

ҚАЗАҚСТАН РЕСПУБЛИКАСЫ
БІЛІМ ЖӘНЕ ҒЫЛЫМ МИНИСТРЛІГІ
Л.Н. ГУМИЛЕВ АТЫНДАҒЫ ЕУАЗИЯ
ҰЛТТЫҚ УНИВЕРСИТЕТІ

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ
РЕСПУБЛИКИ КАЗАХСТАН
ЕВРАЗИЙСКИЙ НАЦИОНАЛЬНЫЙ УНИВЕРСИТЕТ
ИМЕНИ Л.Н. ГУМИЛЕВА

MINISTRY OF EDUCATION AND SCIENCE
OF THE REPUBLIC OF KAZAKHSTAN
L.N. GUMILYOV EURASIAN NATIONAL UNIVERSITY



16-18 маусым
Нұр-Сұлтан, 2022

«TURKLANG 2022»

«Түркі тілдерін компьютерлік өңдеу»
атты X халықаралық конференция
ЕҢБЕКТЕРІ

ТРУДЫ

X Международной конференции
«Компьютерная обработка тюркских языков»

«TURKLANG 2022»

PROCEEDINGS

of the X International Conference
on Computer processing of Turkic Languages

«TURKLANG 2022»

УДК 80/81:004
ББК 81.2:32-973
Т 90

Техникалық редакция:

Ергеш Б.Ж.
Елибаева Г.К.
Турсынова Н.А.

Т 90 ТҮРКІ ТІЛДЕРІН КОМПЬЮТЕРЛІК ӨНДЕУ. X халықаралық конференция: Еңбектері = КОМПЬЮТЕРНАЯ ОБРАБОТКА ТЮРКСКИХ ЯЗЫКОВ. X международная конференция: Труды. / - Нұр-Сұлтан: «Булатов А.Ж.» ЖК, 2022.= Нур-Султан: ИП «Булатов А.Ж.»

ISBN 978-601-326-645-9

Жинақта «Түркі тілдерін компьютерлік өңдеу» атты X халықаралық конференция қатысушыларының баяндамалары енген.

Компьютерлік лингвистика бағыты бойынша оқитын студенттерге, магистранттарға, докторанттарға және мамандарға арналған.

Жинақ «BR11765535» Қазақ тілі мәдениетін арттыру және функцияларды кеңейту бойынша ғылыми-лингвистикалық негіздер мен IT-ресурстарды әзірлеу» бағдарламасы есебінен жарияланды.

В сборнике представлены доклады участников X международной конференции «Компьютерная обработка тюркских языков».

Предназначен для студентов, магистрантов, докторантов и специалистов специализирующихся в областях компьютерной лингвистика.

Сборник издан за счет средств программы BR11765535 «Разработка научно-лингвистических основ и IT-ресурсов по расширению функций и повышению культуры казахского языка».

