

**ҚОЛДАНБАЛЫ ТІЛ БІЛІМІ
ПРИКЛАДНОЕ ЯЗЫКОЗНАНИЕ
APPLIED LINGUISTICS**

Мақала/Статья/Article
FTAMP 16.31.51

<https://doi.org/10.55491/2411-6076-2025-4-93-111>

Нұрсауле Максұтқызы Рсалиева¹, Талшын Қанатқызы Чукаева^{2*}

¹филология ғылымдарының кандидаты, А. Байтұрсынұлы атындағы Тіл білімі институты, Қазақстан, Алматы қ., ORCID: 0000-0003-2065-0505 E-mail: nursauler@mail.ru

^{2*} автор-корреспондент, докторант, А. Байтұрсынұлы атындағы Тіл білімі институты, Қазақстан, Алматы қ., ORCID: 0000-0002-0716-8370 E-mail: greenrose_98@mail.ru

**ҚАЗАҚ ТІЛІН ОҚЫТУДЫҢ КОРПУСТЫҚ НЕГІЗДЕРІ:
АНГЛОФОНДАРҒА БЕЙІМДЕЛГЕН ТӘЖІРИБЕ**

Аңдатпа. Мақалада англофондарға қазақ тілін оқытудың интерактивті корпус базасын – Learner Corpus әзірлеу барысы жан-жақты сипатталады. Зерттеу шетелдік және отандық корпусық лингвистика тәжірибесіне, ағылшынтілді үйренушілерге арналған қолданыстағы оқулықтар мен оқыту материалдарының сапасына салыстырмалы талдау жүргізуден басталады. Талдау нәтижесінде қазіргі оқу құралдарының едәуір бөлігі ағылшынтілді иофондардың когнитивтік, лингвистикалық және коммуникативтік қажеттіліктеріне толық бейімделмегені анықталды: мәтіндердің деңгейге сәйкес келмеуі, жасандылығы, грамматикалық және логикалық қателері, мәдени ақпараттың дәл еместігі сияқты жүйелі кемшіліктер байқалды. Осы негізде Learner Corpus базасына арналған мәтіндерді іріктеу критерийлері айқындалып, 250-ден астам сюжеттік және коммуникативтік мазмұндағы бірегей мәтіндер, лексика-грамматикалық білімдер жүйесі, салғастырмалы грамматикалық материалдар әзірленді. Корпус мазмұнын құрастыру Стивен Крашеннің «i+1» қағидасына, тіл меңгерудің табиғи реттілігіне, қазақ және ағылшын тілдерінің типологиялық айырмашылығына сүйене отырып жүзеге асты. Англофондар жиі жіберетін интерференциялық қателер (фонетикалық, лексикалық, грамматикалық және прагматикалық деңгейлерде) анықталып, олардың корпусқа енгізілетін материалдар арқылы қалай түзетілетіні көрсетілді. Зерттеуде салғастырмалы талдау, теориялық шолу, сапалық мәтіндік талдау, үлгілеу және корпус құрастырудың әдістемелік принциптері қолданылды. Материалдарды таңдау мен құрылымдаудың ғылыми негіздері ұсынылып, фонетикалық-графикалық, лексикалық, грамматикалық және лингвомәдени деңгейлерді қамтитын кешенді жүйе жасалды. Learner Corpus платформасы Оксфорд университетінде әзірленіп жатқан түркі тілдерін оқыту бағдарламаларымен сәйкестендіріліп, 2026-2027 оқу жылынан бастап факультативті құрал ретінде қолдануға арналған прототип ретінде ұсынылады. Мақалада әзірленіп жатқан корпус базасының алғаш рет англофондарға бағытталған әмбебап, көпдеңгейлі, дидактикалық тұрғыдан жүйелі цифрлық ресурс ретінде қарастырылатыны көрсетіледі. Оның теориялық маңызы – қазақ тілін оқытудың корпусық әдістерін жаңа тұрғыдан түсіндіру және интерференциялық құбылыстарды жүйелеу болса, практикалық маңызы – алынған материалдарды оқыту тәжірибесіне тікелей енгізу, оқу үдерісін дараландыру және оқу жүктемесін деңгейлік тұрғыдан оңтайландыру мүмкіндігі.

Тірек сөздер: корпус; корпусық лингвистика; қазақ тілін оқыту; тілді меңгеру; интерференция; интерактивті сценарий

Қаржыландыру көзі: Мақала Қазақстан Республикасы Ғылым және жоғары білім министрлігі Ғылым комитетінің ЖТН BR24993244 «Қазақ тілі ұлттық корпусын Smart-мәтіндер мегажобасы және қазақтілді жасанды интеллект негізі ретінде жетілдіру, ішкорпустарын әзірлеу» (2024-2026 жж.) атты бағдарламалық-нысаналы қаржыландыру жобасы аясында дайындалды.

