УДК 81’33

Җ.Ш. Сөләйманов,
О.А. Невзорова,
А.Р. Гатиатуллин

«TURKLANG» ХАЛЫКАРА КОНФЕРЕНЦИЯСЕ – ТӨРКИ ТЕЛЛӘРНЕ ЭШКӘРТҮЧЕ КОМПЬЮТЕР ТЕХНОЛОГИЯЛӘРЕН ҮСТЕРҮ ДАИРӘСЕ

Целью международной конференций «TurkLang» является создание пространства совместных компьютерных лингвистических исследований для тюркских языков. На конференции представляются новые результаты, связанные с разработкой компьютерных лингвистических приложений для тюркских языков. Международная конференция «TurkLang» проводится ежегодно с 2013 года (Астана (2013), Стамбул (2014 г.), Казань (2015, 2017), Бишкек (2016)).

Ключевые слова: международная конференция, тюркские языки, компьютерная лингвистика, лингвистические модели, лингвистические программы

Актуальлек

Соңгы вакытта төрки телләрне компьютерда эшкәртү мәсьәләләренә багышланган фәнни-гамәли эшләрнең саны елдан-ел арта бара. Әлеге чаралар кысаларында төрки телләрнең мәгълүмати технологияләр өлкәсендә кулланылуын тәэмин итә торган системалар һәм программалар эшләү дә күз уңында тотыла [Bilkiz; Gulila Adongbieke, Mijit Ablimit; Ilgen, Adali, Tantuğ; Makazhanov, Sultangazina, Makhambetov, Yessenbayev; Nart B. Atalay, Kemal Oflazer and Bilge Say; Oflazer; Oorzhak, Khertek; Orhun, Tantuğ, Adalı; Sharipbayev, Razakhova; Zheltov; Сулейманов, 1996; Сулейманов, 1993; Сулейманов 1994]. Бирелдә төп игътибар морфологик анализ программаларын төзүгә һәм текстларга морфологик аннотация ясауга бирелә [Orhun, Tantuğ; Sharipbay, Bekmanova, Yergesh, Mukanova; Дыбо, Шеймович; Желтов; Сиразитдинов]. Бу – төрки телләрнең структур үзенчәлекләре һәм төрки телләрдәге текстларны эшкәртүдә морфологиянең зур роль уйнавы белән аңлатыла. Бастырылган мәкаләләр һәм төзелгән программалар санын анализлап карасаң, бу өлкәдә иң актив эшләүчеләр исемлегенә төрек, уйгур, казакъ һәм татар галимнәренең кергәнлеге күренә [Akba, Uçan, Sezer, Sever; Batuer Aisha and Maosong Sun; Çetin, Amasyali; Gulila Altenbek; Iknur Durgar El-Kahlout and Kemal Oazer; Kaya, Fidan, Toroslu; Razieh Ehsani, Muzaffer Ege Alper, Glsen Eryigit and Esref Adali; Tukeyev; Yıldırım, Tantuğ].

Шул ук вакытта, төрки телләрне саклау, өйрәнү, эшкәртү һәм үстерү өчен программалар төзүдә һәм лингвистик ресурслар туплау өлкәсендә кайбер тискәре тенденцияләр дә күзәтелә.

Алар, беренчедән, фәнни-тикшеренү эшләренең аерым-

аерым башкарылуында һәм белгечләр арасында мәгълүмат һәм технологияләр белән уртаклашу, бердәмлек булмауда чагылыш таба. Бу, аеруча, яңа терминнар ясаганда һәм мәгълүмати технологияләр өлкәсендә яңа концептлар тудырганда мөһим. Хәзерге вакытта компьютер технологияләрендә милли терминнар, нигездә, инглиз теленнән кабул итеп алына. Яңа төшенчәләр төрле ысуллар белән үзләштерелә: термин чит телдән үзгәрешсез алына (computer – компьютер); тәрҗемә ителә (mouse – тычкан), тел кагыйдәләренә фонетик яктан яраштырыла (cash – кәш) яки, сирәк очракта, кулланылышка яңа термин (keyboard – төймәтакта, төймәсар), яисә неологизм (data base – база данных – тамгасар) кертелә.

