

Б.Ж. Ергеш^{1*}, Г.Т. Бекманова¹, Г.О. Сыздыкова¹, А.А. Жумагулова²

¹Л.Н. Гумилев атындағы Еуразия ұлттық университеті, Астана, Қазақстан

²Х.Ж. Садвакасов атындағы Елордалық мектебі, Астана, Қазақстан

E-mail: *b.yergesh@gmail.com

Ахмет Байтұрсынұлының еңбектері бойынша дыбыс жүйесін онтологиялық модельдеу

Аңдатпа. Қазіргі кезде цифрлық форматтағы этносаралық қарым-қатынас тілі ретінде қазақ тілінің қолданыс аясын арттыру және қызметін кеңейту бойынша ғылыми-лингвистикалық негіздер мен IT-ресурстарды әзірлеу өзекті мәселе болып табылады. Сол мәселені шешу үшін орындалып жатқан №BR11765535 грант аясында Ахмет Байтұрсынұлының ғылыми мұрасы және оның еңбектері бойынша барлық құрылымдық қабаттары бар жаңа білім жүйесінен құралған цифрланған материалдарды қосатын «Ахметтану» интеллектуалдық жүйесі әзірленіп жатыр. Мақалада аталған интеллектуалдық жүйеге негіз болатын Ахмет Байтұрсынұлының еңбектері бойынша тілдің құрылымдық деңгейлерінің білімдер базасы негізінде, оның ішінде дыбыс жүйесі мысалында әзірленген онтологиялық модельдеу сипатталған. Ахмет Байтұрсынұлының еңбектері бойынша тілдің дыбыстық жүйесін қолданбалы онтологиялық модельдеу үшін Protégé редакторы қолданылды. Зерттеу нысанына алынған еңбектердегі дыбыс жүйесінің онтологиялық моделі класстардан (түсініктер), олардың қасиеттері мен жеке үлгі даналардан (индивид) тұрады. Табиғи тілді өңдейтін әр түрлі интеллектуалды ақпараттық жүйелер үшін қазақ тілінің ғылыми-лингвистикалық негіздері бөлігінде жаңа нәтижелер Қазақстанда ақпараттық-коммуникациялық технологиялар және қазақ тілінің цифрлану саласының дамуына тікелей әсер етеді.

Түйін сөздер: қазақ тілі, А.Байтұрсынұлының ғылыми мұрасы, тілдің құрылымдық деңгейлері, дыбыс жүйесі, интеллектуалдық жүйе, компьютерлік лингвистика, онтологиялық модель.

DOI: doi.org/10.32523/2616-7263-2023-144-3-52-59

1. Кіріспе

Ақпараттандырудың жаһандану үдерісіне сай қазақ тілін ғылым тіліне айналдыру мүмкіндігін кеңейту қолданбалы тіл біліміндегі жаңа бағыт – компьютерлік лингвистиканың маңызды мәселелерінің бірі саналады. Бұл бағытта қазақ тіл білімінің негізін салушы, ұлт руханиятының көсемі А.Байтұрсынұлының ғылыми мұрасын цифрландыру, соның негізінде «Ахметтану» интеллектуалдық жүйесін құрудың да маңызы зор.

«Ахметтану» интеллектуалдық жүйесінің негізін ұлт ұстазының тілдің құрылымдық деңгейлеріне қатысты ұғымдары, ғылыми тұжырымдар құрайды.

Бұл мақалада қазақ тілінің дыбыс жүйесін Ахмет Байтұрсынұлының «Тіл-құрал» еңбегі бойынша онтологиялық модельдеу мақсаты көзделеді. А.Байтұрсынұлының дыбыс жүйесі бойынша ғылыми тұжырымдарын интеллектуалдық жүйеге енгізу үшін алдымен жүйеге қажет материалдарды жинақтау, дыбыс жүйесі бойынша білім базасын құру үшін жиналған материалдарды талдау, талданған материалдарды концептуалды сызбаға түсіру, жүйелеу және модельдеу міндеттері айқындалды.

ҒЗЖ аясында А. Байтұрсынұлының “Тіл-құрал” еңбегіндегі дыбыс жүйесіне қатысты терминдері мен олардың анықтамалары, дыбыс түрлері, дауысты және дауыссыз дыбыстардың бөлінісі, сипаты, буын және тасымалдың базасы жасалды.