Сілтеме жасау үшін: Рсалиева Н.М., Чукаева Т.Қ. Қазақ тілін оқытудың корпусық негіздері: англофондарға бейімделген тәжірибе. *Tiltany*, 2025. №4 (100). 93-111-бб.

DOI: <https://doi.org/10.55491/2411-6076-2025-4-93-111>

Нұрсауле Максұтқызы Рсалиева¹, Талшын Қанатқызы Чукаева^{2*}

¹кандидат филологических наук, Институт языкознания имени А. Байтұрсынұлы, Казахстан, г. Алматы, ORCID: 0000-0003-2065-0505 E-mail: nursauler@mail.ru

^{2*} автор-корреспондент, докторант, Институт языкознания имени А. Байтұрсынұлы, Казахстан, г. Алматы, ORCID: 0000-0002-0716-8370 E-mail: greenrose_98@mail.ru

**КОРПУСНЫЙ ПОДХОД К ПРЕПОДАВАНИЮ КАЗАХСКОГО ЯЗЫКА:
ОПЫТ АДАПТАЦИИ ДЛЯ АНГЛОФОНОВ**

Аннотация. В статье подробно описывается процесс разработки интерактивной корпусной базы для обучения англофонов казахскому языку – *Learner Corpus*. Исследование начинается с сопоставительного анализа зарубежного и отечественного опыта корпусной лингвистики, а также качества существующих учебников и учебных материалов, предназначенных для англоязычных изучающих. Анализ показал, что значительная часть современных учебных ресурсов недостаточно адаптирована к когнитивным, лингвистическим и коммуникативным потребностям англофонов: выявлены системные недостатки, такие как несоответствие текстов уровню, их искусственность, грамматические и логические ошибки, а также неточности в передаче культурной информации. На этой основе были определены критерии отбора материалов для базы *Learner Corpus* и разработано более 250 оригинальных сюжетных и коммуникативных текстов, лексико-грамматическая система знаний и сопоставительные грамматические материалы. Формирование содержания корпуса осуществлялось с опорой на принцип *i+1* Стивена Крашена, закономерности естественного усвоения языка и типологические различия между казахским и английским языками. Были выявлены типичные интерференционные ошибки англофонов (на фонетическом, лексическом, грамматическом и прагматическом уровнях) и показано, каким образом они устраняются через материалы, включённые в корпус. В исследовании использованы методы сопоставительного анализа, теоретического обзора, качественного текстового анализа, моделирования и методологические принципы построения корпуса. Представлены научные основания отбора и структурирования материалов, сформирована комплексная система, охватывающая фонетико-графический, лексический, грамматический и лингвокультурологический уровни. Платформа *Learner Corpus* согласуется с программами обучения тюркским языкам, разрабатываемыми в Оксфордском университете, и предлагается как прототип факультативного инструмента, который будет использоваться начиная с 2026-2027 учебного года. В статье подчёркивается, что создаваемая корпусная база впервые представляет собой универсальный, многоуровневый, дидактически структурированный цифровой ресурс, ориентированный специально на англофонов. Теоретическая значимость проекта состоит в новом осмыслении корпусных методов обучения казахскому языку и систематизации интерференционных явлений, а практическая – в возможности непосредственного применения разработанных материалов в образовательной практике, индивидуализации обучения и оптимизации распределения учебной нагрузки в соответствии с уровнем владения языком.

Ключевые слова: корпус; корпусная лингвистика; преподавание казахского языка; овладение языком; интерференция; интерактивный сценарий

Источник финансирования: Статья подготовлена в рамках проекта программно-целевого финансирования Комитета науки Министерства науки и высшего образования Республики Казахстан ИРН BR24993244 «Совершенствование Национального корпуса казахского языка как мегапроекта Smart-текстов и основы казахского искусственного интеллекта, разработка его подкорпусов» (2024-2026 гг.).

Для цитирования: Рсалиева Н.М., Чукаева Т.К. Корпусный подход к преподаванию казахского языка: опыт адаптации для англофонов. *Tiltany*, 2025. №4 (100). С. 93-111. (на каз. яз.)

