TurkLang-2020

- ***Syntactic structures for ontological models (on the example of the Uzbek language)***

- **Abdurakhmonova Nilufar, Aripov Mirsaid, Norov Abdusaid**

- Alisher Navo'i Tashkent State University of Uzbek Language and Literature, National University of Uzbekistan named after Ulugbek, Karshi State University

Introduction

- The paper deals with ontological models of syntactic structures of the Uzbek language comparing examples word combinations. According to syntactic models of the word combinations divides two main types: nominal and verbal. Connecting of each words of models has three syntactic ways: agreement, government, adjoinment. In this paper analyzed syntactic relations and models in order to create meta language for NLP and other linguistic technology in the frame of Turkic languages under the project of «AP05132249 Processing of electron thesaurus of Turkic languages for creation multilingual information retrieval system and extracting knowledge».

One of the linguistic properties of languages for natural language processing is syntax.

Syntactic parsing is crucial technology for each application of natural language processing: machine translation, question-answering system, information retrieval system and sentiment analysis, corpus linguistics. Consequently, building of the structure of text and word combinations plays essential role in order to identify the place of parts of speech. Each language has own linguistic peculiarities as according to typological system of languages. For example inflectional and agglutinative, having own ontological classification of parts of speech.

There are two crucial components as constituencies of syntax:

word combination

sentence

- Word combination represents the attitudes and the models of words. Words belong to things and substance, quality, attribute, and action. Things and substance, quality, substance, attribute, and action interconnect each other in word combination, but they cannot apart from independently each other. Syntax of word combination is capability of adjoining of words that estimated as connection ways and schemata (forms) as well as components and forms of word combination associates closely with morphology. It is studied word combination as a part of sentence and postfixes considered as morphological-syntactical category that joining each other's. Word combination comprises semantically and grammatical attitudes of at least two words. One is component of word combination comes as head (governor) and other dependent word. Components interact each other's based on semantically and syntactical rules. Word combination plays role as nominative means of language by headword naming things, substance, quality, substance, attribute, and action interconnect each other's.

- Ontological analysis clarifies the structure of knowledge. Given a domain, its ontology forms the heart of any system of knowledge representation for that domain. Ontology is used for formal and specialized concept and relations that belong to exact domain. Having advantage of ontology in NPL to create metalanguage in the sphere of machine translation (mainly, rule-based machine translation) or other purposes (information retrieval system, text analysis, annotation of text). Thanks to ontology, creating structure of information based on systematical and hierarchical data it aids to ease computational processing of the natural language. Effective way to create of ontology is representing OWL

MORPHOLOGICAL PECULIARIETIES OF UZBEK

- Details of models, formulations and techniques applied into your research (only if applicable)

Tag	Name_English	UZB
WC	Word combination	ukam uchun sotib olmoq
COLC	Collocation	o'z yog'iga qovurilmoq
FP\FCOLC	Free phrase\ Free collocation	xat yozmoq, kuchli iroda, ukamning kitobi
NP	Noun Phrase	bolalarning hammasi, intizomda birinchi, xushbo'y hid
NA	Noun Adjoinment	ona vatan, bebaho sovg'a
NG	Noun Government	ukam uchun sovg'a, senga mukofot
NCS	Noun Collateral subordination	ukamning xati
VP	Verb Phrase	baland uchmoq, kulib gapirmoq
AGRM	Agreement	u o'quvchi, mel keldim
SLP	Singular personal pronouns	Men talabaman
PPL	Plural personal pronouns	Ular talabalar