Компьютер технологияләре белән бәйле күп кенә бер үк яңа терминнарның төрки телләргә төрле сүзләр белән кереп китүен, бу өлкәнең бик тиз үсеп-үзгәреп торуын һәм төрки телләрнең охшашлыгын истә тотканда, компьютер терминнарының уртак сүзлеген туплау бүгенге көндә дә актуаль мәсьәлә булып кала бирә. Бу сүзлек, һичшиксез, төрки телләрдә мәгълүмат туплау, аны уртак кулланылышка кертү юнәлешендә уңайлыклар тудырачак, эшнең нәтиҗәлелеген арттырачак, мәгълүмат эшкәртү программаларын унификацияләштерүгә китерәчәк.

Икенчедән, төрки телләрне эшкәртүче лингвистик модельләр һәм программалар бер-берсен кабатлыйлар, алар 70–80 процентка барлык төрки телләр өчен дә уртак булган өлешләрдән тора. Әлеге кабатлаулардан котылу, уртак проектларда катнашу һәм программалар белән уртаклашу финанс чыгымнарын һәм сарыф ителгән көчне экономияләргә ярдәм итәчәк. Ә бу исә, белгечләрнең тырышлыгын бүгенге көндә чишелмәгән уртак проблемаларны хәл итүгә юнәлтергә мөмкинлек бирәчәк. Мондый эшчәнлек, үз чиратында, төрки телләрнең лексик-грамматик үзенчәлекләренә нигезләнгән яңа технологияләр тудырырга ярдәм итәчәк [Suleymanov, p. 210–213].

Хәзерге этапта төрки телләр даирәсендәге нәзари һәм гамәли тикшеренүләрнең төп юнәлешләре билгеләнде һәм белгечләрнең берләшүе, төрки телләрне эшкәртүдә уртак принциплар һәм ысуллар булдыру зарурлыгы ассызыкланды. Уртак проектлар киләчәктә бу өлкәгә караган эзләнүләрнең нәтиҗәлелеген дә арттырачак, текстларны эшкәртүче күптелле системалар төзергә һәм башка фундаменталь һәм гамәли мәсьәләләрне чишәргә ярдәм итәчәк.

Үткәрү тарихы

Төрки телләрне компьютер технологияләре ярдәмендә эшкәртү мәсьәләләренә багышланган «TurkLang» халыкара конференциясе 2013 елда – Астанада (Казакъстан) [Труды Первой...], 2014 елда – Истанбулда (Төркия) [Proceedings of the International.., 2014], 2015, 2017 елларда – Казанда (Татарстан, Россия Федерациясе) [Proceedings of the International.., 2015; Пятая Международная...], 2016 елда

Бишкәктә (Кыргызстан) [Материалы 4-ой Международной...] үткәрелде. Ул баштан ук белгечләрнең уртак стандартлар төзү өчен тырышлыкларын берләштерүнең, төрки телләрне өйрәнү белән бәйле мәгълүмати һәм технологик ресурсларны уртак бер мәйданга туплауның эффектив механизмы буларак күзалланды.

Эш барышында пленар утырышлар, түгәрәк өстәлләр, программалар белән танышу каралган иде. Конференция кысаларында төрки телләрнең формаль модельләре, компьютер системаларын төрки телләргә яраштыру һәм терминология, электрон корпуслар төзү, текстларны эшкәртү өчен морфологик һәм синтаксик технологияләр, сөйләмне анализлау һәм синтезлау системалары, текстларны эшкәртү өчен семантик системалар, автоматлашкан тәрҗемә системалары, төрки телләргә өйрәтүче интеллектуаль системалар һәм технологияләр, төрки телләрнең лингвистик онтологияләре белән бәйле мәсьәләләр тикшерелде.