DOI: <https://doi.org/10.55491/2411-6076-2025-4-93-111>

Nursaule Rsaliyeva¹, Talshyn Chukayeva^{2*}

¹Candidate of Philological Sciences, A. Baitursynuly Institute of Linguistics,

Kazakhstan, Almaty, ORCID: 0000-0003-2065-0505 E-mail: nursauler@mail.ru

^{2*}Corresponding Author, Doctoral Student, A. Baitursynuly Institute of Linguistics,

Kazakhstan, Almaty, ORCID: 0000-0002-0716-8370 E-mail: greenrose_98@mail.ru

CORPUS-BASED APPROACH TO TEACHING THE KAZAKH LANGUAGE: ADAPTED EXPERIENCE FOR ANGLOPHONES

Abstract. The article provides a detailed account of the development of an interactive corpus-based platform for teaching Kazakh to anglophones – *Learner Corpus*. The study begins with a comparative analysis of international and domestic practices in corpus linguistics, as well as an evaluation of the quality of existing textbooks and teaching materials designed for English-speaking learners. The analysis revealed that a substantial portion of current learning resources is insufficiently adapted to the cognitive, linguistic, and communicative needs of anglophones: systemic shortcomings were identified, including mismatched text difficulty levels, artificiality, grammatical and logical inconsistencies, and inaccuracies in cultural content. Based on these findings, criteria for selecting materials for the *Learner Corpus* were defined, resulting in the creation of more than 250 original narrative and communicative texts, a structured lexical-grammatical knowledge system, and comparative grammatical materials. The construction of the corpus content was guided by Stephen Krashen's *i+1* principle, the natural order of language acquisition, and the typological differences between Kazakh and English. Typical interference errors made by anglophones (at phonetic, lexical, grammatical, and pragmatic levels) were identified, and the ways in which these issues are addressed through corpus materials were demonstrated. The study employs comparative analysis, theoretical review, qualitative textual analysis, modelling, and methodological principles of corpus design. It offers a scientific rationale for selecting and structuring materials and presents a comprehensive system encompassing phonetic-graphical, lexical, grammatical, and linguistic-cultural levels. The *Learner Corpus* platform is aligned with the Turkic language teaching programmes currently being developed at the University of Oxford and is

proposed as a prototype supplementary tool to be used from the 2026-2027 academic year. The article emphasizes that the emerging corpus database is the first universal, multi-level, and pedagogically structured digital resource tailored specifically to anglophones. Its theoretical significance lies in reinterpreting corpus-based approaches to teaching Kazakh and systematizing interference phenomena, while its practical value lies in its direct applicability in teaching practice, enabling individualized instruction and optimising workload distribution according to learners' proficiency levels.

Keywords: corpus; corpus linguistics; Kazakh language teaching; language acquisition; interference; interactive script

The source of financing: The article was prepared within the framework of the program-targeted funding project of the Committee of Science of the Ministry of Science and Higher Education of the Republic of Kazakhstan IRN BR24993244 "Enhancing the National Kazakh language Corpus as a megaproject of Smart-texts and Kazakh artificial intelligence base, development of its subcorpora" (2024-2026).

For citation: Rsaliyeva, N., Chukayeva, T. Corpus-Based Approach to Teaching the Kazakh Language: Adapted Experience for Anglophones. *Tiltanyim*, 2025. No.4 (100). P. 93-111. (in Kazakh)

DOI: <https://doi.org/10.55491/2411-6076-2025-4-93-111>

Кіріспе

Қазақ тілін шетелдіктерге үйрету – тіл білімінің қолданбалы саласындағы өзекті әрі күрделі бағыттардың бірі. Әсіресе, ағылшынтілді инофондар үшін қазақ тілін меңгеруге арналған оқу-әдістемелік құралдардың тапшылығы, бар материалдардың тілді үйренушілердің когнитивтік және лингвомәдени ерекшеліктеріне толық сәйкес келмеуі – бұл саланы түбегейлі жаңғыртуды талап етеді. Соңғы жылдары тіл үйретуде цифрлық технологиялардың, атап айтқанда, корпустық лингвистиканың жетістіктерін қолдану кеңінен тарала бастады. Корпустық әдіс тек тілдік деректерді жинау мен өңдеу құралы ғана емес, сонымен қатар үйренушілердің тілдік қателерін жүйелі түрде тіркеп, сол қателер негізінде оқу мазмұнын бейімдеуге мүмкіндік беретін дидактикалық ресурсқа айналып отыр.

Соңғы кезде қазақ тілін үйренуші англофондар санының артуы байқалады, бұл төменде көрсетілген бірнеше факторға байланысты:

Геосаяси және экономикалық себептер. Қазақстанның халықаралық аренада белсенділігі, әсіресе Батыс елдерімен инвестициялық, дипломатиялық, және инфрақұрылымдық байланыстарының артуы ағылшынтілді шетелдіктердің елге келуіне ықпал етуде (Dave, 2007).

Білім мен академиялық мүдде. Түркі тілдеріне, әсіресе қазақ тіліне деген қызығушылық лингвистер, антропологтар және этнолингвистер арасында артты (Johanson & Csató, 1998).

