

ҚАЗАҚСТАН РЕСПУБЛИКАСЫ ҒЫЛЫМ ЖӘНЕ ЖОҒАРЫ БІЛІМ МИНИСТРЛІГІ

«Л.Н. ГУМИЛЕВ АТЫНДАҒЫ ЕУРАЗИЯ ҰЛТТЫҚ УНИВЕРСИТЕТІ» КЕАҚ

**Студенттер мен жас ғалымдардың
«GYLYM JÁNE BILIM - 2024»
XIX Халықаралық ғылыми конференциясының
БАЯНДАМАЛАР ЖИНАҒЫ**

**СБОРНИК МАТЕРИАЛОВ
XIX Международной научной конференции
студентов и молодых ученых
«GYLYM JÁNE BILIM - 2024»**

**PROCEEDINGS
of the XIX International Scientific Conference
for students and young scholars
«GYLYM JÁNE BILIM - 2024»**

**2024
Астана**

УДК 001

ББК 72

G99

«ǴYLYM JÁNE BILIM – 2024» студенттер мен жас ғалымдардың XIX Халықаралық ғылыми конференциясы = XIX Международная научная конференция студентов и молодых ученых «ǴYLYM JÁNE BILIM – 2024» = The XIX International Scientific Conference for students and young scholars «ǴYLYM JÁNE BILIM – 2024». – Астана: – 7478 б. - қазақша, орысша, ағылшынша.

ISBN 978-601-7697-07-5

Жинаққа студенттердің, магистранттардың, докторанттардың және жас ғалымдардың жаратылыстану-техникалық және гуманитарлық ғылымдардың өзекті мәселелері бойынша баяндамалары енгізілген.

The proceedings are the papers of students, undergraduates, doctoral students and young researchers on topical issues of natural and technical sciences and humanities.

В сборник вошли доклады студентов, магистрантов, докторантов и молодых ученых по актуальным вопросам естественно-технических и гуманитарных наук.

УДК 001

ББК 72

G99

ISBN 978-601-7697-07-5

**©Л.Н. Гумилев атындағы Еуразия
ұлттық университеті, 2024**

9. Sun, Y., Peng, M., Zhou, Y., Huang, Y., & Mao, S. (2019). Application of machine learning in wireless networks: key techniques and open issues. *Ieee Communications Surveys & Tutorials*, 21(4), 3072-3108. <https://doi.org/10.1109/comst.2019.2924243>

10. Islami, K. and Sopia, D. (2022). Artificial intelligence in human resources in the era of society 5.0. *International Journal of Research and Innovation in Social Science*, 06(11), 675-681. <https://doi.org/10.47772/ijriss.2022.61131>

УДК 004.02

САЛЫСТЫРУ КОНСТРУКЦИЯЛАРЫН АНЫҚТАУ КЕЗЕҢДЕРІ

Тогисова Акерке Бакитбековна

togisovakerke@gmail.com,

Л.Н.Гумилев атындағы ЕҰУ, Ақпараттық технологиялар факультеті, Ақпараттық жүйелер кафедрасының 1 курс докторанты, Астана қ., Қазақстан
Ғылыми жетекшісі – Касекеева Айслу Бисеновна

Көптеген адамдар әлеуметтік желі платформаларында белгілі бір объект немесе оқиға туралы пікірлерімен, эмоцияларымен, идеяларымен және сындарымен алмасады.

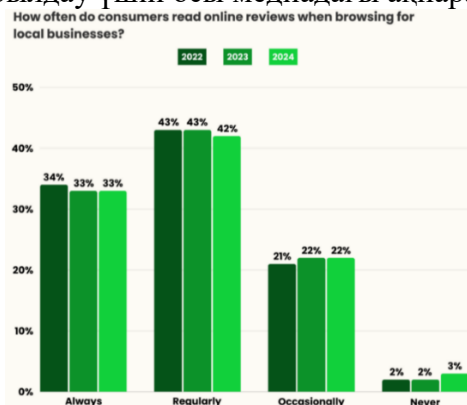
Пікірлер, бағалаулар, көзқарастар және эмоциялар сентименттік талдау мен пікірлерді талдау салаларының зерттеу пәні болып табылады (Sentiment analysis және Opinion mining). Бұл салалардың пайда болуы және қарқынды өсуі интернет қолданушылар және әлеуметтік медиа санының өсуіне сәйкес келеді [11].

