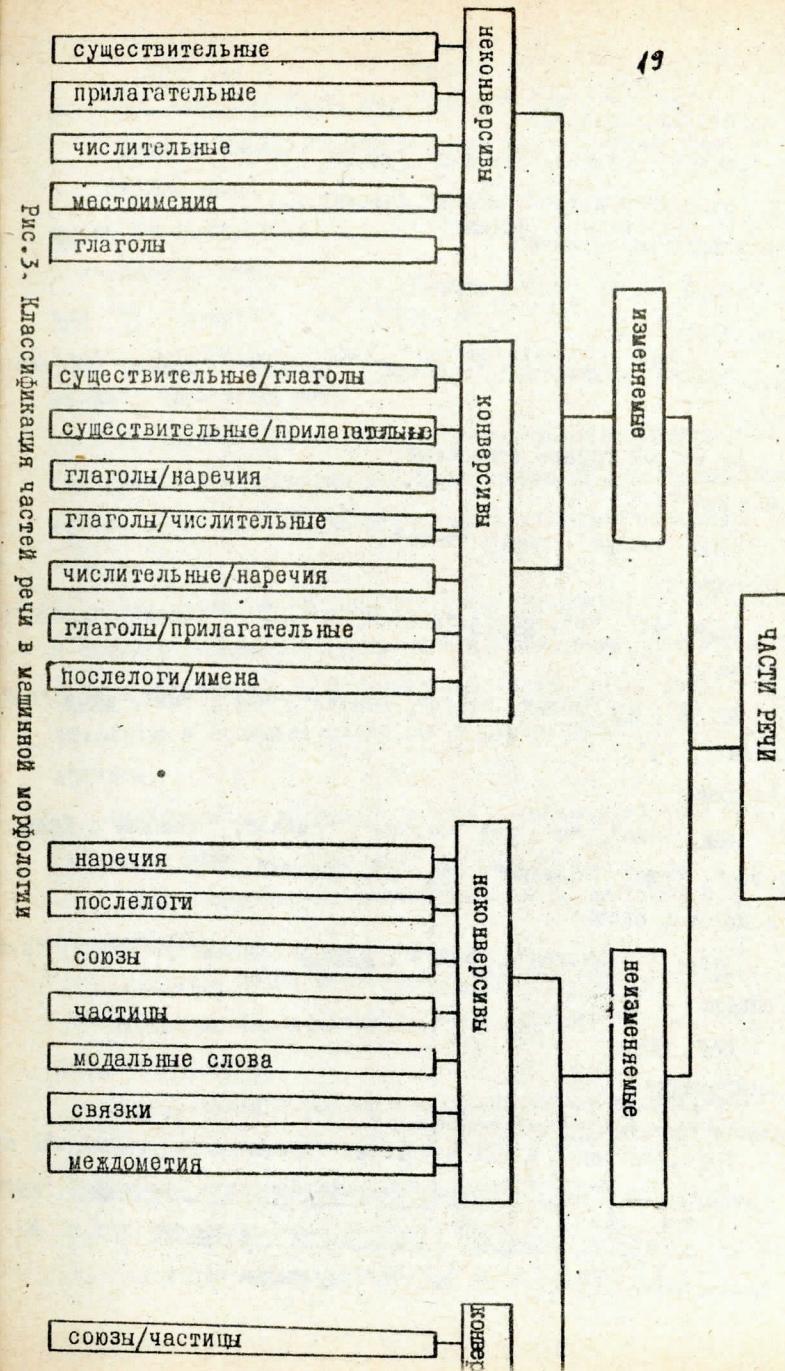
Неизменяемые /части речи, которые охватывают с/ф, не содержащие словоизменительные аффиксы/.

К первой группе относятся:

существительные*: чек "хлеб", инсан "человек"; прилагательные: мави "голубой", берк "крепкий"; числительные: беш "пять", алтынчы "шестой"; местоимения: биз "мы", бу "это"; глаголы: бахмаг "смотреть", көлмек "прийти".

Сюда же следует отнести такие конверсивные классы, как: существительные/глаголы: газ "гусь" /сущев./ // "копай" /глагол/, ат "лошадь" /сущев./ // "бросай" /глагол/; существительные/прилагательные: гоча "старик" /сущев./ // "старый" /прилагат./, кој "зелень" /сущев./ // "зеленый" /прилагат./; глаголы/наречия: дүз "расставь" /глагол/ "ровно", "гладко" /наречие/; глаголы/прилагательные: сары "связывай" /глагол/ // "желтый" /прилагат./, бурма "не крути" /глагол/ " кудрявый" /прилагат./.

* Вуко-образно-подражательные слова по форме словообразования и словоизменения ведут себя как имена существительные. Сочтено целесообразным эту группу слов дать в составе имен существительных, так как при классификации частей речи мы в основном опираемся на слово- и формообразовательную функцию словоформ.



- глаголы/числительные:

мин "садись" /глагол/ "тысяча" /числит./,
гырх "стриги" /глагол/ "сорок" /числит./.

- числительные/наречия:

choх "много" хејли "очень много",

- послелоги/имена:

габаг "перед" /существ./ "впереди", "раньше", "прежде" /после-
лог/.

Ко второй группе относятся:

- наречия ^x:

челд "быстро", ашары "вниз"

- послелоги:

бери "от", "о", соңра "после", "потом", "затем", "далее",

- союзы:

вө "и", да "если", "тоже", "также", "ну", "же", амма "но"
"однако", "а",

- частицы:

ахъ "ведь", "а", лап "совсем", "очень", "слишком", дана
"более", "еще", "больше", бес "а", "разве".

- модальные слова:

дејсан "кажется", "будто", демеки "значит", "есть", "ита

- связи:

иди, имиш

- междометия:

* В азербайджанском языке некоторые наречия могут принимать аффиксы падежа. Напр.: ичеги, ичемире, ичериде, ичериден, кери
керије, кериде, кериден, јахын, јахына, јахыдан и. т.д. Но в
целом словоизменение не присуще наречиям.

ай "эй", "ой", ах "ах", беһ "ба-ба", "ах", оф "оф".

Сюда же следует отнести такие конверсионные классы как:

- союзы/частицы:

анчаг "только", "лишь", "но", "однако", "едва", ки "что"

- послелоги/союзы:

иле "с", "вместе", "по", үчүн "для", "ради", көре "по", "смотря",
"как", "сообразно", "для", "соответственно", етту "для", "ради",
из-за", "впоследствии".

В настоящее время имеется классификация частей речи в рамках машинной морфологии турецкого языка /6,10/. Предложенная нами классификация отличается от вышеуказанной следующими особенностями:

- числительные /порядковые, количественные и разделительные/
и прилагательные относятся не к неизменяемым, как это предлагается в работе /6/, а к изменяемым частям речи, поскольку и чис-
лительные и прилагательные могут присоединять словоизменительные
аффиксы,

- междометия, послелоги и связки рассматриваются как отдель-
ные части речи,

- звуко-образно-подражательные слова включены в состав суще-
ствительных,

- вводятся новые конверсии - глаголы/прилагательные, глаголы/
числительные, числительные/наречия, союзы/частицы, послелоги/со-
юзы,

- предикативы и конверсии существительные/числительные, су-
ществительные/наречия, прилагательные/местоимения, послелоги/ча-
стицы не входят в предложенную классификацию.

В нашей классификации с/ф первой группы, содержащие слово-
изменительные аффиксы, разделяются на следующие подгруппы:

- словоизменительные категории и формы имен /сюда входят категория падежа существительного, категория принадлежности и сказуемости/,

- словоизменительные категории и формы глаголов /сюда входят залоги, наклонения, времена глагола, отрицание, инфинитив, причастие, деепричастие/,

- общие /именные и глагольные/ категории, охватывают категории числа, лица, вопросительности.

В процессе АМА словоформы, относящиеся к неизменяемым частям речи, а также с/ф, содержащие только словообразовательные аффиксы, не анализируются, так как они в качестве самостоятельных лексических единиц включаются в АС основ. При АМА в центре внимания находятся части речи, которые охватывают с/ф, содержащие словоизменительные аффиксы, что в тюркских языках прилагательные, числительные и частично междометия при субстантивации могут принимать словоизменительные аффиксы.

4. Рабочие понятия ММЯ.

Ниже мы будем пользоваться следующими понятиями теории множеств. Под термином множество мы будем понимать совокупность лингвистических объектов /букв, аффиксов, словоформ, предложений и т.д./. Принадлежность какого-либо элемента X к множеству A записывается как $X \in A$. Если каждый элемент X_i множества C принадлежит множеству A , то множество C называется подмножеством множества A . Аналогично, если каждый элемент X_i множества A' принадлежит множеству C' , то множество A' называется подмножеством множества C' .

Произвольные конечные последовательности лингвистических символов X_i будем называть цепочками. Длина цепочки измеряется числом входящих в нее элементов a .

Основное, центральное понятие данной системы – это словоформа. С/ф есть наименьшая формально определимая единица текста, цепочка, состоящая из произвольных конечных последовательностей букв. Длина с/ф измеряется количеством букв, входящих в нее. Отдельные с/ф есть элементы предложения.

Предложение – это цепочка с/ф, расположенных между двумя точками, между точкой и точкой с запятой, между точкой и восклицательным или вопросительными знаками.

В настоящей работе под текстом понимается множество контекто-но-свободных словоформ. При этом синтаксические и семантические связи между с/ф не учитываются.

До сих пор мы рассматривали те элементы текста /речи/, которые имеют четкие формальные границы. Для формального определения этих единиц не требуется никаких дополнительных сведений. Для определения же таких понятий как основа и аффикс необходима лингвистическая информация, обусловленная системой языка. Поэтому после выделения с/ф из предложений мы должны найти "пограничные" зоны, как между основой и цепочкой аффиксов, так и между отдельными аффиксами, что является основной задачей данной работы.

С/ф – образуется конкатенацией основ и словоизменяющих аффиксов. Под термином "основа" понимается корень или сочетание корень – словообразовательный аффикс, которые как равноправные лексические единицы включаются в АС основ. Кроме того, основы могут состоять из сочетания двух и более основ /сложные основы/, также образующих самостоятельные лексические единицы /34,38;121/.

Аффиксы, в отличие от основ, лексического значения не имеют и потому не могут быть единицами АС основ. Присоединяясь к основам, они либо образуют новые основы /лексические единицы/, либо

выражают различные грамматические отношения в предложениях ^x. Исходя из этого, аффиксы делятся на словообразовательные и словоизменительные. При АМА тюркского текста на первом плане стоит обработка словоизменительных аффиксов, так как словообразующие форманты образуют отдельные лексические единицы, которые представлены в АС основ и не подвергаются анализу.

5. Цели и задачи ММТЯ.

Целью построения ММТЯ является разработка теоретических принципов автоматического анализа агглютинативного тюркского текста, построение на их основе математической модели морфологии тюркского текста с последующим представлением ее в виде ВИЛМ. Это целевое задание реализуется на материале азербайджанского языка.

Основная задача заключается в разработке алгоритма распознавания с/ф текста и замены их соответствующими значениями /грамматическими или лексическими/, т.е. той информацией, которая необходима для последующих этапов лингвистического анализа. При этом подразумевается, что между с/ф, составляющими текст, и значениями имеется определенное соответствие.

Эта основная задача реализуется в ходе решения следующих вопросов:

- 1/ отделение текста от текста,
- 2/ выделение предложения из текста,

^x М.Эргин делит аффиксы турецкого языка на две группы: часть из них суть морфемы, которые всегда использовались только как аффиксы, а потому их можно назвать собственно-аффиксами /asıl ekler /; другая часть аффиксов образовалась в результате сложения двух или более аффиксов, либо превращения самостоятельного слова в аффикс /126, 131/.

- 3/ выделение с/ф в предложении,
- 4/ отделение основы от цепочки аффиксов,
- 5/ определение лексико-грамматических признаков основы,
- 6/ выявление грамматических функций каждого выделенного аффикса и приписывание их к признакам основы или переводному эквиваленту основ /ср.: 120/.

Исходя из этих общих и конкретных задач, структура машинной морфологии должна быть задана таким образом, чтобы по чисто формальным признакам можно было бы провести автоматический анализ тюркского текста. Для этого необходимо построить модель процесса АМА /84; 115; 122/.

6. Воспроизводящая инженерно-лингвистическая модель ММТЯ.

При построении модели процесса АМА тюркской словоформы необходимо соблюдать два условия:

1/ каждая лексическая или грамматическая единица языка и речи должна рассматриваться как знаковое образование, имеющее значение /означаемое/ и форму /означающее/.

2/ каждое модельное построение должно быть подвергнуто проверке на его состоятельность, под которым понимается логическая непротиворечивость, лингвистическая осмысленность и достаточная экспланаторность этой модели. Лучшим средством такой проверки является реализация модели на ЭВМ и автоматическая переработка достаточно представительного языкового материала.

С формальной точки зрения процесса АМА - это переход от текста /графической формы речи/ к значению /лексическому или грамматическому/ по определенным правилам, позволяющим сопоставить каждой единице текста /цепочки букв/ соответствующие значения /85; 90/. Модель перехода от текста к значению /от формы к значению/ можно представить следующим образом: $\langle A, F, A' \rangle$

Здесь A - множество с/ф /текот/. $A = \{a_i\}$, где $i = 1 \dots n$ /область определения/, A' - множество значений /грамматических или лексических/ соответствующей с/ф /область значения/ $A' = \{a'_i\}$, где $i = 1, \dots n$

F - композиция функций /совокупность рекурсивных правил/, определенная в A и сопоставляющая для каждого элемента $a \in A$ равна одному $a' \in A'$.

Данную модель можно рассмотреть как отображение множества A в множество A' , т.е. такое соответствие между множествами A и A' , что для всякого $a \in A$ существует $a' \in A'$ /32/.

Множество A' включает все значения, которые соответствуют реальным с/ф. Множество A задано списком /морфемный список/. Морфемные списки состоят из АС основ и списка словоизменяющих аффиксов. Для основ фиксируются лексико-грамматические, а для аффиксов грамматические характеристики.

Единицы морфемного списка /основы и аффиксы/ расположены в соответствии с их длиной. Учитывая количество букв, входящих в цепочку, основы и аффиксы объединяются в \sqrt{N} зон x , причем первая зона включает морфемы, состоящие из x букв, вторая содержит К-1 буквенные морфемы и т.

Алгоритм АМА - как композиция функций F /совокупность рекурсивных правил/.

Как уже было отмечено, в модели /I/ F обозначает композицию функций /совокупность рекурсивных правил/ /63; 94/, которые позво-

ляют перейти от формы $/A/$ к значению $/A'/$. Функция G в данной работе задана алгоритмом. Этот алгоритм можно охарактеризовать как формальную распознавающую грамматику, т.е. совокупность рекурсивных правил, позволяющих задать некоторую характеристику /грамматическую или лексическую/, принадлежащую данному языку.

Отличие данного алгоритма от формальной распознавающей грамматики в том, что если применение правил в такого рода грамматиках произвольно, то в алгоритме задан более жесткий порядок применения отдельных инструкций.

Функцию G можно рассматривать как композицию функций K, P, R, C, Π , полученную с помощью подстановок этих функций друг в друга /59, 50; 108, 2/. Формула, описывающая данную композицию, будет иметь вид: $F/X = \Pi/C/R/P/K/X//$, где функции K, P, R, C, Π имеют следующие значения:

- функция K отделяет текст от текста, предложения из текста, а также выделяет с/ф в предложении,
- функция P сопоставляет с каждой с/ф морфемные списки /АС основ и список аффиксов/ и, таким образом осуществляет расчленения с/ф на основу и аффиксы,
- функция R снимает омонимию основ и аффиксов,
- функция C указывает для каждой морфемы X некоторую информацию /грамматическую или лексическую/,
- функция Π преобразует лексико-грамматическую информацию соответствующей входной с/ф.

Информация о с/ф - это множество некоторых элементарных информаций /дизъюнкция элементарных информаций/. Информация может состоять из одной элементарной информации, являясь однозначной. Сами элементарные информации могут состоять из нескольких приз-

* Зона - это совокупность основ или аффиксов, состоящих из одинакового количества букв. Подзоны объединяют аффиксы с одинаковыми грамматическими характеристиками /позиционные орфографические варианты, алломорфы и т.д./.

наков, связанных друг с другом конъюктивно. Напр., аффикс -ын, -ун, -ин, -ын имеет следующие грамматические характеристики: "родительный падеж" или "принадлежность, 2 лицо, единственное число" или "повелительное наклонение, 2 лицо, единственное число". В данном случае элементарными информацией являются элементы "родительный падеж", принадлежность", "единственное число", "повелительное наклонение", "2 лицо". Информация в целом состоит из пяти элементарных информаций. При работе блока снятия омонимии /функция Р/ выбирается нужная элементарная информация.

7.0. Функция выделения единиц анализа /К/.

Данная функция работает следующим образом: в начале конкретный текст отделяется от множества текстов /перед вводом в начале и в конце текста ставится условный знак, благодаря чему конкретный текст формально выделяется из множества текстов/. После этого происходит выделение предложения из текста. Выделение предложения осуществляется путем поиска знаков препинания, указывающих на конец предложения. Далее, следует выделение с/ф в предложении. При этом учитывается, что с/ф – это цепочка последовательностей букв от пробела до пробела.

7.1. Функция расчленения с/ф /Р/.

При задании функции Р должно быть указано множество морфемных элементов /список/, которое позволяет провести "правильное" расчленение с/ф. Множество морфемных элементов включает последовательности букв, составляющие основы и аффиксы. Разложение с/ф на основы и аффиксы, а также разложение цепочек аффиксов на конкретные "правильные" аффиксы в тюркских языках носит непростой характер, так как с/ф в тюркских языках могут содержать цепочки аффиксов длиной до 10 элементов, сочетание которых выражает до-

вольно сложное значение.

Функция Р – указывает для каждой с/ф единственно правильное для нее расчленение на морфемы. Разложение с/ф может быть представлено как последовательность морфемных элементов с указанием границ между ними. Функция Р не производит никаких изменений в буквенном составе с/ф и лишь расчленяет ее на части.

Итак, функция Р путем последовательного отделения морфемных элементов и сравнения этих элементов со списками морфем производит правильное разложение с/ф.

В данной работе морфемные варианты /алломорфы/ одной морфемы выделяются как самостоятельные морфемы и снабжаются необходимой информацией. Ср. алломорфы аффикса настоящего времени -ыр, -ур, -ир, -үр, а также -жыр, -жур, -жир, -жүр. Можно было бы, конечно, ввести функцию, которая позволяла бы от одной алломорфы перейти к группе алломорф, представляющих данную морфему. Однако, при кодировании мы сочли целесообразным принять более простой и практический способ: задание грамматической информации для каждого алломорфа в отдельности.

7.2. Функция сопоставления информации /С/.

После применения функции Р мы получаем последовательность морфем /основы и аффиксы/, входящих в состав с/ф. Функция С сопоставляет каждой с/ф /X/ некоторую информацию С/X¹/ . Другими словами, для каждой морфемы X определяется некоторая информация У /X/. Р/X/ – это дизъюнкция элементарных информаций. Если морфема a входит в M₁, ..., M_n, где M_i – Р/X_i/ то можно сопоставить ей информацию У(a) = $\bigcup_{i=1}^n$ С(x_i) /108,5/.

Здесь M_i – морфема, полученная после разложения с/ф в результате применения функции Р.

Как уже отмечалось, с/ф в тюркских языках состоят из основы

и цепочки аффиксов: $x_i = \sum_{t=1}^n M_t \{2\}$ После применения функции С получаем $C(x_i) = \sum_{t=1}^n C(M_t) \{8\}$.

Пример 1. Пусть дана с/ф "китабларымыздан". В результате применения функции Р получаем правильное разложение с/ф на морфемы Р /китабларымыздан/ - /китаб/-/лар/-/ымыз/-/дан/.

После этого пользуемся функцией С. По формуле /3/ можно получить: $C(x_i) = C(M_1) \cup C(M_2) \cup C(M_3) \cup C(M_4)$

$$C(M_1) = a_1,$$

$$C(M_2) = \ell_1,$$

$$C(M_3) = \{n_1, n_2, n_3\}$$

$C(M_4) = Z_1$, где элементарные информации $- \ell_1, n_1, n_2, n_3, z_1, a_1$ содержат следующие значения:

a_1 : "существительное"

ℓ_1 : "множественное число"

n_1 : "принадлежность"

n_2 : "1 лицо"

n_3 : "множественное число"

z_1 : "исходный падеж"

В результате получаем: С/Р/китабларымыздан// - С/китаб/ С /лар/ \cup С/ымыз/ \cup С/дан/ - $\{a_1\} \cup \{\ell_1\} \cup \{n_1\} \cup \{n_3\} \cup \{z_1\}$

Здесь кроме грамматического значения /a/ имеется еще и лексическое значение /основа/.

Пример 2. Пусть даны с/ф дарын /подножье/ горы, сатын /продайте/, китабын /твоя/ книга.

После применения функции Р получаем:

Р /дар-ын/,

Р /сат-ын/,

Р /китаб-ын/.

Основы дағ, сат, китаб обозначим соответственно a_1, a_2, a_3 , а аффикс -ин - ℓ . Тогда получаем:

$$x_1 = P / \text{дағ-ын} / - \{a_1, \ell\}$$

$$x_2 = P / \text{сат-ын} / - \{a_2, \ell\}$$

$$x_3 = P / \text{китаб-ын} / - \{a_3, \ell\}$$

После функционирования С получаем:

$$C/x_1/ = C \{a_1\} \cup C \{\ell x_1\}$$

$$C/x_2/ = C \{a_2\} \cup C \{\ell x_2\}$$

$$C/x_3/ = C \{a_3\} \cup C \{\ell x_3\}$$

Здесь:

$$C \{\ell x_1\} = \{y_1\}$$

$$C \{\ell x_2\} = \{k_1, k_2, k_3\}$$

$$C \{\ell x_3\} = \{z_1, z_2, z_3\}$$

где элементы информации $y_1, k_1, k_2, k_3, z_1, z_2, z_3$ имеют следующие значения:

y_1 : родительный падеж

k_1 : повелительное наклонение

k_2 : 2 лицо

k_3 : единственное число

z_1 : принадлежность

z_2 : 2 лицо

z_3 : единственное число

$$\begin{aligned} C(\ell) = & C \{\ell x_1\} \cup \{\ell x_2\} \cup \{\ell x_3\} = \{y_1\} \cup \\ & \cup \{k_1, k_2, k_3\} \cup \{z_1, z_2, z_3\} \end{aligned}$$

Полученная информация {4} показывает, что омонимичный аффикс -ин выражает три типа отношения:

1/ родительный падеж,

2/ принадлежность, 2 лицо, единственное число,

3/ повелительное наклонение, 2 лицо, единственное число.

Как видно из примера, здесь $\kappa_1 = \chi_1$ (κ_1, χ_1 указывают на 2 лицо), $\kappa_3 = \chi_3$ (κ_3, χ_3 - указывают на единственное число).

Если κ_1, χ_1 заменить на t_1 , а κ_3, χ_3 на t_3 , то формула /4/ будет иметь вид:
 $C(L) = \{y_1\} \cup \{\kappa_1, t_1, t_3\} \cup \{z_1, t_1, t_3\} = \{y_1\} \cup \{(k_1 \cup \chi_1)(t_1, t_3)\}$

{5}

Из формулы следует, что морфема -ын указывает или на родительный падеж, или на повелительное наклонение/принадлежность, 2 лицо, единственное число.

7.8. Снятие омонимии основ и аффиксов на морфологическом уровне /функция R/.

Вне контекста, на уровне морфологии, отдельно взятую омонимичную основу невозможно отнести к определенной части речи. Поэтому, если слово омонимично, то в АС основа фиксируется его возможная принадлежность к разным частям речи: *јаз* "пиши", "весна" - глагол/существительное. Точное отнесение данной основы к определенной части речи возможно только на более высоких - синтаксическом и семантическом уровнях, так как "и омонимия и многозначность слова - явление в первую очередь семантическое" /48;57;76/.

Вместе с тем, омонимичность основ, частично может быть устранена и на морфологическом уровне при анализе присоединяющихся к основе аффиксов, которые указывают на принадлежность данной основы к той или иной части речи. Напр.: ат "бросать", "лошадь" в словарном виде выступает и как глагол и как существительное. Присоединяя аффиксы, эта основа может употребляться в следующих формах:

I/ атыр 'бросает', атын 'бросил', атачаг 'будет бросать' и т.д.

2/ атда "на лошади", атымз "наши лошади", атлара "к лошадям" и т.д.

В первом случае ат присоединяет аффиксы времени, что указывает на принадлежность ее к глаголу.

Во втором случае основа ат принимает аффиксы категорий падежа и принадлежности. Таким образом, как видно из примера, омонимичность основ легко устраивается при помощи анализа аффиксов, присоединяющихся к основам.

В тюркских языках существуют омонимичные аффиксы, которые, присоединяясь к глагольным основам и именам, выполняют при этом разные грамматические функции /ср.: подзоны 7.21., 8.0., 8.8., 8.12., 9.0./. Напр.: аффикс -ым, -ум, -им, -тм после глагольных основ обозначает повелительное наклонение, а после существительных - принадлежность: јазым "напишу-ка я", китабым "моя книга". В таких случаях омонимия аффиксов снимается исходя из основы, к которой присоединяется данный аффикс.

Вышеизложенное позволяет дать формальное определение функции R, снимающей омонимию основ и аффиксов в композиции F.

7.8.0. Формальное определение функции R

Здесь необходимо выделить два случая:

- а/ снятие омонимии основ с помощью аффиксов,
- б/ снятие омонимии аффиксов с помощью основ.

Принимаем следующие формальные правила снятия омонимии основ /табл. I/:

Правило I.1. Если основа омонимична /N/V/, а аффикс именной /n/V/, тогда и основа именная /N/V/.