Узган конференцияләрнең эшчәнлегендә дөньяның төрле илләреннән түбәндәге фәнни үзәкләрнең вәкилләре катнашты: Татарстан Фәннәр академиясе (Казан, Татарстан), Казан федераль университеты, М.В. Ломоносов исемендәге Мәскәү дәүләт университеты (Мәскәү), Россия Фәннәр академиясенең Тел белеме институты (Мәскәү), Югары икътисад мәктәбе (Мәскәү), Россия Фәннәр академиясенең Информатика проблемалары институты (Мәскәү), Санкт-Петербург дәүләт университеты (Санкт-Петербург), Новосибирск дәүләт университеты (Новосибирск), Россия Фәннәр академиясенең Уфа фәнни үзәге Тел, әдәбият һәм тарих институты (Уфа, Башкортстан), М.К. Аммосов исемендәге Төньяк-көнчыгыш федераль университет (Якутск, Саха), И.Н. Ульянов исемендәге Чуваш дәүләт университеты (Чебоксар, Чувашия), Тыва дәүләт университеты (Кызыл, Тыва), В.И. Вернадский исемендәге Кырым федераль университеты (Симферополь, Кырым), Синьцзянь университеты (Ырымчы, Кытай), Цинхуа Университеты (Бейджин, Кытай), Баку Евразия университеты (Баку, Азәрбайҗан), Оксбридж академиясе (Баку, Азәрбайҗан), Л.Н. Гумилев исемендәге Евразия милли университеты (Астана, Казакъстан), Назарбаев Университеты (Астана, Казакъстан), әл-Фараби исемендәге Казакъ милли университеты (Алматы, Казакъстан), И. Раззаков исемендәге Кыргыз техник университеты (Бишкәк, Кыргызстан), Кыргызстан милли фәннәр академиясенең Нәзари һәм гамәли математика институты (Бишкәк, Кыргызстан), Н. Исанов исемендәге архитектура, транспорт һәм төзелеш кыргыз дәүләт университеты (Бишкәк, Кыргызстан), Карасаев исемендәге Бишкәк дәүләт университеты (Бишкәк, Кыргызстан), Ош технология университеты (Ош, Кыргызстан), Бөек Витаутас университеты (Каунас, Литва), Истанбул техник университеты (Истанбул, Төркия), Алишер Нәвои исемендәге Үзбәк

теле һәм әдәбияты университеты (Ташкент, Үзбәкстан), Аризона университеты (Тусон, Аризона, АКШ), Суортмор колледжы (Суортмор, Пенсильвания, АКШ).

2018 елда конференцияне Үзбәкстан Республикасының Ташкент шәһәрендә, Алишер Нәвои исемендәге үзбәк теле һәм әдәбияты университетында үткәрү планлаштырыла.

«UniTurk» семинары

Конференция кысаларында ел саен «UniTurk» семинары оештырыла. Анда төрки телләрдәге текстларга унификацияләнгән морфологик аннотацияләр төзү проблемалары тикшерелә. Анализ күрсәткәнчә, төрки телләр үзләренең структурасы һәм тарихи килеп чыгышы ягыннан нык охшаш булсалар да, аларны аннотацияләүнең уртак принциплары һәм ысуллары юк. Киләчәктә бу хәл телләрне чагыштырып өйрәнү юнәлешендә эзләнүләр үткәргәндә, төрки телләрнең параллель корпусларын төзегәндә яки күп телле эшкәртү системалары төзегәндә шактый гына кыенлыкларга китерергә мөмкин.

Әлеге өлкәдә эшләүче белгечләрнең күбесе аннотацияләү өчен уртак теглар системасының кирәклеген аңлауга карамастан, төрки телләрнең корпусларында лингвистик мәгълүматны билгеләү өчен уртак унификацияләнгән формат һаман да эшләнмәгән. Бер үк морфологик категорияләрне төрки галимнәр төрлечә билгели. Төрки телләрнең корпусларын төзүчеләр, нигездә, башка типтагы телләр өчен төзелгән теглар системасын кулланалар. Беренче чиратта, ул һинд-европа телләре системалары, ә алар күп очракта төрки телләрнең үзенчәлекләрен дөрес күрсәтмиләр. Шуңа күрә бүгенге көндә төрки телләр өчен махсус аннотацияләү системасын төзү аеруча актуаль.