УДК 80/81:004
ББК 81.2:32-973

ISBN 978-601-326-645-9

© Л.Н.Гумилев атындағы Еуразия ұлттық университеті, 2022

© Евразийский национальный университет им. Л.Н. Гумилева, 2022

**МӘТІНДЕРДІ МОРФОЛОГИЯЛЫҚ, СИНТАКСИСТІК ЖӘНЕ
СЕМАНТИКАЛЫҚ ӨНДЕУ ТЕХНОЛОГИЯЛАРЫ**

**ТЕХНОЛОГИИ МОРФОЛОГИЧЕСКОЙ, СИНТАКСИЧЕСКОЙ И
СЕМАНТИЧЕСКОЙ ОБРАБОТКИ ТЕКСТОВ**

**TECHNOLOGIES FOR MORPHOLOGICAL, SYNTACTIC AND
SEMANTIC TEXT PROCESSING**

1	<p><i>Сайранбекова А. Д., Бекманова Г.Т.</i> <i>Л.Н. Гумилев атындағы Еуразия ұлттық университеті, Нұр-Сұлтан, Қазақстан</i></p> <p>МАШИНАЛЫҚ АУДАРМА НЕГІЗІНДЕ МӘТІНДЕГІ ЖАҒЫМСЫЗ СЕНТИМЕНТТІ АНЫҚТАУ</p>	125
2	<p><i>Кудубаева С.А., Жусупова Б.Т.</i> <i>Евразийский национальный университет им. Л.Н. Гумилева, Нур-Султан, Казахстан</i> <i>Костанайский региональный университет имени А. Байтурсынова, Костанай, Казахстан</i></p> <p>О ВОЗМОЖНОСТИ УЧЕТА СЕМАНТИЧЕСКОЙ СОСТАВЛЯЮЩЕЙ В СИСТЕМЕ СУРДОПЕРЕВОДА С КАЗАХСКОГО ЯЗЫКА НА КАЗАХСКИЙ ЯЗЫК ЖЕСТОВ</p>	135
3	<p><i>Азат Абдысадыр уулу</i> <i>Канцелярия Администрации Президента Кыргызской Республики, Бишкек, Кыргызстан</i></p> <p>МЕТОД БИНАРНЫХ СВЯЗЕЙ ДЛЯ ВИЗУАЛИЗАЦИИ СМЫСЛА ПРЕДЛОЖЕНИЯ</p>	145
4	<p><i>Khamroeva Shahlo Mirdjanovna</i> <i>Tashkent State university of Uzbek language and literature named Alisher Navai, Tashkent, Uzbekistan</i></p> <p>FINITE STATE MACHINE MODEL OF NOUNS FOR UZBEK LANGUAGE MORPHOLOGICAL ANALYZER</p>	153
5	<p><i>Nasrullayeva A., Mukanova A.</i> <i>Astana International University, Nur-Sultan, Kazakhstan</i></p> <p>RESEARCH AND ANALYSIS OF MECHANISMS FOR DETECTING PROHIBITED CONTENT ON THE INTERNET</p>	165
6	<p><i>Оралбекова І.Т., Ергеш Б.Ж.</i> <i>Л.Н. Гумилев атындағы Еуразия ұлттық университеті, Нұр-Сұлтан, Қазақстан</i></p> <p>ҚАЗАҚ ТІЛІНДЕГІ ҚОНАҚҮЙЛЕР ТУРАЛЫ ШКІРЛЕРДІҢ РЕҢКІН АСПЕКТИЛЕРГЕ ТАЛДАУ</p>	172

UDC 83.00

Khamroeva Shahlo Mirdjanovna

Tashkent State university of Uzbek language and literature

named Alisher Navai

Tashkent, Uzbekistan

hamroyeva81@mail.ru

FINITE STATE MACHINE MODEL OF NOUNS FOR UZBEK LANGUAGE MORPHOLOGICAL ANALYZER

Abstract. This article discusses the Finite State Transducer (FST) models of creating a morphological analyzer of the Uzbek language. The FST method of morphological analysis, the model of the analysis of nouns in the Uzbek language are discussed.

Keywords: FST, model, Uzbek language morphological analyzer, noun, wordform, affixe

УДК 83.00

Хамроева Шахло Мирджановна

Ташкентский государственный университет узбекского языка и

литературы имени Алишера Навои

hamroyeva81@mail.ru

КОНЕЧНАЯ АВТОМАТНАЯ МОДЕЛЬ СУЩЕСТВИТЕЛЬНЫХ ДЛЯ МОРФОЛОГИЧЕСКОГО АНАЛИЗАТОРА УЗБЕКСКОГО ЯЗЫКА

Аннотация. В данной статье рассматриваются модели Finite State Transducer (FST) для создания морфологического анализатора узбекского языка. Обсуждаются метод FST морфологического анализа, модель анализа существительных в узбекском языке.

Ключевые слова: FST, модель, морфологический анализатор узбекского языка, существительное, словоформа, аффикс.

For Turkic languages, which are included in agglutinative languages, the automatic approach of creating a morphological analyzer is more appropriate. The automated approach-based morphoanalyzer has a processing system FST (finite state transducer) and WFST (weighted finite state transducer), which can perform input and output (analysis / synthesis) analysis. The essence of such analyzers is that they follow the rule of "grammatical sequence": the

process is based on the rule of morphological unity of the word. They differ in what units are used:

- 1) show the sequence of morphemes and the required allomorphic sequence;
- 2) rules of sequence of allomorphs.

In constructing a morphological analyzer of natural language, many scientists say [1, 2, 3, 4, 5, 6] that it is an advantage to use the Finite State Machine method. In this article, we discuss the development of a model of the Uzbek language FST (finite state transducer) in the creation of a morphological analyzer of the Uzbek language.

There are key factors for automatic morphological analysis, such as stem, base, prefix, suffix, spelling rules.

To do this, you need to create a database of word-formers in Uzbek (pre-/post-stem), lexical and syntactic suffixes, particles in the form of suffixes.

If all of this forms the database of the morphological analyzer, it is necessary to run the morphological analyzer and develop an analysis model to develop the program.