Нәтижесінде базаға дыбыс жүйесі бойынша төмендегідей мәліметтер енгіздік:

- дыбыс, әріп, буын, тасымал ұғымдарының сипаты;
- дауысты, дауыссыз және жарты дауыстылар, олардың анықтамасы, мысалдары;
- дауысты дыбыстар (толық, аңғал, қымқырулы), шала (ауыз шығысты: босаң, қысаң; мұрын шығысты: босаң, қысаң, тұйық) дауыстылардың түрлері;
- дауыссыз дыбыстар (қатаң, ұяң, ымыралы, ымырасыз);
- буын және тасымал, түрі, анықтамасы, сипаты.

Осы білімдер базасының негізінде дыбыс жүйесінің А.Байтұрсынұлы еңбектері бойынша онтологиялық моделі әзірленді.

2. Әдістер

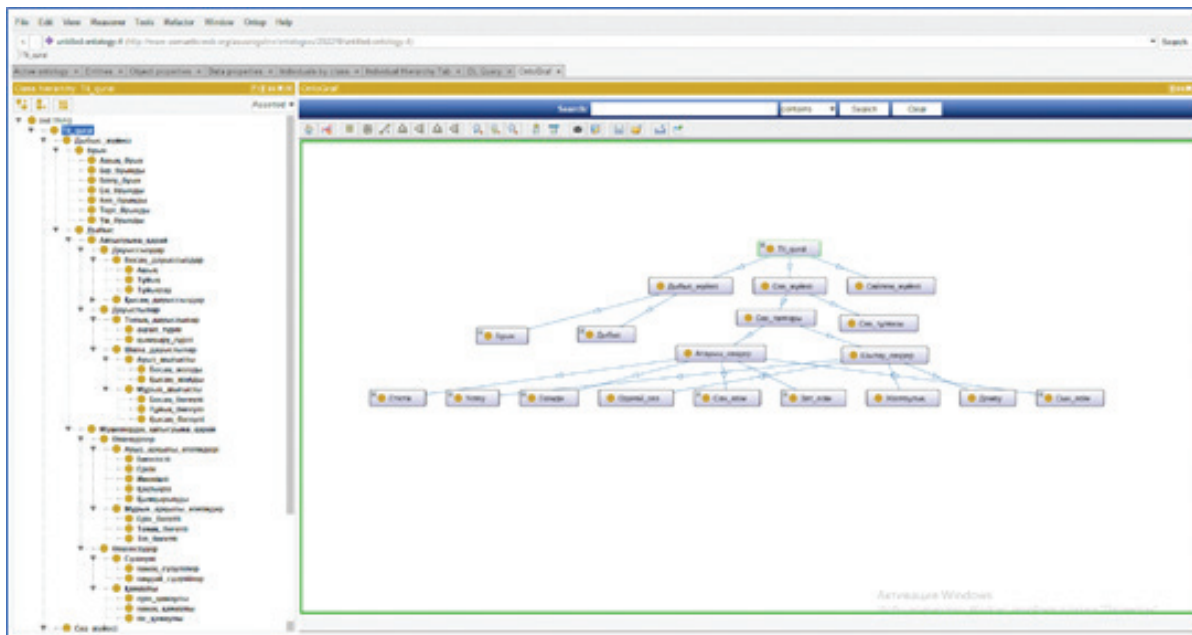
Модельдеу әртүрлі пәндік салалар үшін интеллектуалдық жүйелерді зерттеудің, әзірлеудің ең тиімді құралы болып табылады. Интеллектуалдық жүйелерді құру тәсілдерінің бірі – онтологиялық модельдеу.

Онтология дегеніміз – бұл ұғымдар жиынтығынан және осы ұғымдар туралы тұжырымдар жиынтығынан тұратын концептуалды сызба. Оның негізінде класстарды, қатынастарды, қасиеттерді, функцияларды және индивидтерді (даналарды) сипаттауға болады [1]. Бұл – белгілі бір білім саласын формалды түрде ұсынудың бір түрі. Қазіргі уақытта онтология программалауда, оқытуда, әртүрлі зерттеулерде кеңінен қолданылады. [2-6] Жұмыстарында қазіргі заманғы қазақ тілін математикалық модельдеуде онтологияны сәтті қолданылғанына көз жеткізуге болады.