Правило I.2. Если основа омонимична /N/V/, а аффикс глагольный /V/V/, тогда и основа глагольная /V/V/.

Правило I.3. Если и основа /N/V/ и аффикс /n/V/ омони-

мичны, тогда омонимия основы на морфологическом уровне не снимается. Это верно в том случае, если после основы следует только один аффикс; если же после омонимичного аффикса следуют еще аффиксы, тогда для снятия омонимии и аффикса необходимо обратиться к аффиксу, следующему за омонимичным аффиксом. В данном случае /т.е./ для основы фиксируются ее возможные грамматические характеристики.

Таблица 1. Правила снятия омонимии основ

Омонимичная основа	Аффикс	Снятие омонимии основы /выбор правильного варианта/
НВ	п	Н
НВ	в	В
НВ	пв	НВ

Аналогично основам, омонимия аффиксов может быть снята при помощи следующих правил /таблица 2/.

Таблица 2. Правила снятия омонимии аффиксов

Омонимичный аффикс	Основа	Снятие омонимии аффиксов /выбор правильного варианта/
пв	Н	п
пв	В	в
пв	НВ	пв

Правило 2.1. Если аффикс омонимичен /пв/, а основа именная /Н/, тогда и аффикс именной /п/.

Правило 2.2. Если аффикс омонимичен /пв/, а основа глагольная /В/, тогда и аффикс глагольный /в/.

Правило 2.3. Если и аффикс /пв/ и основа /НВ/ омони-

мичны, то омонимия аффикса, не снимаемая на морфологическом уровне /если после этого аффикса не следует аффикс/. В случае 2.3. для омонимичного аффикса фиксируются его возможные грамматические характеристики. Напр.: аффикс -ым, -ум, -им, -ум /подзона 8.12/ и т.д.

Кроме этого, омонимия на морфологическом уровне может быть снята следующими способами:

7.3.1. Омонимия, снимаемая при помощи сочетания аффиксов.

В некоторых случаях в качестве отдельных формантов фиксируются часто употребляемые сочетания аффиксов. Напр. аффико причастия - јачарымиз состоит из сочетания двух аффиксов -јачаг и -ымиз. Первая часть этого сочетания -јачаг является аффиксом будущего категорического времени глагола, а вторая часть - показателем принадлежности, который присоединяется к именам /существительным/. Сочетание же этих двух аффиксов в целом выступает как формант причастия. Поэтому и в списке это сочетание фиксируется как признак причастия /ср. первая зона, вторая зона, кроме подзоны 2.3., подзоны 3.0., 4.4., 4.5., 5.10/.

Следует отметить, что этот показатель причастия в виде сочетания аффиксов представлен и в описательной грамматике современного азербайджанского языка /ср.: ИИ, 244/.

Что же касается причастия -дүг - аффикс принадлежности, то мы сочли целесообразным разделить данное сочетание на компоненты, поскольку это упрощает процесс АМД и сокращает число аффиксов в списке, не влияя на качество анализа. С этой целью, в списке аффиксов представлен машинный аффикс причастия прошедшего категорического времени -дүг, -дүг, -дик, -диј, -дүј, несмотря на то, что этот аффикс самостоятельно не употребляется /ср.: подзона 7.81/.

При этом в виде исключения, в списке представлены сочетания этого аффикса -лынин, -дуғун, -дијин, -дүйтн - аффико причастия, 2 лица, единственного числа, прошедшего времени /ср.: подзона 5.9/, поскольку отделение омонимичного аффикса -ни, -ун, -ин, -тн усложнило бы процесс анализа.

Точно так же, с целью устранения омонимии используются машинные форманты типа аффико принадлежности - аффиксы падежа /на-пример: -сынын, -сунун, -синин, -стунтн и т.д./. В противном случае, при формальном анализе составные части указанных аффиксов /-сын, -сун, -син, -стн/ ошибочно могут быть отождествлены с аффиксами повелительного наклонения, имеющими аналогичную форму /ср.: 4.6., 5.18., 5.14., 5.15., 6.9., 6.10., 6.11., 6.12/.

Аналогичным образом, аффиксы определенного и неопределенного имперфекта, давнопрошедшего времени также фиксируются в виде сочетания во избежание смешения аффиксов /ср.: подзоны 2.3., 3.8., 3.4., 4.7., 4.8., 4.9., 4.10., 4.12., 5.17., 5.18., 5.19., 5.20., 5.21., 5.22., 6.19., 6.20/.

7.8.2. Омонимия, снимаемая при помощи машинных аффиксов.

Омонимичность аффиксов устраняется также при помощи специальных машинных аффиксов, не представленных в традиционных морфологиях:

а/ -мыр, -мур, -мир, -мтр. Контаминированный вариант сочетания аффиксов отрицания -ма, -ме и настоящего времени -ыр, -ур, -ир, -тр /ср.: подзона 7.24/.

б/ -мисан, -мусан, -мисен, -мусен. Контаминированный вариант сочетания аффикса прошедшего повествовательного времени -мын, -муш, -миш, -мүш с аффиксом 2 лица единственного числа -сан, -се. Здесь первый аффикс утрачивает конечный -ш. Оставшаяся часть этого аффикса совпадает с аффиксом принадлежности -мы, -му, -ми,

-мт, что может при анализе привести к ошибке /ср.: подзона 5.16/.

Аффиксальные формы недостаточного глагола-связки иди и имиш, а также послелог иле представлены в списке как самостоятельные аффиксы: -јда, -јди /подзона 7.27/, -јмыш, -јмиш /подзона 6.14/, -јла, -јле /подзона 7.26/.

7.8.3. Омонимия, не снимаемая на морфологическом уровне.

В следующих случаях омонимия основ и аффиксов не может быть снята на морфологическом уровне:

а/ если омонимичная основа употребляется без аффиксов. Напр.: ат "бросать", "лошадь" и т.д.

б/ в случаях 1.3. и 2.3., т.е. когда и основа и аффикс омонимичны. Напр.: сочетание ат "бросать", "лошадь" - -ым - аффикс повелительного наклонения или принадлежности, I лица, единственного числа может быть переведено двояко: атым "моя лошадь" или "бросать бы мне". Вне контекста дифференциация этих двух значений невозможна.

В указанных выше случаях необходим выход на более высокий уровень анализа – синтаксический и семантический, позволяющий снять омонимию.

В списке представлено 5 омонимичных аффиксов, которые могут присоединяться как к глагольным, так и именным основам. Для омонимичных аффиксов этого типа фиксируются оба признака, причем в четвертой позиции отмечается принадлежность указанного признака к глагольной или именной основе /табл. З/. Так, например, омонимичный аффикс -ым, -ум, -им, -үм имеет код * в виде БАЛФ. После

* Подробно о принятой в работе системе кодирования см.: 2 глава, с. 92. Морфологическая и лексико-морфологическая индексация машинных аффиксов и основ азербайджанского языка/.

снятия омонимии через контекст мы получаем один из указанных в таблице кодов: *ЦАЛГ* или *RALC*. Это указывает, что данный аффикс выражает повелительное наклонение, I лицо, единственное число /*ЦАЛГ*/, если основа, к которой присоединен аффикс, глагольная, или обозначает принадлежность, I лицо, единственное число /*RALC*/, если основа именная. Исходя из индекса в четвертой позиции выбирается соответствующий признак для основы с/ф. Таким образом снимается омонимичность аффиксов данного типа ^x.

На основе вышеизложенного можно построить алгоритм омонимии основ и аффиксов. При этом надо учитывать, что функция Р следует после функции С, т.е. после того, как мы уже получили для каждой выделенной морфемы /основы и аффиксов/ некоторую грамматическую информацию.

Как уже было указано, в модели /I/ композиция функций Р задана алгоритмом. Здесь каждая функция /К, Р, С, Р, П/ соответствует одному этапу алгоритма анализа тюркской с/ф /рис. 4/.

7.4. Алгоритм АМА тюркской словоформы

1. Выделить с/ф из текста

2. Определить длину /N/ с/ф: если $N > x$ перейти к п.6, при отрицательном исходе к п.3.

3. Сравнить с/ф с соответствующей группой слов в АС: если с/ф найдена, перейти к п.4, при отрицательном исходе к п.5.

^x На материале татарского языка разработан фрагмент снятия омонимии аффиксов прилагательных. В татарском языке омонимичные аффиксы могут употребляться в составе прилагательного и наречия. Для различения омонимичных случаев предлагается установление принадлежности с/ф к определенной части речи с проверкой влево и вправо по контексту /III/.

Таблица 3. Омонимичные аффиксы

Зоны и подзоны	Омонимичные аффиксы	Возможные грамматические характеристики	Коды
7.21.	-дыр, -дир, -дур, 1ЛЛФ -дур	повелительн.залог, 3 лицо, единств.число сказуемость, 3 лицо единств.число	ФЛЛГ <i>АЛГС</i>
8.0.	-ja, -је	желательн.накл. дательн.падеж	КФФГ <i>ЩФФС</i>
8.8.	-ын, -ун, -ин, -тн	9+ЛФ повелительн.накл., 2 лицо, единств.число	<i>Ц+ЛГ</i>
		родительн.падеж	<i>ШФФС</i>
		принадлежн., 2 лицо, единств.число	<i>R+ЛС</i>
8.12.	-ым, -ум, -им, -тм	БАЛФ повелительн.накл., 1 лицо, единств.число	<i>ЦАЛГ</i>
		принадлежн., 1 лицо, единств.число	<i>RALC</i>
9.0.	-а, -ә	2ФФФ желательн.накл. дательн.падеж	КФФГ <i>ЩФФС</i>

4. Записать признак слов в резервном поле /РД/ и перейти к п. 7.

5. Выделить с/ф / $N = x$ / и перейти к п.3.

6. Уменьшить длину с/ф на одну букву / $N = N - 1$ /: если $N = 0$, перейти к п.13, при отрицательном исходе - к п. 3.

7. Проверить наличие остатка / ℓ /: если $\ell = 0$, перейти к п.14, при отрицательном исходе к п. 8.

8. Определить длину остатка / ℓ /: если $\ell > y$ перейти к п. 12, при отрицательном исходе к п. 9.

Отделение текста от текста,
предложения из текста и
выделение с/ф в предложении

К

Расчленение с/ф на основу и
на цепочку аффиксов, отде-
ление аффиксов друг от друга

Р

Сравнение выделенных морфем
/основы и аффиксы/ с морфем-
ным списком /список аффиксов
и основ/ и выявление лексико-
грамматических признаков

С

Снятие омонимии основ и аф-
фиксов на морфологическом
уровне

Р

Преобразование лексико-грам-
матической информации, соот-
ветствующей входной с/ф

П

Рис. 4. Этапы алгоритма АМА тюркского текста.

9. Сравнить остаток с соответствующей группой аффиксов: если аффикс найден, перейти к п. 10, при отрицательном исходе - к п. 11.

10. Записать признаки аффикса после признаков слова и пе-
рейти к п. 7.

11. Уменьшить длину остатка / ℓ / на одну букву / $\ell = \ell - 1$ /:
если $\ell = 0$, перейти к п.9, при отрицательном исходе - к п. 18.

12. Выделить остаток / $\ell = y$ / и перейти к п. 9.

13. Записать с/ф в список незнакомых слов и перейти к п.15.

14. Напечатать с/ф с грамматическими признаками и перейти к п. 15.

15. Проверить наличие знака конца текста / α /: если знак есть, перейти к п.16, при отрицательном исходе - к п. 1.

16. Закончить анализ /рис. 5/.

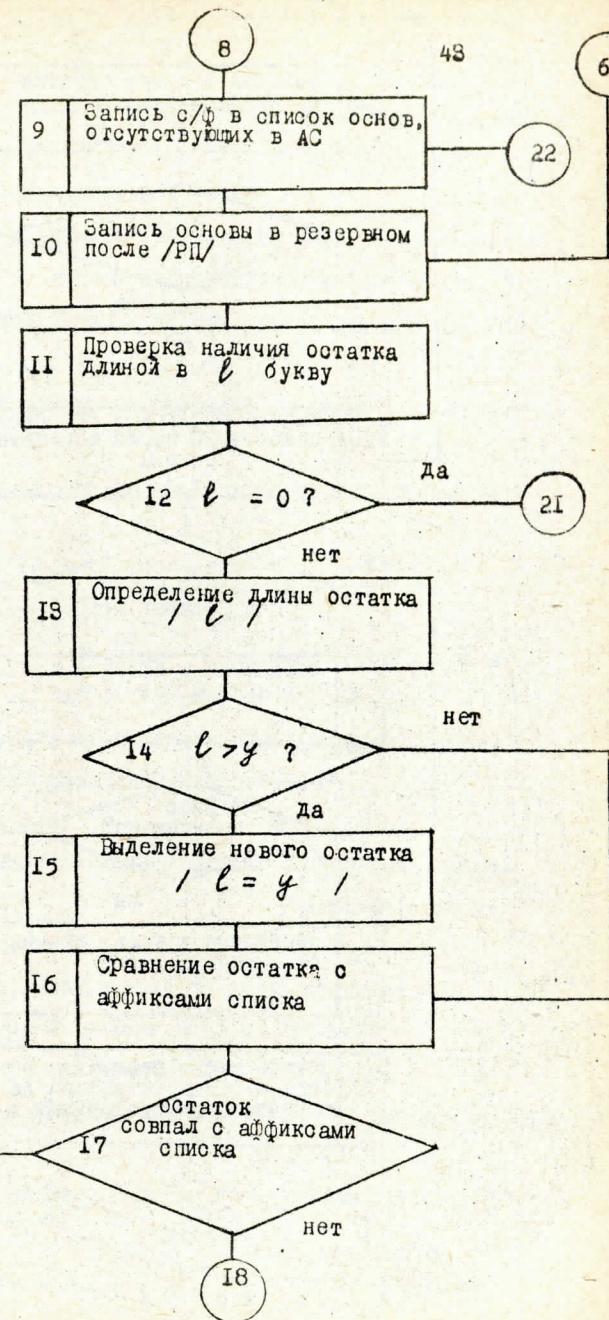
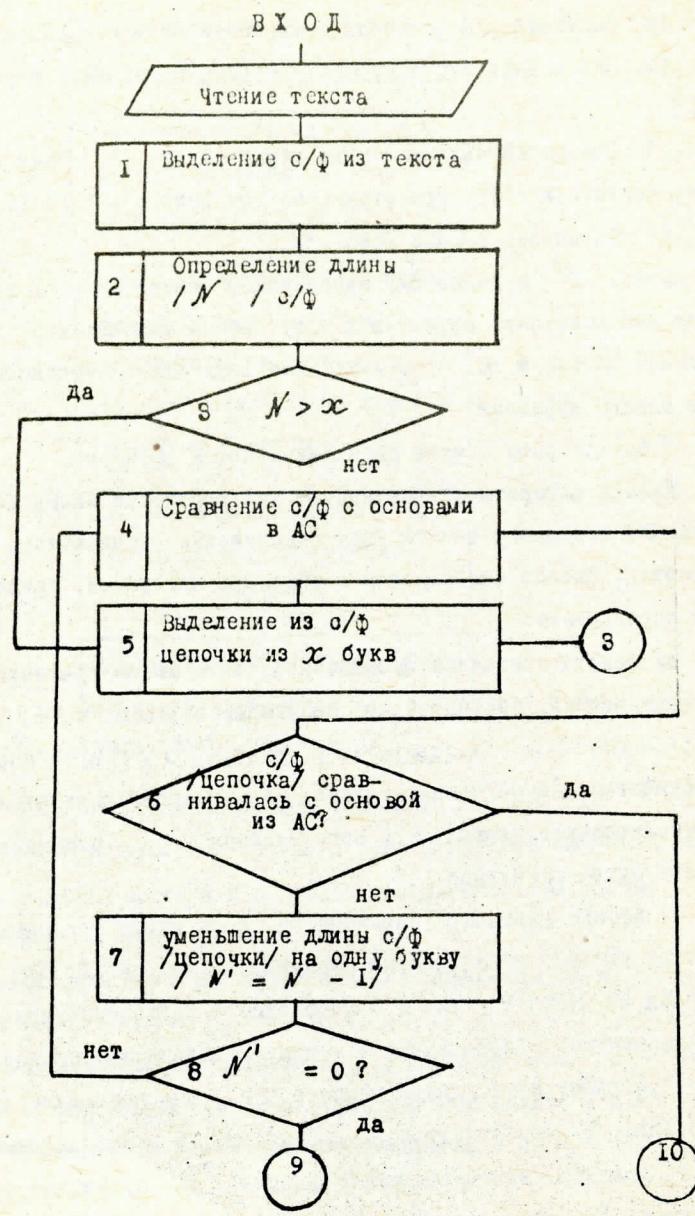
Здесь N - длина с/ф найденной в тексте, ℓ - длина аф-
фикса или сочетания аффиксов в с/ф, X - максимальная длина
туркской основы в АС, Y - максимальная длина тюркского аффик-
са в списке аффиксов.

7.5. Алгоритм снятия омонимии основ и аффиксов.

Данный алгоритм функционирует как самостоятельный блок в си-
стеме АМА тюркского текста /при обнаружении омонимичного аффикса
и основы/. Прежде чем перейти к описанию алгоритма, примем следу-
ющие определения:

а/ основа считается омонимичной, если имеет грамматические
признаки /коды/, обозначенные следующими знаками: 8; 4; 5; 6; 7;
8; %; : - где: 8 - существительное /глагол, 5 - глагол/наречие,
4 - существительное/прилагательное, 6 - глагол/прилагательное,
7 - числительное/глагол, 8 - союз/частица, - отрицание/отгла-
гольное существительное.

б/ аффикс считается омонимичным, если имеет грамматические
признаки /коды/, обозначенные следующими знаками: Д; Ҕ; І; 2;
Я; Т; Б; Ю; Q; 9; : - где Д - необходимое наклонение/причастие,
- деепричастие/причастие, І - понудительный залог/сказуемое,
2 - желательное накл./дательно-напр.пад., Я - принадлежн./вина-
тильн.пад., Т - долженств.накл./причастие, Б - принадлежность,
І лицо, единств. число/повелительн.накл., Ю - будущ.кат.время/
причастие, Q - причастие/прошедш.повеств.время, 9 - родительн.пад./



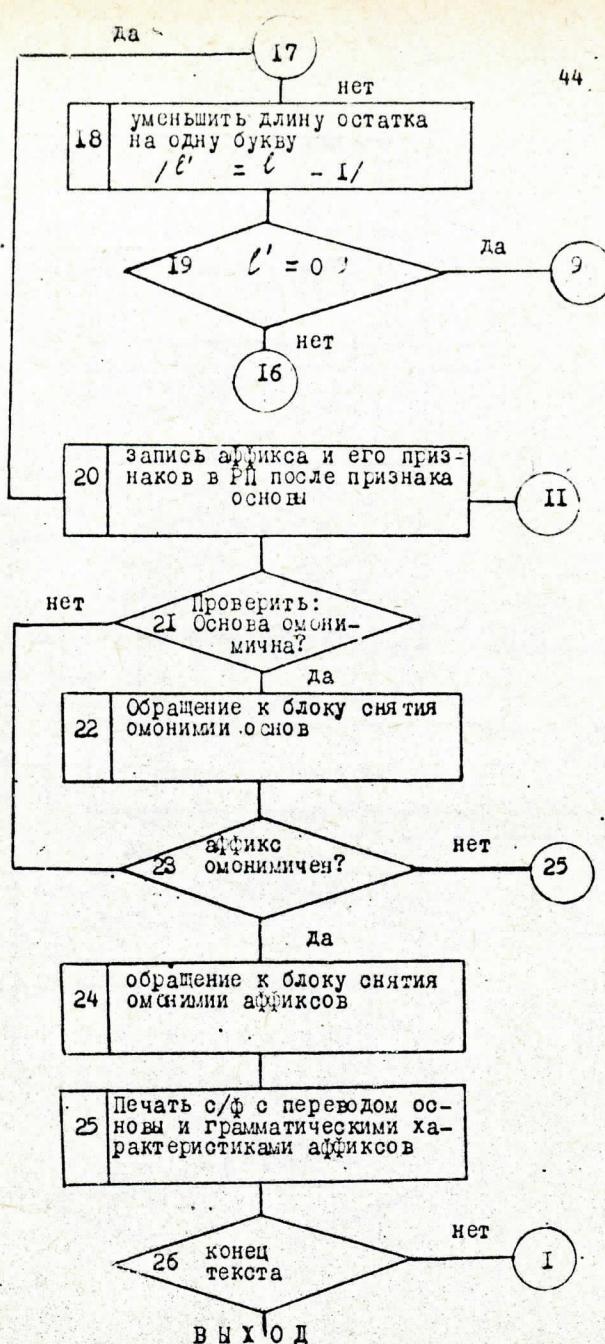


Рис. 5. Принципиальная блок-схема алгоритма AMA татарской словоформы

45

принадлежн., 2/3 лицо, единств.число/повелительн.накл.: - деепричастие/прошедш.время /табл. 15/.

7.5.1. Алгоритм снятия омонимии основ

1. Проверка наличия аффикса, следующего после основы: если аффикс есть, перейти к п. 2, при отрицательном исходе - к п. 7.

2. Выделение аффикса начиная справа после основы и переход к п. 8.

3. Проверить: если аффикс омонимичный, перейти к п. 1, нет - к п. 4.

4. Проверить: если аффикс именной, перейти к п. 5, нет - к п. 6.

5. Основа именная, перейти к п. 8.

6. Основа глагольная, перейти к п. 8.

7. Омонимия основы, не снимаемая на морфологическом уровне.

8. Выдача соответствующей информации об основе /рис. 6/.

В качестве примера проследим анализ с/ф /атымдан/ "от моей лошади". После применения функции Р и С получаем:

Р /атымдан/ = ат-ым-дан

С /ат-ым-дан/ = С /ат/ ∪ С /ым/ ∪ С /дан/ = /"бросай" глагол /
"лошадь" существ./

У /повелит.накл., I лицо, единств.число/

принадлежн., I лицо, единств.число/ У /исходный падеж/

Требуется снять омонимии основы /ат/ в с/ф /атымдан/. При этом алгоритм функционирует следующим образом:

1/ проверяем наличие аффикса, следующего после основы: аффикс есть, переходим к п. 2*

2/ выделяем аффикс, начиная справа после основы: /-ым/ и переходим к п. 5.

3/ проверяем по табл. 9 и, выявив омонимичность аффикса,

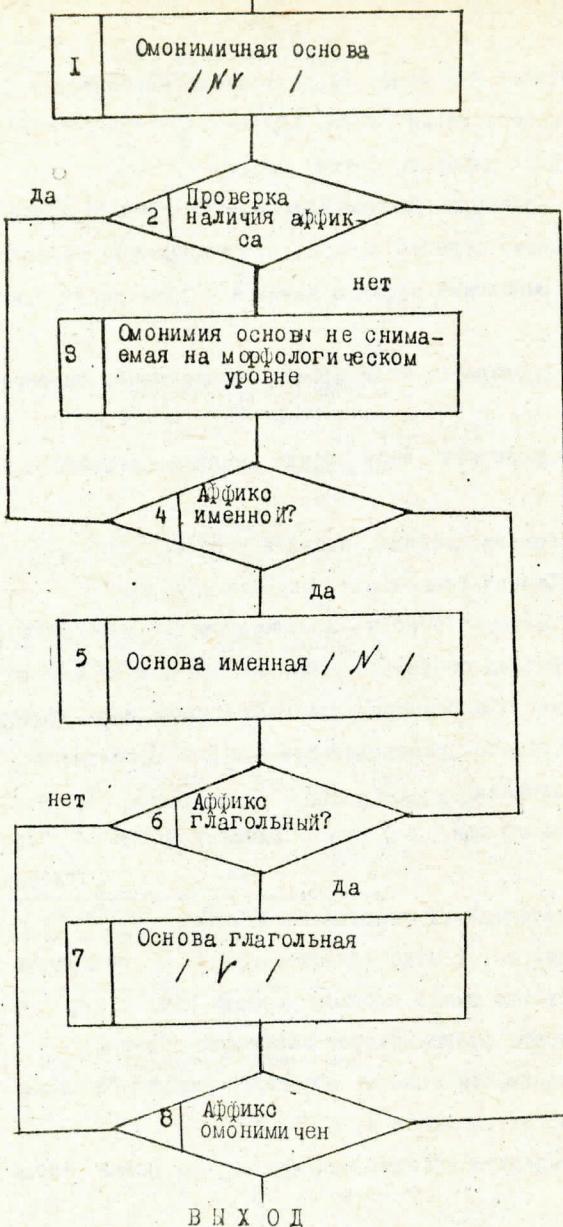


Рис. 6: Блок снятия омонимии основ

переходим к п. 1.

4/ проверяем наличие следующего аффикса: аффикс есть, переходим к п. 2.

5/ выделяем следующий аффикс /-дан/ - и переходим к п. 3,

6/ проверяем по табл. 3: аффикс не омонимичен, переходим к п. 4,

7/ проверяем: аффикс именной, переходим к п. 5,

8/ основа ат в с/ф атымдан именная.**7.5.2. Алгоритм снятия омонимии аффиксов.**

1. Проверить: если основа, к которой присоединен данный аффикс омонимична, перейти к п. 5, нет - перейти к п. 2.

2. Проверить: если основа именная, перейти к п. 3, при отрицательном исходе к п. 4.

3. Аффикс именной, перейти к п. 6.

4. Аффикс глагольный, перейти к п. 6.

5. Омонимия аффикса не снимаемая на морфологическом уровне.