Hence, for the morphological analysis of the word “*daraxtlar*” at the initial stage, the following information is required to be in the database:

- 1) Base (information indicating to which category the basis belongs);
- 2) Suffix (information representing the type).

Based on this information, the word form of “*daraxtlar*” are analyzed as follows: daraxt - ; -s is the plural form.

When Ashref Adali wrote about the creation of a morphological analyzer of Turkish and English, he distinguished the following database for the analyzer, which can identify the series of stems, additional and grammatical interpretation of the word²:

- 1) dictionary;
- 2) a series of suffixes;
- 3) spelling rules.

To do this, it is determined what suffixes a particular group of words can take. In Table 1 we present the suffixes that can be associated with a particular word and their position.

1.1. Finite state machine (FSM) models of the Uzbek language for morphological analysis for the Uzbek language.

It is designed to show boundary and transition states in processes involving the Restricted Case Machine (FSM). SDM and its special forms are widely used in the construction of various linguistic analyzers. To understand the Limited Case Machine (FSM), we need to know its basic terms.

1. Initial state: Indicates the initial state of the limited state machine. It is indicated by an arrow coming from an unknown place.

2. Acceptance status: A status indicating that the **FSM** has successfully completed its mission.

3. Recipient and Recognizers: Indicates whether the application has been accepted or not.

4. Transition motion: the transition from one state to another.

5. Converter: forms the output state using the input and motion used. There are two types of converters: (a) **FSM** uses only input motions; the output depends on the situation; (b) (**FSM**) uses only input actions: output depends on input and status.

1.2. Spelling rules. Spelling rules also play an important role in creating a morphological analyzer. Although Uzbek is an agglutinative language among Turkic languages, there are some cases of inflection. Flexion cases occur as sound changes. For the analysis of word forms on the basis of the morphological analyzer should reflect the typical cases of sound changes, all sound changes in the Uzbek language. There are three main types of sound changes in the Uzbek language:

1. Sound changes:

The vowel at the end of a word changes with the addition of an suffix:

1) When the suffix -v, -q, -qi is added to verbs ending in a vowel, the vowel **a** is pronounced as **o** and written as: *sayla – saylov, sina – sinov, aya – ayovsiz; so‘ra – so‘roq, bo‘ya – bo‘yoq; o‘yna – o‘ynoqi, saura – sayroqi;*

2) When most verbs ending in the vowel **i** are followed by the suffix -v, -q, the vowel is pronounced and written as **u**: *o‘qi – o‘quvchi, qazi – qazuvchi, sovi – sovuq.* However, in some verbs ending in the vowel **i**, when the suffix -q is added, the vowel **i** is pronounced and written as **i**: *og‘ri – og‘riq, qavi – qaviq.*

When the possessive suffix is added to multi-syllable words ending in **k**, **q**, as well as to certain syllables such as **bek**, **yo‘q**, the consonant **k** becomes the consonant **g**, the consonant **q** becomes the consonant **g’**, and is written as follows: *tilak – tilaging, yurak – yuragim, kubok – kubogi, bek – begi; tayoq – tayog‘i, qoshiq – qoshig‘i, yaxshiroq – yaxshirog‘i, yo‘q – yo‘g‘i.* But in multi-syllable mastery words, when the possessive suffix is added to single-syllable plural words, the sound **k**, **q** is actually pronounced and written: *ishtirok – ishtiroki, ocherk – ocherki, erk – erki, huquq – huquqim, ravnaq – ravnaqi, yuq – yuqi.*

2. Sound drop:

With the addition of the following suffixes, the sound in the word structure drops:

1) When the possessive suffix is added to some words, such as o‘rin, qorin, burun, o‘g‘il, bo‘uin, ko‘ngil, and the suffix **-il**, which forms the relative form to verbs such as qayir, ayir, is added to ikki, olti, yetti words - **ov**, The vowel in the second syllable is not pronounced or written when the suffix **-ala** is added: *o‘rin – o‘rnim, qorin – qorni, burun – burning, o‘g‘il – o‘g‘ling, ko‘ngil – ko‘ngli, yarim – yarmi; qayir – qayril, ulug‘ – ulg‘ay, sariq – sarg‘ay, ikki – ikkov, ikki – ikkala, yetti – yettov;*

3. Sound increase:

With the addition of the following suffixes, the sound content of the word increases:

2) It is pronounced with the sound **n** when the suffixes **-da, -dan, -day, -dagi, -ga, -gacha, -cha** are added to those pronouns **u, bu, shu, o‘sha**, and it is written as follows: unda, bunday, shunda, o‘shancha; the possessive suffixes to these pronouns are added as follows; buningiz, o‘shanisi;