Онтология тілдің құрылымдық деңгейлерін Ахмет Байтұрсынұлының еңбектері бойынша модельдеу үшін де қолданылды. Бұл жағдайда онтологиялық модельдеу есептің мақсатына сәйкес құрастырылды. Тілдің құрылымдық деңгейлерінің А.Байтұрсынұлы еңбектері бойынша қолданбалы онтологиялық моделі Protégé ортасында құрылды. Себебі бұл орта түсініктерді, нақты нысандарды сипаттауға мүмкіндік береді. Сонымен қатар онда көптеген операторлар жиынтығы бар. Мұндай онтологияны құру логикалық нәтижеге қол жеткізуге мүмкіндік береді.

Ахмет Байтұрсынұлының еңбектері бойынша тілдің құрылымдық деңгейлерінің онтологиялық моделі класстардан (түсініктер), олардың қасиеттері мен бөлек үлгі нұсқалардан (индивидтер) тұрады. Онтологияны түзу барысында ең бірінші класстар құрылады.

Ахмет Байтұрсынұлының “Тіл-құрал” еңбегі негізінде құрылған класстар «Class hierarchy» терезесінде келтірілді (1-сурет).



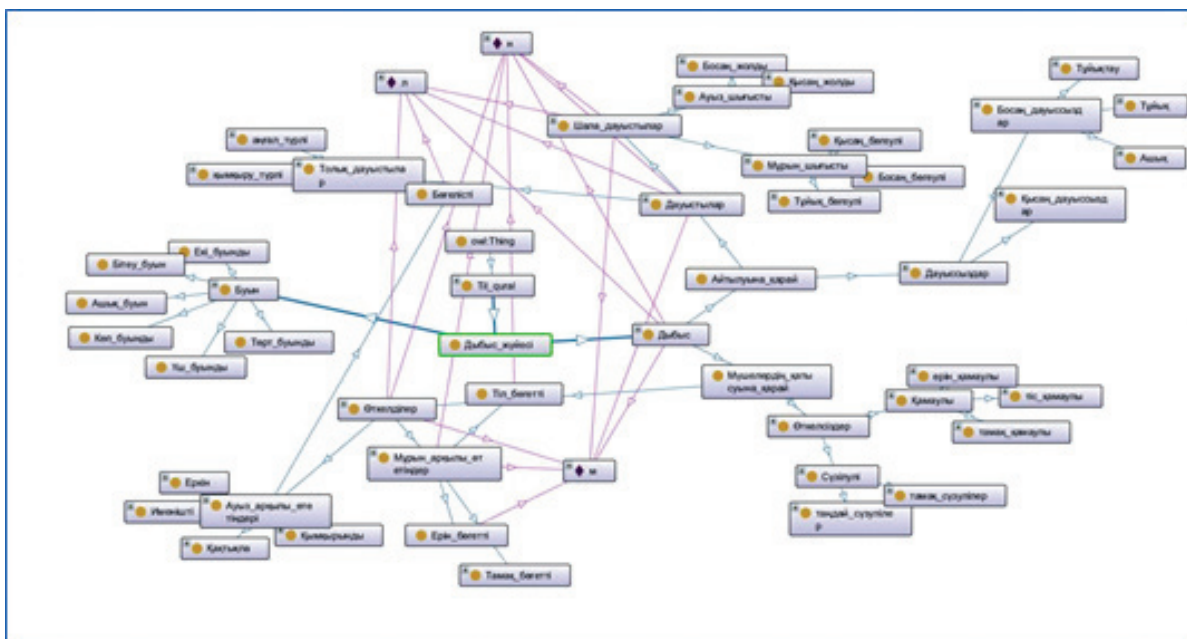
Сурет 1. Ахмет Байтұрсынұлының “Тіл-құрал” еңбегі негізінде құралған онтологиялық модель

Класс дегеніміз – бірдей белгілерді (атрибуторлар), операцияларды, қатынастарды және семантиканы бөлетін нысандардың сипаттамасы. Ал кез келген нысан класстың үлгі нұсқасы болып табылады.