PREDICATIVE AGREEMENT RELATIONS BETWEEN UZBEK AND KAZAK

Dependent	Head									
	N, Adj, Pro, NUM, ADV + personal endings		V+Present / Future simple		V+Past simple				Condition	
	UZB	KZ	UZB	KZ	UZB	KZ	UZB	KZ	UZB	KZ
			-a / -y /-moqchi	-a, -e, -й	-b / -ib	-п, -ып, -iп	-di	-ti, -ты		
1 personal singular	-man	-мын, -мiн, -бын, -бiн, -пын, -пiн	-man	-мын, -мiн, -бын, -бiн, -пын, -пiн	-man	пын, -пiн	-m	-м	-y,-yin, -ay,-ayin	-йын, -йiн, -айын, -ейiн
2 personal singular	-san	-сың,-сiң	-san	-сың,-сiң	-san	-сың,-сiң	-ng, -ing	-ң	0	0
2 personal singular formal	-siz	-сыз, -сiз	-siz	-сыз, -сiз	-siz	-сыз, -сiз	-ngiz, -ingiz	-ңыз, -ңiз	-ngiz, ingiz	-ңыз, -ңiз, -ыңыз, -iңiз
3 personal singular	0	0	-di	-ды, -дi	-ti	ты, -тi	0	0	-sin	-сын, -сiн
1 personal plural	-miz	-мыз, -мiз, -быз, -бiз, -пыз, -пiз	-miz	-мыз, -мiз, -быз, -бiз, -пыз, -пiз	-miz	пыз, -пiз	-k	-қ, -к	-aylik	-йық, -йiк, -айық, -ейiк
2 personal plural	-sanlar	-сыңдар,-сiңдер	-sanlar	-сыңдар,-сiңдер	-sanlar	-сыңдар,-сiңдер	-nglar /-inglar	-ңдар, -ңдер	-ishgin	-ңдар, -ңдер, -ыңдар, -iңдер
2 personal plural formal	-sizlar	-сыздар, -сiздер	-sizlar	-сыздар, -сiздер	-sizlar	-сыздар, -сiздер	-ingiz / -ingizlar	-ңыздар, -ңiздер	-inglar, -ingiz	-ңыздар, -ңiздер, -ыңыздар, -iңiздер
3 personal plural	0	0	0	-di /-dilar	-lar	ты, -тi	-lar	0	-sin, -sinlar	-сын, -сiн

Nominal adjointment:

- Noun+Noun=> temir uskuna <=> Noun+Noun=> iron equipment
- Adj.+Noun=> qulay imkoniyat <=>Adj.+Noun=> suitable opportunity
- PNoun+Noun=>hamma ishtirokchilar <=>PNoun +Noun=> all participants
- Num.Noun=> birinchi kun <=> Num.+Noun=>the first day
- Gerund+Noun=> o'qiyotgan qiz <=> Gerund+Noun=> reading girl
- Infinitive+ Noun=>nishonlash kuni <=> Gerund +Noun=> celebrating day
- Adv.+Noun=> sekin harakat <=>Adj.+Noun=> slow movement
- (Noun+dagi)+ Noun=> devordagi rasm<=>Noun+be +Prep.+ Noun=> the picture is on the wall
- (Infinitive+dagi)+ Noun=> ishlashdagi g'ayrat <=> Noun+Prep.+Gerund=> enthusiasm in working
- (Adv.+dagi)+ Noun=> yuqoridagi qavat <=> Adv+Noun=>upper floor
- ↓PNoun+↓Adv. +Gerund+Noun=> (kimgadir) (sekin) o'qib berayotgan qiz <=>Noun+Question word+be+Ving(Adv.) (to smb.) girl who is reading (slowly) to smb.

- Noun|PNoun{ni, ga, da, dan}+Gerund+Noun=> maktabga ketayotgan qiz <=> Noun+Question word+be+Ving(Ad) +Prep.+Noun=> the girl who is going to school
- Adj.+Gerund|Past participle+Noun=> yaxshi o'qigan bola <=> Adj. | Adv+Gerund | Participle+Noun => well educated boy
- Adv.+Gerund+Noun=>tez kelgan lahza <=> Adv.+Gerund+Noun =>fast coming time
- (Noun+day|dek)+Adj.=> oyday oppoq<=>Adj.+like+Noun=> white like the moon
- (Noun+dagi)+Adj.=> sinfdagi a'lochi<=> Adj.+Prep.+Noun=>the smart in the classroom
- Adj.+Num.=> mo'jizaviy yetti<=> Adj.+Num.=> marvelous seven
- (Noun+dagi)+Num.=> rasmdagi bir<=>Num+Prep+Noun=>one in the picture
- Noun+Infinitive=>kitob o'qish (lekin *ism qo'yish, nonushta qilish* bu so'z birikmasi emas, ko'makchi fe'lli so'z qo'shilmasi) <=> reading a book
- Adj.+Infinitive=> qulay joylashish<=> Adj.+Gerund=>convenient placing
- Adv.+Infinitive=> tez yeyish<=> Infinitive+ Adv. => to eat fast