3) Possessive suffixes are added to words ending in **o, o, u, e** vowels as follows:

a) A lot of possessive suffixes **-m, -ng, -si; -miz, -ngiz, -si** (or **-lari**) are added without sound: bobom, bobong, bobosi, bobomiz, bobongiz, bobosi (yoki bobolari); orzum, orzung, orzusi; orzumiz, orzungiz, orzusi;

b) When the first and second person possessive suffixes are added to the words **parvo, obro, mavqe, mavzu, avzo**, a vowel is added and written as follows: parvoyim, parvoying; parvoyimiz, parvoyingiz; obro‘yim, obro‘ying; obro‘yimiz, obro‘yingiz; The possessive suffix of the third person is added to the words **parvo, avzo, obro’**, **mavqe** in the form **-yi**, and to the words xudo, mavzu in the form of **-si**: as avzoyi, mavzusi (as dohiy, **-si** is added to the word ending a “-y” consonant in the third person: like a dohiysi);

4) When the suffixes **-ni, -ning, -niki** are added to the pronouns men, sen, the sound **n** in the suffix is not pronounced or written: as meni, mening, meniki; seni, sening, seniki.

1.3. Database of morphemes. Another component required to run a morphological analyzer is a database of morphemes. Additions in the Uzbek language are divided into three groups in terms of function:

1. Word formers.
2. Syntactic formers.
3. Vocabulary formers.

In order to distinguish word-forming suffixes in the process of morphological analysis, a complete list of them should be provided. Below is a list of word-formative suffixes that are actively used in the Uzbek language.

Table 2.

Noun form suffixes

Noun - noun form suffixes (ad yapım ekleri)				
	ek	teg	açıklama	yasalma (gövde) örneğ
1	-bin:	Ŷ 1	noun - noun	Folbin
2	-bon:	Ŷ 1	noun - noun	darvozabon, soyabon, tarozibon, xazinabon
3	-boz:	Ŷ 3	noun - noun	masxaraboz, qimorboz, dorboz
4	-voy:	Ŷ 4	noun - noun	novvoy (nonvoy)
5	-gar// -kar:	Ŷ 5	noun - noun	zargar, savdogar, da'vogar, miskar
6	-garchilik:	Ŷ 6	noun - noun	yog'ingarchilik, odamgarchilik
7	-gin	Ŷ 7	noun - noun	jahongir, fazogir
8	-goh:	Ŷ 8	noun - noun	oromgoh, saylgoh, sayrgoh, qarorgoh, ziyoratgoh, bazmgoh
9	-go'y:	Ŷ 9	noun - noun	kalimago'y, maslahatgo'y
10	-diq// -dik	Ŷ 10	noun - noun	o'rindiq
11	-don:	Ŷ 11	noun - noun	guldon, kuldon, qalamdon
12	-don	Ŷ 12	noun - noun	muhrdor, chorvador
13	-dosh:	Ŷ 13	noun - noun	sinfdosh, kursdosh, maslakdosh
14	-do'z:	Ŷ 14	noun - noun	etikdo'z, mahsido'z, kashtado'z
15	-zor:	Ŷ 15	noun - noun	olmazor, gulzor, olchazor
16	-iston:	Ŷ 16	noun - noun	guliston, go'riston, O'zbekiston
17	-kash:	Ŷ 17	noun - noun	aravakash, qalamkash, suratkash
18	-kor:	Ŷ 18	noun - noun	ganchkor, paxtakor, sholikor, san'atkor
19	-kov:	Ŷ 19	noun - noun	go'rkov
20	-lik/liq:	Ŷ 20	noun - noun	bolalik, vaqtichog'lik, do'stlik, boshliq
21	-loq:	Ŷ 22	noun - noun	O'tloq, qumloq, toshloq
22	-noma:	Ŷ 23	noun - noun	taklifnoma, tabriknoma, pandnoma
23	-navis:	Ŷ 24	noun - noun	tarixnavis, voqeanavis, romannavis
24	-paz	Ŷ 25	noun - noun	oshpaz, kabobpaz, somapaz
25	-soz	Ŷ 26	noun - noun	soatsoz, kemasoz
26	-fuwsh:	Ŷ 27	noun - noun	baliqfurush, nosfurush
27	-xon:	Ŷ 28	noun - noun	kitobxon, she'rxon
28	-xona:	Ŷ 29	noun - noun	darsxona, mehmonxona, ishxona,