6. Выдача соответствующей информации об омонимичном аффиксе /рис. 7/. Напр.: с/ф китабы "моja книга" после применения функций Р и С получает вид:Р /китабым/ = китаб-ым

С /китаб-ым/ = С/китаб/ ∪ С/-ым/ - /китаб "книга", существ./прилагательн./ ∪ /-ым - принадлежн., 2 лицо, единств.число/повелительн.накл., I лицо, единствен.число/.

Снятие омонимии аффикса -ым осуществляется следующим образом:

1/ находим, что основа, к которой присоединен данный аффикс, не омонимична и переходим к п. 2,

2/ находим, что основа именная, и переходим к п. 3.

3/ определяем, что аффикс именной: /-ым - принадлежность, I лицо, единствен.число/.

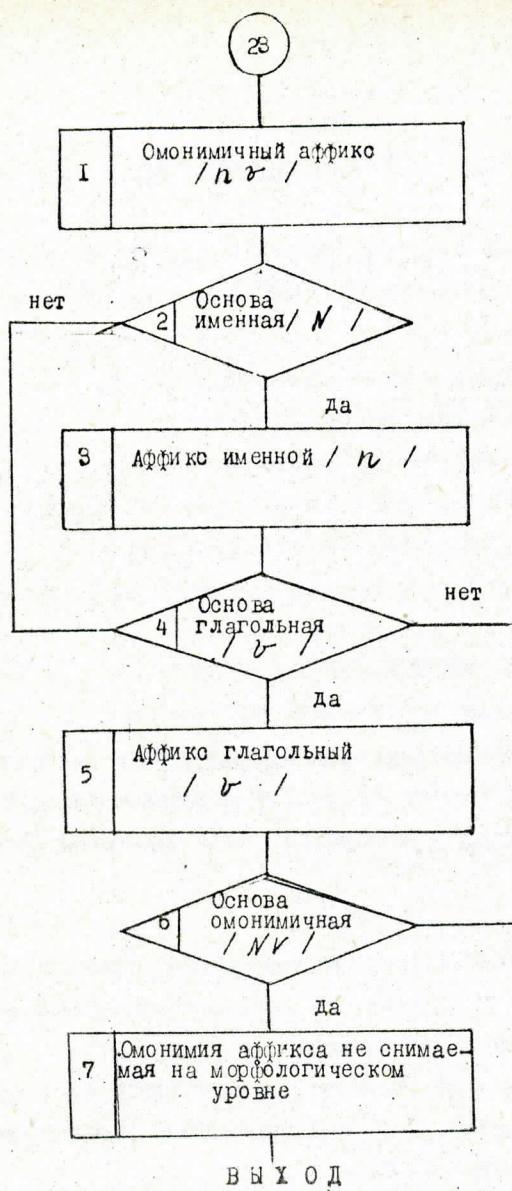


Рис. 7. Блок снятия омонимии аффиксов

7.6. Преобразования информации /функция П/.

Функция П, как уже указывалось, преобразует информацию соответствующую реальной с/ф /Ф/Х/. В результате получаем лексико-грамматическую информацию, соответствующую значению с/ф.

Напр.: /атларымыздан/ - П/Р/С/Р /атларымыздан///= П/Р/С/ат+ларымыз+дан/-П/Р / "бросай", "лошадь" /- множественное число + принадлежн., 1 лицо, множественное число + исходный падеж + П /лошадь + множественное число + принадлежность, 1 лицо, множественное число + исходный падеж /от наших лошадей/.

ГЛАВА II

РЕАЛИЗАЦИЯ ВОСПРОИЗВОДСТВА ИНЖЕНЕРНО-
ЛИНГВИСТИЧЕСКОЙ МОДЕЛИ ММТЯ НА МАТЕ-
РИАЛЕ АЗЕРБАЙДЖАНСКОГО ЯЗЫКА

I. Описание банка лингвистических данных

Для реализации воспроизведения инженерно-лингвистической модели ММТЯ на материале азербайджанского языка необходимо построить базу лингвистических данных, поскольку "любая экспериментальная или промышленная переработка текста не может быть осуществлена без того, чтобы в автомат не было бы предварительно введено описание определенного участка объективной действительности, представленное в виде упорядоченных множеств таких лингвистических и энциклопедических сведений, с которыми могли бы сопоставляться лингвистические и информационные единицы, которые описывает данный участок внешнего мира и которые с достаточно большой вероятностью встречаются в перерабатываемых в ЛА текстах. Точность, надежность и быстрота переработки текста во многом зависит объемом этих сведений и тем, как организовано их хранение в базе данных ЛА" /84,65/.

В данном разделе описываются организация и хранение машинных основ и аффиксов, их грамматические характеристики, функции и специфические особенности.

Следует отметить, что в связи с АПТ нецелесообразно использование АС с/ф, так как для тюркских языков процент покрытаемости текста словоформами намного ниже чем для языков флексивно-аналитического и флексивно-синтетического типа. Исследования на материале тюркских языков показывают, что "для обеспечения 70 проц. покрытия словоформами наугад взятого текста достаточный объем выборки в агглютинирующем языке должен почти в 5 раз превосходить

объем выборки флексивно-аналитического языка и более чем в два раза быть больше, чем выборка флексивно-синтетического языка" /12,37/. Увеличение процента покрытаемости тюркского текста словоформами возможно только путем увеличения объема с/ф, что практически в настоящее время не выполнимо, если учесть, что "английские научно-технические подъязыки используют не более 18 тыс. разных словоформ, во флексивных языках их число достигает 20-25 тыс., в агглютинативных, например, в казахском, число разных Ch с/ф поднимается до 60 тыс. /52,13/.

Эти обстоятельства приводят к необходимости учитывать особенности тюркских языков при разработке систем МП и АПТ, так как "в настоящее время разработка программ МП и АР ориентирована на западно-европейские языки" /52,16/.

Именно поэтому в связи с автоматической переработкой тюркского текста, ЧС основ имеет определенные преимущества. Указывая на это, К.Б.Бектаев отмечает, что ЧС основ "увеличивает эффективность заполняемости текста, т.е. одну и ту же покрытаемость можно достичь меньшим числом наиболее частных слов, чем словоформ. Объем словаря, особенно для агглютинативных языков, сильно сокращается, т.е. количество различных слов ЧС намного меньше количества различных словоформ в ЧС, составленных по одному и тому же объему текстов" /12,55/.

С учетом вышеуказанного, АС основ построен на базе ЧС основ азербайджанского языка, составленного на материале газетных текстов. Выборка текстов производилась из газет "Коммунист", "Бакы", "Азербайчан көнчлери" за 1976 год по следующим темам: "Социально-экономическая информация /информация по экономике, сельскому хозяйству, новости науки и техники/", "зарубежная информация", "искусство и культура", "спортивная информация".

I.I. Машинные основы азербайджанского языка, распределение их по зонам и характерные особенности.

АС основ включает около 2000 высокочастотных основ, взятых из частотного словаря азербайджанской газетной лексики, составленного сотрудниками Института языкознания АН Азерб.ССР и дополненного общеупотребительной лексикой.

В зависимости от длины машинные основы АС распределяются по зонам. Весь массив разбит на 16 зон. Первая зона включает основы, состоящие из 16 букв, во вторую входят 15-буквенные основы, в третью - 14 буквенные и т.д. Внутри зон основы располагаются в алфавитном порядке. В АС учитываются наблюдаемые в основах некоторые морфонологические изменения, зафиксированные в азербайджанской орфографии.

В качестве основы для имени традиционно принимается форма именительного падежа единственного числа јолдаш "товарищ", а для глагола - форма императива көтүр "возьми". При этом формы инфинитива јазмаг, кетүрмек рассматриваются как сложные образования, состоящие из основ јаз, кетүр, хранящихся в АС, и инфинитивных морфем /-маг, -мек и т.д./, помещаемых в списке аффиксов /табл.4/.

Выделение машинных основ азербайджанского языка определяется морфологическими особенностями данного языка, которые необходимо учитывать при анализе.

* Таблица 4. Фрагмент азербайджанско-русского АС основ

Исходная форма слова и основа	Принадлежность к частям речи	Перевод
јолдаш	существительное	товарищ
кемек	существительное	помощь
кемеј	существительное	помощь
кетүр	глагол	возьми
јаз	существительное глагол	писать весна

Формализация морфонологических и фонетических изменений, возникающих на стыке основы и аффиксов рассматривалась в ряде работ по тюркологии. В частности, К.А.Велиевой разработаны формальные правила образования морфологически и орфографически правильных азербайджанских с/ф, образуемых путем присоединения аффиксов к корням слов /или к неразложимым основам/, т.е. правил порождения азербайджанских с/ф /28; 24,3/. Автором изложены методы формального описания правил расстановки морфем, построена схема автоматического порождения слова. Практическим результатом работы явился машинный эксперимент по автоматическому порождению азербайджанских с/ф, который был поставлен в институте кибернетики АН Азерб.ССР. Аналогичная работа выполнена Н.К.Якубовой на материале узбекского языка /122, 139-151; 128/. Реализованная на ЭМ система автоматического синтеза азербайджанских глагольных форм с учетом морфологических правил разработана также В.Я.Пинесом /74/.

Следует отметить, что если при синтезе необходима полная формализация всех правил стыка основ и аффиксов, то при анализе морфонологические изменения учитываются в АС основ и списке аффиксов. Коротко остановимся на некоторых из них:

а/ в азербайджанском, как и в других тюркских языках, во многих случаях наблюдаются изменения глагольных основ и недифференцированность основ и аффиксов ^x. Конечный согласный многослож-

^x Как указывает А.Н.Кононов, "в силу динамической природы аффиксальных морфем, структура тюркского слова в части деления на корневую и аффиксальную морфемы чрезвычайно трудно поддается анализу" /56, 120/. Именно с этим явлением связаны некоторые спорные вопросы развития и формирования нечленимых сложных ос-

ных глагольных основ на -т, а также глаголов ет, кет при присоединении аффикса, начинающегося с гласного, т.е. в интервокальном положении заменяется своими звонкими коррелятами д. Например, ет-, едир, едек и т.д., кет, кедек, кедир и т.д., гајт - гајыдир, гајыдаг и т.д. Поэтому глагольные основы данного типа в ММТЯ фиксируются в виде двух вариантов: ет/ед, кет/кед, гајт/гајыд и т.д. Звонкие варианты ед, кед, гајыд и т.д. являются здесь типично машинными основами, имеющими один и тот же грамматический признак, что и реальные основы /табл. 5/.

Таблица 5. Распределение машинных основ по зонам
/фрагмент/

Зоны	Исходная форма слова или машинная основа	Принадлежность к части речи	Перевод или значение
XII	гајт	глагол	вернуться, возвращаться
	гајыд		
XIII	ешит	глагол	слушать, услышать
	ешид		
XIV	кет	глагол	идти, уходить
	кед		
XV	ет	глагол	делать, сделать
	ед		

б/ в азербайджанском языке узкие гласные -и, -у, -и, -т конечного закрытого слога в двусложных словах выпадают при присоединении к слову аффиксов принадлежности, родительного, дательно-направительного и винительного падежей. Например, слово /шекил

продолжение сноски со стр. 58

нов и непроизводных или малопроизводных аффиксов тюркских языков. Еще В.Радлов называл этот процесс гиперагглютинацией, указывая на то, что иногда корень и аффикс, прочно спаиваясь, совершенно утрачивают способность к делению /54, I20/.

"картина"/ при склонении /в родительном, дательно-направительном и винительном падежах/ и при присоединении аффиксов принадлежности приобретает следующие формы: шекил - шоклин, шэкле, шекли, шеклим и т.д.

Аналогичное явление наблюдается и в других тюркских языках. Так, А.П.Поцелуевский пишет: "Ассимиляция и выпадение звуков обогащает морфологию туркменского языка новыми формами путем превращения ряда самостоятельных слов в формальные элементы речи и способствует переходу туркменского языка от явлений строгого агглютинативного типа к переходным агглютинативно-фузионным и вполне фузионным, ведущим в конечном счете, к явлениям флексии" /89, 51/.

Эту закономерность можно было бы обобщить и алгоритмизировать. Однако выпадение закрытого гласного во втором слоге характерно только для некоторых слов указанной структуры. Например, в словах лепир "след", демир "железо", чениз "приданое", өсир "пленик", "пленица" и многих других с присоединением аффиксов принадлежности, родительного, дательно-направительного и винительного падежей гласный во втором слоге не выпадает. Поэтому слова, в которых наблюдается выпадение гласного, целесообразно дать в виде списка /табл. 6/.

в/ конечные согласные -г и -к в многосложных словах при присоединении аффиксов, начинающихся с гласных, подвергаются замене г г к ж. Например, аяг "нога" - аягым, аяга, аяры, чиочек "цветок" - чицеје, чицеји и т.д.

Иключение составляют заимствования из европейских языков, оканчивающиеся на -олог и -ик и слова арабского происхождения на -иг и -ик. Например: биолог, технолог, физик, фабрик, электрик и т.д. /европ./, мустентиг, тәсдиғ, тебрик, шарик и т.д. /араб./.

В связи с тем, что глагол в словнике представлен без формального признака инфинитива аффикс -маг, -мәк, количество слов, оканчивающихся на -г, -к было значительно уменьшено, что позволило отразить в списке оба варианта основы /табл. 6/.

Таблица 6. Распределение машинных основ в АС основ /фрагмент/

Зоны	Исходная форма слова или машинная основа	Принадлежность к части речи	Перевод или значение
XII	шәкил	существительное	картина
	синиф	существительное	класс
	чиҹек	существительное	цветок
	чиҹәј	существительное	цветок
XIII	ајаг	существительное	нога
	ајағ	существительное	нога
	шәкл	существительное	картина
	синф	существительное	класс

В целях алгоритмического устранения конверсионной омонимии в словаре, при необходимости, указывается возможная принадлежность основы к разным частям речи /табл. 7/.

Таблица 7. Размещение омонимичных основ в АС

Зоны	Исходная форма слова или машинная основа	Принадлежность к части речи	Перевод или значение
XIV	газ	существительное	гусь, газ
		глагол	копать
	јаз	существительное	весна
		глагол	писать

Прочная грамматическая информация при основах не фиксируется, так как в тюркских языках основы приобретают новые грамматические характеристики при помощи аффиксов.

I.2. Машинные аффиксы азербайджанского языка, их грамматические характеристики при помощи аффиксов, распределение по зонам и подзонам.

Описания тюркских аффиксов с теоретическими целями и с целью АПТ содержатся в целом ряде работ /II; 25; 24; 33, 163-170; 74; 76; 93; 118; 128; 124 и т.д./.

В работе И.И.Ревзина и Г.Д.Юлдашевой описывается расположение и классификация глагольных аффиксов современного узбекского литературного языка /93/. Авторами были определены семь порядков и глагольные аффиксы распределены по ним в виде таблицы. На основе этой таблицы, при помощи фонологических правил, регулирующих выбор алломорфа в зависимости от фонетических условий, и морфологических правил, зависящих от выбора соответствующего порядка, построена порождающая грамматика глагольных аффиксов с/ф узбекского языка.

В диссертации В.Я.Линеса /74/ дается описание морфологической структуры азербайджанских глагольных форм и применение построенной модели в системе азербайджанского английского АС, предлагается теория упорядочения, близкая к изложенной в статье И.И. Ревзина и Г.Д.Юлдашевой /93/. Автор в основу своей работы берет тот факт, что в тюркских языках существует высокая степень упорядоченности аффиксов разного значения относительно основы и относительно друг друга. Впервые автором разработана математическая модель грамматики порядков с использованием понятия отношений, которая успешно применяется для описания структуры личных форм азербайджанского глагола. Представленный "механизм" порождения

цепочек аффиксов являясь фрагментом грамматики азербайджанского языка порождает все грамматически правильные глагольные формы, причем "не только наблюдаемые формы, но и те, которые принципиально допустимы, но не употребляются в речевой практике", что можно считать одним из важных достоинств работы.

Учитывая специфику процесса АА тюркского текста, аффиксы азербайджанского языка в данной работе распределяются по другому принципу.

Аналогично основам машинные аффиксы также распределяются по зонам в зависимости от их длины. При этом на первом плане стоит обработка только словоизменяющих аффиксов азербайджанского языка, поскольку словообразование форманты, как указывалось выше, образуют отдельные лексические единицы, которые включаются в АС и не подвергаются морфологическому анализу.

К числу категорий, образуемых с помощью словоизменяющих аффиксов, относятся категории отрицания, наклонения, залога, причастия, деепричастия, времен глаголов, падежа, сказуемости и принадлежности, относящиеся, прежде всего к существительным и субстантивным словам.

Первая зона словоизменительных аффиксов содержит аффиксы, состоящие из 9-ти букв, вторая - из 8-ми букв и т.д. В списке для каждого аффикса задается поле грамматических признаков, включающее четыре позиции, где указывается основное морфологическое значение и дополнительные морфологические характеристики. В подзонах отражены также все возможные позиционные варианты /алломорфы/ тех или иных грамматических элементов.

Если аффикс имеет только один морфологический признак, то этот признак указывается в первой позиции, а в остальных позициях ставятся нули. Например, для аффикса -ды, -ду, -ди, -ду в первой

позиции записывается его основная морфологическая характеристика - "прошедшее категорическое время", а в остальных трех позициях ставятся нули, так как данный аффикс не имеет других грамматических признаков.

В списке аффиксов для многозначных аффиксов фиксируются их возможные грамматические признаки. Дело в том, что на уровне морфологии точная идентификация с/ф, образованных, например, при помощи аффиксов -ачаг, -ечек, -ачағ, -ечәј невозможна. Так, с/ф көләчек является формой будущего категорического времени "он придет" и вместе с тем, причастием "будущий". Различение аномалических форм с аффиксом -ачаг и его алломорфами возможно только на синтаксическом уровне охуначаг китаб "книга, которая будет прочитана". Поэтому для указанного аффикса фиксируются обе морфологические характеристики.

Информация при таком аффиксе, как -ын, -ун, -ин, -үн содержит уже три признака, поскольку этот аффикс, присоединяясь к именам, образует формы родительного падежа или принадлежности 2-го лица единственного числа, а присоединяясь к глаголам - форму повелительного наклонения /ср.табл. I3, подзона 8.8/.

Учитывая чередование к > ј, г > ғ аффиксы, оканчивающиеся на -к, -ғ, представлены двумя вариантами /ср.табл. I3, подзона 7.1/.

Во избежание смешения аффиксов в некоторых случаях в качестве отдельных формантов фиксируются часто употребляемые сочетания аффиксов /ср.табл. I3, подзона 5.13 и т.д./.

Некоторые аффиксы в составе с/ф в результате выпадения гласных или согласных приобретают упрощенную форму. Например, аффикс отрицания -ма, -ме в сочетании с показателем настоящего времени утрачивает гласный: алм/а/ырсыныз "не получаете", кәлм/а/ирсән "не приходишь" и т.д. В результате возникает контаминированный

элемент -мир, -мир, -мур, -мүр /отрицание – настоящее время/, который целесообразно дать в качестве самостоятельной единицы слова варя аффиксов /ср.табл. Г3, подзона 7.4./.

Аффикс прошедшего повествовательного времени -мыш, -муш, -миш, -мүш может употребляться и без конечного -ш: көрмүшсөн "ты видел". Первая часть этого аффикса -мы, -му, -ми, -мү совпадает с теждественным по форме аффиксом вопроса. Поэтому данная форма аффикса -мыш фиксируется не самостоятельно, а в сочетании с показателем второго лица /ср.табл.Г3, подзона 5.16/.

Недостаточные глаголы – связки иди и имиш представлены в списке основ. Однако они выступают /с соответствующими фонетическими изменениями/ в форме аффиксов. Напр.:

алмалы иди "должен был получить",

алмалыидн

көлмел иди "должен был прийти",

көлмелиди

көлмел имиш "говорят, должен был прийти"

көлмелиджимиш.

Сказанное относится и к послелогу иля. Поэтому мы сочли целесообразным включить в список машинных аффиксов аффиксальные варианты указанных слов.

Ниже подробно рассматриваются машинные аффиксы азербайджанского языка, которые распределены по изложенному выше принципу, и их грамматические характеристики.

Первая зона /аффиксы, состоящие из 9-ти букв/.

Подзона 1.0. -јачагымиз, -јөчејимиз: сочетания аффиксов, имеющие следующие грамматические характеристики – причастие – I лицо – множественное число – будущее категорическое время. Присоединением к аффиксам -ачаг, -ечек, -јачаг, -јачаг, -јөчеј,

-јечек аффиксов принадлежности, образуются аффиксы причастий, которые сообщают о совершении действия в будущем времени. Например, башлајачагымиз иш "работа, которую мы начнем".

Подзона 1.1. -јачагныз, -јөчејиниз. Грамматические признаки: причастие – 2 лицо – множественное число – будущее категорическое время: охујачагныз китаб "книга, которую вы прочтете".

Подзона 1.2. -јачаглари, -јечеклери. Грамматические признаки: причастие – 3 лицо – множественное число – будущее категорическое время: охујачаглари китаб "книга, которую они прочтут".

Вторая зона /аффиксы, состоящие из 8-и букв/.

Подзона 2.0. -ачагымиз, -өчејимиз. Грамматические признаки: причастие – I лицо – множественное число – будущее категорическое время: јазачагымиз мектуб "письмо, которое мы напишем".

Подзона 2.1. -ачагныз, -өчејиниз. Грамматические признаки: причастие – 2 лицо – множественное число – будущее категорическое время: јазачагныз мектуб "письмо, которое вы напишете".

Подзона 2.2. -ачаглари, -өчејлери. Грамматические признаки: причастие – III лицо – множественное число – будущее категорическое время: јазачаглари мектуб "письмо, которое они напишут".

Аффиксы причастий первой зоны присоединяются к основам, оканчивающимся на гласный, а второй зоны – на согласный.

Подзона 2.3. -јачагыш, -јечекмеш. Грамматические признаки: будущее-прошедшее время /будущее в прошедшем/ – пересказательность охујачагмыш "он оказывается /как стало известно из косвенных источников/ собирался прочитать".

Третья зона /аффиксы, состоящие из 7-и букв/.

Подзона 3.0. -јачагым, -јөчејим. Грамматические признаки: причастие – I лицо – единственное число – будущее категорическое время јазачагым мектуб "письмо, которое я напишу".

Подзона 3.1. -јачарын, -јәчәйин. Грамматические признаки: причастие – I лицо – единственное число – будущее категорическое время јазачағын мектүб "письмо, которое ты напишешь".

Подзона 3.2. -дыглары, -дуглари, -дикләри, -дүкләри. Грамматические признаки: причастие – II лицо – множественное число – прошедшее время јаздыглары мектүб "письмо, которое они написали".

Подзона 3.3. -јачагды, -јечәкди. Грамматические признаки: будущее-прошедшее время /будущее в прошлом/ – охујачагды "он собирался прочитать".

Подзона 3.4. -ачагмыш, -әчәкмиш. Грамматические признаки: будущее-прошедшее время будущее в прошлом – пересказательность јазачагмыш "он оказывается, собирался написать".

Четвертая зона /аффиксы, состоящие из 6-и букв/.

Подзона 4.0. -јаркән, -јөркән, -јүркән, -јыркән, -јиркән, -јүркән. Грамматический признак: деепричастие, обозначающее действие одновременное с действием спрягаемого глагола охујаркән "в тот момент, когда читал".

Эти форманты, как ряд других показателей деепричастия, образованы контаминацией временных аффиксов со связкой иң.

Аффиксы деепричастия в тюркских языках связывают один глагол с другим, причем, в составе предложения при одновременном употреблении, первый глагол /глагол с аффиксами деепричастия/ дополняет действие, обозначенное последующим глаголом. Деепричастие не принимает аффиксов категории лица, времени, числа.

Подзона 4.1. -мышкән, -мишкән, -мушкән, -мүшкән. Грамматические признаки: деепричастие, обозначающее предшествование в прошлом јазмышкән "написав когда-то, в прошлом".

Подзона 4.2. -јандән, -јөндән. Грамматические признаки: деепричастие, обозначающее временное значение охујандан сонра "с

тех пор, как после того как читал".

Подзона 4.3. -дырлар, -дурлар, -дирләр, -дүрләп. Грамматические признаки: сказуемость – III лицо – множественное число төле-бәдирләр "они студенты".

Подзона 4.4. -ачарым, -әчејим. Грамматические признаки: причастие – I лицо – единственное число – будущее категорическое время јазачатым мектүб "письмо, которое я напишу".

Подзона 4.5. -ачарын, -әчејин. Грамматические признаки: причастие – II лицо – единственное число – будущее категорическое время јазачағын мектүб "письмо, которое ты напишешь" /ср.: З-я зона, подзона 3.1/.

Подзона 4.6. -сиңдан, -сундан, -синдән, -стиндән. Грамматические признаки: принадлежность – II лицо – единственное число – исходный падеж атасиңдан "от его отца".

Подзона 4.7. -јачары, -јәчәји. Грамматические признаки: причастие – III лицо – единственное число – будущее категорическое время охујачары китаб "книга, которую она прочтет".

Подзона 4.8. -мишмыш, -мушмуш, -мишмиш, -мүшмүш. Грамматические признаки: давнопрошедшее время – пересказательность јазмы-мыш "оказывается он написал когда-то".

Подзона 4.9. -јырмыш, -јурмуш, -јирмиш, -јүрмүш. Грамматические признаки: определенный имперфект – пересказательность оху-јүрмүш "он, оказывается, читал".

Подзона 4.10. -ачагды, -әчәкди. Грамматические признаки: будущее-прошедшее время јазачагды "он собирался написать".

Подзона 4.11. -јармыш, -јөрмиш. Грамматические признаки: неопределенный имперфект – пересказательность охујармыш "он, оказывается, имел обыкновение читать в прошлом".

Подзона 4.12. -јибмүш, -јубмуш, -јисмиш, -јүбмүш. Граммати-

ческие признаки: давнопрошедшее время - пересказательность охујубуш "он, оказывается, читал когда-то".