Корпусларны аннотацияләү өчен уртак система булмауның сәбәпләре, бер яктан, объектив фәнни эчтәлекле проблемаларга бәйле булса (тел системасының катлаулылыгы, төрле телләрнең грамматик структураларындагы үзенчәлекләр, морфонологик процессларның аермалылыгы), икенче яктан, оештыру мәсьәләләренә кайтып кала (координацияләүче бер үзәкнең булмавы; корпусларда грамматик һәм семантик аннотацияләр төзүнең стандартлары һәм уртак лингвистик мәгълүмат базаларының юклыгы).

Төрки телләр корпусларының грамматик аннотацияләү системаларын унификацияләүгә багышланган «UniTurk» фәнни-гамәли семинарында катнашучылар, төрки тел белемендә иң мөһим мәсьәләләрнең берсе итеп, лингвистик мәгълүматны билгеләү өчен шундый стандартлар тудыруны атыйлар. Әлеге стандартлар элекке һәм яңадан төзелүче электрон корпусларны кулланучылар өчен бердәм мәгълүмати даирә итеп коруның ышанычлы нигезе булыр иде.

Фәнни нәтиҗәләр: уртак грантлар, хезмәттәшлек турындагы килешүләр, яңа катнашучылар

«TurkLang» конференциясе башка төр хезмәттәшлекне үстерүгә дә ярдәм итә. Мәсәлән,

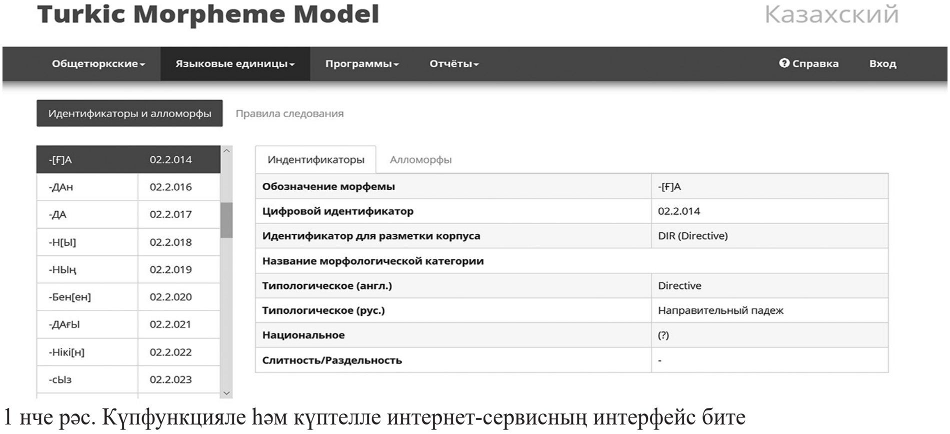
Татарстан Фәннәр академиясенең Гамәли семиотика институты хезмәткәрләре, Л.Н. Гумилев исемендәге Евразия милли университетының Шәкли фәһем институты, Истанбул техник университетының компьютер инженериясе һәм информатика факультеты, И. Раззаков исемендәге Кыргыз техник университеты хезмәткәрләре һәм Алишер Нәвои исемендәге Үзбәк теле һәм әдәбияты университеты хезмәткәрләре белән берлектә, «Күптелле эзләү системалары өчен төрки телләрнең электрон тезаурусларын төзү» исемле уртак халыкара проектта катнашалар.