2024 жылдың қаңтар айында Meltwater және We Are Social бірлесе жариялаған Digital 2024 Global Overview Report бойынша әлемде мобильді құрылғыларды қолданушылар саны 5,61 млрд, интернетті қолданушылар саны 5,35 млрд, ал әлеуметтік желілерді қолданушылар саны 5,04 млрд-қа жеткен [8].

Пікірлер – тұтынушылардың өнім немесе қызмет бойынша тәжірибелері туралы жазған бағалаулары.

Пікірлер бағалау, тәжірибенің жазбаша сипаттамасы және фотосуреттер немесе бейнелер сияқты кең ауқымды ақпаратты қамтуы мүмкін. Пікірлер өнім немесе қызметтің мүмкіндіктері, бағасы және сапасы сияқты нақты ақпаратты қамтуы мүмкін.

Шешім қабылдау қажет болған кезде біз басқалардың пікірлерін білгіміз келеді. Сол сияқты кәсіпорындар мен ұйымдар әрқашан өз өнімдері мен қызметтері туралы тұтынушылардың пікірлерін білгісі келеді. Жеке тұтынушылар өнімді сатып алмас бұрын оны басқа пайдаланушылардың пікіріне немесе саяси сайлауда дауыс беру шешімін қабылдас бұрын саяси кандидаттар туралы басқалардың пікірлеріне жүгінеді. Тұтынушылардың пікірлерін білу маркетинг, қоғаммен байланыс және саяси науқан компаниялары үшін маңызды болып табылады. Интернеттегі әлеуметтік медианың (мысалы, шолулар, форум, блогтар және әлеуметтік желілердегі хабарламалар, комментарийлер) қарқынды өсуімен жеке адамдар мен ұйымдар шешім қабылдау үшін осы медиадағы ақпаратты көбірек пайдаланады.



1 сурет. Brightlocal компаниясының жүргізген зерттеу нәтижесі

Тұтынушылардың шешіміне басқа тұтынушылардың ақпарат құралдарында жарияланған пікірлер мен эмоциялары әсер етеді. Brightlocal компаниясының жергілікті онлайн шолу сауалнамасы тұтынушылардың 75 % шешім қабылдау үшін онлайн пікірлерді оқитынын көрсетеді [9].

Пікірлерді талдау бизнес үшін қоғамдық пікірден құнды ақпарат алудың маңызды құралы болып табылады. Бұл тұтынушыларға қызмет көрсетуге, брендті басқаруға, өнімді әзірлеуге, өнім немесе қызмет сапасын жақсартуға және нарықты зерттеуге көмектеседі.

Әлеуметтік желілердегі пікірлерді талдау қазіргі сентиментті талдау (Sentiment analysis) мен пікірді талдау (Opinion mining) салаларының маңызды зерттеу объектісіне айналды [5].

Сентиментті талдау және пікірді талдау – NLP-де қолданылатын ең танымал әдістер. Олар тұтынушылардың өнімдер, қызметтер және т.б. туралы пікірлерін талдау және түсіну үшін пайдаланылады.

Табиғи тілді өңдеу (NLP) – бұл машиналарға адам тілін әртүрлі формаларда түсінуге және қолдануға мүмкіндік беретін зерттеу саласы.

Сентиментті талдау мен пікірді талдау жиі бір-бірін алмастыратын ұғымдар ретінде қолданылғанымен, екі әдіс арасында айырмашылықтар бар. Негізгі айырмашылығы, сентиментті талдау мәтін бөлігін алдын ала анықталған көңіл күй санаттарына (оң, теріс немесе бейтарап) жіктеуге, ал пікірді талдау мәтіннен субъективті ақпаратты алуға бағытталған.

Сентиментті талдау (немесе пікірді талдау деп те аталады) – адамдардың өнімдер, қызметтер, ұйымдар, жеке тұлғалар, мәселелер, оқиғалар, тақырыптар және олардың атрибуттары сияқты нысандарға қатысты пікірлерін, көңіл-күйлерін, бағалауларын, бағалауларын, көзқарастарын және эмоцияларын талдайтын зерттеу саласы [5].

Сентиментті талдау және пікірді талдау терминдері алғаш рет 2003 жылы Nasukawa Tetsuya және Kushal Dave жариялаған мақалаларында айтылған (сәйкесінше). Сентиментті және пікірді талдау туралы Bo Pang және Lillian Lee, Bing Liu жазған еңбектері бар.