Verbal adjoinment:

- sifat+fe'l=> yaxshi o'qimoq–reading well
- ravish+fe'l=> astoydil o'qimoq–studying hard
- ravishdosh+fe'l=> kulib gapirmoq–speaking with smiling
- Adj.+Verb=> yaxshi o'qimoq<=>V+Adv=>read well
- Adv.+Verb=> astoydil o'qimoq<=>V+Adv=>study hard
- Gerund+Verb=> kulib gapirmoq<=>V+Prep.+Gerund=>speak with smiling
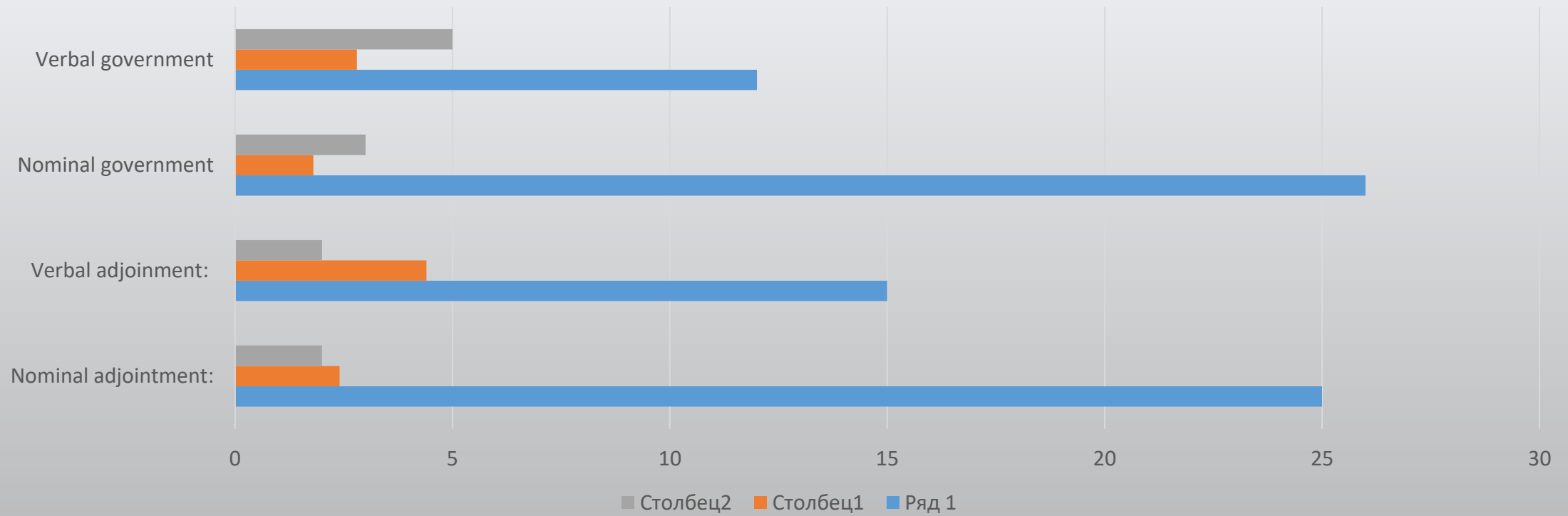
Nominal government:

- Noun+ dan+Noun=> Andijondan kelish <=>Noun+Prep.+Noun=> the letter from Andijan.

- Noun+↓dan ham |↓dan koʻra+Adj.+↓roq=>onadan mehribon <=> Adj. + than + Noun=> kinder than mother .