Татарстан Фәннәр академиясе соңгы ике ел эчендә И. Раззаков исемендәге Кыргыз дәүләт техник университеты һәм Алишер Нәвои исемендәге Үзбәк теле һәм әдәбияты университеты белән төрки телләрне компьютерлаштыру өлкәсендә инструменталь һәм лингвистик ярдәм күрсәтү буенча хезмәттәшлек турында килешүләр төзеде. Әлеге эшчәнлек кысаларында бердәм мәгълүмати даирә төзү мәсьәләләрен хәл итү дә каралган. Килешүләрне төзүгә конференциядә төрки телләрне компьютерда эшкәртүнең яңа модельләрен һәм технологияләрен күрсәтү этәрде.

Моннан тыш, хәзерге вакытта бердәм мәгълүмати даирә төзү максаты белән бер төркем махсус программалар эшләнә. Алар галимнәргә әлеге өлкәдә уртак эшләр башкару мөмкинлеген тудырачак. Шундый программаларның берсе – Татарстан Фәннәр академиясенең Гамәли семиотика институтында эшләнә торган күпфункцияле һәм күптелле интернет-сервис (1 нче рәс.). Төрки морфема моделе нигезендә төзелгән әлеге интернет-сервис, бер яктан, төрки телләрне компьютерда эшкәртү мөмкинлекләрен биргән программалар каталогын, ә икенче яктан, шул программалар белән эшләү мөмкинлекләрен тудырган веб-сайтны тәшкил итә.

Оештыру мәсьәләләре

Конференцияне оештыручы төп оешмалар – Татарстан Фәннәр академиясе, Казан федераль университеты, Л.Н. Гумилев исемендәге Евразия милли университеты, Истанбул техник университеты, Россиянең Шәкли фәһем