29	-xo'r	Ŷ 30	noun - noun	merosxo'r
30	-cha	Ŷ 31	noun - noun	qalamcha
31	-chak//-choq:	Ŷ 32	noun - noun	o'yinchoq
32	-chi:	Ŷ 33	noun - noun	ishchi, temirchi, terimchi, gulchi, bosqinchi
33	-chilik	Ŷ 34	noun - noun	hunarmandchilik, o'zbekchilik, dehqonchilik
34	Ham-	Ŷ 35	noun - noun	hamqishloq, hamshahar, hamyurt
Noun formers from the base of another word category				
35	-a (1):	Ŷ 36	imitation	qahqah, sharshar, g'arg'ar, jizz,
36		Ŷ 37	adjective - noun	bo'z, quyuq
37	-a	Ŷ 38	adjective - noun	xarob, vayron
38	-ak:	Ŷ 39	imitation	bizbiz, pirpir, guldir, var, qar, xur
39	-archilik:	Ŷ 40	adjective - noun	Och
40	-at:	Ŷ 41	verb - noun	ko'chat, o'lat
41		Ŷ 42	adjective - noun	ko'kat
42	-vchi//-uvchi:	Ŷ 43	verb - noun	o'quvchi, yozuvchi, uchuvchi, aniqlovchi, to'kjuruvchi
43	-garchilik	Ŷ 44	adjective - noun	namgarchilik, xafagarchiik, xunobgarchilik, sharmandagarchilik
44	-gi//-ki// - qi//-g'i//-g'u:	Ŷ 45	verb - noun	sezgi, sevgi, supurgi, kukju, turtki sanchqi, tomizg'i, tuyg'u.
45	-gach//-kich// -qich//-g'ich:	Ŷ 46	verb - noun	kulgich, o'tkazgich, ko'rsatkich, yoritqich, tutqich, ochqich, to'g'nag'ich, chizg'ich, o'chirg'ich
46	-gin//-qin// -kin//-gun// -qun:	Ŷ 47	verb - noun	tizgin, surgun, tolqin, to'sqin, quvg'in, yong'in, uchqun
47	-dak//-doq:	Ŷ 48	verb - noun	yugurdak, kekirdak, qovurdoq, qo'ndoq
48	-diq	Ŷ 49	verb - noun	qoldiq, topildiq, hordiq
49	-ik:	Ŷ 50 (1)	verb - noun	ko'rik, teshik (noun - noun va adjective - noun), kekirik

50	-ik:	Ŷ 50(2)	verb - noun	bilik, bitik
51	-ik:	Ŷ 50(3)	adverb - noun	ko'pik
52	-ildoq:	Ŷ 51	imitation	hiqildoq, chirildoq
53	-imlik:	Ŷ 52	verb - noun	o'simlik, ichimlik
54	-in//-un:	Ŷ 53	verb - noun	yig'in, yog'in, ekin, tiqin, tugun, tutun
55	-indi// undi//ndi:	Ŷ 54	verb - noun	chiqindi, yuvundi, chirindi, cho'kindi, kuyundi, yig'indi
56	-it:	Ŷ 55	verb - noun	chiqit
57	-ich:	Ŷ 56	verb - noun	cho'mich, cho'kich, o'pich, bog'ich
58	-ish:	Ŷ 57	verb - noun	Qarg'ish
59	-iq//-uq:	Ŷ 58	verb - noun	chaqiriq, kesatiq, yutqiziq, chopiq, yutuq
60	-k:	Ŷ 59	verb - noun	ko'rik, elak, tilak, kurak, bezak, to'shak
61	-kilik//gilik:	Ŷ 60	verb - noun	ichkilik, ko'rgilik
62	-lik	Ŷ 61(1)	numerativ-noun	birlik, to'rtlik
63		Ŷ 61(2)	Pronoun - noun	o'zlik
64		Ŷ 61(3)	adverb - noun	tezlik, sekinlik, birgalik
65		Ŷ 61(4)	Modal - noun	borliq, yo'qlik
66	-m//-im//-um:	Ŷ 62(1)	verb - noun	to'plam, ho'plam, chidam, tishlam, kechirim, qo'nim, terim, chiqim, bitim, bosim, unum, tuzum
67		Ŷ 62(2)	imitation	qultum
68	-ma:	Ŷ 63	verb - noun	surma, o'sma, tortma
69	-mak//-moq:	Ŷ 64	verb - noun	yemak, ilmoq, chaqmoq, topishmoq
70	-mashoq:	Ŷ 65	verb - noun	bekinmashoq, quvlashmashoq
71	-mish:	Ŷ 66	verb - noun	noun - nounmish, kechmish, qilmish.
72	-movchilik:	Ŷ 67	verb - noun	anglashilmovchilik, kelishmovchilik, yetishmovchilik
73	-on:	Ŷ 68	verb - noun	qiron, to'zon
74	-os:	Ŷ 69	imitation	uvvos, chuvvos
75	-ot	Ŷ 70	Adjective-noun	ma'lumot, mushkulot xarobot
76	-at:	Ŷ 71	adjective - noun	she'riyat, madaniyat, majburiyat