Пятая зона /аффиксы, состоящие из 5-и букв/.

Подзона 5.0. -магда, -мәкдә. Грамматические признаки: настояще дательное время јазмагдадыр "он пишет, находится в процессе писания".

Подзона 5.1. -јараг, -јерек. Грамматический признак: деепричастие, обозначающее предшествование, образ действия охујарааг "читая".

Подзона 5.2. -јанда, -јәндә. Грамматические признаки: деепричастие, обозначающее одновременность с действием спрягаемого глагола охујанда "в тот момент, в процессе, когда он читал".

Подзона 5.8. -аркән, -әркән, -ыркән, -иркән, -уркән, -үркән. Грамматические признаки: деепричастие /ср. подзону 4.0./: јазаркән "в тот момент, когда писал".

Подзона 5.4. -дыгда, -дикдә, -дугда, -дүкдә. Грамматические признаки: деепричастие, обозначающее одновременность с действием спрягаемого глагола: охудугда "когда он читал".

Подзона 5.5. -дыгча, -дугча, -дикче, -дүкче. Грамматические признаки: деепричастие, обозначающее продолжительное действие, одновременное с действием спрягаемого глагола охудугча "читая, продолжая читать".

Подзона 5.6. -јинча, -јунча, -јинче, -јүнче. Грамматические признаки: деепричастие, обозначающее продолжение действия до окончания действия основного глагола охујунча "пока читал".

Подзона 5.7. -мадан, -мәден. Грамматические признаки: деепричастие, обозначающее временное отношение, образ действия дин мәден "молча".

Подзона 5.8. -андан, -әндән. Грамматические признаки: дее-

причастие /ср.: подзону 4.2./, кәләнден "с тех пор, как ты вернулся".

Подзона 5.9. -дырын, -дијин, -дурун, -дүјүн. Грамматические признаки: причастие - II лицо - единственное число - прошедшее время јаздырын мәктүб "письмо, которое ты написал".

Подзона 5.10. -ачары, -өчеји. Грамматические признаки: причастие /ср. подзону 4.7./ - јазачары мәктүб "письмо, которое он напишет".

Подзона 5.11. -јаачаг, -јөчек, -јачағ, -јөчөј. Грамматические признаки: будущее категорическое время или причастие будущего времени план тутулачаг "план будет составлен", тутулачаг план "план, который будет составлен".

Подзона 5.12. -сыныз, -синиз, -сунуз, -стүнүз. Грамматические признаки: сказуемость - II лицо - множественное число охумалысыныз "вы должны читать".

Подзона 5.13. -сынын, -синин, -сунун, -стүнүн. Грамматические признаки: принадлежность - II лицо - единственное число - родительный падеж онун бачысынын китабы "книга его сестры".

Подзона 5.14. -сында, -сунда, -синде, -стүнде. Сочетание аффиксов. Грамматические признаки: принадлежность - II лицо - единственное число - местный падеж онун чантасында "в его сумке".

Подзона 5.15. -ындан, -ундан, -индән, -үндән. Сочетание аффиксов. Грамматические признаки: принадлежность - II лицо - единственное число - исходный падеж онун евиндән "из его дома".

Подзона 5.16. -мысан, -мусан, -мисән, -мүсән. Сочетание аффиксов -мыш, -муш, -миш, -муш. /прошедшее повествовательное время/ с аффиксами -сан, -сән /второе лицо - единственное число/. Такие сочетания -мышсан, -мушсан, -мисән, -мүсән могут употребляться и без согласного -ш, в результате чего получаются формы

-мисан, -мусан, -мисен, -мусен. Грамматические признаки: прошедшее повествовательное /перфективное/ время – I лицо – единственное число алмышсан "ты купил".

Подзона 5.17. -мишди, -мушду, -мишди, -мушду. Грамматический признак: давнопрошедшее время јазымшды "писал раньше, давно".

Подзона 5.18. -ирмыш, -урмуш, -ирмиш, -тумыш. Грамматические признаки: определенный имперфект – пересказательность јазирмыш "оказывается, он писал /в тот момент/".

Подзона 5.19. -армыш, -өрмиш /ср. подзону 4.II/: јазармыш "он оказывается, имел обыкновение писать в прошлом".

Подзона 5.20. -јарды, -јерди. Грамматические признаки: неопределенный имперфект охујарды "он имел обыкновение читать в прошлом".

Подзона 5.21. -јирды, -јурду, -јирди, -јүрдү. Грамматический признак: определенный имперфект охујурду "он читал /в тот момент/".

Подзона 5.22. -ыбыш, -убумш, -ибыш, -тумыш. /ср. подзону 4.II/ јазыбыш "он оказывается писал когда-то".

Четвертая зона /аффиксы, состоящие из 4-х букв/.

Подзона 6.0. -јасы, -јеси. Грамматические признаки: необходимое наклонение или причастие.

Необходимое наклонение выражает действие, которое необходимо выполнить охујасыjam "мне необходимо читать", јазыласы мәктүб "письмо, которое необходимо написать".

Подзона 6.1. -малы, -мәли. Грамматические признаки: повествовательное наклонение или причастие јазмалыjam "я должен написать".

Подзона 6.2. -араг, -өрек. Грамматический признак, деепричастие /ср. подзону 5.I./ дурараг "встав".

Подзона 6.3. -анда, -әнде. Грамматический признак: деепри-

частие /ср.подзону 5.2./ - јазанда "в тот момент, в процессе, когда писал".

Подзона 6.4. -инча, -унча, -инче, -үнче. Грамматический признак: деепричастие /ср.: подзону 5.6./ алинча "в тот момент, когда я получил, получив только что".

Подзона 6.5. -јали, -јели. Грамматический признак: деепричастие, обозначающее действие, протекавшее до совершения действия спрягаемого глагола охујали "после того, как прочитал".

Подзона 6.6. -ачаг, -өчек, -ачағ, -өчәj. /ср. подзону 5.II/: китаб охуначаг "книга будет прочитана", охуначаг китаб "книга, которая будет прочитана".

Подзона 6.7. -ымыз, -умуз, -имиз, -түмүз. Грамматические признаки: принадлежность – I лицо – множественное число китабымыз "наша книга".

Подзона 6.8. -ыныз, -унуз, -иниз, -түнүз. Грамматические признаки: принадлежность – II лицо – множественное число китабыныз "ваша книга".

Подзона 6.9. -ынын, -унун, -инин, -түнүн. Сочетание аффиксов, состоящее из аффикса принадлежности и аффикса родительного падежа. Грамматические признаки: принадлежность – III лицо – единственное число – родительный падеж.

Следует отметить, что здесь приходится учитывать омонимию целого ряда морфокомплексов. Это явление, обусловленное морфологическими причинами, наблюдается при присоединении к основам, оканчивающимся на согласный, аффиксов принадлежности II/III лица единственного числа и следующих за ними падежных показателей. Так, например, морфокомплекс -ынын имеет двойное членение: малы-ынын "его товара" и малынын "твоего товара" и т.д. Поэтому, для морфокомплексов такого типа фиксируется принадлежность к

П/III лицу.

Подзона 6.10. -ында, -унда, -инде, -унде. Грамматические признаки: принадлежность - П/III лицо - единственное число - местный падеж дайтеринде "в твоей /его/ тетради".

Подзона 6.11. -сына, -сұна, -синә, -стнә. Сочетание аффиксов, состоящее из аффикса принадлежности /-сы/ и аффикса дательного падежа /-а, -ө/, также соединительного согласного "н".

Грамматические признаки: принадлежность - III лицо - единственное число - дательный падеж атасына "его отцу".

Подзона 6.12. -сыны, -сұну, -сини, -стнү. Сочетание, состоящее из аффикса принадлежности и аффикса винительного падежа, а также соединительного согласного /-сы-н-ы/. Грамматические признаки: принадлежность - II лицо - единственное число - винительный падеж атасыны "его отца".

Подзона 6.13. -јчын, -јчун, -јчин, -јчтн. Аффиксальный вариант послелога учтн, употребляемый в разговорном языке и речи персонажей произведений художественной литературы. Грамматические признаки: значения причины, цели и т.д. Более подробно ср.: 107, с. 291-294 атасытн "для его отца".

Подзона 6.14. -јмыш, -јмуш, -јмиш, -јмтш. Аффиксальный вариант недостаточного глагола-связки имиш. Грамматические признаки: прошедшее - пересказательность төлебејміш "оказывается, /он/ был студентом".

Подзона 6.15. -јчан, -јчен. Аффиксальный вариант послелога -чан, -чен. Грамматические признаки: обозначает временной или пространственный предел совершения действия.

Подзона 6.16. -масы, -мәси. Сочетание аффикса. Грамматические признаки: отглагольное существительное - принадлежность - П/III лицо - единственное число онун келмеси "его приезд".

Подзона 6.17. -дырт, -дурт, -дирт, -дүрт. Плеонастическое сочетание двух аффиксов побудительного залога. Действие совершается по поручению, просьбе, или приказу первого объекта через посредство второго и третьего лица /ср.: подзону 7.21./ - јаздырт-маг "поручить /кому-либо/ заставить писать".

Подзона 6.18. -инчи, -унчу, -инчи, -тнчт. Формальный признак порядковых числительных: бешинчи "пятый".

Подзона 6.19. -арды, -әрди. Грамматический признак: неопределенный имперфект јазарды "он /она/ имел/а/ обыкновение писать вообще в прошлом".

Подзона 6.20. -ырды, -урду, -ирди, -үрдт. Грамматический признак: определенный имперфект јазырды "он писал /в тот момент/!"

Седьмая зона /аффиксы, состоящие из 5-х букв/.

Подзона 7.0. -асы, -еси /ср.: подзону 6.0./, јазасыјам "мне необходимо писать".

Подзона 7.1. -маг, -мек, -мағ, -мәј. Аффикс инфинитива. В азербайджанском языке инфинитив имеет все категории глагола, кроме наклонения - время, лицо, залог, вид. Инфинитив образуется присоединением к основе глагола аффикса -маг, -мағ, -мек, -мәј, при этом глагольные основы склоняются и принимают аффиксы принадлежности. В отличие от традиционных словарей, в АС основ глагол представлен без аффикса инфинитива јазмаг "писать".

Во втором лице единственного числа повелительное наклонение в отличие от других наклонений глагола, не имеет формальных признаков и не принимает никаких аффиксов. В словаре основ глагол представлен без аффикса инфинитива, в повелительном наклонении второго лица, единственного числа кеть "иди".

Подзона 7.2. -јыб, -јуб, -јиб, -јүб. Аффикс деепричастия. Действие относится к прошедшему повествовательному времени, обоз-

начает состояние, образ действия, причину или цель совершения действия. Как омонимичный аффикс также обозначает действие совершенное в прошедшем повествовательном времени дурмајыб "не вставая".

Подзона 7.3. -алы, -ели. /ср.: подзону 6.5./ - јазалы "послого как писал".

Подзона 7.4. -јар, -јер. Грамматические признаки: будущее некатегорическое время, обозначающее предположение совершения действия охујарам "я буду читать". В сочетании с формой на -маз, -мез образует деепричастие, указывающее на действие, за которым тотчас следует другое действие. Определение второй функции этого аффикса возможно только на синтаксическом уровне.

Подзона 7.5. -чаг, -чек. Грамматические признаки: деепричастие, указывающее на начало совершения действия спрягаемого глагола алчаг "получив".

Подзона 7.6. -маз, -мез. Грамматические признаки: отрицание будущее некатегорическое время - II лицо единственного числа оху-
маз "он не будет читать".

Подзона 7.7. -јан, -јен. Омонимичный аффикс. Грамматические признаки: причастие настоящего времени или деепричастие охујан "читаюй".

Подзона 7.8. -мыш, -муш, -миш, -муш. Омонимичный аффикс. Грамматические признаки: обозначает прошедшее повествовательное /перфективное/ время или причастие/ јазмышам "я уже написал".

Подзона 7.9. -јыр, -јур, -јир, -јүр. Грамматические признаки: настоящее время охујур "читает".

Подзона 7.10. -нын, -нун, -нин, -нүн. Грамматические признаки: родительный падеж гапынын "от твоей двери".

Категория падежа в азербайджанском языке охватывает сущес-

7

вительное, прилагательное, числительное, местоимение, инфинитив. При склонении имен существительных в родительном, дательном и винительном падежах происходит чередование конечных гласных -г и -к, г -е, к -и. Это фонетическое явление при АМЛ строго учитывается.

Подзона 7.11. - дан, -ден. Грамматический признак: исходный падеж дағдан "с горы".

Подзона 7.12. -лар, -лэр. Грамматический признак, множественное число китаблар "книги".

Подзона 7.13. -јам, -јем. Грамматические признаки: сказуемость - I лицо - единственное число тәлебәјем "я студент".

Подзона 7.14. -јым, -јум, -јим, -јтм. Грамматические признаки: повелительное наклонение - I лицо - единственное число оху-
јум "читать бы мне".

Подзона 7.15. -јаг, -јак. Грамматические признаки: повелительное наклонение - I лицо - множественное число /охујаг "чи-
тать бы нам".

Подзона 7.16. -миз, -муз, -миз, -мүз. /ср.: подзону 6.6./ чантамыз "наша сумка".

Подзона 7.17. -низ, -нуз, -низ, -нүз. /ср. подзону 6.8./ - чантаныз "ваша сумка".

Подзона 7.18. -сан, -сан. Грамматические признаки: сказуемость - II лицо - единственное число тәлебәсән "ты студент".

Подзона 7.19. -јин, -јун, -јин, -јүн. Грамматические признаки: повелительное наклонение - II лицо - множественное число хујун "читайте".

Подзона 7.20. -сын, -сун, -син, -стн. Грамматические признаки: повелительное наклонение - III лицо - единственное число урсун "пусть он встанет".

Подзона 7.21. -лыр, -дур, -дир, -дүр. Грамматические признаки: сказуемость – III лицо – единственное число или побудительный залог јазды́маг "заставить писать".

Подзона 7.22. -ыны, -уну, -ини, -тнү. Сочетание аффиксов принадлежности и винительного падежа. Грамматические признаки: принадлежность – II/III лицо – единственное число – винительный падеж китабыны "твою /его/ книгу".

Подзона 7.23 -нна, -уна, -инө, -тнө. Сочетание аффиксов принадлежности и дательно-направительного падежа. Грамматические признаки: принадлежность – II/III лица – единственное число – дательно-направительный падеж китабына "твоей /его/ книге".

Подзона 7.24. -мир, -мур, -мир, -мүр. Контаминированное сочетание аффиксов отрицания /-ма/ и настоящего времени /-ыр/. Грамматические признаки: отрицание – настоящее время алмыр "он не получает".

Подзона 7.25. -јла, -јле. Аффиксальная форма послелога иље. Грамматические признаки: послелог или наречие образа действия елијлә "рукой".

Подзона 7.26. -јди, -јди. Аффиксальная форма недостаточного глагола-связки иди. Грамматические признаки: связка–прошедшее время тәлебәди "был студентом".

Подзона 7.27. -чин, -чун, -чин, -чүн. /ср.: подзону 6.13./ – сөнинчин "для тебя".

Подзона 7.28. -чан, -чэн. /ср.: подзону 6.15./ – евәчен "до дома".

Подзона 7.29. -јым, -јул, -јил, -јүл. Грамматические признаки: страдательный залог дејilmək "быть оказанным".

По формальным признакам страдательный и возвратный залоги в азербайджанском языке практически не различаются в рамках мор-

фологического анализа, принадлежность глагола к страдательному или возвратному залогу нередко определяется по контексту.

Подзона 7.30. -нчи, -нчи.. /ср.: подзону 6.18/ – икинчи "второй".

Подзона 7.31. -дыг, -дур, -диј, -дүј. Грамматические признаки: причастие – прошедшее время. Этот аффикс самостоятельно не употребляется, выступая лишь в сочетании с аффиксами принадлежности.

Подзона 7.32. -јыг, -јик, -јут, -јүк. Грамматические признаки: сказуемость – I лицо – множественное число тәлебәјик "мы студенты".

Восьмая зона /аффиксы, состоящие из 2-х букв/.

Подзона 8.0. -ја, -је. Омонимичный аффикс. Грамматические признаки: желательное наклонение или дательно-направительный падеж көрек охуја "читал бы он", гапыја "к двери". Желательное наклонение глагола имеет простую и сложную форму и два времени – будущее и прошедшее Охуја идин "читал бы ты". Желательное наклонение прошедшего времени образуется при помощи служебных глаголов иди и имиш.

Подзона 8.1. -са, -сө. Грамматический признак: условное наклонение көлсөм "если я приду".

Сложная форма условного наклонения образуется при помощи служебных глаголов иди и имиш көлсө идим "если бы я пришел". Формальным признаком условности в азербайджанском языке является также связка исө.

Подзона 8.2. -ыб, -уб, -иб, -үб. /ср.: подзону 7.2./ – алыб "получил, получив".

Подзона 8.3. -ар, -эр. /ср.: подзону 7.4./ – јазарам "я буду писать".

Подзона 8.4. -ан, -эн. /ср.: подзону 7.7./ - јазан "пишу-
щий".

Подзона 8.5. -ди, -ду, -ди, -ду. Грамматический признак: прошедшее категорическое время јаздим "я написал".

Подзона 8.6. -ир, -ур, -ир, -ур. /ср.: подзону 7.9./.

Подзона 8.7. -ма, -ме. Грамматические признаки: отрицание или отглагольное существительное јазма "не пиши", јазма "пиша-
ние".

Подзона 8.8. -ин, -ун, -ин, -ун. Омонимичный аффикс. Грамма-
тические признаки: 1. родительный падеж дејтерин "твоей тетрадь",
2. принадлежность - II лицо - единственное число дејтерин "твоя
тетрадь", 3. повелительное наклонение јазын "пишите".

Подзона 8.9. -ни, -ну, -ни, -ну. Грамматические признаки:
винительный падеж арын "пчелу".

Подзона 8.10. -да, -де. Грамматический признак: местный па-
деж евинде "в доме твоем /его/".

Подзона 8.11. -ам, -эм. Грамматические признаки: сказуе-
мость - I лицо - единственное число муеллихем "я учитель".

Подзона 8.12. -ым, -ум, -им, -ум. Омонимичный аффикс. Гра-
матические признаки: принадлежность - I лицо единственное число
или повелительное наклонение китабым "моя книга", јазым "писать
бы мне".

Подзона 8.13. -аг, -ек. /ср.: подзону 7.15/ - јазаг "писать
бы нам".

Подзона 8.14. -сы, -су, -си, -су. Грамматические признаки:
принадлежность - II лицо - единственное число атасы "его отец".

Подзона 8.15. -мы, -му, -ми, -му. Грамматический признак:
вопросительность алдымы "получил ли он?".

Подзона 8.16. -ла, -ле. /ср. подзону 7.25./ - китабла "с

книгой".

Подзона 8.17. -кы, -ку, -ки, -ку. Грамматический признак:
принадлежность ахшамы "вечерний", "относящийся к вечеру".

Подзона 8.18. -на, -на. Грамматические признаки: дательно-
направительный падеж - принадлежность П/Ш лица евине "/его/
тво-
ему дому".

Подзона 8.19. -нл, -ул, -ил, -ул. /ср.: подзону 7.29./ -
дејүлмек "быть избитым".

Подзона 8.20. -ыш, -уш, -иш, -үш. Грамматический признак:
взаимно-совместный залог јазышмаг "переписываться".

Подзона 8.21. -нг, -уг, -ик, -ук. Грамматические признаки:
сказуемость - I лицо - множественное число /ср.: подзону 7.32/ -
тәлебејик "мы студенты".

Девятая зона /аффиксы, состоящие из I-й буквы/.

Подзона 9.0. -а, -е. Омонимичный аффикс. Грамматические
признаки: желательное наклонение или дательно-направительный
падеж китаба "к книге", јаза "писал бы он".

Подзона 9.1. -м. Грамматические признаки: принадлежность -
I лицо - единственное число чантам "моя сумка".

Подзона 9.2. -ы, -у, -и, -ү. Омонимичный аффикс. Граммати-
ческие признаки: винительный падеж или принадлежность - II лицо -
единственное число китабы "его книга".

Подзона 9.3. -г, -к. Грамматические признаки: сказуемость -
I лицо - множественное число алдыг "мы получили".

Подзона 9.4. -н. Грамматические признаки: принадлежность -
II лицо - единственное число чантан "твоя сумка".

Подзона 9.5. -ш. /ср.: подзону 8.20/ - мелашмек "блеять",
"совместно блеять".

Распределение перечисленных аффиксов по зонам и по подзонам

их грамматические признаки и коды в сжатой форме представлены в таблице 8.

Таблица 8. Распределение аффиксов по зонам и подзонам, их характеристики и коды

Зоны и подзоны	Аффиксы	Грамматические характеристики	Коды	
			1	2
I.0.	-јачарымыз, -јөчөјимиз	Причастие, I лицо, множественное число, будущее категорическое время	E AJW	
I.I.	-јачарыныз, -јөчөјиниз	Причастие, 2 лицо, множественное число, будущее категорическое время	E + XW	
I.2.	-јачаглары, -јөчеклери	Причастие, 3 лицо, множественное число, будущее категорическое время	E AJW	
2.0.	-ачарымыз, -өчөјимиз	Причастие, I лицо, множественное число, будущее категорическое время	E AJW	
2.1.	-ачарыныз, -өчөјиниз	Причастие, 2 лицо, множественное число, будущее категорическое время	E + XW	
2.2.	-ачаглары, -өчеклери	Причастие, 3 лицо, множественное число, будущее категорическое время	E AJW	
2.3.	-јачагмыш, -јөчекмис	Будущее-прошедшее/будущее/в прошедшем/, пересказательность	X фф	

1	2	3	4
3.0.	-јачарым, -јөчөјим	Причастие, I лицо, единственное число, будущее категорическое время	E ALW
3.1.	-јачарын, -јөчөјин	Причастие, 2 лицо, единственное число, будущее категорическое время	E + LW
3.2.	-дыглары, -дүглары	Причастие, 3 лицо, множественное число, прошедшее категорическое время	E LЖF
3.3.	-јачагды, -јөчекди	Будущее-прошедшее /будущее в прошедшем/) фф
3.4.	-ачагмыш, -өчекмис	Будущее-прошедшее /будущее в прошедшем/, пересказательность	X фф
4.0.	-јаркен, -јөркен -јирке, -јиркен, -јүркен	Деепричастие	= фф
4.1.	-мышкен, -мушкен, -мишкен, -мүшкен	Деепричастие	= фф
4.2.	-јандан, -јөндөн	Деепричастие	= фф
4.3.	-дирлар, -дурлар, -дирлер, -дүрлөр	3 лицо, множественное число	Л фф
4.4.	-ачарым, -өчөјим	Причастие, I лицо, единственное число, будущее категорическое время	E ALW
4.5.	-ачарын, -өчөјин	Причастие, 2 лицо, единственное число, будущее категорическое время	E + LW

I	2	3	4
4.6.	-сындан, -сундан, -синден, -стиден	Принадлежность, 3 лицо, единственное число, исходный падеж	RЛХ
4.7.	-мышмыш, -мушмуш -мишмиш, -мушмуш	Давнопрошедшее, пересказательность	*ХФФ
4.8.	-јырмыш, -јурмуш, -јирмаш, -јурмаш	Определенный имперфект, пересказательность	EФФФ
4.9.	-аагды, -өчекди,	Будущее-прошедшее	*ФФФ
4.10.	-јармыш, -јермиш	Неопределенный имперфект, пересказательность	JХФФ
4.II.	-јачары, -јөчеји	Причастие, 3 лицо, единственное число, будущее категорическое время	EЛШ
4.12.	-јубмыш, -јубмуш, -јибмиш, -јубмуш	Давнопрошедшее, пересказательность	*ХФФ
5.0.	-магда, -мекде	Настоящее длительное время	1ФФФ
5.1.	-јараг, -јерек	Деепричастие	=ФФФ
5.2.	-јанда, -јенде	Деепричастие	=ФФФ
5.3.	-аркен, -еркен, -ыркен, -уркен, -иркен, -трукен	Деепричастие	=ФФФ
5.4.	-дыгда, -дикде, -дугда, -дукде	Деепричастие	=ФФФ
5.5.	-дыгча, -дугча, -дикче, -дукче	Деепричастие	=ФФФ
5.6.	-јынча, -јунча, -јинче, -јинче	Деепричастие	=ФФФ

I	2	3	4
5.7.	-мадан, -мәден	Деепричастие	=ФФФ
5.8.	-андан, -ендән	Деепричастие	=ФФФ
5.9.	-дағын, -дурун, -дијин, -дүгүн	Причастие, 2 лицо, единственное число, прошедшее время	E+ЛF
5.10.	-ачары, -өчеји	Причастие, 3 лицо, единственное число, будущее категорическое время	EЛШ
5.II.	-јачаг, -јечек, -јачас, -јечеј	Будущее категорическое время/ причастие	10ФФФ
5.12.	-сынз, -синиз, -сунуз, -сунз	Сказуемость, 2 лицо, множественное число	J+ХФ
5.13.	-сынны, -сунун, -синин, -стунн	Принадлежность, 3 лицо, единственное число, родительный падеж	RЛШ
5.14.	-сында, -сунда, -синде, -стиде	Принадлежность, 3 лицо, единственное число, местный падеж	RЛЧ
5.15.	-ындан, -ундан, -инден, -тиден	Принадлежность, 2 лицо, единственное число, исходный падеж	R+ЛХ
5.16.	-мисан, -мусан, -мисан, -мүсен	Прошедшее повествовательное время, 2 лицо, единственное число	b+ЛФ
5.17.	-мышын, -мушду, -миши, -мүшди	Давнопрошедшее	*ХФФ
5.18.	-ырмыш, -урмуш, -ирмиш, -тұрмаш	Определенный имперфект, пересказательность	JХФФ
5.19.	-армыш, -өрмиш	Неопределенный имперфект, пересказательность	JХФФ
5.20.	-јарды, -јөрди	Неопределенный имперфект	JФФФ