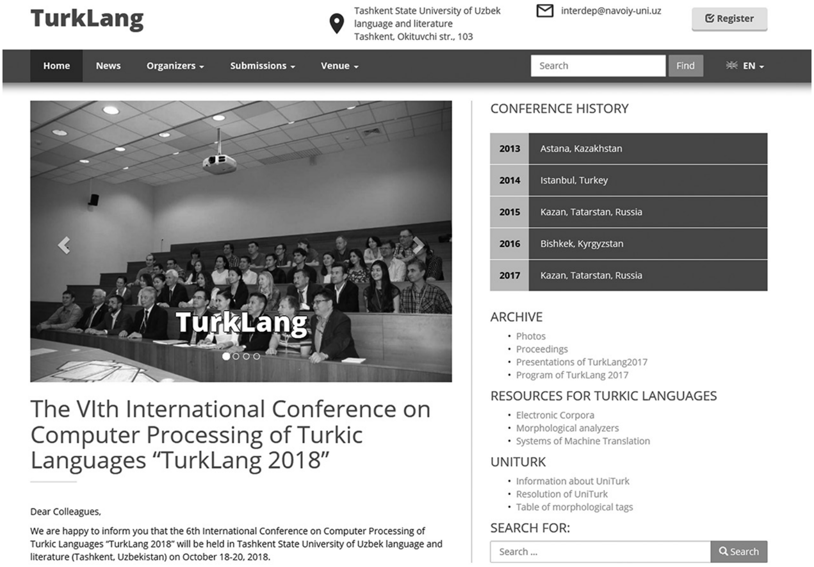


1 нче рәс. Күпфункцияле һәм күптелле интернет-сервисның интерфейс бите

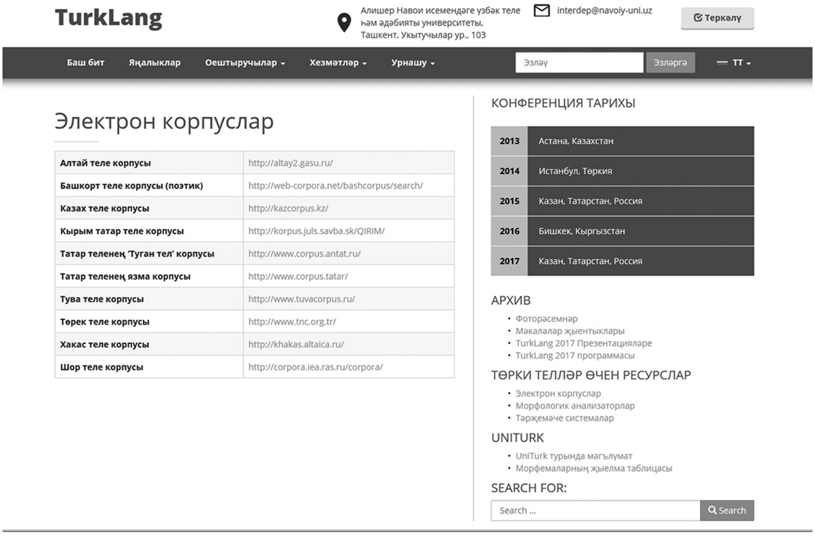
ассоциациясе. Конференциягә кагылышлы актуаль мәгълүмат махсус www.turklang.tatar сайтында даими чагылдырыла (2 нче рәс.). Сайтта конференция кысаларында 5 ел эчендә тупланган барлык мәкаләләр җыентыклары да урнаштырылган.

Сайтта, конференциянең үзе турындагы мәгълүматтан тыш, төрки телләр өчен ясалган компьютер программалары һәм электрон ресурслар турында да мәгълүмат бирелгән (3 нче рәс.).

Конференциянең сайты күптелле итеп эшләнде. Анда инглиз,



2 нче рәс. Конференция сайты



3 нче рәс. Төрки телләрнең электрон корпуслары турындагы сайт бите

татар, казакъ, кыргыз, төрек, үзбәк, саха һәм рус телләрендәге битләр бар. Интерфейс телләре саны һәм сайтта урнаштырылган ресурслар саны даими рәвештә артып барачак.

5 ел эчендә конференциягә бирелгән мәкаләләр һәм авторлар санына анализ ясалды (1 нче таблица). Бу анализ катнашучыларның һәм мәкаләләрдә күрсәтелгән төрки телләр санының әкренләп артуын күрсәтә. Иң күп мәкаләләр саны Казанда үткән конференцияләрдә булган (2015 елда – 49, 2017 елда – 46).

Нәтиҗә

Оештыручылар һәм катнашучылар «TurkLang» конференциясен төрки телләрне эшкәртү, өйрәнү, үстерү өчен уртак лингвистик ресурслар, системалар һәм технологияләр төзү, уртак лингвистик эзләнүләр башкару даирәсе итеп күрәләр һәм шул юнәлештә эшләүне дәвам итәләр. Конференция сайтын төрки телләрдәге ресурсларны һәм төрки телләрне эшкәртүче программаларны урнаштыру өчен ачык платформага әйләндерү актуаль мәсьәлә булып санала (анда мәгълүмат базалары, терминологик һәм аңлатмалы сүзлекләр, тезауруслар урнаштыру каралган) һәм анда, беренче чиратта, морфологик һәм синтаксик анализаторлар, автоматлашкан тәрҗемә һәм Интернетта мәгълүмат эзләү системалары тупланырга тиеш.

1 нче таблица

	2013		2014		2015		2016		2017	
Телләр	Авт.	Док.	Авт.	Док.	Авт.	Док.	Авт.	Док.	Авт.	Док.
Азәрбайҗан	0	0	0	0	1	1	0	0	2	2
Башкорт	2	1	1	1	5	2	0	0	4	1
Казакъ	60	34	18	7	22	12	27	8	26	11
Каракалпак	0	0	0	0	1	1	0	0	0	0
Кыргыз	5	3	0	0	3	1	6	5	9	6
Кырым татар	0	0	0	0	1	1	0	0	0	0
Татар	14	8	11	6	24	11	3	2	14	11
Тыва	0	0	0	0	4	2	0	0	7	2
Төрек	1	1	17	8	8	2	1	1	1	1
Үзбәк	9	6	0	0	1	1	1	1	3	4
Уйгур	0	0	0	0	7	2	0	0	0	0
Хакас	0	0	0	0	3	1	0	0	0	0
Чуваш	0	0	0	0	9	5	0	0	6	5
Саха (якут)	0	0	0	0	5	4	0	0	3	2
Гомумтөрки	0	0	0	0	3	3	2	1	2	2
Барлыгы	91	53	47	22	87	49	38	18	77	46
Телләр саны	5		4		15		6		11	

Әдәбият

Akba F., Uçan A., Sezer E., Sever H. Assessment of Feature Selection Metrics for Sentiment Analysis: Turkish Movie Reviews // In 8th European Conference Data Mining. 2014. P. 180–184.