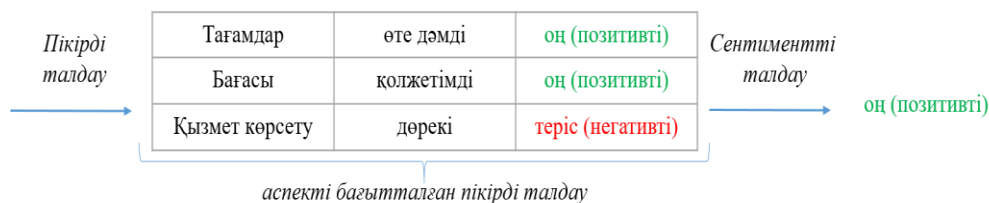
Bing Liu өз кітабында [5] сентиментті талдау және пікірді талдау синонимдер ретінде, бір-бірін алмастыратын ұғымдар ретінде қолданатынын жазған.

Bing Liu [5] сентименттік талдау негізгі үш деңгейде зерттеледі деп жазған: құжат деңгейі, сөйлем деңгейі, субъект және аспект деңгейі.

Abbasi Moghaddam S. [17] диссертациялық жұмысында пікірлерді талдаудың классификациясын жасаған.

- Құжат деңгейіндегі пікірді талдау
- Субъективтіліктің классификациясы
- Сентимент классификациясы
- Пікір сапасы мен пайдалылығын бағалау
- Спам анықтау
- Құжат деңгейіндегі пікірді талдау
- Пікірлерді іздеу және алу
- Пікір сұрағына жауап беру
- Пікірді қорытындылау (қорытынды пікір)
- Салыстырмалы сөйлемдердегі пікірді анықтау
- Фраза (сөз тіркестері, сөздер) деңгейіндегі пікірді талдау
- Аспект-бағытталған пікірді талдау

Мейрамханада ұсынылған тағамдар өте дәмді және бағалары қолжетімді, бірақ қызметкерлері дөрекі.



2 сурет. Сентиментті талдау және пікірді талдау

Объект және оның аспектілері туралы оң немесе теріс пікірлерді тікелей білдіруден басқа, ұқсас субъектілерді салыстыру арқылы да пікір білдіруге болады. Мұндай пікірлерді салыстырмалы пікірлер деп атайды.

Объектіні немесе қызмет түрін бағалаудың ең маңызды тәсілдерінің бірі оны ұқсас нысанмен немесе қызмет түрімен тікелей салыстыру болып табылады [6].

Салыстыру – екі (немесе одан да көп) объектілердің әртүрлі қасиеттерін (ұқсастығын, айырмашылығын, артықшылықтары мен кемшіліктерін) сандық немесе сапалық салыстыру процесі [3]. Салыстыру арқылы объектілердің, құбылыстардың және процестердің сапалық және сандық сипаттамалары анықталады.

Салыстырмалы пікірлер әдеттегі пікірлерге қарағанда әр түрлі семантикалық мағынаға ие, сонымен қатар синтаксистік формалары да әр түрлі. Мысалы, әдеттегі тұрақты пікір сөйлемі «Бұл телефонның дыбыс сапасы керемет» және әдеттегі салыстырмалы пікір сөйлемі «Samsung Galaxy дыбыс сапасы iPhone 14-ке қарағанда жақсырақ». Бұл салыстырмалы сөйлемде екі телефонның дыбыстық сапасы жақсы немесе нашар екендігі айтылмайды, тек екі смартфонның дыбыстық сапасына қатысты салыстырмалы пікірін білдірген [5].

Салыстыру – бағалаудың ең маңызды әдістерінің бірі. Бағалау екі жолмен жүргізілуі мүмкін: тікелей және салыстыру. Тікелей бағалау өнім және олардың ерекшеліктері туралы жалпы оң немесе теріс пікір береді. Ал, салыстыру өнімді және олардың ерекшеліктерін басқа ұқсас өнімдермен және олардың ерекшеліктерімен салыстырады.

Пікірлерді мен сентиментті талдаулар жеткіліксіз, себебі бұл адамдардың белгілі бір өнімдер немесе қызметтерге жалпы қалай қарайтынын көрсетеді. Бұл дұрыс емес пікірлерге әкеледі. Ал салыстырмалы пікірлер салыстыратын өнімдер немесе қызметтер арасындағы ұқсастықтар мен айырмашылықтар туралы нақтырақ ақпаратты ашады [11]. Сол себепті пікірлердегі салыстыру конструкцияларын анықтау пікір талдаудағы маңызды проблемалардың бірі. Мәтінде екі немесе одан да көп объектілерді салыстыру конструкциялары (сөйлемдер) арқылы беріледі. Кейбір деректер көздерінде салыстыру конструкцияларын компаративті конструкциялар немесе салыстыру сөйлемдері деп те көрсетіледі.

