

Жоба туралы қысқаша ақпарат

Жоба аты	AP15473441 «Қазақ тілі мәтіндер корпусындағы сөзжасамдық белгіленімдер әзірлемесі»
Жоба өзектілігі	<p>Бүгінгі таңда өзекті болып отырған машиналық аударма, тілдік элементтерді автоматты тану және сөйлеу тілін синтездеу, орфография мен грамматиканы тексеру құралдары, заманауи жиілік сөздіктер т.б. ұлттық корпустардан алынады.</p> <p>Қазақ тілінің ұлттық корпусының қызметін күшейту үшін оған енгізілетін лингвистикалық белгіленім түрлерін көбейту, оның лингвистикалық базасын дайындау, программалық мүмкіндіктерін қарастыру – өзекті мәселе.</p>
Жоба мақсаты	Қазақ тіліндегі сөзжасамдық белгіленімдерді корпуста енгізудің тілдік базасын жасау және программалық әзірлемесін дайындау.
Жоба міндеттері	<ul style="list-style-type: none"> - корпустағы белгіленім ұғымына, белгіленім түрлеріне қазақ тіліндегі «туынды сөз» ұғымына және сөзжасамдық тәсілдерге сипаттама беру; - қосымшалардың сөзжасамға және формажасамға қатысын айқындау; - түрік тілінің ұлттық корпусындағы сөзжасамдық белгіленімдер енгізу әзірлемесіне талдау жасау; - орыс тілінің ұлттық корпусындағы сөзжасамдық белгіленімдер енгізу әзірлемесіне талдау жасау; - сөздерді корпуста автоматты танудың алгоритмін түзуге негіз болатын лингвистикалық нұсқаулықтар ретінде қос сөздер тізімін, біріккен сөздер тізімі мен сөздігін, әр сөз табы бойынша сөзжасамдық қосымшалар кестесін жасау; - сөзжасамдық белгіленімдер болып табылатын қос сөздер, біріккен сөздер және қосымшалы туынды сөздерді корпуста автоматты тану әзірлемесін дайындау.
Күтілетін және қол жеткізілген нәтижелер	<ul style="list-style-type: none"> - корпустағы белгіленім ұғымына, белгіленім түрлеріне қазақ тіліндегі «туынды сөз» ұғымына және сөзжасамдық тәсілдерге сипаттама беріледі; - қосымшалардың сөзжасамға және формажасамға қатысы айқындалады; - түрік тілінің ұлттық корпусындағы сөзжасамдық белгіленімдер енгізу әзірлемесіне талдау жасалады; - орыс тілінің ұлттық корпусындағы сөзжасамдық белгіленімдер енгізу әзірлемесіне талдау жасалады; - сөздерді корпуста автоматты танудың алгоритмін түзуге негіз болатын лингвистикалық нұсқаулықтар ретінде қос сөздер тізімі, біріккен сөздер тізімі мен сөздігі, әр сөз табы бойынша сөзжасамдық қосымшалар кестесі жасалады; - сөзжасамдық белгіленімдер болып табылатын қос сөздер, біріккен сөздер және қосымшалы туынды сөздерді корпуста автоматты тану әзірлемесі дайындалады. - жарияланымдарды жүзеге асыру: зерттеу нәтижелері Жоғары Аттестациялау Комиссиясы ұсынған отандық ғылыми басылымдарда және шетелдік ғылыми журналдар мен конференцияларда мақала түрінде жарыққа шығады:

	<p>- Web of Science деректер базасында Arts and Humanities Citation Index немесе Social Science Citation Index-те индекстелетін немесе Scopus деректер базасында CiteScore бойынша кемінде 35 (отыз бес) процентилі бар ғылыми басылымдарда кемінде 2 (екі) мақала</p> <p>Немесе:</p> <p>- рецензиялаудан өткен және кемінде екі профессор ғылыми атағының және/немесе ғылым докторы ғылыми дәрежесінің иегерлері жариялауға ұсынған постдокторанттың үлесі 6 б. п. кем емес 1 (бір) кем емес монография;</p> <p>- рецензияланған шетелдік және (немесе) БҒСҚК ұсынған отандық басылымдарда 5 (бес) мақаладан және (немесе) шолулардан кем емес.</p>
<p>Зерттеу тобы мүшелерінің аты-жөні, идентификаторлары (Scopus Author ID, Researcher ID, ORCID, бар болса) және сәйкес профильдерге сілтемелер</p>	<p>Пірманова Күнсұлу Қамбарбекқызы, Постдокторант, https://orcid.org/0000-0003-3783-3199</p>
<p>Жарияланымдар тізімі (URL, DOI көрсетілген)</p>	<ol style="list-style-type: none"> 1. <u>Пірманова К.Қ., Ақар А., Мырзаханова Г.К.</u> Ways of transferring partially word-forming affixes in the national corpus of the kazakh language. TILTANYM № 1 (89) 2023 https://www.tiltanym.kz/jour/article/view/799/493 2. <u>Пірманова К.Қ., Тореханова А., Ашимова Н.</u> Theory and methodology of the world's national linguistic corpora. Абылай хан атындағы Қазақ халықаралық қатынастар және әлем тілдері университеті ХАБАРШЫСЫ № 2 (69), 2023 https://bulletin-philology.ablaikhan.kz/index.php/j1/article/view/713 3. <u>Пірманова К.Қ., Ақар А., Шаймерденова Н.Ж., Шалабаева И.М.</u> Methods of transmission full word-formation affixes in the national corpus of the kazakh language. TILTANYM № 3, 2023 https://www.tiltanym.kz/jour/article/view/1151 4. <u>Пірманова К.Қ., Карбозова Б.</u> О программном обеспечений базы данных электронного корпуса казахских текстов. Publisher.agency: Proceedings of the 1st International Scientific Conference «Foundations and Trends in Research» (December 1-2, 2022). Copenhagen, Denmark, 2022. 233p. ISBN 978-2-684-08428-6. DOI 10.5281/zenodo.7396104 https://ojs.publisher.agency/index.php/FTR/article/view/420 5. <u>Пірманова К.Қ., Заурбекова Г., Тоқмырзаев Д.О.</u> Derivatives and word formation methods (structural aspect of word formation designations). POLISH JOURNAL OF SCIENCE. №67 (2023), ISSN 3353-2389 6. <u>Пірманова К.Қ., Заурбекова Г., Аязбекова Қ.</u> The concept of mark and types of marks in the corpus. The scientific heritage (Budapest, Hungary). ISSN 9215 — 0365 № 123 (123) (2023)

	7. <u>Пирманова К.Қ.</u> , Татиева С., Казезова А. Разработка программы словообразовательных разметок в корпусе текстов казахского языка. Znanstvena misel journal The journal is registered and published in Slovenia. ISSN 3124-1123, №83/2023
Патент туралы ақпарат	-