77	-ch// -j// -inch:	Ŷ 72	verb - noun	sevinch, quvonch, yupanch, ilinj, qo'rqinch
78	-cha:	Ŷ 73(1)	verb - noun	tushuncha
79	-cha:	Ŷ 73 (2)	adjective - noun	qizilcha, olacha
80	-chak	Ŷ 74	verb - noun	belanchak, ovunchoq, taqinchoq
81	-chi	Ŷ 75(1)	adjective - noun	qiziqchi
82	-chi	Ŷ 75(2)	verb - noun	suyunchi, tilanchi, tomchi
83	-chilik:	Ŷ 76(1)	undov	haybarakallachi
84	-chilik:	Ŷ 76(2)	adjective - noun	pishiqchiiik, arzonchilik
85	-chiq:	Ŷ 77	adverb - noun	ko'pchilik, ozchilik
86	-shunos:	Ŷ 78	verb - noun	suyanchiq, yopinchiq

These suffixes come in handy in distinguishing word-forming suffixes from other types of suffixes in the process of morphological analysis.

Various lexical and syntactic suffixes are also actively used in the Uzbek language. We present these additions in Table 3 below.

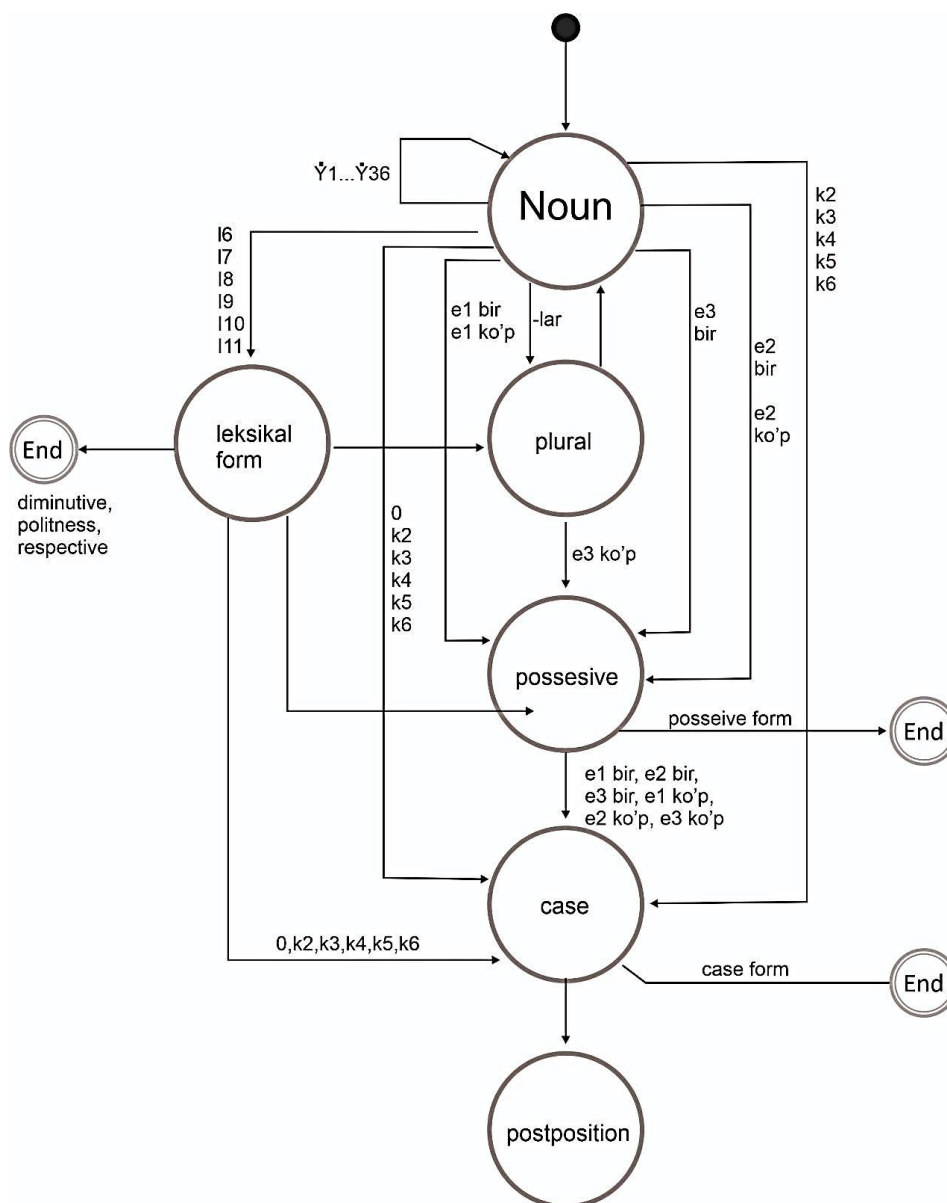
Table 3.