- Noun+dan+Infinitive=>ustozdan soʻrash <=> Gerund+Prep.+Adj.+Noun =>asking from the master.

- Gerund+ {Noun}dan +Infinitive => bilgandan soʻrash <=> Gerund + Prep. + Adj. +Noun=> asking from educated person.

- Gerund+dan+Adj.=>koʻrgandan gumon<=>Adj.+{Prep.}+Gerund=>doubtful of seeing.

- Noun | PNoun + dan + Num.=>hammadan birinchi <=> Num.+Prep.+PNoun|Noun=> the first all of them.

- Noun | PNoun + dan +Adj.=>hammadan ustun <=> Adj. + Prep. + Noun | PNoun=> the best of all of them.

- Adj. + ↓lar+dan+Num.=> aʼlochilardan ikkitasi<=>Num.+Prep.+Adj.=> two of the smarts.

- Adv. + dan+Adv.=> kechagidan erta<=>Adv.+than+Adv. => earlier than yesterday.

- Adv. + dan+Infinitive=>koʻpdan bilish<=> Infinitive +Prep.+Adj.=>to know from many (people).

- Num.+dan+Num.=>yuztadan bittasi<=>Num.+Prep.+one out of hundred.

- Noun+ga+Noun=> vatanga muhabbat <=>Noun+to+Noun=>love to homeland.

- PNoun+ga+Noun=> hammaga doʻst <=> Noun+to+PNoun=>friend to everybody.

- Gerund+ga+Noun=>oʻqiyotganga omad<=>Noun+to+Gerund+Noun=> luck to studying man.

- Infinitive+ga+Noun=>oʻqishga mehr <=>Noun+to+Infinitive=>love to study.

- Infinitive|Noun+ga+Infinitive=>oʻqishga intilish<=> Gerund|Noun +to+ Infinitive=> trying to study.

- Noun+ga+Adv.=>bayramga yaqin<=>Adv.+to+Noun=> close to holiday.

- Noun+da+Noun=>yozuvdv xato<=>Noun+Prep.+Noun=> mistake in writing.

- Noun+da+Num.=>tartibda birinchi <=>Num.+Prep.+Noun=>the first of order.

- Noun|PNoun+da+Adj.=>menda koʻp<=>PNoun|Noun+have+Adj.=> I have many | much.

- Adj.+ni+Infinitive=> qahramonni eslash <=>Infinitive+Noun=>to remember hero.

- Noun+ni+Infinitive=>farzandni sogʻinish<=>Gerund+Noun=> missing the child.

Verbal government

- Noun+ga+Verb=>maktabga bormoq <=>Infinitive+Prep.+Noun=>to go to school.
- Noun+ga+Infinitive=>daftarga yozmoq<=>Infinitive+Noun=>to write notebook.
- Noun|Pronoun+dan+Verb=>universitetdan qaytmoq<=> Verb +Prep. +Noun=>return from the university.
- Noun|Pronoun +ni+Verb=>hikoyani o'qimoq <=>Infinitive +Noun=> to read story.
- Noun +ni+ravishdosh=>ishni bajarib<=>Gerund+Noun=>doing work.
- Noun+da+Verb=>maktabda o'qimoq <=>Infinitive+Prep.+Noun=>to study at school.
- Noun+da+Gerund=>osmonda uchib kelayotgan=>Gerund +Prep.+Noun=> flying in the sky.

Syntactic structures of Uzbek

Usage protégé program input all linguistic data of structures, and then it becomes easy to split up components of sentence according to input ontological hierarchy.

Conclusion

This work has been implemented by project as mentioned above. Creation of Meta language for Turkic languages is crucial in order to NLP and other special purpose text analysis. All classes and subclasses with attributes input Protégé program to work further work. In spite of diversity of languages, there is commonness of grammatical rules among the Turkic languages. Entities inputted in Protégé as classes including object properties, data properties, individuals, annotation etc. In the result, ontology grammatical rules of Turkic languages (Uzbek, Kazakh, Tatar, Turkish, Kyrgyz) could be used for computational language processing in perspectives.

Thank you for attention!