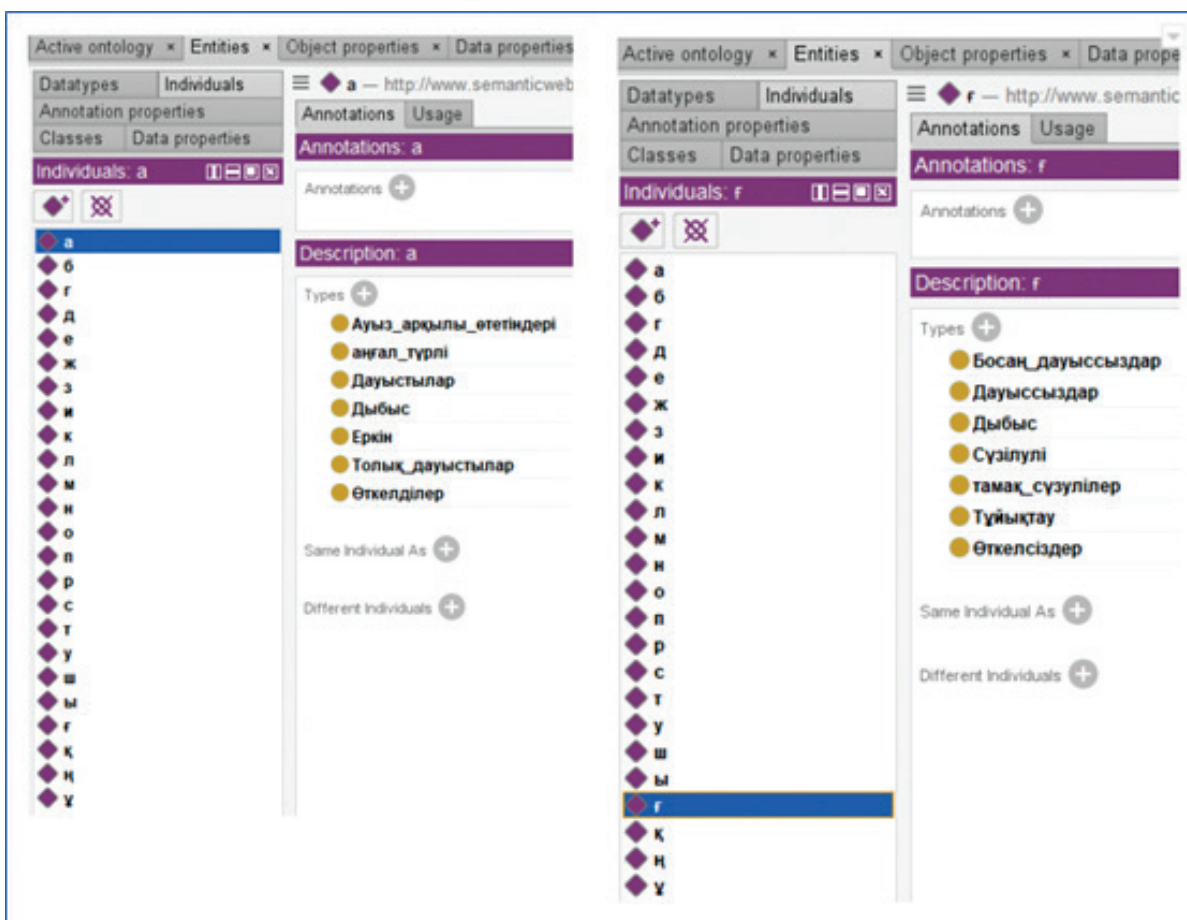
Мұнда THING класы – пәндік саланың барлық нысанын қамтитын жиынтық класс. Пәндік саланың класстары осы класстың ішкі класы болады.

А.Байтұрсынұлы ұсынған дыбыс жүйесіне қатысты терминдер мен олардың анықтамасы, дыбыс түрлері, дауысты және дауыссыз дыбыстардың бөлінісі, сипаты, буын жүйесі базасының онтологиялық моделі ғалымның 1927 жылы Қызылорда қаласында жарық көрген «Дыбыстарды жіктеу туралы» мақаласы [6] негізінде құрастырылды.