Көптілділік пен мәдени қызығушылық. Сирек тілдерді меңгеруге деген жаһандық қызығушылық, тілдік саналуандылыққа ұмтылу қазақ тіліне деген сұранысты арттырды (Crystal, 2000).

Fulbright, DAAD, Bolashak. Халықаралық академиялық бағдарламалар арқылы келетін студенттер жергілікті тілдік ортаға бейімделу үшін қазақ тілін үйренуде (Fulbright, 2022).

Онлайн ресурстардың көбеюі. Цифрлық құралдар – қосымшалар, подкасттар мен бейне сабақтар – қазақ тілін шетелдіктер үшін қолжетімді ете түсті (Warschauer, 2000).

Еңбек нарығындағы трендтер. Халықаралық компаниялардың Қазақстанда жұмыс жүргізуі тіл меңгеруді кәсіби қажеттілікке айналдырады (Bissenova, 2017).

Қазақ диаспорасының ұрпақтары. Heritage speakers өз этникалық тілі мен мәдениетін қайта жаңғыртуға ұмтылады (Polinsky & Kagan, 2007).

Ғылыми-зерттеу мақсаттары. Қазақ мәдениеті мен қоғамын зерттейтін ғалымдар үшін тіл – зерттеу нысанына апарар құрал (Fierman, 2006).

TikTok, YouTube, Instagram. Қазақ тіліндегі контентке қызығушылық әлеуметтік медидағы көптілді аудиторияның назарын аударады (Leppänen & Elo, 2016).

Көшу және қоныстану ниеті. Қазақстанды жаңа өмір сүру, кәсіп бастау алаңы ретінде таңдағандар тілдік ортаға бейімделуге мүдделі (Cohen, Duncan, & Thulemark, 2015).

Мақалада ағылшынтілді инофондарға қазақ тілін оқытудың арнайы әзірленген интерактивті негізі – Learner Corpus құрудың теориялық алғышарттары мен әдістемелік қырлары жан-жақты қарастырылады. Зерттеудің басты мақсаты – Learner Corpus базасын құрастыру арқылы қазақ тілін шет тілі ретінде үйреніп жатқан англофондардың нақты тілдік қажеттіліктері мен типтік қателіктерін анықтап, сол деректерге негізделген оқыту материалдарын жүйелеу. Бұл бағытта шетелдік және отандық тәжірибелерге салыстырмалы шолу жасалып, қазіргі таңда қолданылып

жүрген оқу құралдарының мүмкіндіктері мен кемшін тұстары бағамдалады. Соның нәтижесінде оқытудың корпустық моделін қалыптастыру барысында мәтіндердің бірегейлігі, лексика-грамматикалық бірліктердің оқыту үдерісінде қарапайымнан күрделіге қарай прогрессиялануы, үйренушілердің тілдік интерференциясын азайту тетіктері басты назарға алынады.

Мақалада сипатталып отырған Learner Corpus базасы бір жағынан, тілді меңгерудің интерактивті моделін қамтамасыз етсе, екінші жағынан, тілдік деректерді жүйелі талдау арқылы оқытудың икемді әрі дараланған формаларын ұсынуға бағытталған. Мұндай базаның құрылуы қазақ тілін оқыту үдерісіндегі жаңа сапалық кезеңнің бастамасы болмақ. Сонымен қатар бұл бастама қазақ тілін екінші тіл ретінде оқытудың ғылыми-әдістемелік негіздерін нығайтып, халықаралық аудиторияға бейімделген, заманауи талаптарға сай білім ресурстарын жасауға мүмкіндік береді. Осыған орай, мақалада қазақ тілін үйретудің жаңа тәсілі ретінде корпустық әдісті қолданудың ғылыми негізі, қолданбалы мүмкіндігі және педагогикалық маңыздылығы кеңінен сараланады.

Материал және әдістер

Зерттеу барысында салыстырмалы талдау, теориялық шолу, корпус лингвистикасына тән сапалық контент-талдау, сондай-ақ оқыту материалдарын жобалауға арналған үлгілеу әдісі қолданылды. Салыстырмалы талдау ағылшынтілді инофондарға арналған қазақ тілі оқулықтарын, халықаралық Learner Corpus үлгілерін (British National Corpus, Cambridge Learner Corpus және басқалары), сондай-ақ қазақ тілін шет тілі ретінде оқытуда қолданылатын отандық оқу құралдарын жүйелі салыстыруға мүмкіндік берді.