Салыстыру конструкциялары – қандай да бір объект, атрибут немесе оқиғаны Р өз алдына емес, қандай да бір Q объектіге, атрибутқа немесе оқиғаға қатысты сипатталатын конструкциялар. Белгілі бір Z атрибуты салыстыру критерийі V мәнін қабылдайды (Р – салыстыру объектісі - салыстырылатын нәрсе, Q – салыстыру стандарты немен салыстырылады, Z – салыстыратын атрибуты - салыстырылатын нәрсе, V – салыстыру атрибутының мәні). [12]

Салыстырмалы сөйлем әдетте бірнеше субъектінің ұқсастығына немесе айырмашылығына негізделген қатынасты білдіреді. Тіл мамандары ұзақ уақыт бойы салыстыруларды зерттеп келеді.

Салыстыру конструкциялары көптеген лингвистика ғалымдарының зерттеу нысаны болған: орыс тілінде Н. А. Широкова [1963], М. И. Черемисина [1971, 1976]; ағылшын тілінде R. Ultan [1972], L. Stassen [1984], M. Haspelmath, O. Buchholz [1998], R. Dixon [2005, 2008]; саха (якут) тілі Ю. Васильев [1986]; алтай тілдері Л. Тыбыкова [1988, 1989]; хақас тілі Э.

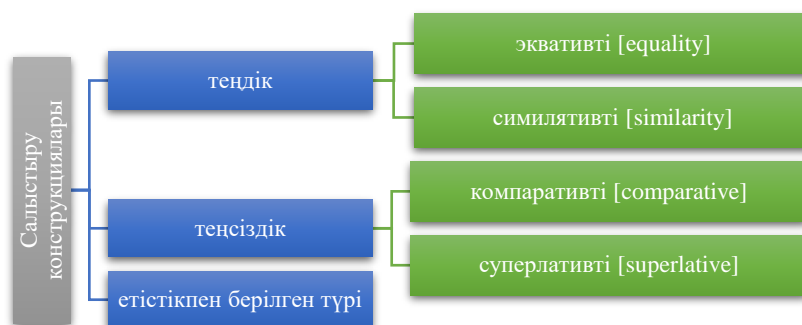
Кыржинакова [2010]; тыва тілінде Л. А. Шамина [2001]; қазақ тілінде Т. Қоңыров [1985], Ф. Ахметжанова [1982], Б. Ескельдиева [2014; 2016], С.Ж. Тәжібаева зерттеген [10].

Қазақ тілінде салыстырмалы конструкцияларды ғалымдардың алғашқысы болып филология ғылымдарының докторы Қоңыров Тұрағұл (1935-1995 жж.) зерттеген.

Т.Қоңыров өз еңбектерінде қазақ тіліндегі салыстыруларды салыстырмалы мағына түрлеріне қарай семантикалық топқа бөледі: теңдік, теңсіздік, күшейту, кеміту, жалған салыстыру, метафоралық салыстыру, растау, теріс (қарама-қайшы) салыстыру [13].

Қазақ тіл білімінде салыстырмалы қатынастарды сипаттауға арналған қазіргі терминология, оның ішінде канондық салыстырмалы конструкциялар термині алғаш рет ғылыми айналымға мақалаларда және ең дәйекті түрде Б.Е. Ескельдиеваның диссертациялық зерттеуінде енгізілді [4].

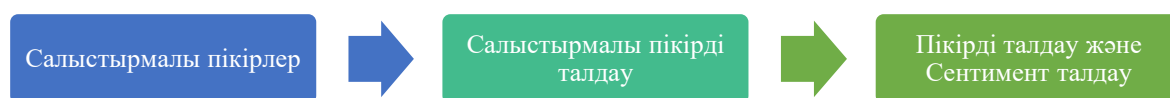
Б.Е. Ескельдиеваның өз еңбегінде салыстыру конструкцияларын негізгі үш топқа бөліп жіктейді: теңдік мағыналы, теңсіздік мағыналы және етістікпен берілген түрі (3 сурет).



3 сурет. Салыстыру конструкциялардың классификациясы

Мәтіннен субъективті ақпаратты анықтау және алу үшін табиғи тілді өңдеуді (Natural Language Processing), мәтінді талдау және компьютерлік лингвистика әдістері пайдаланылады. Салыстырмалы пікірді анықтау – салыстырмалы түрде көрсетілген ақпаратты анықтау және алумен айналысатын пікірді талдаудың ішкі саласы (19-03-008).