А.Байтұрсынұлы ұсынған дыбыс жүйесінің онтологиялық моделінен үзінді 2,3-суреттерде көрсетілген.



Сурет 2. А.Байтұрсынұлы ұсынған дыбыс жүйесінің онтологиялық моделінен үзінді



Сурет 3. А.Байтұрсынұлы ұсынған дыбыстардың сипаттамасынан үзінді

3. Нәтижелер және талқылау

Мақалада А. Байтұрсынұлының “Тіл-құрал” еңбегінде ұсынған дыбыс жүйесіне қатысты терминдері мен олардың анықтамалары, дыбыс түрлері, дауысты және дауыссыз дыбыстардың бөлінісі, сипаты, буын және тасымал жүйесінің базасы және жиналған білім базасының негізінде әзірленген онтологиялық моделдері сипатталды.

А.Байтұрсынұлы ұсынған дыбыс жүйесі бойынша құрылған онтологиялық модель арқылы Protégé редакторында Sparql сұраныстар тілін қолданып сұрақтарға жауап алу мүмкіндігі де бар (4-сурет).

SPARQL query

```

PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX owl: <http://www.w3.org/2002/07/owl#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
SELECT ?subject ?class
WHERE {
?subject rdfs:subClassOf ?class.
?class rdfs:label "Дауыссыз дыбыстар"@kz
    
```

subject	class
Босаң_дауыссыздар	Дауыссыздар
Қысаң_дауыссыздар	Дауыссыздар

Сурет 4. Sparql сұраныстар мысалы

Қорыта келгенде, А.Байтұрсынұлының ғылыми мұрасын цифрландыру мақсатында А.Байтұрсынұлы ұсынған дыбыс жүйесіне қатысты терминдер мен олардың анықтамалары, дыбыс түрлері, дауысты және дауыссыз дыбыстардың бөлінісі, сипаты, буын және тасымал жүйесінің базасы жасалып, онтологиялық моделдері әзірленді. Базаға енгізілген деректер А.Байтұрсынұлының «Тіл тағылымы» [8], «Тіл-құрал» [7] еңбектерінен алынды.

Құрылған онтологиялық модель қолданушыға Ахмет Байтұрсынұлының еңбектеріндегі дыбыс жүйесі туралы мол ақпарат алу мүмкіндігін береді. Ғалымның дыбыс жүйесі туралы ғылыми тұжырымдарын, қалыптастырған ұғымдары мен терминдерін жинақтау, жүйелеу, талдау нәтижесінде әзірленген білім базасы, онтологиялық модельдер «Ахметтану» интеллектуалды жүйесінің маңызды құрамдас бір бөлігі саналады.

А. Байтұрсынұлының ғылыми мұрасын цифрлық жүйеде қалың көпшілікке қолжетімді ету – аса қажетті жұмыстардың бірі. Бұған дейін жасаңды интеллект, компьютерлік лингвистика салаларында Ахмет Байтұрсынұлының еңбектері бойынша тілдің құрылымдық деңгейлерінің білімдер базасы әзірленіп, модельденген жоқ. Алынған нәтижелер ағарту және тіл үйрену, интеллектуалды оқыту, білімді бағалау жүйелерінде, тілге байланысты қосымшаларда, электрондық грамматикалық анықтамалықтар әзірлеуде, көптілді электрондық тезаурустарда, мультимедиялық сұрақ-жауап жүйелерінде, кол-орталықтарда, дауыстық көмекшілерде, смарт үйлерде, және т.б. қолданыс тұрғысынан нақты практикалық маңыздылыққа ие.

4. Қорытынды

А. Байтұрсынұлының еңбектеріндегі дыбыс жүйесі бойынша цифрланған білімдер базасын әзірлеу, сондай-ақ онтологиялық модельдер құру негізінде келесі нәтижелер алынды:

– Ахмет Байтұрсынұлының еңбектеріндегі тілдің дыбыс жүйесі бойынша цифрландырылған материалдар;

- Ахмет Байтұрсынұлының мұрасы бойынша тілдің дыбыс жүйесінің білім базасы;
- Ахмет Байтұрсынұлы ұсынған дыбыс жүйесінің онтологиялық моделі.