Теориялық шолу корпустық әдістің ғылыми негіздеріне, тілдік интерференция теориясына, сонымен қатар Krashen, Lewis, McEneaney сияқты зерттеушілердің еңбектеріне сүйеніп жүргізілді. Контент-талдау әдісі 550-ден астам қолданыстағы оқу мәтіндерін талдауда қолданылды. Бұл кезеңде мәтіндерді іріктеу нақты критерийлерге сүйенді: тілдік және логикалық қателердің болмауы, мәтіннің деңгейге сәйкестігі, сюжеттік және коммуникативтік тартымдылық, мәдени шынайылық, сондай-ақ ұсынылатын лексика-грамматикалық минимумның сақталуы. Аталған критерийлер бастапқыда «Нәтижелер мен талқылау» бөлімінде сипатталса да, олардың зерттеу әдістемесінің бір бөлігі екені анық көрінгендіктен, осы бөлімге ауыстырылды.

Үлгілеу әдісі Learner Corpus базасының құрылымын жасауға және англофондарға бейімделген материалдарды әзірлеуге қолданылды. Бұл әдіс корпус құрамындағы материалдардың мазмұндық, лексикалық және грамматикалық прогрессиясын анықтап, болашақ цифрлық базаның ішкі логикасын құруға мүмкіндік берді.

Зерттеу материалдары бірнеше компоненттен тұрады. Бірінші компонент – ағылшынтілді аудиторияға арналған арнайы жазылған мәтіндер. Қазіргі уақытта A1 және A2 деңгейлеріне арналған 200-ден астам бірегей, сюжеттік және коммуникативтік мәтін әзірленген. Бұл мәтіндер жоғарыда аталған критерийлер бойынша іріктеліп, агглютинативті жүйеге тән грамматикалық формаларды, қазақ тілінің фонетикалық ерекшеліктерін және мәдени мағыналарды біртіндеп енгізуге бағытталған.

Екінші компонент – лексика-грамматикалық білімдер жүйесі. Корпус базасында тақырыптық бағыт бойынша бөлінген лексикалық минимум, мысалдар және жаттығулар топтамасы жинақталған. Лексикалық минимум әрбір деңгейдің коммуникативтік мақсаттарын ескере отырып құрылды, ал грамматикалық база қазақ және ағылшын тілдерінің салғастырмалы талдауына сүйенеді.

Үшінші компонент – қазақ-ағылшын салғастырмалы грамматикасының негізі. Бұл бөлім англофондар жиі қателесетін құрылымдарды анықтап, тілдік интерференцияның типтік көріністерін жүйелеуге бағытталған. Жинақталған материал қазіргі қолданыстағы оқулықтардан алынған кемшіліктерді түзетуге, сондай-ақ аналитикалық-флективті тілден агглютинативті жүйеге өтуді жеңілдетуге мүмкіндік береді.

Әлемдік Learner Corpus үлгілері де зерттеу әдістемесінің бір бөлігі ретінде қарастырылды. British National Corpus және Cambridge Learner Corpus базасындағы құрылымдық шешімдер, деңгейге сәйкестік принциптері, грамматикалық прогрессияны ұйымдастыру тәсілдері қазақ тіліне бейімдеу мақсатында талданды.

Әдебиетке шолу

Тілді оқыту мен тіл меңгертудің теориялық негіздерін және оны корпустық әдіс арқылы жүзеге асырудың ғылыми-әдістемелік қағидаттарын зерттеген ғалымдардың еңбектері тіл біліміндегі іргелі бағыттардың бірі ретінде қарастырылады. Бұл салада жүргізілген зерттеулер екі үлкен бағытқа бөлінеді: тілді үйретудің психолінгвистикалық және педагогикалық негіздерін анықтаған классикалық еңбектер және тіл үйренушілерге бағытталған арнайы корпустарды әзірлеуге арналған жұмыстар.

Алғашқы бағытта Стивен Крашеннің *Input Hypothesis* теориясы (1982) маңызды орын алады (Crashen, 1982). Ол тіл меңгеруде мағыналы, түсінікті әрі күрделірек кірістің (comprehensible input) шешуші рөл атқаратынын дәлелдеді. Бұл идея кейіннен оқыту бағдарламалары мен материалдарын құрылымдауда басшылыққа алынды. Крашеннің еңбектері тілді меңгеру мен оқытудың табиғи жолдарын сипаттай отырып, корпустық тілтанымға негіз болатын тілді тұтас құрылым ретінде тануға мүмкіндік берді.

Майкл Льюистің *Lexical Approach* тұжырымдамасы да тіл меңгерудің лексикалық бірліктер арқылы жүретінін атап өтті (Lewis, 1993). Бұл көзқарас корпустық әдіске сәйкес келіп, жиі қолданылатын тіркестер мен құрылымдарды нақты деректер арқылы іріктеуге негіз болады.