I	2	3	4
5.21.	-јирды, -јурду,	Определенный имперфект	Г ф ф
	-јирди, -јурдү		
5.22.	-ыбмыш, -убмуш,	Давнопрошедшее, пересказатель-	*Х ф
	-ибмиш, -тбмущ	ность	
6.0.	-јасы, -јеси	Необходимостное наклонение/ причастие	Д ф ф
6.1.	-малы, -мели	Долженствовательное наклонение/ причастие	Т ф ф
6.2.	-араг, -әрек	Деепричастие	= ф ф
6.3.	-анда, -әнде	Деепричастие	= ф ф
6.4.	-ынча, -инче,	Деепричастие	= ф ф
	-унча, -тнчө		
6.5.	-јала, -јели	Деепричастие	= ф ф
6.6.	-ачаг, -өчек,	Будущее категорическое время/ причастие	Ю ф ф
	-ачаг, -өчөј		
6.7.	-ымыз, -имиз,	Принадлежность, I лицо, множественное венное число	R+Ж ф
	-умуз, -тмуз		
6.8.	-иниз, -иниз,	Принадлежность, 2 лицо, единственное число	R+Л Ч
	-унуз, -тнуз		
6.9.	-ынин, -инин,	Принадлежность, 2/3 лицо, единственное число, родительный падеж	R+Л Ш
	-унун, -тнун		
6.10.	-ында, -инде,	Принадлежность, 2/3 лицо, единственное число, местный падеж	R+Л Ч
	-унда, -тнде		
6.11.	-сына, -сине,	Принадлежность, 2/3 лицо, единственное число, дательный падеж	R+Л Ш
	-сұна, -стне		
6.12.	-сыны, -сини,	Принадлежность, 2/3 лицо, единственное число, винительный падеж	R+Л З

I	2	3	4
	-суну, -стнту	венное число, винительный падеж	P ф ф
6.18.	-јчын, -јчин,	Послелог	
	-јчун, -тчун		
6.14.	-јмыш, -јмис	Связка, прошедшее время, пересказательность	И в ф ф
6.15.	-јчан, -јчен	Послелог	P ф ф
6.16.	-масы, -меси	Отглагольное существительное, принадлежность, 3 лицо, единственное число	—
6.17.	-дирт, -дирт.	Понудительный залог	Р ф ф
	-дурт, -дүрт		
6.18.	-инчи, -инчи,	Порядковое числительное	1 ф ф
	-унчу, -тнчт		
6.19.	-арды, -әрди	Неопределенный имперфект	Г ф ф
6.20.	-ырды, -ирди,	Определенный имперфект	Г ф ф
	-урду, -труду		
7.0.	-асы, -еси	Необходимое наклонение/причастие	Д ф ф
7.1.	-маг, -мек,	Инфинитив	Й ф ф
	-маг, -меј		
7.2.	-јиб, -јиб,	Деепричастие/прошедшее время	: ф ф
	-јуб, -јүб		
7.8.	-алы, -әли	Деепричастие	= ф ф
7.4.	-јар, -јэр	Будущее некатегорическое время/ причастие	Д и ф ф
7.5.	-маз, -мез	Будущее некатегорическое время/ отрицание/причастие	Д- ф
7.6.	-чаг, -чек	Деепричастие	= ф ф

I	2	3	4
7.7.	-јан,-јен	Причастие/деепричастие	Дфф
7.8.	-мыш,-миш,	Прошедшее повествовательное время/ причастие	Qфф
7.9.	-јир,-јир.	Настоящее время	Вфф
7.10.	-нин,-нин,	Родительный падеж	Шфф
	-нун,-нун		Xфф
7.11.	-дан,-ден	Исходный падеж	Xфф
7.12.	-лар,-лер	Множественное число	Xфф
7.13.	-јам,-јем	Сказуемость, I лицо, единственное число	ГЛф
7.14.	-јым,-јим,	I лицо, единственное число, повелитель-	АЦф
	-јум,-јут	ное наклонение	
7.15.	-јаг,-јек	I лицо, множественное число, повелитель- ное наклонение/желательное наклонение	АЖЧф
7.16.	-мыз,-миз,	Принадлежность, I лицо, множественное	РАЖф
	-муз,-муз	число	
7.17.	-сан,-сен	2 лицо, единственное число, сказуемость	+ЛГф
7.18.	-ныз,-низ,	Принадлежность, 2 лицо, множественное число/сказуемость	Р+Жф
7.19.	-јин,-јин.	2 лицо, единственное число, повелитель- ное наклонение	Ц+Жф
7.20.	-сын,-син,	3 лицо, единственное число, повелитель-	УЛф
	-сун,-стн	ное наклонение	
7.21.	-дир,-дир.	3 лицо, единственное число, сказуемость/	1Лф
	-дур,-дур	побудительный залог	
7.22.	-југ,-јик,	Сказуемость, I лицо, множественное число	ГАЖф

I	2	3	4
7.23.	-ны,-ни,	Принадлежность, 2/3 лицо, единственное	RЛЛ3
	-уну,-тн	число, винительный падеж	
7.24.	-ына,-иње,	Принадлежность, 2/3 лицо, единственное	РЛЛШ
	-уна,-тне	число, дательно-направительный падеж	
7.25.	-мир,-мир,	Отрицательная форма настоящего време-	-Вфф
	-мур,-мур	ни	
7.26.	-јла,-јле	Послелог/наречие образа действия	Рфф
7.27.	-јды,-јди	Связка, прошедшее время	ИРфф
7.28.	-чин,-чин,	Послелог	Рфф
	-чун,-чун		
7.29.	-чан,-чен,	Послелог	Рфф
7.30.	-јул,-јил,	Страдательный залог	Лфф
	-јул,-јул		
7.31.	-нчи,-нчи	Порядковое числительное	Гфф
7.32.	-ды,-диј,	Причастие, прошедшее время	ЕFф
	-дур,-дур		
8.0.	-ја,-је	Желательное наклонение/дательный падеж	2фф
8.1.	-са,-се	Условность	Гфф
8.2.	-ыб,-иб,	Деепричастие/прошедшее повествователь-	1фф
	-уб,-уб	ное время	
8.3.	-ар,-ер	Будущее некатегорическое время/ причастие	ДWф
8.4.	-ан,-ен	Причастие/деепричастие	Рфф
8.5.	-ди,-ди,	Прошедшее категорическое время	Fфф
	-ду,дт		
8.6.	-ир,-ир,	Настоящее время	Вфф
	-ур,-ур		

I	2	3	4
8.7.	-ма,-ме	Отрицание/отглагольное существительное	9 ффф
8.8.	-ын,-ин,	Родительный падеж/принадлежность, 2/3	9 ффф
	-ун,-тн	лицо, единственное число/повелительное наклонение	
8.9.	-ны,-ни,	Винительный падеж	3 ффф
	-ну,-нү		
8.10.	-да,-де	Местный падеж	2 ффф
8.11.	-ам,-эм	Сказуемость, I лицо, единственное число	АЛФ
8.12.	-ым,-им,	Принадлежность, I лицо, единственное число/повелительное наклонение	Бфф
8.13.	-аг,-ек	Повелительное наклонение, I лицо, множе- ственное число/желательное наклонение, I лицо, множественное число	АЖЦФ
8.14.	-сы,-си,	Принадлежность, 3 лицо, единственное	РЛФ
	-су,-ст	число	
8.15.	-мы,-ми,	Вопросительность	Эфф
	-му,-мт		
8.16.	-ла,-ле	Послелог	Рфф
8.17.	-кы,-ки,	Принадлежность	Рфф
	-ку,-ку		
8.18.	-на,-не	Дательный падеж, принадлежность, 2/3 лицо	Щфф
8.19.	-нл,-ил,	Страдательный залог	<фф
	-ул,-ул		
8.20.	-ыш,-иш,	Взаимно-совместный залог	>фф
	-уш,-тш		
8.21.	-ыг,-ик,	Сказуемость, I лицо, множественное число	ГАЖФ
	-уг,-ук		

I	2	3	4
9.0.	-а,-ө	Желательное наклонение/дательный падеж	2 ффф
9.1.	-м	Принадлежность, I лицо, единственное число	РАЛФ
9.2.	-ы,-и,-у	Винительный падеж/принадлежность, 3 лицо, единственное число	Яфф
9.3.	-г,-к	I лицо, множественное число	АЖФ
9.4.	-н	Принадлежность, 2 лицо, единственное число	+Лфф
9.5.	-ш	Взаимно-совместный залог	>фф

2. Морфологическая и лексико-морфологическая индексация ма-
шинных основ и аффиксов азербайджанского языка.

Кодирование грамматической и лексико-грамматической инфор-
мации осуществляется по схеме, разработанной в общесоюзной группе
"Статистика речи" /2,19-50; 22/. В основе данной индексации пред-
лагается классификация частей речи и грамматических категорий, пред-
ставленная в I главе, которая дает возможность осуществлять не-
противоречивое кодирование основ и аффиксов.

Таблица 9. Коды частей речи, охватывающие с/ф., которые
неприсоединяются к аффиксам

Части речи	Коды	Примеры
Послелоги	р	еввел, иле, тчн
Союзы	с	вө, нем
Частицы	у	бес, каш
Модальные слова	б	демек, шубнесиз
Связки	и	иди, имиш
Междометия	!	ај, оғ

Таблица I0. Коды частей речи, охватывающих с/ф, которые присоединяют аффиксы

Части речи	Коды	Примеры
Имя существительное	С	әлке, черек
Имя прилагательное	П	јакшы, сојуг
Имя числительное	Ч	беш, бешинчи
Местоимение	М	мен, бу
Глагол	Г	алмаг, вермәк
Наречие	Н	јаваш, ашары

Словообразовательные аффиксы /словообразовательные категории и формы имен и глаголов/ не кодируются. Содержание их с/ф рассматриваются как самостоятельные лексические единицы. При этом фиксируется лишь принадлежность к определенной части речи и лексическое значение.

Словоизменяющие аффиксы /словоизменяющие категории и формы имен и глаголов/ при АМА тюркского текста являются непосредственным объектом внимания, так как работа алгоритма АМА в основном заключается в приведении с/ф к словарному виду и анализе словоизменяющих аффиксов в составе с/ф.

Таблица II. Коды словоизменительных категорий имен и форм имен

Словоизменяющие категории и формы имен	Коды	Примеры
<u>Категория падежа</u>		
Родительный	Ш	адамын, Гапынын
Дательно-направительный	Щ	адама-Гапыја
Винительный	З	адамы, Гапыны
Местный	Ч	адама, Гапыда
Исходный	Х	адамдан, Гапыдан
Категория принадлежности	Р	китабым, дәфтерим
Категория склоняемости	Ж	китабдыр, дәфтердир

Таблица I2. Коды словоизменяющих категорий и форм глагола

Словоизменяющие категории и формы глагола	Коды	Примеры
I	2	3
<u>Залоги глагола</u>		
Страдательный	<	јазылмаг, атылмаг
Взаимно-совместный	>	јазышмаг, атышмаг
Понудительный	Ф	јаздиртмаг, атдирамаг
<u>Наклонение глагола</u>		
Повелительное	Ц	јазым, атым
Изъявительное	Ж	охујурам, атырам
Желательное	К	охујам, атам
Долженствовательное	У	охумалијам, атмалијам
Необходимое	Н	охујасијам, атасијам
Условное	Г	охусам, атсам
Продолжительное	/	охумагдајам, јазмагдајам
<u>Времена глагола</u>		
Прошедшее категорическое	F	охудум, јаздым
Прошедшее повествовательное	б/	охумушам, јазмышам
Настоящее	В	охујурам, јазырам
Будущее категорическое	И	охујачагам, јазачагам
Будущее исказиторическое	О	охујарам, јазарам
Давно прошедшее	*	охумушду, јазмыши
Пересказательность	Х	јазачагмыш, охумутмыш
Неопределенный имперфект	Г	охујармыш, јазармыш
Определенный имперфект	Л	охујурду, јазырлы
Будущее-прошедшее)	охујачагмыш, јазачагмыш
Отрицание	—	охумарам, јазмарым

I	2	3
Инфинитив	Й	охумаг, јазмаг
Причастие	Е	охујан, јазан
Деепричастие	=	охујуб, јазуб

Таблица I3. Коды общих /именных и глагольных/ категорий

Категория и формы	Коды	Примеры
Категория числа		
единственное	~	китаб, дефтер
множественное	Ж	китаблар, дефтерлэр
Категория лица		
1 лицо	А	охујарам, охујуруг
2 лицо	+	охујурсан, охујурсунуз
3 лицо	І	охујур, охујурлар
Вопросительность	Э	охујурму? калирми?

Кодирование омонимичных основ и аффиксов /омонимия, не снимаемая на морфологическом уровне/:

Таблица I4. Коды омоними основ

Омонимичные основы /возможные грамматические характеристики/	Коды	Примеры
Существительное/глагол	3	газ, јаз
Существительное/прилагательное	4	гоча, көзел
Глагол/наречие	5	дүз, дик
Глагол/прилагательное	6	сары, бурма, горхаг
Числительное/глагол	7	мин, гырх
Союз/частица	8	анчаг, ки
Отрицание/отглагольное существительное	%	охума, јазма

Таблица I5. Коды омонимии аффиксов /омонимия, не снимаемая на морфологическом уровне/

Омонимичные аффиксы /возможные грамматические характеристики	Коды	Примеры
Необходимое накл./причастие	Д	-асы/-еси, -јасы, -јеси/
Деепричастие/причастие	Д	-ан/-ен, -јан, -јен/
Понудительный залог/сказуемость	І	-дир/-дир, -дур, -дур/
Целательное накл./дательно-направительный падж	2	-а/-е, -ја, -је/
Принадлежность/винительный падж	Я	-и/-у, -т, -и/
Долженство.накл./причастие	Т	-мали/-моли/
Принадлежность, I лицо, единственное число/повелительное наклонение	Б	-ым/-им, -ум, -тм/
Будущее категорическое время/причастие	И	-ачаг/-өчек, -јачаг, -јечек
Причастие/прш. повест. время	О	-мыш/-миш, -муш, -мтш/
Родительный падж/принадлежность, 2/3 лицо, единственное число/повелительное наклонение	9	-ин/-ин, -ун, -тн/
Деепричастие/прошедшее повеств. возвратительное время	:	-јиб/-јиб, -јуб, -јуб/; -миш/-миш, -муш, -мтш/

3. Кодирование графем

В связи с тем, что АШУ ЭМ работает на основе русской и латинской графики, возникает необходимость замены специфических графем тюркских языков соответствующими русскими и латинскими литерами. Для удобства кодировки текстов и для универсальности программ по АШУ тюркских языков К.Б.Бектаевым предложены единые обозначения идентичных фонем различных тюркских языков /12, 20/. Автор при этом указывает и недостатки обозначения одной буквы двумя символами /12, 22/.

При разработке систем АА азербайджанского языка пришлось модифицировать и дополнять предложенную К.Б.Бектаевым единую транскрипцию. Дело в том, что при замене одной буквы двумя символами, алломорфы одного аффикса могут попасть в разные зоны и подзоны, что может привести к "хаосу" в списке основ и аффиксов. Напр., аффикс деепричастий -јынча, -јунча, -јинче, -јтнче распределился бы в результате по разным подзонам. Алломорфы данного аффикса - јынча, -јунча, -јинче /-јынч'a, -јунч'a, -јинч'e/ разместились бы в шестой зоне, а алломорф -јтнче /-јунч'e/ в седьмой, тогда как они имеют единую грамматическую характеристику и их целесообразно было бы поместить в одну подзону. С учетом вышеуказанного специфические азербайджанские графемы заменены следующими знаками /табл. 16/.

Таблица 16. Кодирование азербайджанских графем

Специфические азербайджанские буквы	Г К Ө Ч Ө Ь Џ Ү
По транскрипции К.Б.Бектаева	Г' К' Ө' Ч' З' Х' И' Ү'
Принятая транскрипция при АА азербайджанского языка	Ғ Ҳ Ҷ Ҵ Ҹ ҹ һ

4. Адаптация общетюркского алгоритма АА к азербайджанскому языку.

Адаптация общетюркского алгоритма к азербайджанскому языку состоит в следующем:

а/ прежде всего определяется длина с/ф. В зависимости от длины с/ф сравнивается с соответствующей зоной машинных основ. Поскольку максимальная длина единиц АС основ не превышает 16 букв,

то, если длина с/ф превышает 16, на следующем этапе путем отсечения букв с конца с/ф выделяется ее часть длиной $N = 16$ букв.

б/ если в зоне, включающей машинные основы длиной $N = 16$ букв, с/ф не найдена, то на последующих этапах анализа ее длина последовательно уменьшается на одну букву и полученная часть сравнивается с АС до тех пор, пока не будет найдена. Следует отметить, что если в результате последовательного уменьшения с/ф на одну букву основа не найдена, то данная с/ф записывается в список незнакомых слов /это можно объяснить двумя причинами: или основа не зафиксирована в АС, или при перфорации допущена ошибка/.

в/ для найденной машинной основы фиксируются присущие ей грамматические характеристики /или переводной эквивалент/.

г/ далее следует проверка на наличие остатка в с/ф /после выделения в ней машинной основы/. Максимальная длина машинных аффиксов в нашем списке не превышает 9 букв. Поэтому, если длина остатка /сочетание аффиксов/ превышает 9 букв / $\ell > 9$ /, то на следующем этапе, как и в случае основ, выделяется левая часть остатка длиной $\ell = 9$ букв и сравнивается с аффиксами, находящимися в соответствующей зоне.

Если аффикс не найден, то, как и при анализе основ, осуществляется последовательное уменьшение указанного сочетания букв путем отсечения одного знака справа и оставшаяся часть сравнивается с соответствующей зоной аффиксов. Эта операция повторяется вплоть до выделения всех аффиксов, и выделенным аффиксам приписываются соответствующие им грамматические признаки. Если же конечным результатом отсечения букв является $\ell = 0$, то анализ заканчивается записью с/ф в список незнакомых слов.

В качестве иллюстрации проследим, как идет АА с/ф чагырдырымызлардансызымы? - "вы из тех, которых мы визывали?"

Прежде всего определим длину этой с/ф. Поскольку с/ф состоит из 25 букв, т.е. $N > 16$, то на следующем этапе алгоритма она сравнивается с соответствующими по длине группами слов и основ АС.

Наша цепочка не представлена в АС. Поэтому на последующих этапах анализа ее длина последовательно уменьшается на одну букву и полученная часть сравнивается с АС до тех пор, пока не будет приведена к словарному виду чагыр. После этого слово и его грамматический признак /"глагол"/ записываются в резервное поле.

Далее следует проверка на наличие остатка в с/ф после выделения в ней канонической формы чагыр. В нашем случае остаток состоит из 20 букв: -дынынзлардансынызы. Поскольку такого аффикса в списке нет, то, как и при анализе основы, осуществляется последовательное уменьшение указанного сочетания букв путем отсечения одного знака справа и сравнение оставшейся части с соответствующими аффиксами списка. Эта операция повторяется вплоть до выделения причастного аффикса -лыр, представленного в списке.

Аналогичным образом выделяются и остальные аффиксы, которым приписываются соответствующие грамматические признаки: -дар- /мн.ч./, -мыз- /принадлежность I л.мн.ч./, -дан- /исходный падеж/, -сынзы- /сказуемость, II л.мн.ч./. В результате на выходе алгоритма получаем каноническую форму /основу/ вместе с ее грамматическим признаком основы и перечисление аффиксов вместе с их грамматическими характеристиками /ср.рис.8/.

5. Проверка объяснительной силы модели анализа.

Построенная модель автоматического морфологического анализа была подвергнута трем видам проверки:

I. Построенный алгоритм на основе данной модели реализован на ЭВМ ЕС-1050. Полученные результаты показали достаточно высо-

SLOVO1='ЧАГҮРДЫҒЫНЫЗЛАРДАНСЫНЫЗЫ'	N=	25;
ЧАГҮРДЫҒЫНЫЗЛАРДАНСЫНЫЗЫ		
ЧАГҮР ОСНОВА СРАВИЛ		
ЧАГҮР ГЛАГОЛ		
АФФ. НАЙДЕН	ДМФ	
INFAF='ЕГЕБ';		
ДМФ ПРИЧАСТИЕ	ПРОЧЕД.КАТЕГ.ВРЕМЯ	
АФФ. НАЙДЕН	ИМЫЗ	
INFAF='РАЖБ';		
ИМЫЗ ПРИНАДЛЕЖНОСТЬ	1-Е ЛИЦО	
АФФ. НАЙДЕН	ЛАР	
INFAF='ХЕББ';		
ЛАР ИМОХ.ЧИСЛО		
АФФ. НАЙДЕН	ДАН	
INFAF='ХЕББ';		
ДАН ИСХОДНЫЙ ПАД		
АФФ. НАЙДЕН	СЫНЫЗ	
INFAF='+КЭББ';		
СЫНЫЗ 2-Е ЛИЦО	ИМОД.ЧИСЛО	
АФФ. НАЙДЕН	НЯ	
INFAF='ЗЕББ';		
НЯ ВОПРОСИТЕЛЬН.		
НЧ ПОСЛЕДН. АФФ.		

Рис.8. Машинный результат автоматического морфологического анализа.

кую экспланаторность модели. С/ф, анализируемых компьютером, оказались вполне корректными, были получены необходимые грамматические характеристики, соответствующие реальным с/ф. Фрагмент машинного результата представлен на рис. 8.

2. Данная модель была применена и для других тюркских языков. При этом оказалось, что модель анализа может быть легко адаптирована к любому тюркскому языку. Для этого необходимо установить длину основ и аффиксов, зафиксированных в АС основ и аффиксов данного языка.

В частности, данная модель анализа реализована, кроме азербайджанского языка, и на материале турецкого языка.

3. Данная модель анализа включена в реальные системы турецко-русского и азербайджанско-русского машинного перевода и составляет начальный блок этих систем, экспериментальная реализация на ЭВМ этих систем показала состоятельность данной модели в связи с автоматической переработкой тюркских текстов.

Экспериментальное опробование обоих алгоритмов показало практически 100 процентов правильных решений.

ГЛАВА III

ЛЕКСИКО-МОРФОЛОГИЧЕСКИЙ МП ТҮРКСКОГО ТЕКСТА НА РУССКИЙ ЯЗЫК

I. Проблема лексико-морфологического МП с тюркских языков

Как известно, начиная с 60-х гг. исследования в области МП носили преимущественно теоретический характер. Это имело свои объективные причины: прежде всего, для практической реализации не было необходимых технических средств. ЭВМ имели малый объем памяти и небольшую скорость, и потому не отвечали требованиям автоматического перевода. С другой стороны, исследователи, работающие в этой отрасли, не имели достаточной подготовки. Очень слаба была связь и взаимопонимание между языковедами, математиками, программистами и кибернетиками. И хотя специалистами, занимавшимися МП, был разработан целый ряд формализованных описаний грамматики, вопрос о МП все-таки оставался нерешенным и запутанным. Выдвигались теоретические гипотезы, но забывалась прикладная сторона, не уделялось нужного внимания согласованию практики с теорией /Б1; 58; II4/. Дело дошло до того, что в 1961 году на заключительном заседании конференции, посвященной МП, было предложено снять с повестки дня вопрос о МП как практической задаче и усилить исследования в области разработки теоретических вопросов МП, в результате чего, несмотря на то, что наша страна являлась пионером МП, в настоящее время, "наше отставание от передовых западноевропейских и американских коллективов можно оценить здесь в 8-10 лет" /Б1, 7/.

В последние же годы наблюдается новая волна исследований в области МП. Вопрос о состоянии МП в нашей стране рассматривался

специально созданной временной научно-технической комиссией Государственного Комитета Совета Министров СССР по науке и технике. Тщательно обсуждены и принты решения о переходе к практической реализации МП на уровне теоретических и экспериментальных исследований /28, 48/. Для координации и руководства работ по построению практических систем МП создан Всесоюзный центр перевода /ВЦП/. Проходивший в ноябре 1979 года в г. Москве международный семинар, посвященный машинному переводу /МП/, еще раз подтвердил, что эта проблема, которая имеет большое социальное и научное значение, и теперь находится в центре внимания ученых /70/.

При пословном /текстическом/ переводе в память ЭВМ вводится двуязычный словарь /например, русско-азербайджанский АС/. При каждом русском слове указывается адрес эквивалентного азербайджанского слова. При помощи этого адреса в машине разыскивается переводный эквивалент азербайджанского слова. При этом основное внимание уделяется скорости перевода, качество же последнего бывает на низком уровне. Практически получены удовлетворительные результаты пословного перевода для некоторых групп языков. С другой стороны, известно, что "переход от одного языка к другому не есть простое, механическое "наклеивание" одних "ярлыков" на место других, т.е. на одни и те же заранее данные сами по себе ясно выделенные мысли" /99, 47/. Синонимия, омонимия и многозначность слов усложняет пословный перевод. При пословном переводе сохраняется тот же порядок слов, что и в исходном языке. По этой причине пословный перевод для языков с различными типологическими особенностями /например, русского и азербайджанского/ нельзя признать удовлетворительным. Естественно, что на первичном этапе развития МП речь должна была идти о пословном переводе. Но принять пословный перевод конечным научным достижением, остано-

виться на этом и отказаться от переводов с лингвистическим анализом было бы ошибочно.