Осыған дейін ашық қолданыста бұл салада Ахмет Байтұрсынұлының мұралары бойынша тілдің құрылымдық деңгейлері, оның ішінде дыбыс жүйесінің білімдер базасы әзірленіп, модельденген жоқ. Бұл жұмыс орындалып жатқан бағдарлама аясында алғаш рет жасалып, жарияланып отыр.

Зерттеудің алдағы маңызды міндеттерінің бірі Ахмет Байтұрсынұлының мұрасы бойынша тілдің барлық құрылымдық қабаттарына қатысты білімнің онтологиялық моделін құру, цифрланған жаңа мәліметтер қосу және түзету функциясы бар қазақ тілінде қолданушылық интерфейстен тұратын «Ахметтану» интеллектуалдық жүйесін әзірлеу болып табылады.

Растау

Бұл зерттеуді Қазақстан Республикасы Білім және ғылым министрлігінің Ғылым комитеті қаржыландырады (№BR11765535 грант).

Әдебиеттер тізімі

- 1 Gruber T. R. A Translation Approach to Portable Ontology Specifications // Knowledge Acquisition. – 1993. – Vol.5(2). – С. 199-220.
- 2 Sharipbay A., Yergesh B., Razakhova B., Yelibayeva G., Mukanova A. Syntax parsing model of Kazakh simple sentences // ACM International Conference Proceeding Series. - 2019. - art. no. a54, doi: 10.1145/3368691.3368745.
- 3 Yelibayeva G., Mukanova A., Sharipbay A., Zulkhazhav A., Yergesh B., Bekmanova G. Metalanguage and Knowledgebase for Kazakh Morphology // Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). – 2019. – Vol. 11619 LNCS. - P. 693 – 706, doi: 10.1007/978-3-030-24289-3_51.
- 4 Mukanova A., Yergesh B., Yelibayeva G., Bekmanova G. Applying the Ontological Approach to Electronic Guide Development //2022 International Conference on Engineering & MIS (ICEMIS). - Istanbul, Turkey, 2022. – P. 1-4, doi: 10.1109/ICEMIS56295.2022.9914299.
- 5 Yelibayeva G., Yergesh B., Bekmanova G., Razakhova B., Sharipbay A., Mukanova A. Modelling of Verb Phrases of the Kazakh Language // 2022 International Conference on Engineering & MIS (ICEMIS). – 2022. - P. 1-3, doi: 10.1109/ICEMIS56295.2022.9914015.
- 6 Omarbekova A., Sharipbay A., Barlybaev A. Generation of Test Questions from RDF Files Using PYTHON and SPARQL // Journal of Physics: Conference Series. - 2017. Vol. 806 (1). - art. no. 012009, doi: 10.1088/1742-6596/806/1/012009.
- 7 Шәріпбай А.Ә., Разахова Б.Ш., Муканова А.С., Ергеш Б.Ж. Қазақ тілі грамматикасының математикалық және онтологиялық моделдері мен электрондық тезаурусы. Монография. - Нұр-Сұлтан: «Бұлатов А.Ж.» ЖК. – 2020. – 218 с.
- 8 Шарипбай А.А., Муканова А.С., Ергеш Б.Ж., Разахова Б.Ш., Елибаева Г.К. Разработка онтологической модели грамматики казахского языка // Вычислительная обработка казахского языка: сборник научных трудов / под редакцией Рахимовой Д.Р. – Алматы: Қазақ университеті, 2020. – Глава 1. – С.6-32.
- 9 Байтұрсынұлы А. Тіл – құрал. – Алматы: Сардар. – 2009. - 348 б.
- 10 Байтұрсынұлы А. Тіл тағылымы. – Алматы: Ана тілі. – 1992. - 448 б.