Салыстырмалы пікірді анықтау (COM) – бұл бәсекелес өнімдердің артықшылықтары мен кемшіліктерін салыстыруға бағытталған пікірді талдаудың ерекше түрі.

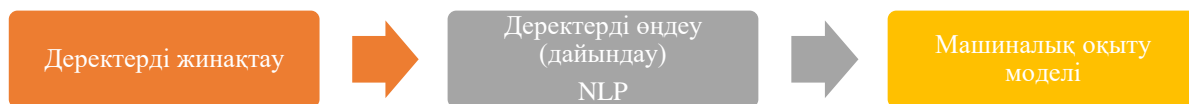


4 сурет. Салыстырмалы пікірді анықтау

Салыстырмалы пікірлер әдеттегі пікірлерден ерекшеленеді және көбінесе салыстырмалы кілттік сөздер, екі салыстырылатын нысан немесе салыстырылатын аспектілер сияқты бірнеше маңызды элементтерден тұрады.

Салыстырмалы конструкциялар бойынша компьютерлік лингвистика саласында бір қатар зерттеулер жүргізілді. Зерттеушілер салыстырмалы конструкцияларды анықтау үшін NLP мен машиналық оқыту әдістерін қолданған.

Салыстыру конструкцияларын анықтау келесі кезеңдерден тұрады: (1) салыстырмалы пікірді анықтауға қатысты үш деректер жиынтығын алу, (2) әртүрлі алдын ала өңдеу қадамдарын қолдану арқылы мәтінді тазалау, (3) әртүрлі машиналық оқыту классификаторлары арқылы салыстырмалы пікірді талдау, (4) пайдаланылған классификаторлардың нәтижелерін салыстыру және (5) тиімді нәтижелермен жіктеуішті ұсыну (4 сурет).



5 сурет. Салыстыру конструкцияларын анықтау

Алғашқылардың бірі болып 2006 жылы Nitin Jindal мен Bing Liu текстік құжаттардағы салыстыру сөйлемдерді анықтаудың [6] мақаласын жариялады. Мақалада тіл білімінің зерттеушілері ең алдымен салыстырмалы конструкциялардың синтаксисі мен семантикасын анықтауға ден қойғанын, ал компьютерлік лингвистика саласында (тексті талдау және деректерді талдау) салыстыру конструкцияларын автоматты анықтауға қатысты зерттеу жұмыстардың болмағанын жазады.

Қазіргі кезде салыстырмалы пікірлерді анықтау бойынша зерттеулер тек ағылшын, қытай және корей тілдерінде зерттелді.

Қазақ тілінде сентименттік талдау бойынша зерттеулер жүргізіліп, нәтижелері [14, 15, 16] мақалаларда жарияланған.

Алайда, компьютерлік лингвистика саласында қазақ тілінің салыстыру конструкциялары бойынша зерттеулер әлі жүргізілмеген.

Қазіргі уақытта әлеуметтік медиа ақпараттың тамаша көзі болып табылады және олардан салыстырмалы пікірлерді алу құнды білім алу үшін өте перспективті бағыт болып көрінеді.

Қолданылған әдебиеттер тізімі

1. Kushal Dave, Steve Lawrence, David M. Pennock. Mining the peanut gallery: Opinion extraction and semantic classification of product reviews. In Proceedings of WWW, p. 519–528, 2003.
2. Nasukawa Tetsuya, Jeonghee Yi. Sentiment analysis: capturing favourability using natural language processing. Proceedings of the KCAP03, 2nd International Conference on knowledge capture, 2003.
3. Кохановский В.П. Философия и методология науки. – Ростов-на-Дону: «Феникс», 1999. – 576 с.
4. Ескельдиева Б.Е. Сравнительные конструкции в казахском и в тюркских языках Сибири. Диссертациялық жұмыс, 2016 жыл, Астана қ.
5. Bing Liu. Sentiment Analysis and Opinion Mining, Morgan & Claypool Publishers, May 2012
6. Nitin Jindal and Bing Liu. 2006. Identifying comparative sentences in text documents. In Proceedings of the 29th annual international ACM SIGIR conference on Research and development in information retrieval (SIGIR '06). Association for Computing Machinery, New York, NY, USA, 244–251. <https://doi.org/10.1145/1148170.1148215>
7. Pang Bob, Lee Lillian. «Opinion Mining and Sentiment Analysis». Foundations and Trends in Information Retrieval 2 , no. 1-2 (2008): 1--135.
8. DIGITAL 2024: GLOBAL OVERVIEW REPORT <https://datareportal.com/reports/digital-2024-global-overview-report>
9. Local Consumer Review Survey 2024: Trends, Behaviors, and Platforms Explored <https://www.brightlocal.com/research/local-consumer-review-survey/>
10. Тажибаева, С. Ж. Исследование сравнительных конструкций в современной казахской лингвистике / Предложение как единица языка и речи : Материалы Всероссийского науч. симпозиума с межд. участием, посвященного 95-летию со дня рождения М.И. Черемисиной, Новосибирск, 08–11 октября 2019 года. – Новосибирск: Федеральное государственное бюджетное учреждение науки Институт филологии Сибирского отделения Российской академии наук, 2019. – С. 53-59.
11. Varathan Kasturi & Giachanou Anastasia & Crestani Fabio. (2016). Comparative Opinion Mining: A Review. Journal of the Association for Information Science and Technology.