Идея автоматического анализа текстов появилась в 50-х годах XX века в рамке проблем ИЛ. Автоматический лингвистический анализ /АЛ/ - это переход от текста /обычной графической формы/ к смыслу /на основе сведений о данном языке/ по определенным инструкциям. Алгоритм анализа текста - это совокупность сведений и правил, позволяющих осуществить переход от текстовой /графической/ формы к смыслу.

МП с лингвистическим анализом открывает новый перспективный этап в области автоматизации перевода.

В настоящее время отдельным вопросам АЛ в связи с МП /в частности, АМ/ посвящены некоторые исследования /I; 7; 64; 69; 70; 97; 98; 99; 100; 101; 106; 109/.

В тюркологии также разработаны отдельные фрагменты МП с тюркских языков /65; 105; 8; 37; 39; 67; 68; II; III; I 28/.

Большой интерес представляет предварительный алгоритм распознавания глагольных групп в азербайджанском языке для научно-технического стиля, разработанный И.В.Михайловой /71/. При распознавании глаголов автор опирается на тот факт, что в азербайджанском языке сказуемые в предложении занимают конечную позицию перед точкой. Если глагол находится в середине фразы, то после него следует запятая. Если глагол составной, то в этом случае сказуемое выявляется по начальным буквам. Затем, в зависимости от глагола, предыдущее слово выдается на печать вместе с выделенным глаголом. Каждое слово в предложении проверяется на наличие аффикса инфинитива. При обнаружении инфинитива выделяются после-логи, модальные глаголы, могущие следовать за ним.

В диссертации П.К.Байрамовой /8/ рассматривается проблема

перевода на татарский язык придаточных определительных предложений с относительным /союзным/ словом "который" и экспериментальная оценка на ЭЦМ некоторых особенностей этих предложений. В исследовании указываются три типа основных способа перехода от одного языка к другому на уровне конструкции:

1/ перевод русского придаточного предложения татарским простым предложением с определительным оборотом,

2/ перевод татарским сложноподчиненным предложением,

3/ перевод татарским сложносочиненным предложением.

Автор особо обращает внимание на изучение структуры и семантики определительных предложений с местоимением "который" и использование причастных оборотов при переводе придаточных определительных предложений.

П.К.Байрамовой на основе анализа способов перевода удалось найти формальные признаки причастного способа перевода и аналитического способа для МП, на основе которых можно полностью передавать содержание русских определительных предложений с относительным словом "который". Автором был проведен эксперимент на ЭМ М-20, давший возможность получить информацию о способе перевода данного предложения и соотнести при переводе местоимение "который" с определенным словом.

Следует отметить, что вышеизложенные работы являются фрагментами МП с тюркских языков. Для системного решения проблемы МП с тюркских языков необходимо разработать проблему АЛА тюркского текста. Только после полного решения данной проблемы с применением грамматики можно было бы получить качественные образцы МП. При этом на первый план выходит установление соответствий между грамматическими формами – одна из сложных проблем в системе МП /38; 47; 74/.

При МП с языков разного типа и структуры проблема установления грамматических соответствий приобретает более усложненный характер. Напр.: аналитические формы будущего времени в английском языке, образующиеся с помощью вспомогательных глаголов *shall/will* + инфинитив переводятся на азербайджанский язык формами, образуемыми аффиксальным способом "*I shall read - мен охуячағам*" "Я буду читать".

В работе В.Я.Динеса /74/ описывается машинная реализация алгоритма азербайджанско-английских глагольных соответствий в виде азербайджанско-английского автоматического словаря глагольных форм. Как указывает автор, этот словарь является бесконтекстным – для каждой азербайджанской глагольной формы выдается ряд возможных эквивалентов, и эта многозначность не снимается. Важной чертой данного словаря является тот факт, "что осуществляется не перевод на английский язык лексического значения, выраженного основой азербайджанского глагола, а выдача той формы английского глагола, которая соответствует цепочке аффиксов, присоединенных к указанной основе. Причем это достигается посредством сочетания только определенных глагольных словоформ – вспомогательных и знаменательных. Никакая дополнительная информация при "переводе" не выдается, но при необходимости она может быть легко получена" /74/.

Используя правила установления соответствий между азербайджанскими и английскими глагольными формами, при помощи алгоритма МП с азербайджанского языка можно получить образцы лексико-морфологического МП азербайджанской глагольной с/ф на английский язык.

Проследим, как можно перевести на английский язык с/ф кетмелисен "ты должен идти".

В условном АС основ можно найти, что кет, – соответствует

английскому глаголу *(to) go*

В списке аффиксов можно найти, что аффикс подзоны 6.1. -малы, -мели соответствует английскому модальному глаголу *must* а аффикс подзоны 7.17. -сан, -сан указывает на личное местоимение II лица единственного числа В результате получаем, что форме кетмелисан "ты должен идти" в английском языке соответствует следующее сочетание морфем: *go + must + you*

Дальнейшее преобразование этого сочетания позволяет получить правильную цепочку: *You must go*

Аналогичным образом, можно перевести и с/ф охујашагам "буду читать" *I shall read*

На тех же основаниях можно построить общую схему лексико-морфологического МП тюркской словоформы на русский язык /рис. 9/.

2. Общая структура азербайджанско-русского лексико-морфологического МП.

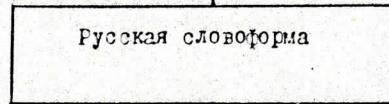
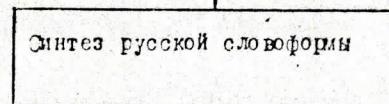
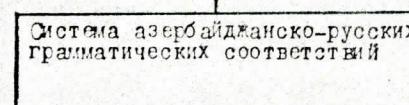
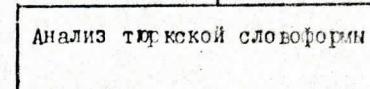
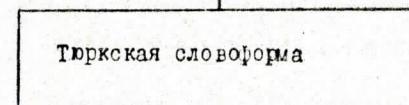
При построении системы азербайджанско-русского лексико-морфологического МП использована модель ИМЯ.

Система азербайджанско-русского МП состоит из следующих компонентов:

- АС основ,
- АС общеупотребительных оборотов /ACSO/
- список глаголов, функционирующих как служебный глагол /2-й компонент оборотов/,
- список слов, функционирующих как I-й компонент оборотов,
- алгоритм и блок-схема азербайджанско-русского МП на морфологическом уровне с элементами синтаксического анализа.

Как известно, качественный МП может быть осуществлен, естественно, только после полного лингвистического анализа, который помимо морфологического уровня имеет также синтаксический и се-

В В О Д



В Н В О Д

Рис. 9. Общая схема лексико-морфологического МП тюркской словоформы

манический уровня и подуровни. АИА является лишь первичным этапом анализа, и образцы МП, полученные на основе морфологического уровня, будут менее качественными и слишком "грубыми" /80; 54; 58; 60/. Поэтому в данном алгоритме учтены и элементы синтаксического анализа, что в значительной степени повышает качество перевода МП /табл. 17/.

Таблица 17. МП после подключения синтаксического блока

Словосочетание	МП на уровне морфологического анализа	МП после подключения синтаксического блока
тедрис ет/мек/	преподавание, обучение-делать, сделать	преподавать, учить
несаб ет/мек/	счет, расчет-делать, сделать	считать, подсчитать
баш вер/мек/	голова, головной-давать, отдавать	случиться, возникать
үрек вер/мек/	сердце, душа-давать, отдавать	ободрять, подбодрять
дерд чек/мек/	горе, скорбь-тянуть, таскать	горевать
гол чек/мек/	рука, ветвь-тянуть, таскать	подписать, расписываться
баш чек/мек/	голова, головной-тянуть, таскать	наведываться, навестить
баш гош/маг/	голова, головной-запрягать, присоединять	вмешиваться, обращать

При переводе основ фиксируется два наиболее употребительных эквивалента в данном подъязыке /газетная и общеупотребительная лексика/.

Омонимичные слова представлены в виде одной словарной единицы уз "лицо", "плавать". Следует отметить, что отдельно взятое омонимичное слово типа уз вне контекста невозможно отнести к определенной части речи. Однако во многих случаях омонимичность основ легко снимается при анализе аффиксов /ср.: глава I, 7.8./.

8. Особенности АС оборотов

АС оборотов включает около 2000 наиболее употребительных словосочетаний типа имя-глагол. Все обороты распределены по 20 группам в зависимости от глаголов, выступающих в качестве второго компонента сочетания. Первая группа включает обороты, в которых вторым компонентом является служебный глагол ет/ед "делать, сде-лать", во второй группе вторым компонентом сочетаний является служебный глагол ол "быть" и т.д.

Указанные глаголы даны отдельным списком, причем возможно расширение списка за счет проявления глагольных основ, что не влияет на функционирование синтаксического блока. Данный список составлен на базе АС основ на материале азербайджанского газетного текста и включает следующие глаголы ал, ет/ед, гоj, вер, ол, кер, чал, вур, кел, чек, сал, еле, душ, олун, лечир, газан, ко-стер, чыхар.

Кроме того, в памяти ЭВМ хранятся первые компоненты сочета-ний по группам /учитывая принадлежность к глаголам, данным в отдельном списке/. Внутри групп первые компоненты распределены исходя из их длины в именительном падеже, единственном числе, так как в составе оборотов они употребляются в этой форме /табл. 18/.

Глаголы фиксируются в форме императива, как и в АС основ. Надо отметить, что и АС основ и АС оборотов при необходимости можно расширить. Это не влияет на функционирование алгоритма

/19,20/.

Таблица 18. Фрагмент азербайджанско-русского АС оборотов

Основы или первые компоненты оборотов	Перевод или значение оборота, вторым компонентом которого является глагол ет/ед
азад	освобождать
арзу	желать, мечтать
гəсd	покушаться, посягать
зəнн	полагать, предполагать

Отдельные слова, в зависимости от глагольных групп, с которыми они соотнесены, переводятся по-разному. Например, слово тəныш в АС основ представлено в значении "знакомый, знакомец". Это же слово в АС оборотов представлено различным образом: в группе ет/ед как "познакомить, ознакомить", в группе дл как "познакомиться" /ср.табл. 19/.

Таблица 19. Перевод основ в отдельности и в составе оборотов

Исходная форма слова или основа	Перевод или значение в АС	Перевод или значение в АС оборотов	
		группа ет/ед	группа ол
тəныш	знакомый знакомец	ознакомить познакомить	познакомиться
hесаб	счет, расчет	считать, подсчитать	считаться, приводиться в расчет
дахил	внутренняя часть, внутренность	вводить, включать	вступать, входить
наранат	беспокойный суевийский	беспокоить тревожить	беспокоиться, тревожиться

Таким образом, например, если при переводе слово /рəгс/ найдено в АС основ, то оно переводится как "танец, пляска". Если же это слово найдено в группе служебного глагола ет/ед "делать, сделать", оно будет переведено как "танцевать", "плясать".

Самостоятельность слов или их принадлежность к обороту выясняется при работе элементарного синтаксического блока совместно с алгоритмом АМА азербайджанского языка.

4. Тезаурус в системе АЛТ

Качественный МП на лексическом уровне обработки текста, не может быть достигнут без распознавания смысла многозначных слов и словосочетаний. При этом особое внимание следует уделить устранению многозначности многокомпонентных словосочетаний. Анализ текстов по общественно-политической тематике, на перевод которых ориентирована система турецко-русского МП, выявил широкое распространение в них сложных терминологических словосочетаний. Опыт показывает, что морфологосинтаксический анализ, являющийся одним из основных этапов в системе турецко-русского МП, не всегда позволяет успешно решить проблему снятия многозначности таких словосочетаний, т.к. в зависимости от сочетания ЛЕ в структуре последних они могут выражать разные значения. В качестве примера рассмотрим следующие словосочетания:

1/ Урасир истəһсал үасителəри

- современные средства производства

2/ Бeјнəлхалг өмек тəшкилаты

- международная организация труда

Если в тексте речь идет о современных средствах производства или о международной организации труда, то в данном случае сочетания средства производства и организация труда корректины, а современное производство и международный труд некорректины, они

несут совершенно иной смысл.

Однозначный выбор из множества синтаксически связанных элементов таких, между переводными эквивалентами которых имеет место семантическая связь, можно провести по аналогу существующих систем МП с помощью тезауруса, который в первую очередь предназначен для выявления отношений между семантическими единицами описываемой им предметной области /ПО/, вторая функция тезауруса – эксплицитное представление семантических отношений, отражающих структуру связей в анализируемой области знания. Будучи семантической моделью исследуемой ПО, тезаурус позволяет установить смысл языковой единицы /слова, словосочетания/ в тексте.

Возможности и пути использования тезауруса в системе турецко-русского МП нами будут изучены на отраслевом тезаурусе по общественно-политической тематике, в котором заданы семантико-статистические отношения, характеризующие реальные связи этой конкретной области. Для формирования смысловой системы терминологии использованы дефиниции соответствующих терминов.

В качестве модели тезауруса выбрана семантическая сеть, вершины которой сопоставлены лексическим значениям слов, а ребра – семантическим парадигматическим отношениям между словами. При построении семантической сети принята следующая методика. Ребро графа сопоставлено отношению непосредственной семантической производности, т.е. отношению между словом и его непосредственным семантическим компонентом /или обратному отношению – между словом и его непосредственным семантическим дериватом/. Считается, что слово, участвующее в определении другого слова, является непосредственным компонентом последнего. Таким образом, если одно слово входит в состав определения другого, то соответствующие вершины соединены другой, направленной из первой вершины во вто-

рю. Такой упрощенный критерий облегчает возможность автоматического построения семантической сети.

Семантически связанными считаются слова, связанные отношением непосредственной или опосредствованной семантической производности, а также слова, имеющие общий семантический компонент.

Граф непосредственной семантической производности, представляющий структуру ПО, отражаемой в тезаурусе, позволяет выявить для каждой ЛЕ ее семантические связи, как непосредственные, так и опосредственные. При этом отношение непосредственной семантической производности допускает любую дифференциацию семантических отношений. В соответствующем тезаурусе могут быть отношения "род-вид", "целое-часть" /иерархии/, ассоциативные отношения, отношения эквивалентности /синонимии/ и т.д. в зависимости от содержания и назначения АС.

Разрабатываемый отраслевой тезаурус по общественно-политической тематике в качестве семантического фильтра позволит в системе турецко-русского МП осуществить корректное соответствие синтаксически связанным элементам входного /турецкого/ языка их переводные эквиваленты на выходном /русском/ языке.

5. Алгоритм лексико-морфологического МП азербайджанского текста на русский язык с элементами синтаксического анализа.

1. Выделение с/ф начиная с конца предложения.
2. Определение длины с/ф / N /: если $N > 16$, перейти к п. 6, при отрицательном исходе к п. 3.
3. Сравнение с/ф с соответствующей группой АС: если с/ф найдена, перейти к п. 7, при отрицательном исходе к п. 3.
4. Уменьшение длины с/ф на одну букву / $N = N - 1$ /: если $N = 0$, перейти к п. 5, при отрицательном исходе к п. 3.
5. Записать с/ф в список незнакомых слов и перейти к п. 19.

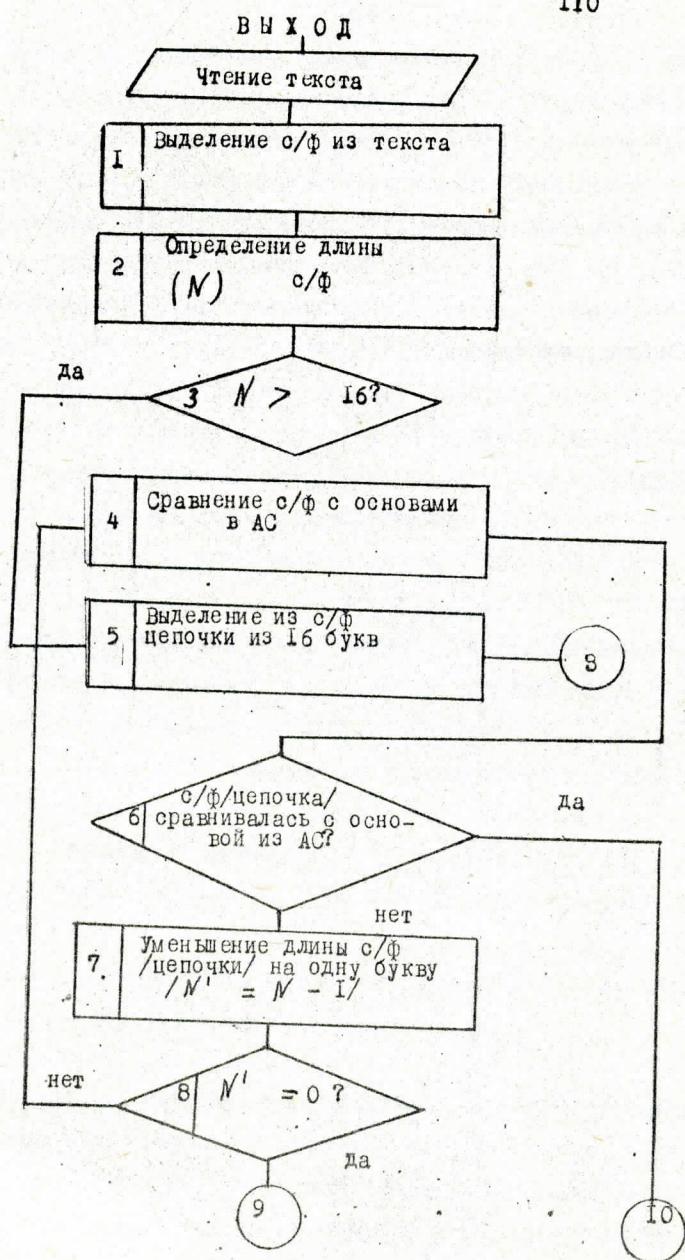
6. Выделить с/ф / N - 16/ и перейти к п. 3.
7. Сравнение с/ф со списком глаголов: если глагол найден перейти к п. 9, при отрицательном исходе к п. 8.
8. Записать перевод слова в РП и перейти к п. 12.
9. Выделить в предложении с/ф, предшествующую данному глаголу и перейти к п. 10.
10. Сравнение выделенной с/ф со списком, соответствующим данному глаголу: если слово найдено, перейти к п. 11, при отрицательном исходе к п. 2.
11. Записать перевод словосочетания в РП и перейти к п. 12.
12. Проверка наличия остатка: если $\ell = 0$, перейти к п. 16, при отрицательном исходе к п. 13.
13. Определение длины остатка / ℓ /: если $\ell > 9$, перейти к п. 17, при отрицательном - к п. 14.
14. Сравнение остатка с соответствующей группой аффиксов: если аффикс найден, перейти к п. 18, при отрицательном - к п. 15.
15. Уменьшение длины остатка на одну букву / - - 1/: если $\ell = 0$, перейти к п. 5, при отрицательном исходе к п. 14.
16. Напечатать перевод слова и перейти к п. 19.
17. Выделить остаток / $\ell = 9$ / и перейти к п. 14.
18. Записать признаки аффикса в РП после перевода слов и перейти к п. 12.
19. Проверить конец текста, при положительном исходе перейти к п. 20, при отрицательном к п. 1.
20. Конец /рис. 10/.

На рис. 10 показана принципиальная блок-схема алгоритма азербайджанско-русского лексико-морфологического И.П.

В качестве иллюстрации проследим процесс автоматического анализа словосочетания несаб ет/мек/. Прежде всего выделяем с/ф

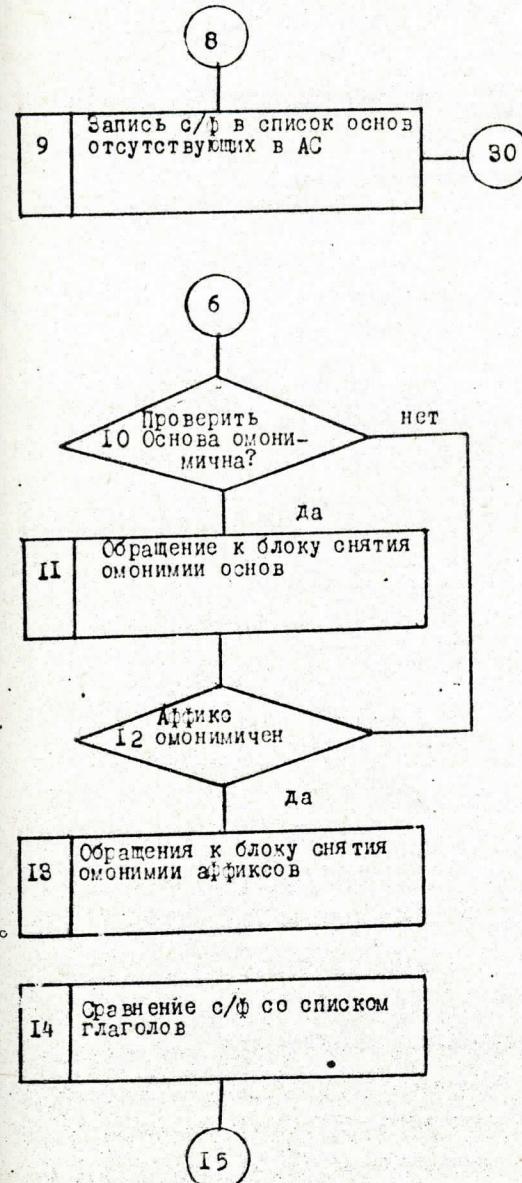
начиная с конца /с правой стороны/ предложения. Определим длину с/ф. Поскольку эта с/ф состоит из пяти букв ет/мек/ т.е. $N > 16$, то на следующем этапе алгоритма она сравнивается с соответствующими по длине группами слов в АС основ /для упрощения процесса предполагаем, что в словаре глагол представлен в инфинитиве/. Находим, что слово переводится как "делать, сделать". Далее следует проверить: - является оно самостоятельной лексической единицей или компонентом оборота. С этой целью данное слово сравнивается со списком глаголов. Находим, что оно представлено в списке глаголов-вторых компонентов оборотов. На следующем этапе выделяется в предложении с/ф, предшествующая данному глаголу, и сравнивается со списком группы слов, соответствующих данному глаголу. Это слово представлено в списке, и поэтому сочетание переводится как оборот "считать, подсчитывать".

На рис. 11 показан машинный результат лексико-морфологического И.П. словосочетаний несаб ет/мек/ "считать, подсчитывать", нем саб ол/маг/ "считаться", "приниматься в расчет".



110

III



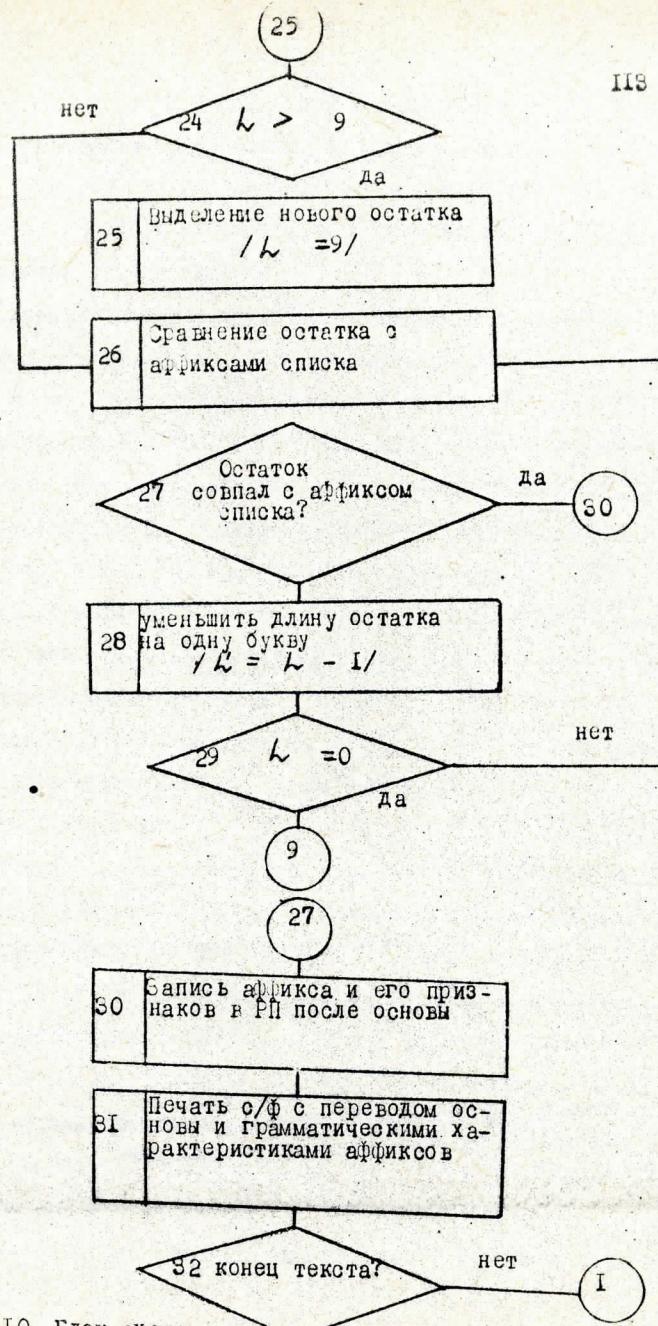
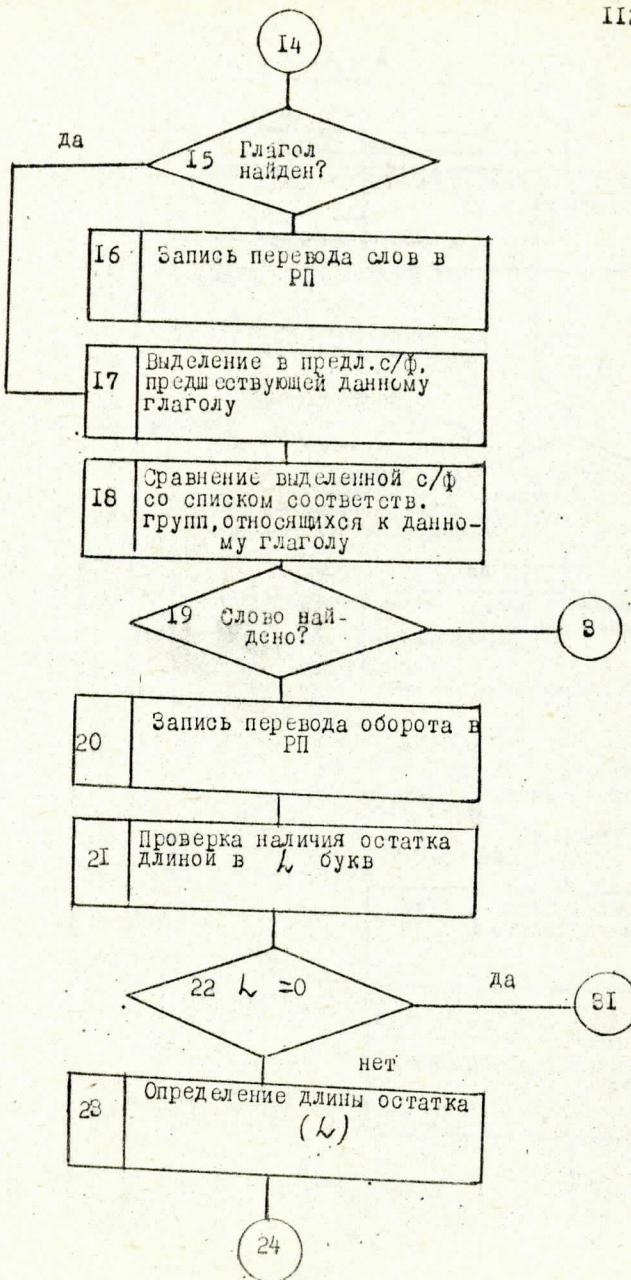


Рис. 10. Блок-схема лексико-морфологического МП азербайджанского текста на русский язык с элементами синтаксического анализа.

```

>>>>>>>>>>>>>>>>>>>
ГЕСАБ ЕТ
ЕТ
=====
ОСНОВА:
ЕТ ДЕЛАТЬ СДЕЛАТЬ
АФИКСЫ:
АФФ. В СЛОВЕ НЕТ
=====
ГЕСАБ ЕТ СЧИТАТЬ, ПОДСЧИТЫВАТЬ

>>>>>>>>>>>>>>>>>>
ГЕСАБ ОЛ
ОЛ
=====
ОСНОВА:
ОЛ БЫТЬ, СТАТЬ
АФИКСЫ:
АФФ. В СЛОВЕ НЕТ
=====
ГЕСАБ ОЛ СЧИТАТЬСЯ; ПРИНИМАТЬСЯ В РАСЧЕТ
>>>>>>>>>>>>>>>>>

```

Рис. II. Машинный результат лексико-морфологического ИП азербайджанского текста на русский язык

ЗАКЛЮЧЕНИЕ

Автоматизация переработки текстов на тюркских языках, обладающих развитой аффиксальной системой словообразования и словоизменения, невозможна без создания специальной машинной морфологии тюркских языков /ММТЯ/.