Онтологическое моделирование звуковой системы по наследию Ахмета Байтурсынулы

Б.Ж. Ергеш^{1*}, Г.Т. Бекманова¹, Г.О. Сыздыкова¹, А.А. Жумагулова²

¹Евразийский национальный университет имени Л.Н. Гумилева, Астана, Казахстан

²Столичная Школа имени Х.Ж. Садвакасова, Астана, Казахстан

Аннотация. В настоящее время разработка научно-лингвистических основ и IT-ресурсов по расширению деятельности и использования казахского языка как языка межэтнического общения в цифровом формате является актуальной задачей. Для решения этой задачи в рамках гранта №BR11765535 разрабатывается интеллектуальная система «Ахметтану», которая включает базу данных по научному наследию Ахмета Байтурсынулы и новые системы знаний по всем структурным ярусам языка в свете его учений. В статье описываются онтологические модели, разработанные на основе базы знаний структурных уровней казахского языка на примере звуковой системы, по учению Ахмета Байтурсынулы. Данные модели являются основой для разработки интеллектуальной системы. Для прикладного онтологического моделирования звуковой системы языка по наследию Ахмета Байтурсынулы был использован редактор Protégé. Онтологическая модель звуковой системы в работе состоит из классов (понятий), их свойств и отдельных экземпляров (индивидов). Новые результаты научно-лингвистических основ казахского языка для различных интеллектуальных систем, обрабатывающих естественный язык, оказывают непосредственное влияние на развитие информационно-коммуникационных технологий и цифровизацию казахского языка в Казахстане.

Ключевые слова: казахский язык, научное наследие А. Байтурсынулы, структурные уровни языка, звуковая система, интеллектуальная система, компьютерная лингвистика, онтологическая модель.

Ontological modeling of the sound system based on the heritage of Akhmet Baitursynuly

B. Yergesh¹, G. Bekmanova¹, G. Syzdykova¹, A. Zhumagulova²

¹L.N. Gumilyov Eurasian National University, Astana, Kazakhstan

²Kh. Sadvakasov Capital School, Astana, Kazakhstan

Abstract. At present, the development of scientific and linguistic foundations and IT resources to expand the activities and use of the Kazakh language as a language of interethnic communication in digital format is an actual task. To solve this problem, within the framework of grant No. BR11765535, the "Akhmettanu" intellectual system is being developed, which includes a database on the scientific heritage of Akhmet Baitursynuly and new knowledge systems for all structural layers of the language in the light of his teachings. The article describes ontological models developed on the basis of the knowledge base of the structural levels of the Kazakh language on the example of the sound system, according to the teachings of Akhmet Baitursynuly. These models are the basis for the development of an intelligent system. Protégé editor was used for applied ontological modeling of the language sound system following the heritage of Ahmet Baitursynuly. The ontological model of the sound system in the work consists of classes (concepts), their properties and instances (individuals). New results of the scientific and linguistic foundations of the Kazakh language for various intelligent systems that process natural language have a direct impact on the development of information and communication technologies and the digitalization of the Kazakh language in Kazakhstan.

Keywords: Kazakh language, scientific heritage of A. Baitursynuly, structural levels of language, sound system, intelligent system, computational linguistics, ontological model.

References

- 1 Gruber T. R. A Translation Approach to Portable Ontology Specifications, Knowledge Acquisition, 5(2), 199-220 (1993).
- 2 Sharipbay A., Yergesh B., Razakhova B., Yelibayeva G., Mukanova A. Syntax parsing model of Kazakh simple sentences, ACM International Conference Proceeding Series, art. no. a54 (2019), doi: 10.1145/3368691.3368745.