I	VOCABULARY FORMERS			
I.5	Number form			
1)	singular	-Ø	s1	Indicates that it is singular
2)	plural	-lar	s2	Indicates that there are many
I.6	Decreasing	-choq// chak	kich1	It means to diminution and to caress
		-kay	kich2	It means to diminution and to caress
		-ak	kich3	It means to diminution and to caress
I.7	Personal attitude	-jon	shm1	Positive attitude
		-xon	shm2	Positive attitude
		-oy	shm3	Positive attitude
		-loq/ aloq	shm4	Positive attitude, caress
		-gina/ kina/ qina	shm5	Positive attitude, caress
		-boy	shm6	Positive attitude, respect
		-bek	shm7	Positive attitude, respect
I.8	Dependence	-niki	q	Dependence
I.9	Place and	-dagi	o'-p	Means place and time

	time			
I.10	Limit	-gacha/ kacha/ qacha	ch	Time and place and sometimes mean boundaries between different objects
I.11	Assimilation	-dek	o'x1	means assimilation
		-day	o'x2	means assimilation
II.4	Kelishik			
1)	Case 1 (Bosh kelishik)	-Ø	k1	Connects words
2)	Case 2 (tushum kelishigi)	-ning	k2	Indicates that an object or event belongs to another object or person
3)	Case 3 (qaratqich kelishigi)	-ni	k3	Represents the ability to completely take an action of an object, perceived from the added base
4)	Case 4 (jo'nalish kelishigi)	-ga/ ka/ qa	k4	Represents the direction of motion, object, place, time, and so on
5)	Case 5 (o'rin-payt kelishigi)	-da	k5	Action represents the point where an object have to be, the object, the place, the time
6)	Case 6 (chiqish kelishigi)	-dan	k6	Harakatning chiqish nuqtasi, obyekt, o'rin, payt, sabab ma'nolarini anglatadi
II.5	Possession			
1)	I Person singular	-m/ im	e1bir	Means I Person singular and possession
2)	II Person singular	-ng/ ing	e2bir	Means II Person singular and possession
3)	III Person singular	-i/ si	e3bir	Means III Person singular and possession
4)	I Person plural	-miz/ imiz	e1ko'p	Means I Person plural and possession
5)	II Person plural	-ngiz/ ingiz	e2ko'p	Means II Person plural and possession
6)	III Person plural	-i(lari)/ si(lari)	e3ko'p	Means III Person plural and possession

2. About Uzbek FSM models

2.1. Morphological analysis of word group noun. Word-formation, lexical and syntactic suffixes are added to the nouns in the Uzbek language. In the Uzbek language, noun-forming suffixes from different stems are shown in Table 2. Possession and case suffixes to the noun are set out in Table 3. In the Uzbek language, the forms of words in the noun category are usually arranged in the following sequence: Base + word-former + number + possession + case. We have given these suffixes in the tables above, which are necessary in the analysis of the Uzbek language horse phrase based on FSM models. The rules of spelling formed as a result of the addition of a horse are given in paragraph 1.2. The FSM model of the analysis based on these appendices is given in Figure 1 below.



The first introductory part of the analysis is marked with a black dot. To make the model clear, the tags shown in Table 2-3 of the appendices are

written between actions. When suffixes are involved in the transition from one grammatical form to another, they are indicated by an arrow, cases of non-indicative change of grammatical meaning are indicated by an empty arrow. We explain the diagram below.