Корпус арқылы оқытудың әдістемелік негізін қалаған ғалымдардың бірі – Тони МакЭнери мен Ричард Ксан (Tony McEneery & Richard Xiao, 2006). Олар *corpus-based language learning* (корпусқа негізделген тіл үйрену) мен *data-driven learning (DDL)* (ақпаратқа негізделген тіл үйрену) ұғымдарын енгізіп, тіл үйренушілердің шынайы тілдік мәліметтермен тікелей жұмыс істеуі арқылы интерференцияны азайтып, тілдік интуицияны қалыптастыратынын дәлелдеді. Олардың еңбектерінде *BNC (British National Corpus)* және басқа да тілдік корпустардың тіл үйретуде қолданылу әдістері сипатталған.

2006-2025 жылдар аралығындағы зерттеулер корпус арқылы тіл үйретудің тиімділігін бірнеше бағытта нақтылап берді: бір жағынан, жүйелі шолулар корпусқа негізделген оқытудың жалпы нәтижелілігін дәлелдеп, сөздік қоры, коллокацияны меңгеру, жанрлық сезімталдық пен жазылым сапасының жақсаратынын көрсетсе (Lusta, Demirel, & Mohammadzadeh, 2023; Li, Noordin, Ismail, & Cao, 2025), екінші жағынан, нақты дағдыларға бағытталған жұмыстар DDL-дің (data-driven learning) сөз тіркестері мен коллокацияларды үйретуде, әсіресе EFL контекстінде, айтарлықтай артықшылығы бар екенін көрсетті (Sun, 2023; He & Xie, 2024).

Лексиканы үйрету саласында Ма мен Меі (2021) корпустық құралдардың сөздік минимумды дәл іріктеуге, жиілік пен үлгілік қолданыстарды көрсетуге мүмкіндік беретінін, ал кейінгі эмпирикалық жұмыстар бұл тәсілдің дәстүрлі жаттауға қарағанда оқушының автономиясы мен стратегияларын жақсырақ дамытатынын атап өтеді. Соңғы жылдары корпусқа негізделген тілдік педагогика тек үйренушіге емес, мұғалімге бағытталған зерттеулермен де күшейді: Ма, Tang және Lin (2022) екі қадамды тренинг моделі арқылы TESOL мұғалімдерінің корпус сауаттылығы мен корпусқа негізделген сабақ жобалау құзыретін айтарлықтай арттыруға болатынын көрсетсе, Fatt мен Leńko-Szymańska (2024) корпустарды мұғалім даярлығына енгізудің үш бағытты (тілдік сана, DDL-бағдарланған оқыту, толық CL курстары) шеңберін ұсынады.

Жас үйренушілермен жұмыс та жеке бағыт ретінде қалыптасып келеді: Crosthwaite (2024) орта мектеп оқушыларына пассив дауысты үйретуде жарты жылдық DDL эксперименті арқылы рецептивтік әрі продуктивтік білімнің өскенін дәлелдесе, Sprivey (2023) жапон бастауыш сынып оқушыларының корпусқа негізделген EFL тапсырмаларына деген көзқарасын зерттеп, бастапқы техникалық қиындықтарға қарамастан, мұндай тапсырмалардың қызығушылық пен метатілдік рефлексияны арттыратынын көрсетеді. Жалпы алғанда, қазіргі зерттеулер корпусқа сүйенген оқыту грамматикалық дәлдік пен сөздік байлықты ғана емес, оқушы автономиясын, метатілдік ойлауды және мұғалімнің кәсіби құзыреттерін бір мезгілде дамытуға қабілетті екенін, алайда бұл үшін мұғалімге бағытталған жүйелі тренинг, қолжетімді интерфейстер мен жас ерекшелігіне және контекстке бейімделген корпус құралдары қажет екенін айқын көрсетіп отыр (Emir, 2024; Basal, 2024; Boulton & Vyatkina, 2024).

Әлемдік корпус құрастыру тәжірибесінде Learner Corpus – тіл үйренушілердің нақты тілдік

өнімдерінен (эссе, әңгіме, сөйлеу мәтіндері т.б.) құралған корпус. Ол автоматты түрде өңделіп, лингвистикалық талдау жасауға бейімделеді.

Мұндай корпусстың негізгі міндеттері:

- Тілді үйренушілердің типтік қателіктерін талдау (мысалы, интерференция себебінен туындаған қателер т.б.);

- Шет тілі ретіндегі сөйлеу мен жазу ерекшеліктерін анықтау;

- Оқу бағдарламаларын, оқулықтарды, грамматикалық түсіндірмелерді жетілдіру;

Корпус құрамына:

- Жазбаша мәтіндер (оқушылардың эсселері, хаттары т.б.)