- 3 Yelibayeva G., Mukanova A., Sharipbay A., Zulkhazhav A., Yergesh B., Bekmanova G. Metalanguage and Knowledgebase for Kazakh Morphology, Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 11619 LNCS, 693 – 706 (2019), doi: 10.1007/978-3-030-24289-3_51.
- 4 Mukanova A., Yergesh B., Yelibayeva G., Bekmanova G. Applying the Ontological Approach to Electronic Guide Development, 2022 International Conference on Engineering & MIS (ICEMIS), 1-4, (2022), doi: 10.1109/ICEMIS56295.2022.9914299.
- 5 Yelibayeva G., Yergesh B., Bekmanova G., Razakhova B., Sharipbay A., Mukanova A. Modelling of Verb Phrases of the Kazakh Language // 2022 International Conference on Engineering & MIS (ICEMIS), 1-3 (2022), doi: 10.1109/ICEMIS56295.2022.9914015.
- 6 Omarbekova A., Sharipbay A., Barlybaev A. Generation of Test Questions from RDF Files Using PYTHON and SPARQL, Journal of Physics: Conference Series, 806 (1), art. no. 012009 (2017), doi: 10.1088/1742-6596/806/1/012009.
- 7 Sharipbay A.A., Razakhova B. Sh., Mukanova A. S., Yergesh B. Zh. Qazaq tili grammatikasynyn matematikalyq zhane ontologiyalyq modelder men elektrondyq tezaurusy [Mathematical and ontological models and electronic thesaurus of the Kazakh language grammar], Monograph (IP “Bulatov A. ZH.”, Nur-Sultan, 2020, 218 p.) [in Kazakh].
- 8 Sharipbay A. A., Mukanova A. S., Yergesh B. Zh., Razakhova B. Sh., Yelibayeva G. K. Razrabotka ontologicheskoi modeli grammatiki kazahskogo yazyka [Development of ontological models of grammar of the Kazakh language], Vychislitel'naya obrabotka kazahskogo yazyka: sbornik nauchnykh trudov pod redakciei Rahimovoi D.R. [The computational processing of the Kazakh language: scientific works / under the editorial board Rakhimova D. R.] (Kazakh University, Almaty, 2020, p. 6-32) [in Russian].
- 9 Baitursynuly A. Language is a tool [Til-qural] (Sardar, Almaty, 2009, 348 p) [in Kazakh].
- 10 Baitursynuly A. Language lessons [Til tagylymy]. (Ana tili, Almaty, 1992, 448 p) [in Kazakh].

Авторлар туралы мәлімет:

Б.Ж. Ергеш – PhD, Электрондық оқытуды және құжат айналымын сүйемелдеу секторының менеджері, Л.Н.Гумилев атындағы Еуразия ұлттық университеті, Сәтпаев көш., 2, Астана, Қазақстан.

Г.Т. Бекманова – PhD, т.ғ.к., қауымдастырылған профессор, цифрлық даму және қашықтықтан оқыту департаментінің директоры, Л.Н.Гумилев атындағы Еуразия ұлттық университеті, Сәтпаев көш., 2, Астана, Қазақстан.

Г.О. Сыздықова – ф.ғ.д., қазақ тіл білімі кафедрасының профессоры, Л.Н.Гумилев атындағы Еуразия ұлттық университеті, Сәтпаев көш., 2, Астана, Қазақстан.

А.А. Жумагулова – PhD, қазақ тілі және әдебиет мұғалімі, Х.Ж. Садвакасов атындағы Елордалық мектебі, Ақын Сара көш., 21, Астана, Қазақстан.

В. Yergesh – PhD, e-learning support and document management sector manager, L.N. Gumilyov Eurasian National University, 2 Satpayev str., Astana, Kazakhstan.

G. Bekmanova – PhD, Candidate of Technical Sciences, Associate Professor, Director of Digital Development and Distance Learning Department, L.N. Gumilyov Eurasian National University, 2 Satpayev str., Astana, Kazakhstan.

G. Syzdykova – Doctor of Philology, Professor, L.N. Gumilyov Eurasian National University, 2 Satpayev str., Astana, Kazakhstan.

A. Zhumagulova – PhD, Kazakh Language and Literature Teacher, Kh. Sadvakasov Capital School, 21 Akyn Sara st., Astana, Kazakhstan.

Б.Ж. Ергеш – PhD, менеджер сектора сопровождения электронного обучения и документооборота, Евразийский национальный университет имени Л.Н. Гумилева, ул. Сатпаева, 2, Астана, Казахстан.

Г.Т. Бекманова – PhD, к.т.н., ассоц. профессор, Директор департамента цифрового развития и дистанционного обучения, Евразийский национальный университет имени Л.Н. Гумилева, ул. Сатпаева, 2, Астана, Казахстан.

Г.О. Сыздықова – д.ф.н., профессор кафедры казахского языкознания, Евразийский национальный университет имени Л.Н. Гумилева, ул. Сатпаева, 2, Астана, Казахстан.

А.А. Жумагулова – PhD, учитель казахского языка и литературы, Столичная Школа имени Х.Ж. Садвакасова, ул. Ақын Сара, 21, Астана, Казахстан.