Noun is the base. The condition indicated by a black dot indicates a word coming from outside for analysis. This provides access to the FSM. It is considered an artificial noun with $\hat{Y}1 \dots \hat{Y}35$ suffixes in the base. Even if a noun-former comes after the base, the noun is made. For example: *fol+bin*, *darvoza+bon*, *soya+bon*, *zar+gar*, *savdo+gar*, *nam+garchilik*, *xafa+garchilik*, *sez+gi*, *sev+gi*. The drawing shows that a new noun was made from the noun's core. The noun can be in a state of transition from the base to several forms. These can be noun cases in the plural, lexical form, possessive, and accusative forms.

Plural: The form *base+lar* makes the plural forms of the noun: *bola+lar*, *olma+lar*. The word can end in this form or take the forms of possession and agreement, so there are three exits from this form:

- 1) plural noun: *qizlar*, *gullar*;
- 2) a plural noun that takes the form of a case: *qizlarning*, *qizlarni*, *qizlarga*, *qizlarda*, *qizlardan*;
- 3) noun in possession and plural: *qizlarim*, *qizlaring*, *qizlari*, *qizlarimiz*, *qizlaringiz*, *qizlari*.

Possession: status *e1bir*, *e2bir*, *e3bir*, *e1kop*, *e2kop*, *e3kop* take suffixes have many symbols and have two outputs:

- 1) A noun in the form of possession: *kitobim*, *kitobing*, *kitobi*, *kitobimiz*, *kitobingiz*, *kitoblari*;
- 2) A noun in the form of possession and case: *kitobim*, *kitobimning*, *kitobimni*, *kitobimga*, *kitobimda*, *kitobimdan*.

The state of the lexical form of the noun: the lexical form of the noun is formed by suffixes marked with tags I6, I7, I8, I9, I10, I11. For example: *odamniki*, *tog'agi*, *hovligacha*, *bo'taloq*, *onaxon*, *ukajon*. There will be 3 exit points from this case:

- 1) The lexical form of the noun: *uydagi*, *opajon*;
- 2) The lexical form of the possessive noun: *uyimizdagi*, *onaxonimiz*;
- 3) The lexical form of the plural noun: *uydagini*, *onajonimni*.

The state of the noun in the **form of a case** is formed in the following conditions.

The first condition. A direct case suffix is added to the base: *base + 0 / k2 / k3 / k4 / k5 / k6*. For example: *xona*, *xonaning*, *xonani*, *xonaga*, *xonada*, *xonadan*.

The second condition. To the base that takes the lexical form, an case suffix is added: *base + I6 / I7 / I8 / I9 / I10 / I11 + 0 / k2 / k3 / k4 / k5 / k6*.

For example: *bolajon*, *bolajonning*, *bolajonni*, *bolajonga*, *bolajonda*, *bolajondan*.

The third condition. A case suffix is added to the base that takes the form of possession: base + e1bir / e2bir / e3bir / e1kop / e2kop / e3kop + 0 / k2 / k3 / k4 / k5 / k6. For example: *kitobim*, *kitobimning*, *kitobimni*, *kitobimga*, *kitobimda*, *kitobimdan*.

There can be two output states from the contract form. A noun in the form of a case and a noun that receives an auxiliary after the contract.

References

1. Тузов В.А. Морфологический анализатор русского языка //Вестник СПбГУ, сер. 1. 1996. Вып. 1 (N15). – С. 41-45.
2. Формальные модели и программные инструменты компьютерной обработки татарского языка / Р.Р.Гатауллин, А.Р.Гатиатуллин, О.А.Неврозова, Д.Р.Мухамедшин, Д.Ш.Сулейманов, Б.Э.Хакимов, А.Ф.Хусаинов. – Академия наук РТ, Институт прикладной семиотики АН РТ. – Казань: Академии наук, 2019. – 260 с. – С.39-40.
3. Eşref Adalı. Türkçe Doğal dil İşlemi. – Ankara, 2020. – 754 s.
4. Gelbukh A. F. Effectively realizable morphologic model of inflective language [Эффективно реализуемая модель морфологии и эктивного языка]. Научно-техническаяИнформация [Scientific and Technical Information], series 2, №1, 1992. – P. 24-31.
5. Kemal Oflazer. Two-level Description of Turkish Morphology. Literary and Linguistic Computing, – Vol. 9, No 2, – 1994.
6. Çağrı Çöltekin (2014) A Set of Open Source Tools for Turkish Natural Language Processing In: Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14) Ed. by N. Calzolari et al. 1079-1086.