12. Ю.П. Князев. Грамматическая семантика: Русский язык в типологической перспективе. М.: Языки славянских культур, 2007.
13. Коныров Т.К. Структурно-семантическая природа сравнения в казахском языке [Текст] : (компаративные конструкции): учеб. пособие / Т. К. Коныров – Алма-Ата : Мектеп, 1985. - 215 с.
14. Yergesh Banu, Bekmanova Gulmira, and Sharipbay Altynbek. ‘Sentiment Analysis of Kazakh Text and Their Polarity’. 1 Jan. 2019 : 9 – 15
15. A. Nugumanova, Y. Baiburin and Y. Alimzhanov, "Sentiment Analysis of Reviews in Kazakh With Transfer Learning Techniques," *2022 International Conference on Smart Information Systems and Technologies (SIST)*, Nur-Sultan, Kazakhstan, 2022, pp. 1-6, doi: 10.1109/SIST54437.2022.9945811.
16. Оралбекова И. Т., Ергеш Б. Ж. Аспектно-ориентированный сентимент анализ отзывов о гостиницах (на материале казахского языка) [Электронный ресурс] // Родные языки и культуры в современном изменяющемся мире. 2022.
17. Abbasi Moghaddam S. Aspect-based opinion mining in online reviews. – 2013.

КАСКАДНЫЙ МАШИННЫЙ ПЕРЕВОД КАЗАХСКО-ТУРЕЦКОЙ РЕЧИ ПО ПОЛНОМУ НАБОРУ ОКОНЧАНИЙ

Толеубеков Жандос

zhandoslp@gmail.com

Магистрант казахского национального университета

имени аль-Фараби, Алматы, Казахстан

Научный руководитель - У.А. Тукеев

Аннотация. Исследования машинного перевода речи в речь для тюркских языков практически отсутствуют из-за трудностей создания параллельных речевых корпусов для обучения нейронных моделей. Поэтому актуальной проблемой является создание синтетического параллельного корпуса для исследования машинного перевода речи тюркских языков. В работе предложена технология формирования параллельных корпусов тюркской речи на основе морфологической модели CSE. Для данной технологии формирования параллельных корпусов тюркской речи используется каскадная схема: речь-текст, текст-текст и текст-речь. Особенностью этой схемы является то, что для фазы «текст-текст» используется реляционная модель перевода, основанная на морфологической модели CSE (Complete Set of Endings).

Ключевые слова: автоматическое формирование, речь, синтетическая, параллельные корпуса, казахско-турецкий язык.

Введение. Большинство тюркских языков относятся к языкам с небольшими электронными ресурсами. Проблема компьютерной обработки естественных языков с небольшими электронными ресурсами является одной из наиболее острых проблем в области нейронного машинного перевода.

Для тюркских языков исследования машинного перевода речи в речь практически отсутствуют из-за трудностей создания параллельных речевых корпусов для обучения нейронных моделей. Создание машинного перевода речи по каскадной схеме (СТТ – ТТТ – ТТС), где S – речь (speech), Т – текст (текст), также сложно из-за отсутствия ТТТ (Text-To-Text) для большинства тюркских языков. Поэтому в данной работе предлагается изучить подход построения машинного перевода речи по каскадной схеме формирования синтетических параллельных речевых корпусов, который может быть использован для изучения нейромашинного перевода речи в речь. Каскадную схему предлагается решать с помощью реляционных моделей для фазы ТТТ на основе морфологической модели на полной системе окончаний (модели CSE).