Batuer Aisha and Maosong Sun. A Uyghur Morpheme Analysis Method based on Conditional Random Fields, International Journal of Asian Language Processing, 2009, 19(2). P. 69–83.

Bilkiz. Uyghur Corpus experimental study of part of speech tagging, Xinjiang fifth Annual Youth Symposium, Xinjiang People's Publishing House in October (in Chinese), 2004.

Çetin M., Amasyali M.F. Active Learning for Turkish Sentiment Analysis // In Innovations in Intelligent Systems and Applications (INISTA), IEEE International Symposium, 2013.

Gulila Adongbieke, Mijit Ablimit. Research on Uyghur word segmentation, Journal of Chinese information processing, 18(6): 61–65 (in Chinese), 2004.

Gulila Altenbek. Automatic Morphological Tagging of Contemporary Uighur Corpus. The IEEE International Conference on Information Reuse and Integration, 2006. P. 557–560.

Iknur Durgar El-Kahlout and Kemal Oazer. Initial explorations in English to Turkish statistical machine translation. In Fifth International Conference on Information Technology, Las Vegas, NV, 2008.

Ilgen B., Adali E., Tantuğ A.C. A Comparative Study to Determine the Effective Window Size of Turkish Word Sense Disambiguation Systems, 28th International Symposium on Computer and Information Sciences, Paris, France, 2013.

Kaya M., Fidan G., Toroslu I. Sentiment Analysis of Turkish Political News // In Web Intelligence and Intelligent Agent Technology (WI-IAT), IEEE/WIC/ACM International Conferences, V. 1. 2012. P. 174–180.

Makazhanov A., Sultangazina A., Makhambetov O., Yessenbayev Zh. Syntactic annotation of Kazakh: following the universal dependencies guidelines. A report // Proceedings of the International Conference «Turkic Languages Processing: TurkLang-2015». Kazan: Academy of Sciences of the Republic of Tatarstan Press, 2015. P. 338–350.

Nart B. Atalay, Kemal Oflazer and Bilge Say. The Annotation Process in the Turkish Treebank. In Proceedings of the EACL Workshop on Linguistically Interpreted Corpora. Budapest, 2003.

Oflazer K. Two-level Description of Turkish Morphology. Literary and Linguistic Computing. Vol. 9. No 2. 1994.

Oorzhak B., Khertek A. Development of semantic mark-up for the corpus of Tuvan language // Proceedings of the International Conference «Turkic Languages Processing: TurkLang-2015». Kazan: Academy of Sciences of the Republic of Tatarstan Press, 2015. P. 351–362.

Orhun M., Tantuğ A.C., Adalı E. Morphological Disambiguation Rules For Uyghur Language // IEEE International Conference on Software Engineering and Service Sciences (ICSESS), Beijing, China, 2010.

Proceedings of the International Conference «Turkic Languages Processing: TurkLang-2015». Kazan: Academy of Sciences of the Republic of Tatarstan Press, 2015. 488 p.

Proceedings of the International Conference on Turkic Language Processing (TURKLANG-2014). (Istanbul, November 6–7, 2014). Istanbul: Özkaracan Matbaacılık-Bağcılarr, 2014. 135 p.

Razieh Ehsani, Muza_er Ege Alper, Glsen Eryigit, and Esref Adali. Disambiguating main POS tags for Turkish // In Proceedings of ROCLING – Conference on Computational Linguistics and Speech Processing. Association for Computational Linguistics and Chinese Language Processing (ACLCLP), Taiwan, 2012.