- Ауызша тіл (диалогтар, интервью)

- Аннотацияланған қателіктер (қате түрі, орны, түсіндірмесі)

- Сөзжасам, морфология, синтаксиске қатысты белгіленім кіреді.

Әлемдік learner corpus-тар – шет тілі ретінде меңгерілетін тілдерге қатысты үйренушілердің тілдік өнімдерін жинақтап, оларды жүйелі түрде зерттеуге арналған арнайы корпустар.

Ең танымал learner corpus-тардың бірі – *International Corpus of Learner English* (ICLE). Бұл корпус Бельгиядағы Лувен Католик университетінің *Corpus Linguistics* жобасы аясында әзірленіп, ана тілі әртүрлі болып келетін студенттердің ағылшын тіліндегі жазба мәтіндерін қамтиды. ICLE корпусында жиырмадан астам елдің жазба эсселері сақталған. Ол грамматикалық ауытқуларды, сөз тіркестеріндегі қателіктерді, құрылымдық интерференцияны анықтауда жиі қолданылады (Granger S. et al., 2020).

Cambridge Learner Corpus (CLC) – Кембридж университеті баспасы мен *Cambridge Assessment* ұйымының бірлескен жобасы. Бұл корпус ағылшын тілін үйренушілердің халықаралық емтихан жұмыстарынан алынған жазбаша тапсырмаларын қамтиды. CLC-нің ерекшелігі – материалдардың деңгей бойынша (A1 – C2) сараланып, әрбір қате мен құрылым талданып, тегтелгендігінде. Бұл оқулық әзірлеушілерге тіл үйренушілердің нақты қиындықтарын ескере отырып, бейімделген тапсырмалар құрастыруға мүмкіндік береді.

Louvain International Database of Spoken English Interlanguage (LINDSEI) – ағылшын тілін екінші тіл ретінде меңгерушілердің ауызша сөйлеу тілін қамтитын корпус. Ол да Лувен университетінде құрастырылған және тіл үйренушілердің ауызша тіліндегі интерференциялық белгілер мен лексикалық таңдауларын зерттеуге бағытталған. Мұнда тіл деңгейі, ана тілі, сөйлеу жағдайы ескеріледі (Gilquin G. et al., 2010).

EF-Cambridge Open Language Database (EFCAMDAT) – әлемнің 100-ден аса елінен 1 миллионнан астам студенттің ағылшын тіліндегі жазба жұмыстарын қамтитын ауқымды ашық база. Бұл корпус арқылы жаһандық тіл үйрену үрдістері, деңгейлік ерекшеліктер, жиі қайталанатын қателер динамикасы зерттеледі. Ашықтық пен кең масштабтылық – бұл жобаның басты артықшылығы.

Lang-8 Learner Corpus Жапонияда құрылған әлеуметтік жазба платформасы негізінде жасалған. Бұл база өзге тілді меңгерушілердің күнделікті жазбалары мен оларды тексерген тіл иелерінің түзетулерін қамтиды. Мұндай деректер емін-еркін стильде жазылған мәтіндерге тән қателіктерді зерттеу үшін маңызды.

JEFLC Corpus (*Japanese EFL Learner Language Corpus*) жапон тілді үйренушілердің ағылшын тіліндегі мәтіндерінен құралған. Бұл база жапон оқушыларының грамматикалық даму траекториясын зерттеуге бағытталған.

TICLE (*Turkish International Corpus of Learner English*) түріктілділердің ағылшын тілінде жасаған жазылым қателерін зерттеу үшін құрылған. Мәтін түрі эсселер мен баяндама мәтіндерін қамтиды. Бұл база мұғалімдер мен зерттеушілерге контенттің арнайы маркерленген нұсқасын ұсынады.

HKUST Learner Corpus (*Hong Kong University of Science and Technology*) қытайтілді оқушылардың ағылшын тілінде жасаған жазбаша жұмыстарынан және транскрипцияланған ауызша мәліметтерден құралған. Бұл база интерлингвистикалық құрылымдарды зерттеуге арналған.

SweLL (*Swedish Learner Language Corpus*) – шведтілді үйренушілердің тілін зерттеу

мақсатында құрылған база. Құрамында швед тіліндегі емтихандық эсселер бар, олар CEFR (Common European Framework of Reference for Languages – Тілдерді меңгерудің жалпыеуропалық құзыреттер шеңбері) деңгейіне қарай бөлінген. Зерттеу нысаны: морфология, синтаксис, қателер, лексика.

Korean Learner Corpus (KLC) корей тілін екінші тіл ретінде үйренушілер үшін құрылған. База корей тілін үйрену барысында туындайтын грамматикалық және лексикалық қиындықтарды зерттейді.