В работе рассмотрена проблема создания воспроизводящей инженерно-лингвистической модели машинной морфологии для тюркских языков на материале азербайджанского языка и реализация ее на ЭВМ. Предложенная модель позволяет провести полный морфологический анализ и получить необходимые грамматические характеристики для последующих этапов лингвистического анализа.

Одним из важнейших компонентов ММТЯ является алгоритм автоматического морфологического анализа. При построении алгоритма АМА основная задача заключается в получении оптимального способа автоматического отделения в составе с/ф аффиксов от основы и друг от друга. АМА функционирует на основе сравнения текстовых словоформ с единицами банка лингвистических данных по длине буквенных цепочек.

Данный вариант алгоритма, как показал эксперимент, является наиболее оптимальным для агглютинативных языков.

Указанная особенность АМА обуславливает структуры АС основ и оборотов, а также списка словоизменительных аффиксов.

Единицы ММТЯ /машинные основы и аффиксы/ распределяются по зонам, включающим единицы с одинаковым количеством букв, и подзонам, содержащим варианты аффиксов, т.е. показатели, тождественные по грамматическим признакам, но различающиеся фонетически /графически/.

Предложенный алгоритм АМА тюркского текста может быть ис-

пользован для анализа словоформ любого агглютинативного языка. Данный алгоритм реализован на материале азербайджанского языка на ЕС-ЭВМ.

АС основ азербайджанского языка включает около 2000 высокочастотных лексических единиц, отобранных из частотного словаря азербайджанской газетной лексики. Весь массив разбит на 16 зон в зависимости от длины составляющих единиц. Внутри каждой зоны основы располагаются в алфавитном порядке.

АС оборотов включает более 2000 наиболее употребительных словосочетаний типа имя - глагол. Все обороты распределены по группам в зависимости от глаголов, выступающих в качестве второго компонента сочетания.

Первая группа включает обороты, в которых вторым компонентом является служебный глагол ет/ед "делать, сделать", во второй группе вторым компонентом сочетания является служебный глагол од "быть" и т.д.

Азербайджанские словоизменительные аффиксы в зависимости от их длины распределены по девяти зонам. К первой зоне относятся аффиксы, состоящие из девяти букв, во вторую зону включаются 8-буквенные аффиксы и т.д.

Во всех трех списках учитывается зарытие графемного состава единиц, связанное с морфологическими особенностями азербайджанского языка.

Алгоритм АМА применен при разработке системы лексико-морфологического МП азербайджанского текста на русский язык. Данная система включает также элементы синтаксического анализа, что в значительной мере улучшило качество перевода.

АДД - автореферат докторской диссертации

АЗНИИТИ - Азербайджанский Научно-Исследовательский Институт Научно-технической информации

АЛА - автоматический лингвистический анализ

АМА - автоматический морфологический анализ

АПТ - автоматическая переработка текста

АПИИЯ - Азербайджанский Педагогический Институт Иностранных Языков

АС - автоматический словарь

АСОО - автоматический словарь общеупотребительных оборотов

АСУ - автоматизированная система управления

БИЛМ - воспроизводящая инженерно-лингвистическая модель

БИИТИ - Всесоюзный Институт Научной и Технической Информации

ВЯ - Вопросы языкоznания

ИАН Известия АН ССР, серия литературы и языка

ИИ - искусственный интеллект

ИЛОСИЯ - инженерная лингвистика и оптимизация преподавания иностранных языков.-Об.научных работ.Л.,1976,1980 гг.

ИПС - информационно-поисковая система

ЛА - лингвистический автомат

ИГИИЯ - Кинский Государственный Педагогический Институт иностранных языков

МТЯ - машинная морфология тюркских языков

МО - машинная основа

МП - машинный перевод

МПАИП - машинный перевод и автоматизация информационных процессов.М.,1975.

НТИ - научно-техническая информация. Ежемесячный научно-технический сборник.

СТРААТ - сборники: Статистика речи и автоматический анализ текста.

С/У - словоупотребление

С/Ф - словоформа

ШСОПИЯ - Школа-семинар по оптимизации преподавания иностранных языков с помощью технических средств /тезисы докладов и сообщений/.

ОЛА - обучающий лингвистический автомат.

ЛИТЕРАТУРА

1. Автоматический перевод: сб.статьй /под ред. и с предисл. О.С.Кулагиной и др./.-М.: Прогресс, 1971.-368 с.
2. Автоматическая переработка текста: Системный подход к решению проблем машинного перевода /отв.ред. Е.С.Тарасова/. - Кипинев: Штиинца, 1978, - 93 с.
3. Азербајҹан дилинин грамматикасы, I һиссө /морфология/. Бакы: Азербајҹан ССР Елмлөр Академијасынын нөшри, 1960, - 384 с.
4. Азербајҹанча-русча лүгәт /тертиб едәни Ҳ.Ә.Әзизбәјов/.-Бакы: Азеришр, 1965.- 417 с.
5. Арзиколов Ҳ.А., Пицровский Р.Г., Попеску А.Н., Хажинская М.В. Автоматизированная система тазаурусного аннотирования научно-технического документа.-НТИ, сер. 2, № 12, 1978, с. 12-20.
6. Бабанаров А. Разработка принципов построения словарного обеспечения турецко-русского машинного перевода: АКД.-Л.:Ленингр. гос.ун-т им. А.А.Ханова, 1981.- 17 с.
7. Базмаджян Р.А., Джанполадян Т.К., Тер-Микаелян Т.М. Основном варианте морфологического анализа и методе его реализации на специализированной машине.-В кн.: Проблемы кибернетики, вып. 15 - М.: Физматгиз, 1965, с. 209-233.
8. Байрамова П.К. Перевод на татарский язык русских придаточных определительных предложений с относительным словом "который" и экспериментальная оценка на ЭЦВМ некоторых вопросов перевода этих предложений: АКД.-Л., 1966.-16 с.
9. Баскаков Н.А. Двойственная природа слова и проблема частей речи /на материале тюркских языков/.-В кн.: Вопросы теории частей речи.-Л.: Наука, Ленингр. отд., 1968, с. 287-299.
10. Баскаков Н.А. Морфологическая структура слова в тюркских

языках различных типов.-В сб.: Морфологическая структура слова в языках различных типов.-М.Л.: Наука, 1968, с. 76-83.

11. Баскаков Н.А. Морфологическая структура слова и части речи в тюркских языках. Сов.востоковедение, 1957, № 1, с. 72-85.

12. Бектаев К.Б. Статистико-информационная типология тюркского текста.-Алма-Ата: Наука, 1978.-182 с.

13. Бектаев К.Б., Беляева Л.Н., Королева Т.Н., Марчук І.Н., Пиотровский Р.Г., Садчикова П.В., Чайковская И.И., Шингарева Е.А. Опыт машинного перевода английских и японских текстов.-НТИ, сер.2, № 5, 1981, с. 26-51.

14. Бектаев К.Б., Миразбеков С. Ет ст кт и форма өзгерткі ш аффикстер и н статистикалық талдаңы.-В сб.: Статистика казахского текста.-Алма-Ата: Наука, 1973, с. 136-196.

15. Бектаев К.Б., Миразбеков С. Сөз формалары.-В сб. Статистика казахского текста.-Алма-Ата: Наука, 1973, с. 197-224.

16. Белоногов Г.Г., Новоселов А.П. Автоматизация процессов накопления, поиска и обобщения информации.-М.:Наука, 1979.-256 с.

17. Бельская И.К. Язык человека и машина.-М.: Изд-во моск. ун-та, 1969.- 410 с.

18. Бельская И.К. О принципах построения словаря для машинного перевода.- ВЯ, № 3, 1959, с. 89-94.

19. Беляева Л.Н., Лукьянова Е.М., Пиотровский Р.Г. Машинная лексикография в автоматическом русском словаре.- В кн.: Вопросы лексикологии, лексикографии и прикладной лингвистики.- М.: АН ССР, Ин-т языкоznания, 1976, с. 28-43.

20. Билан В.Н., Боркун М.Н., Пиотровский Р.Г. Методы автоматического анализа и синтеза текста /учебн.пособие по курсу "Теория общ. и машинного перевода" для студентов 3-4 курсов язвуз/.-Минск: Мин.гос.пед.ин-т иностряз., 1977.-110 с.

21. Борисевич А.Д., Гочаренко В.В., Гречишнина В.И., Добрускина З.М., Кочеткова В.К., Крисевич В.С., Нехай С.А., Пиотровский Р.Г., Штирбу Т.А., Ястребова С.В. Кодирование грамматической информации в машинном словаре.-В кн.: Статистика текста, т.2: Автоматическая переработка текста.-Минск:Изд.БГУ, 1970, с.331-614.

22. Велиева К.А. Алгоритмическое описание правил порождения азербайджанских словоформ.-учен.зап.азерб.гос.ун-та им.С.М.Кирова,сер.яз. и лит.-Баку, № 5, 1968, с. 58-66.

23. Велиева К.А. Формальное описание синтеза азербайджанского слова: АКД.-М., 1971.- 20 с.

24. Велиев А., Велиева К.А., Махмудов М.А. Кибернетика и языкоzнание. В сб.: Лексико-морфологические структуры тюркских языков. Изд. АГУ им.С.М.Кирова,Баку, 1981,стр. 26-56.

25. Велиева К.А., Мамедова М.Г., Махмудов М.А., Пинес В.Я., Рахманов Дж.А. Принципы построения системы лексико-грамматического машинного перевода с тюркских языков. Международная конференция. "Теория и практика научно-технического перевода", М., 1985. Тезисы докладов стр. 133-134.

26. Велиева К.А., Махмудов М.А., Пинес В.Я. Проблемы автоматической переработки текстов на азербайджанском языке. Тезисы докладов семинара-совещания "Республиканская система научно-технической информации и опыт организации информационного обеспечения народного хозяйства. АЗНИИИТИ, 1985, с. 85-87.

27. Вертель Е.В., Горлин Г.Л. Обработка на ЕС 54М текстов естественных языков.-Учен.зап.Тартуского ун-та, вып. 502, 1979, с. 153-156.

28. Володин А.П., Храковский В.С. Типология морфологических категорий глагола /на материале агглютинативных языков/.- В кн.: Типология грамматических категорий.-М.:Наука, 1975, с. 170-196.

29. Вокуа В. О деятельности центра исследований по математическому переводу при Национальном центре научных исследований, Франция.— В сб.: Автоматический перевод.—М.: Прогресс, 1971, с. 819–838.
30. Герасимов В.Н., Марчук Ю.Н. Современное состояние машинного перевода.—В сб.: Машинный перевод и автоматизация информационных процессов.—М.: ВНИТИ, 1975, с. 5–18.
31. Гладкий А.В. и др. Элементы математической лингвистики.—М.: Наука, 1969.—192 с.
32. Глисон Г. Введение в дескриптивную лингвистику.—М.: ИЛ, 1959.—486 с.
33. Грамматика азербайджанского языка /под общей редакцией М.Ш.Ширалиева и Ə.В.Севоргяна/.—Баку; ӘДМ, 1971.—418 с.
34. Действующие системы ИП и автоматические словари: Обзорная информация /Перевод научн.-техн.лит., сер. 2. Машинный перевод и автоматизация информационных процессов/.—1979.—70 с.
35. Дәмирчизадә Ә.Ә. Әтасир Азәрбајҹан дили /I հայտ/.—Бакы: Маариф, 1972.—308 с.
36. Дрейзин Ф.А. Об одном способе синтаксического анализа простого предложения.—Научн.труды Ташк.Јн-та.Новая серия, вып. 208: Матем.науки, кн. 28, 1962, с. 76–81.
37. Дрейзин Ф. Таблица татарско-русских соответствий. В сб.: Машинный перевод.—М.: Труды ИТИ и ВТ АН СССР, вып. 2, 1961, с. 304–313.
38. Дрейзин Ф. и Рашитов Р. Принципы синтаксического анализа татарской фразы.—В сб.: Машинный перевод.—М.: Труды ИТИ и ВТ АН СССР, вып. 2, 1961, с. 295–308.
39. Дрейфус Х. Чего не могут вычислительные машины: Критика искусственного разума /пер.с англ./.—М.: Прогресс, 1978.—534 с.

40. Жеребин В.М. Принципы построения системы китайско-русских грамматических соответствий.—М.: ИТИ и ВТ АН СССР, 1963.—25 с.
41. Закиев М.З. Классификация частей речи и аффиксов в тюркских языках.—ВЛ, № 6, 1973, с. 3–8.
42. Зореф М.Г. Машинные основы и машинная морфология в немецко-русском автоматическом словаре: АКД.—1972.—28 с.
43. Зореф М.Г. Машинная морфология и ее применение в немецко-русском автоматическом словаре.—В кн.: СТРААТ.—Л.: Наука, Ленингр.отд., 1975, с. 245–259.
44. Зубов А.В. Обработка на ЕС ЭИМ текстов естественных языков.—Минск: Вышэйша школа, 1977.—173 с.
45. Зубов А.В. Переработка текста естественного языка в системе "человек-машина".— В кн.: СТРААТ.—Л.: Наука, Ленингр.отд., 1971, с. 286–484.
46. Иванов Вяч.Вс. Машинный перевод и установление соответствий между языковыми системами.—В сб.: Машинный перевод.—М.: Труды ИТИ и ВТ АН СССР, вып. 2, 1961, с. 47–69.
47. Ингве В.Г. Синтаксис и проблема многозначности.—В сб.: Машинный перевод.—М.: ИЛ, 1957, с. 208–304.
48. Исхаков Ф.Г. Лексико-грамматическая классификация слов и части речи в тюркских языках.—В кн.: Исследования по сравнительной грамматике тюркских языков.—М.: Изд. АН СССР, 1956, с. 72–77.
49. Исхакова Х.Ф. Автоматический синтез форм существительного в татарском языке.—ИТИ, сер. 2, вып. 8, 1968, с. 26–27.
50. Исхакова Х.Ф. Исследования в области формальной морфологии тюркских языков /на материале татарского литературного языка в сопоставлении с турецким и узбекским/: АКД.—1972.—25с.
51. Кенесбаев С.К., Пиотровский Р.Г., Бектаев К.В. Инженер-

ная лингвистика и тюркология.-Советская тюркология, № 6, 1970,
с. 5-16.

52. Коваль Н.А. Опыт построения словаря основ и таблиц
окончаний украинского глагола для анализа при машинном переводе.
- В кн.: Применение новых методов в изучении языка, вып. I.-
Днепропетровск: Днепропетровский гос.ун-т. им. 500-летия воссое-
динения Украины с Россией, 1969, с. 125-132.

53. Коверин А.А., Баклушин А.В. Общая схема грамматического
анализа и синтеза текстов в ходе машинного перевода.-В кн.: Модели-
рование языка и машинный перевод.-Иркутск: Иркутск.пед.ин-т.,
1979, с. 50-55.

54. Кононов А.Н. О природе тюркской агглютинации.-ВТ, № 1,
1976, с. 5-17.

55. Кононов А.Н. О фузии в тюркских языках.-В кн.: Структура
и история тюркских языков.-М.: Наука, 1971, с. 108-120.

56. Коптилов В.В. О формальном разграничении омонимии и поли-
семии.-В кн.: Прикладная лингвистика и машинный перевод.-Киев:
Изд. Киевского ун-та, 1962, с. 76-78.

57. Котов Р.Г. Лингвистика и современное состояние машинно-
го перевода в стране.-ВЯ, № 5, 1976, с. 37-59.

58. Кузнецов О.П., Адельсон-Вельский Г.М. Дискретная матема-
тика для инженера.-М.: Энергия, 1980.-842 с.

59. Кулагина О.С. Исследования по машинному переводу.-М.:
Наука, 1979.-820 с.

60. Лингвистическое обеспечение в системе автоматического
перевода третьего поколения /Предвар.публ. Е.Д. Адресян, И.И. Богу-
славский и др./.- М.: АН ССР, Научн.совет по компресн.проблеме
"Кибернетика", 1978.- 47 с.

61. Лингвистические проблемы "искусственного интеллекта".

Сборник научно-аналитических обзоров, АН ССР, Ин-т научн.информ.
по обществ.наукам.-М.: ИИОН, 1970.- 250 с.

62. Мальцев А.И. Алгоритмы и рекурсивные функции.-М.: Наука,
1965.-391 с.

63. Марчук Е.Н. Определение роли существительного и глагола
в английском языке при грамматическом анализе для машинного пе-
ревода.-НТИ, сер. 2, № 5, 1965, с. 39-43.

64. Марчук Е.Н. Методы моделирования перевода, изд. "Наука",
М., 1985.- 200 стр.

65. Махмудов М.А. Разработка систем формального морфологиче-
ского анализа тюркской словоформы. Вопросы тюркского языкоznания,
изд. "Наука", Алма-Ата, стр. 118-125.

66. Махмудов М.А., Пинес В.Я. База лексико-грамматических
данных в системе азербайджанско-русского ИП. Тезисы докладов на-
учно-технической конференции "Проблемы развития и совершенствова-
ния республиканской системы научно-технической информации", АЗНИ
ИИТИ, Баку, 1985 с.

67. Махмудов М.А., Пинес В.Я. Некоторые аспекты построения
систем ИП в тюркских языках. Международный семинар по машинному
переводу.М., ВЦП, 1980, с.

68. Международный семинар по машинному переводу /Москва, но-
ябрь, 1975г./: Тез.докл., Ред.кол.: Марчук Е.Н. /отв.ред./.

69. Международный семинар по машинному переводу /Москва,
ноябрь, 1979г./: Тез.док., Ред.кол.: Марчук Е.Н. /отв.ред./. и др.
-М.: ВЦП, 1979.- 195 с.

70. Михайлова И.В. Некоторые предварительные данные по вы-
делению глагольной группы в азербайджанском языке.-В сб. Стати-
стическое и информационное изучение тюркских языков /тезисы док.
и сообщений/.-Алма-Ата: АН Казахской ССР, Ин-т языкоznания, 1969,
с. 72.

71. Масир Азербајҹан дилинин морфологијасы.-Бакы: В.И. Ленин ад. АДГИ, 1961, 191 с.
72. Нильсон Н. Искусственный интеллект: Методы поиска решений. Пер. с англ. В.Л. Стефанюка, под. ред. С.В. Фомина.-М.: Мир, 1978. - 270 с.
73. Пинес В.Я. Некоторые вопросы автоматического перевода и тюркские языки:-Советская тюркология, № 3, 1974, с. 100-107.
74. Пинес В.Я. Моделирование структуры азербайджанских глагольных форм в связи с проблемой автоматического словаря: АКД.-М., АН СССР, Ин-т востоковедения, 1970.- 19 с.
75. Пинес В.Я., Ревзин И.И. Описание морфологической структуры азербайджанского глагола с помощью грамматики порядков. В сб.: Статистическое и информационное изучение тюркских языков /тезисы докладов и сообщений/.-Алма-Ата: АН Казахской ССР, Ин-т языкоznания, 1969.- с. 19.
76. Пиотровская А.А. Автоматическое приведение именных словоупотреблений к канонической форме.-ИТИ, сер. 2, № 1, 1977, с. 32-36.
77. Пиотровская А.А. Машинная морфология русского глагола.- В кн.: СТРААТ, 1972.-Л.: Наука, 1973, с. 260-277.
78. Пиотровская А.А., Пиотровский Р.Г. Математические модели диахронии и текстообразования.-В кн.: СТРААТ, 1974.-Л.: Наука, 1974, с. 361-40.
79. Пиотровский Р.Г. Инженерная лингвистика: Теория-эксперимент-реализация.-Изв. АН СССР, сер. лит. и языка, т. 57, № 1.-1978, с. 10-19.
80. Пиотровский Р.Г. Отраслевой вероятностный машинный перевод.-В сб.: Статистика текста 2 /Автоматическая переработка текста.-Минск: Изд. ВГУ, 1970, с. 5-82.

81. Пиотровский Р.Г. Инженерная лингвистика и лингвистические автоматы.-Вестник АН СССР, № 9, М., 1979, с. 112-119.
82. Пиотровский Р.Г. Текст, машина, человек.-Л.: Наука, Ленинград. отд., 1975.-327 с.
83. Пиотровский Р.Г. Инженерная лингвистика и теория языка.-Л.: Наука, Ленинград. отд., 1979.- 112 с.
84. Пиотровский Р.Г., Аникина Н.В., Аполлонская Т.А., Билан В.Н., Боркун М.Н., Лесоухин М.М., Шингарева Е.А. Формальное распознание смысла текста.-В кн.: СТРААТ, Л.: Наука, Ленинград. отд., 1980, с. 5-51.
85. Пиотровский Р.Г., Бектаев К.В. Машинный перевод: теория-эксперимент-реализация.-ВИ, № 5, 1977, с. 27-33.
86. Пиотровский Р.Г. Лингвистические аспекты "искусственного разума".-ВЯ, № 3, 1981, с. 27-39.
87. Попов З.В., Фердан Г.Р. Алгоритмические основы интеллектуальных роботов и искусственного интеллекта.-М.: Наука, 1976.-455с.
88. Поцелуевский А.П. Избранные труды. Ред. кол. А.Н. Кононов /отв. ред./ и др.-АН ТССР, Ин-т языка и лит. им. Махтумкули.-Ашхабад: Шым, 1975.-327 с.
89. Проблемы кибернетики /сб. статей/. Под ред. А.А. Ляпунова, вып. М.: Гизматгиз, 1961.-404 с.
90. Рафаэль В. Думашы: компьютер /Пер. с англ. под ред. В.Л. Стефанюка, с предисл. Д.А. Постолова/-М.: Мир, 1979.-107 с.
91. Ревзин И.И. Модели языка. М.: Изд. АН СССР, 1962.-191 с.
92. Ревзин И.И., Ілдашева Г.Д. Грамматика порядков и ее использование.-ВЯ, № 1, 1969, с. 42-56.
93. Роджерс Х. Теория рекурсивных функций и эффективная вычислимость. Пер. с англ., под. ред. В.А. Успенского.-М.: Мир, 1972.-624с.