Sharipbay A.A., Bekmanova G., Yergesh B., Mukanova A. Synchronized liner tree for morphological analysis and generation of the Kazakh language. Proceedings of the international conference «Turkic languages processing», TurkLang 2014, Istanbul, P. 113–117.

Sharipbayev A.A., Razakhova B.Sh. Mathematical models of syntactical rules of Kazakh language subject to semantics of parts of the sentence // The 4th Congress of the Turkic World Mathematical Society. Baku, 2011. P. 463.

Suleymanov D.Sh. Natural Cognitive Mechanisms in the Tatar language // In the Collection of the Vienna Proceedings of the Twentieth European Meeting in Cybernetics and Systems Research. Edited by Robert Trappel. Vienna, Austria, 6–9 April, 2010. P. 210–213.

Tukeyev U. Automaton models of the morphology analysis and the completeness of the endings of the Kazakh language // Proceedings of the International Conference «Turkic Languages Processing: TurkLang-2015». Kazan: Academy of Sciences of the Republic of Tatarstan Press, 2015. P. 91–100p.

Yıldırım E., Tantuğ A.C. The feasibility analysis of re-ranking for N-best lists on English-Turkish machine translation, IEEE International Symposium on Innovations in Intelligent Systems and Applications, Albena, Bulgaria, 2013.

Zheltov P. Morphological annotation system for the national corpus of the Chuvash language // Proceedings of the International Conference «Turkic Languages Processing: TurkLang-2015». Kazan: Academy of Sciences of the Republic of Tatarstan Press, 2015. P. 328–330.

Дыбо А.В., Шеймович А.В. Автоматический морфологический анализ для корпусов тюркских языков // Филология и культура. 2014. № 2.

Желтов П.В. Морфологический анализатор чувашского языка. Материалы Международной конференции студентов и аспирантов по фундаментальным наукам «Ломоносов 2002», М., 2002.

Материалы 4-ой Международной конференции по компьютерной обработке тюркских языков «TurkLang 2016» // Известия кыргызского государственного технического университета им. И.Раззакова, 2016 №2(38), С. 9–130.

Сиразитдинов З.А. Алгоритмическая грамматика словоизменения башкирского языка // [Электронный ресурс]. URL: http://mfbl.ru/bashdb/algram/ algram.htm (дата обращения: 19.03.2018).

Сулейманов Д.Ш. К описанию татарских падежей через структурно-функциональную модель морфем (статья на англ. яз.) // Proceedings of the LP'96 Conference on Languages Typology (Prague, 22-24 August, 1996). Prague, 1996. P. 388–395.

Сулейманов Д.Ш. К разработке эффективной морфологии среднетюркского языка // Сб. трудов межд. конф. Проблемы создания среднетюркского языка Ортатюрк. Ташкент, 1993. С. 85–86.

Сулейманов Д.Ш. Регулярность морфологии татарского языка и типы нарушений в языке Серия: Интеллект. Язык. Компьютер. Вып. 1. Казань: Изд-во Казан. ун-та, 1994. С. 77–106.

Труды Первой международной конференции «Компьютерная обработка тюркских языков». Астана: ЕНУ им. Л.Н. Гумилева, 2013. 345с.

Пятая Международная конференция по компьютерной обработке тюркских языков «TurkLang 2017»: Труды конференции: в 2-х томах.– Казань: Издательство Академии наук Республики Татарстан, 2017. Т. 1. 300 с.: Т. 2. 327 с.

Сөләйманов Җәүдәт Шәүкәт улы,
техник фәннәр докторы, профессор,
Татарстан Фәннәр академиясе академигы

Невзорова Ольга Авенировна,
техник фәннәр кандидаты, Казан (Идел буе) университетының
Хисаплау математикасы һәм мәгълүмати технологияләр
институты доценты

Гатиатуллин Айрат Рафиз улы,
техник фәннәр кандидаты, Казан (Идел буе) университетының
Хисаплау математикасы һәм мәгълүмати технологияләр
институты укытучысы