94. Самойлович А.И. Краткая учебная грамматика современного османского-турецкого языка.-Л.: Ленингр.ин-т живых восточных языков, 1926.-154 с.
95. Севорян З.В. Аддиксы именного словообразования в азербайджанском языке /опыт сравнительного исследования/.-М.: Наука, 1966.- 437 с.
96. Совель И.В. Вопросы реализации систем автоматического анализа текстов естественных языков: АКД.-1980.- 16 с.
97. Смирницкий А.И., Ахманова О.С. О лингвистических основах преподавания иностранных языков /проблема новой языковой системы в связи с вопросом о единстве языка и мышления/. Иностранные языки в школе.-М.: Государственное учебно-педагогическое изд.мин. просв.РСФСР, № 3, 1954, с. 41-51.
98. Слейгл Дж. Искусственный интеллект: подход на основе эвристического программирования.Пер. с англ.-М.: Мир, 1973.-319 с.
99. Статистическо-информационное изучение тюркских языков /тезисы докл. и сообщений/, ред.кол.: В.А.Исенгалиева /отв.ред./ и др. - Алма-Ата: АН Казахской ССР, ин-т языкоznания, 1969.-94с.
100. Статистика речи и автоматический анализ текста /ред. кол.: Пиотровский Р.Г. /отв.ред./ и др.-Л.: Наука, 1971.-464 с.
101. Статистика речи и автоматический анализ текста, 1972 /ред.кол.: Пиотровский Р.Г. /отв.ред./ и др.-Л.: Наука, 1973.-339с.
102. Статистика речи и автоматический анализ текста /ред. кол.: Пиотровский Р.Г. /отв.ред./ и др. -Л.: Наука, 1974.-404 с.
103. Статистика речи и автоматический анализ текста /ред. кол.: Пиотровский Р.Г. /отв.ред./ и др.-Л.: Наука, 1980.-259 с.
104. Тамарашвили Л.Р. Морфологический анализ для грузинского языка /в целях автоматич.перевода/: АКД, 1965.-39 с.
105. Уинстон П. Искусственный интеллект /пер. с англ. В.Л.

- Степанюка, под.ред. и с предисл. Д.А.Поспелова/-М.: Мир, 1980,-519 с.
106. Фитиалов С.Я. О построении Формальной морфологии в связи с машинным переводом.-М.: АН ССР, ин-т научн.информ. Доклады на конференции по обработке информации, машинному переводу и автоматич.чтению текста, вып. 2, 1961.- 24 с.
107. Халитова Н.А. Начальный этап работы над машинным переводом /МП/ с татарского на русский.-В кн.: Проблемы тюркологии и истории востоковедения /материалы конференции, ред. М.З.Завиев и др./.-Казань: Изд.зап.Казан.ун-та, т. I22, кн. 4, 1962, с. 266-274.
108. Халитова Н.А., Бакирова Р.А., Гимадутдинова Р.У. Морфологический анализ при машинном переводе с татарского языка на русский.-Учен.зап. Казан.ун-та, т. I22, кн. 4, 1962, с. 98-105.
109. Халитова Н.А., Заидуллина И., Минуллина Д. Машинный перевод прилагательных с татарского языка на русский.-Учен.зап.Казан.ун-та, т. I24, кн. 2, 1964, с. 146-152.
110. Хант З. Искусственный интеллект.-М.: Мир, 1978.- 558с.
111. Խոյեյնզադե Մ. Մասիր Ազերբայջան դիլ /մօրֆոլոգիա/, Бакы, Маариф, 1973.- 358 с.
112. Чхайдзе М.П. Еще раз о мире и правде машинного перевода.-ВЯ, № 5, 1976, с. 50-57.
113. Чжео Қань-День. Модели в лингвистике и модели вообще.- В кн.: Математическая логика и ее применение.-М.: Наука, 1965, с. 281-292.
114. Чәфәров С. Ազերբայջան դիլində սөzdüzəldicili və səzdejiş-diiriçi şəkilçilər, Бакы: Маариф, 1968.-107 с.
115. Шевченко С.М. Алгоритм морфологического анализа японского текста.-М.: Ин-т точной механики и вычислите.техники АН ССР,

О ГЛАВЛЕНИЕ

ВВЕДЕНИЕ
ГЛАВА I. МАШИННАЯ МОРФОЛОГИЯ ТЮРКСКИХ ЯЗЫКОВ
I. Место морфологического анализа в системе АМТ
2. О необходимости создания машинной морфологии тюркских языков
3. Формальная классификация частей речи и грамматических категорий в машинной морфологии
4. Рабочие понятия ММТЯ
5. Цели и задачи ММТЯ
6. Воспроизводящая инженерно-лингвистическая модель ММТЯ
7. Алгоритм АМА - как композиция функций F /совокупность рекурсивных правил/
7.0. Функция выделения единиц анализа /K/
7.1. Функция расчленения с/ф /F/
7.2. Функция сопоставления информации /C/
7.5. Снятие омонимии основ и аффиксов на морфологическом уровне /Функция P/
7.8.0. Формальное определение Функции P
7.8.1. Омонимия, снимаемая при помощи сочетания аффиксов
7.8.2. Омонимия, снимаемая при помощи машинных аффиксов
7.8.3. Омонимия, не снимаемая на морфологическом уровне
7.4. Алгоритм АМА тюркской словоформы
7.5. Алгоритм снятия омонимии основ ■ аффиксов
7.5.1. Алгоритм снятия омонимии основ
7.5.2. Алгоритм снятия омонимии аффиксов

1964.- 80 с.

116. Щербак А.М. о морфологической структуре слова в тюркских языках.-В кн.: Морфологическая структура слова в языках различных типов.-М.-Л.,1965, с. 267-270.

117. Щербак А.М. Очерки по сравнительной морфологии тюркских языков /имя/, АН СССР, Ин-т языкоznания.-Л.: Наука,Ленингр. отд.,1977.-191 с.

118. Щербак А.М. Очерки по сравнительной морфологии тюркских языков /глаголов/, Л.: Наука,Ленингр.отд.,1981.- 184 с.

119. Щербак А.М. Способы выражения грамматических значений в тюркских языках.-ВЯ, № 1, 1957, с. 18-26.

120. Шрейдер Е. А. О понятии "математическая модель языка".-М.: Знание,1971.-68 с.

121. Якубова Н.К. Правила формального синтеза узбекских словоформ.-В кн.: Системы и уровни языка.-М.: Наука,1969.с. 189-151.

122. Якубова Н.К. Правила формального синтеза узбекских словоформ.-Ташкент: ФАН, 1979.-122 с.

123. Bangıoglu T. Türkçenin grameri - Istanbul;
Baha matbaası, 1974, с. 319-321.

124. Ergin M.. Türk dilbilgisi. — Зюдя:
Narodna prosвета, 1967. — 386с.

125. İmer K. Türkiye Türkçesinde kökler. — Ankara:
Türk dil kurumu yayınları, 1976, с.70.

126. Deliyeva K. A., Mahmudov M. A., Pines B. Y.
Automatic machine translation. Symposium on
Automatic compilation of dictionaries. Tallinn,
November, 25-27, 1985, Summaries, P. 81.

7.6. Преобразования информации функция /П/

ГЛАВА П. РЕАЛИЗАЦИЯ ВОСПРОИЗВОДСТВА ИНЖЕНЕРНО-ЛИНГВИСТИЧЕСКОЙ МОДЕЛИ ИМПЛ НА МАТЕРИАЛЕ АЗЕРБАЙДЖАНСКОГО ЯЗЫКА

- I. Описание банка лингвистических данных
- I.I. Машинные основы азербайджанского языка, распределение их по зонам и характерные особенности
- I.2. Машинные аффиксы азербайджанского языка, их грамматические характеристики и распределения по зонам и подзонам
- 2. Морфологическая и лексико-морфологическая индексация машинных основ и аффиксов азербайджанского языка
- 3. Кодирование графем
- 4. Адаптация общетюркского алгоритма к азербайджанскому языку
- 5. Проверка объяснительной силы модели анализа

ГЛАВА П. ЛЕКСИКО-МОРФОЛОГИЧЕСКИЙ ИП ТИРКСКОГО ТЕКСТА НА РУССКИЙ ЯЗЫК

- I. Проблема лексико-морфологического ИП с тюркских языков
- 2. Общая структура азербайджанско-русского лексико-морфологического ИП с элементами синтаксиса
- 3. Особенности АС обзоров
- 4. Тезаурус в системах АИТ
- 5. Алгоритм лексико-морфологического ИП тюркского /азербайджанского/ текста на русский язык с элементами синтаксического анализа

ЗАКЛЮЧЕНИЕ

СОКРАЩЕНИЯ

ЛИТЕРАТУРА

ПРИЛОЖЕНИЯ

Приложение I

МАШИННЫЕ РЕЗУЛЬТАТЫ АИА АЗЕРБАЙДЖАНСКОЙ

СЛОВОЗОМЫ

IEN0461 IHEVSEB
***GO DOES NOT EXIST BUT HAS BEEN ADDED TO DATA SET
DIAGNOSTIC MESSAGE DI

IEN0461 WARNING - SYMBOL PRINTED IS AN UNRESOLVED EXTERNAL REFERENCE
MARKED FOR RESTRICTED NO-CALL OR NEVERCALL.

IEF3731 STEP /LKED / START 88192.1734
IEF3741 STEP /LKED / STOP 88192.1737 CPU 8MIN 09.06SEC MAIN
//GO,SYSSIN DO OSN=PUBLIC,SOURCE(PIOTTX),DISP=SHR

САТЫЛДЫГАЗ
ОСНОВА:
САТ ГЛАГОЛ
АФИКСЫ:
ИЛ СТРАДАТ.ЗАЛОГ
ДЫГАЗ ДЕЕПРИЧАСТИЕ

ОХУНАДАН
ОСНОВА:
ОХУ ГЛАГОЛ
АФИКСЫ:
НАДАН ДЕЕПРИЧАСТИЕ

ЧАЛАЗАҒЫНЫЗ
ОСНОВА:
ЧАЛ СУЩЕСТВ./ГЛАГОЛ
АФИКСЫ:
АЗАҒЫНЫЗ ПРИЧАСТИЕ 2-Е ЛИЦО

ИЕДДИНЗИ
ОСНОВА:
ИЕДДИ ЧИСЛИТЕЛЬНОЕ
АФИКСЫ:
НЗИ ПОРЯДКОВ.ЧИСЛИТ.

БЕШИНЗИ
ОСНОВА:
БЕШ ЧИСЛИТЕЛЬНОЕ
АФИКСЫ:
ИНЗИ ПОРЯДКОВ.ЧИСЛИТ.

ВУРУШУР
ОСНОВА:
ВУР ГЛАГОЛ
АФИКСЫ:
ИЛ ВЗАЙМО-СОВМСТ.ЗАЛОГ

ПЕШ
ОСНОВА:
ПЕШ ИМЯ СУЩЕСТВИТА.
АФИКСЫ:
АФ. В СЛОВЕ НЕТ
ИМЯ СУЩЕСТВИТА.
МЭКТЭБЛЯРИН
ОСНОВА:
АФИКСЫ:
МЭКТЭБ ИМЯ СУЩЕСТВИТА.
ЛЭР ИМОГ.ЧИСЛО
ИНЗ ПРИНАДЛЕЖНОСТЬ 3-Е ЛИЦО

ИМИ
ОСНОВА:
ИМИ ЧИСЛИТЕЛЬНОЕ
АФИКСЫ:
АФ. В СЛОВЕ НЕТ
ИМЯ СУЩЕСТВИТА.
ИИЗ
ОСНОВА:
ИИЗ ЧИСЛИТЕЛЬНОЕ
АФИКСЫ:
АФ. В СЛОВЕ НЕТ
ИМЯ СУЩЕСТВИТА.
МИН
ОСНОВА:
МИН ЧИСЛИЛ./ГЛАГОЛ
АФИКСЫ:
АФ. В СЛОВЕ НЕТ
ИМЯ СУЩЕСТВИТА.
НЭӨЗР
ОСНОВА:
НЭӨЗР ИМЯ СУЩЕСТВИТА.
АФИКСЫ:
АФ. В СЛОВЕ НЕТ
ИМЯ СУЩЕСТВИТА.

VSHZ
ОСНОВА:
VSHZ ИМЯ ПРИЛАГАТЕЛ.
АФИКСЫ:
АФ. В СЛОВЕ НЕТ
ИМЯ ПРИЛАГАТЕЛ.

ГӘБҮЛ
ОСНОВА:
ГӘБҮЛ ИМЯ СУЩЕСТВИТА.
АФИКСЫ:
АФ. В СЛОВЕ НЕТ
ИМЯ СУЩЕСТВИТА.
ЕДИЛНИШ
ОСНОВА:
ЕД ГЛАГОЛ
АФИКСЫ:
ИЛ СТРАДАТ.ЗАЛОГ

АФИКСЫ:
***. НЕ НАДЕН В ТАБЛ.

ТИХ
ОСНОВА:
АФИКСЫ:
***. НЕ НАДЕН В ТАБЛ.

АИЗ
ОСНОВА:
АИЗ ИМЯ СУЩЕСТВИЛ.

АФИКСЫ:
***. В СЛОВЕ НЕТ

ДЛУЖДУР
ОСНОВА:
ЭЛ ГЛАГОЛ
ЧУЧ БУДУЧ ЧЕКАТ.ВРЕМЯ
ДУР ПОНУДИТ, ЗАЛОГ/СКАЗУЕ 3-Е ЛИЦО ЕД.ЧИСЛО

АГЫРДЫГАНЫЗЛАРДАНСЫНЫЗМИ-

ОСНОВА:
ЧАГЫР ГЛАГОЛ
ДЫГ ПРЧАСТИЕ ПРОШЕД.КАТЕГ.ВРЕМЯ
ЧМЫЗ ПРИНАДЛЕЖНОСТЬ 1-Е ЛИЦО МНОЖ.ЧИ
ЛАР МНОЖ.ЧИСЛО
ДАН ИСХОДНЯЯ ПАД
СЫНЫЗ 2-Е ЛИЦО МНОЖ.ЧИСЛО СКАЗУЕ
МЫ ВОПРОСИТЕЛЬН.

ДАНЫШМАГАЛАМ
ОСНОВА:
АФИКСЫ:

ДАНЫШ ГЛАГОЛ
НАГДА ПРОДОЛЖ.НАКЛ.
ДА ЖЕЛАТЕЛЬН.НАКЛ./ДАТ.НАПР.ПА
Н ПРИНАДЛЕЖНОСТЬ 1-Е ЛИЦО ЕД.ЧИСЛО

АЛАРКИН
ОСНОВА:
АФИКСЫ:

АЛ ГЛАГОЛ/ПРИЛАГТ.
АРКИН ДЕЕПРИЧАСТИЕ

КОНЕЦ ФАЙЛД

МЭКТЭБЛЭРНИН
ОСНОВА:
МЭКТЭЕ ИМЯ СУЩЕСТВИЛ.
АФИКСЫ:
ЛЭР МНОЖ.ЧИСЛО
ДИНН ПРИНАДЛЕЖНОСТЬ 2-Е ЛИЦО ЕД.ЧИСЛО

ХИСУСИ
ОСНОВА:
ХИСУСИ ИМЯ ПРИЛАГАТЕЛ.

АФИКСЫ:
АФ. В СЛОВЕ НЕТ

ЧЭКИСИ
ОСНОВА:
ЧЭКИ ИМЯ СУЩЕСТВИЛ.

АФИКСЫ:
СИ ПРИНАДЛЕЖНОСТЬ 3-Е ЛИЦО ЕД.ЧИСЛО

ОН
ОСНОВА:
ОН ЧИСЛИТЕЛЬНОЕ

АФИКСЫ:
АФ. В СЛОВЕ НЕТ

ДОРД
ОСНОВА:
ДОРД ЧИСЛИТЕЛЬНОЕ

АФИКСЫ:
АФ. В СЛОВЕ НЕТ

БЕШ
ОСНОВА:
БЕШ ЧИСЛИТЕЛЬНОЕ

АФИКСЫ:
АФ. В СЛОВЕ НЕТ

ФАИЗ
ОСНОВА:
ФАИЗ ИМЯ СУЩЕСТВИЛ.

АФИКСЫ:
АФ. В СЛОВЕ НЕТ

ТЭҮКИЛ
ОСНОВА:
ТЭҮКИЛ ИМЯ СУЩЕСТВИЛ.

АФИКСЫ:
АФ. В СЛОВЕ НЕТ

ЕДИРДИСЭ
ОСНОВА:
ЕД ГЛАГОЛ
ИР НАСТОЯЩЕЕ ВРЕМЯ
ДИ ПРОШЕД.КАТЕГ.ВРЕМЯ

ПЕРЕД

ОСНОВА:

ПЕШ ИМЯ СУЩЕСТВИЛ.

АФФИКСЫ:
АФФ. В СЛОВЕ НЕТ

ИЗКТЭБЛЭРИНЗ
ОСНОВА:

ИЗКТЭБ ИМЯ СУЩЕСТВИЛ.

АФФИКСЫ:
ЛЭР ИМЕР.ЧИСЛО

ИМЕ ПРИНАДЛЕЖНОСТЬ 3-Е ЛИЦО

ИКИ
ОСНОВА:
ИКИ ЧИСЛИТЕЛЬНОЕ

АФФИКСЫ:
АФФ. В СЛОВЕ НЕТ

ИУЗ
ОСНОВА:
ИУЗ ЧИСЛИТЕЛЬНОЕ

АФФИКСЫ:
АФФ. В СЛОВЕ НЕТ

МИН
ОСНОВА:
МИН ЧИСЛИЛ./ГЛАГОЛ

АФФИКСЫ:
АФФ. В СЛОВЕ НЕТ

ИЗЭЗР
ОСНОВА:
ИЗЭЗР ИМЯ СУЩЕСТВИЛ.

АФФИКСЫ:
АФФ. В СЛОВЕ НЕТ

ВЭНZ
ОСНОВА:
ВЭНZ ИМЯ ПРИЛАГАТЕЛ.

АФФИКСЫ:
АФФ. В СЛОВЕ НЕТ

ГЭБҮЛ
ОСНОВА:
ГЭБҮЛ ИМЯ СУЩЕСТВИЛ.

АФФИКСЫ:
АФФ. В СЛОВЕ НЕТ

ЕДИЛНИШ
ОСНОВА:
ЕД ГЛАГОЛ

АФФИКСЫ:
ИЛ СТРАДАТ, ЗА ТОГ
ИУ БУЗУЛ НЕКАТ.ЭРДИ

УНУДАР
ОСНОВА:

УНУД ГЛАГОЛ
АР БУДУЧ.КАТЕГ.ВРЕМЯ

ИЧЛЭННИЙДИ
ОСНОВА:
АФФИКСЫ:

ИЧЛ СУЩЕСТ./ГЛАГОЛ
МИН БУДУЧ НЕКАТ.ВРЕМЯ
ДИ ПРОШЕД.КАТЕГ.ВРЕМЯ

СЭСИНДЭН
ОСНОВА:
АФФИКСЫ:
СЭС ИМЯ СУЩЕСТВИЛ.
ИМ ПРИНАДЛ.1 Л.ЕД./ПОВЕЛ.НАКЛ.
ДЭН ИСХОДНЫЙ ПАД

НАЯННЫ
ОСНОВЫ НЕТ В СПИСК

САМЫЯ
ОСНОВА:
АФФИКСЫ:
САМЫ МЕСТОИМЕНИЕ
ЯА ЖЕЛАТЕЛЬН.НАКЛ./ДАТ.НАПР.ПА

ОРАДА
ОСНОВА:
АФФИКСЫ:
ОРА МЕСТОИМЕНИЕ
ДА МЕСТНЫЙ ПАД.

БЕШИ
ОСНОВА:
АФФИКСЫ:
БЕШ ЧИСЛИТЕЛЬНОЕ
И ПРИНАДЛ./ВИНН.ПАД.

ТЭКЗЭ
ОСНОВА:
АФФИКСЫ:
АФФ. В СЛОВЕ НЕТ

ДОГГУЗУНЗУ
ОСНОВА:
АФФИКСЫ:
ДОГГУЗ ЧИСЛИТЕЛЬНОЕ

БЕШ
ОСНОВА:

БЕШ ЧИСЛИТЕЛЬНОЕ

АФФИКСЫ:
АФФ. В СЛОВЕ НЕТ

МУНАДЭМ
ОСНОВА:

НИН ЧИСЛИТЛ./ГЛАГОЛ

АФФИКСЫ:

ДЭМ ИСХОДНАЯ ПАД

ЧОХ
ОСНОВА:

ЧОХ ЧИСЛИТЕЛЬНОЕ

АФФИКСЫ:
АФФ. В СЛОВЕ НЕТ

ИХТИСАСЛЫ
ОСНОВА:

ИХТИСАСЛЫ ИМЯ ПРИЛАГАТЕЛ.

АФФИКСЫ:
АФФ. В СЛОВЕ НЕТ

ИЧЧИ
ОСНОВА:

ИЧЧИ ИМЯ СУЩЕСТВИТА.

АФФИКСЫ:
АФФ. В СЛОВЕ НЕТ

ЧЕНГЭРЭРИЛНИЙДИР
ОСНОВА:

ЧЕНГЭР ГЛАГОЛ

АФФИКСЫ:

ИЛ СТРАВАТ.ЗАЛОГ

НИН БУДУН ЧЕКАТ.ВРЕМЯ

ДИР ПОНУДИУ.ЗАЛОГ/СКАЗУЕЗ-Ё ЛИЧО

ОМУНГУ
ОСНОВА:

ОН ЧИСЛИТЕЛЬНОЕ

АФФИКСЫ:

УНГУ ПОРЯДКОВ.ЧИСЛИТ.

СОНВРВАЛИЙК
ОСНОВА:

СОНВРВАЛИЙК ИМЯ СУЩЕСТВИТА.

АФФИКСЫ:
АФФ. В СЛОВЕ НЕТ

ВО
ОСНОВА:

ВВ 2003

АФФИКСЫ:

МАШИННЫЕ РЕЗУЛЬТАТЫ ЛЕКСИКО-МОРФОЛОГИЧЕСКОГО
МП АЗЕРБАЙДЖАНСКОГО ТЕКСТА НА РУССКИЙ ЯЗЫК

Приложение 2

ИАТИСАСАИ

ОСНОВА: ИМЕЮЩИЙ СПЕЦИАЛЬНОСТЬ
АФФИКСЫ:

АФФ. В СЛОВЕ НЕТ

ИШЧИ

ОСНОВА:

ИШЧИ РАБОТНИК

АФФИКСЫ:
АФФ. В СЛОВЕ НЕТ

УНДАЭРИЛЛИДИР

ОСНОВА:

УНДАЭР ПОСЛАТЬ, ОТПРАВЛЯТЬ

АФФИКСЫ:
ИЛ СТРАДАТЬ, ЗАЛОГ

НИЧИ ДАВНОПРОШЕДШЕЕ ПЕРЕСКАЗАТЕЛ,

АФФ. НЕ НАЙДЕН В ТАБЛ.

ОМУНХҮ

ОСНОВА:

ОН ДЕСЯТЬ

АФФИКСЫ:

ХҮНХҮ ПОРЯДКОВ. ЧИСЛИТ.

БЕШИДЛИК

ОСНОВА:

БЕШ ПЯТЬ

АФФИКСЫ:
ИЛ СТРАДАТЬ, ЗАЛОГ

АФФ. НЕ НАЙДЕН В ТАБЛ.

СЭНЭРЭЛИДИК

ОСНОВА:

СЭНЭРЭЛИДИК РАЦИОНАЛЬНОСТЬ, ПОЛЕЗНОСТЬ

АФФИКСЫ:
АФФ. В СЛОВЕ НЕТ

ЗЭ

ОСНОВА:

ЗЭ ИДА

АФФИКСЫ:
АФФ. В СЛОВЕ НЕТ

КЕЛВИИД

ОСНОВА:

КЕЛВИИД КЧЕСТВО, ПОЛОЖЕНИЕ, ОБСТОЯТЕЛЬСТВО

ОСНОВА:
ВЕР ДАВАТЬ, ОТДАВАТЬ

АФФИКСЫ:

АФФ. В СЛОВЕ НЕТ

СЭС ВЕР ГОЛОСОВАТЬ

>>>>>>>>>>>>>>>>>

ЗЭАБ ВЕР

ВЕР

ОСНОВА:
ВЕР ДАВАТЬ, ОТДАВАТЬ

АФФИКСЫ:

АФФ. В СЛОВЕ НЕТ

ЗЭАБ ВЕР МУЧИТЬ, УГНЕТАТЬ

>>>>>>>>>>>>>>>

ЗЭАБ ЧЭК

ЧЭК

ОСНОВА:
ЧЭК ТЯНУТЬ, ТАСКАТЬ

АФФИКСЫ:

АФФ. В СЛОВЕ НЕТ

ЗЭАБ ЧЭК ТОМИТЬСЯ, СТРАДАТЬ

>>>>>>>>>>>>>>>

ТАННЫ ОЛ

ОЛ

ОСНОВА:
ОЛ БЫТЬ, СТАТЬ

АФФИКСЫ:

АФФ. В СЛОВЕ НЕТ

ТАННЫ ОЛ ПОЗНАКОМИТЬСЯ

>>>>>>>>>>>>>>>

ТАННЫ ЕДИН

ЕДИН

ОСНОВА:

ЕТ ДЕЛАТЬ СДЕЛАТЬ

АФФИКСЫ:

АФФ. В СЛОВЕ НЕТ

ДИ ПРОШЕД. КАТЕГ. ВРЕМЯ

И ПРИНАДЛЕЖНОСТЬ 1-Е ЛИЦО ЕД. ЧИСЛО

ТАННЫ ЕТ ПОЗНАКОМИТЬ, ОЗНАКОМИТЬ ПРОШЕД. КАТЕГ. ВРЕМЯ ПРИНАД

>>>>>>>>>>>>>>>

ИБ 2253

Формат 60x84¹/16. Бумага типографская № 1.
Печать офсетная. Усл. печ. л. 8,37.
Уч.-изд.л. 5,9. Заказ 421 Тираж 400.
Договорная цена 2 руб.

Издательство "Элм".
370143. Баку, пр. М.Азизбекова, 31.
Академгородок, Главное здание.
Типография АН Азербайджанской Республики.
Баку, пр. Азизбекова, 31.