Arabic Learner Corpus (ALC) Араб тілін үйренушілердің жазба қателерін зерттеу мақсатында жинақталған. Бұл база *Qatar University* тарапынан әзірленген және Еуропа мен Азия елдеріндегі оқушылардың араб тіліндегі қателерін саралайды.

Көптеген *learner corpus*-тар қателерді аннотациялаумен (error-tagging) толықтырылған. A1 – C2 деңгейлері CEFR бойынша бөлінеді. Кейбір корпустар ауызша тіл материалдарын (диалог, презентация, интервью) да қамтиды.

Біз әзірлеп жатқан база дәстүрлі *learner corpus*-тардан түбегейлі өзгеше. Әлемдік тәжірибеде қалыптасқан *learner corpus*-тардың басым бөлігі тіл үйренушілердің жазба немесе ауызша мәтіндеріндегі типтік қателерді жинақтап, оларды лингвистикалық тұрғыда жіктеуге бағытталған. Мұндай корпус құрылымында қателерді тіркеу, тегтеу, интерференциялық құбылыстарды сипаттау, соның негізінде оқу материалдарын бейімдеу басты назарда болады. Яғни, олардың мақсаты – үйренушінің тілдік өнімі арқылы жиі кездесетін грамматикалық немесе лексикалық ауытқуларды анықтап, тіл меңгерудің өзекті (яғни шешімі табылуы тиіс) қиындықтарын жүйелеу.

Ал біз ұсынып отырған база, керісінше, дайын тілдік материалдар жинағынан емес, тілді үйренуге арналған арнайы әзірленген, толықтай бірегей контентке негізделеді. Мұнда басты назар үйренушінің қатесін тіркеу емес, оның тілдік дағдыларын саналы, сатылы түрде қалыптастыруға аударылады. Бұл интерактивті платформа лексика-грамматикалық білімдерді кезең-кезеңімен меңгертуді, тілдік бірліктерді қолдану аясына қарай ұсынуды, сондай-ақ тілдік интуицияны қалыптастыруға мүмкіндік беретін жаттығуларды қамтиды.

Базаның құрылымында оқытуға арналған арнайы жасалған мәтіндер мен мәтіндік тапсырмалар, грамматика мен лексиканы түсіндіретін бейнематериалдар, сондай-ақ дағдыларды бекітуге арналған интерактивті жаттығулар жүйесі қамтылады. Мұндай тәсіл үйренушіге тек тілдік білім беріп қана қоймай, оны қарым-қатынас құралы ретінде меңгертуге жол ашады. Сонымен қатар база тілдік құзыреттілікті кезең-кезеңімен дамыту қағидатына сүйенеді, яғни қарапайым құрылымдардан бастап күрделендірілген сөйлем үлгілеріне дейін бірізді әрі жүйелі түрде үйретіледі.

Сәйкесінше, А. Байтұрсынұлы институтында әзірленетін базаның мақсаттары мен міндеттері мынадай:

- қазақ тілін шет тілі ретінде үйренушілердің (әсіресе англофондардың) тілдік интерференциясын, қателерін, сөйлеу ерекшеліктерін зерттеу;
- тілді оқыту материалдарын әзірлеуде эмпирикалық база ретінде пайдалану;
- қазақ тіліне арналған автоматтандырылған деңгейлік тапсырмалар, грамматикалық тренажерлер мен электронды оқулықтар әзірлеуге көмектесу.

Нәтижелер және талқылау

Қазіргі кезге дейін корпус базасын жинақтау жұмыстары бірнеше бағыт бойынша жүргізілуде:

- қазіргі уақытта қолданылып жүрген инофондарға қазақ тілін оқыту құралдарына шолу жасау, артықшылықтары мен кемшіліктерін анықтау;
- *Learner Corpus* базасына арналған бірегей мәтіндер жазу;
- корпус базасына арналған лексика-грамматикалық білімдер жүйесін жинақтау;
- қазақ-ағылшын салғастырмалы грамматикасын құрастыру жұмысын бастау.

I. Ағылшынтілді инофондарға қазақ тілін оқытудың қазіргі жағдайы

Ағылшын және қазақ тілдері – гетерогенді, яғни құрылымдық және типологиялық жағынан әртүрлі тілдер. Ағылшын тілі – аналитикалық, ал қазақ тілі – агглютинативті тілдер қатарына

жатады, сондықтан олардың грамматикалық жүйесі мен сөйлем құрылымы, сондай-ақ дыбыстық жүйесі түбегейлі ерекшеленеді. Осындай айырмашылықтар тіл үйрену үдерісінде тілдік интерференцияға алып келеді. Сол себепті қазақ тілін ағылшын тілді аудиторияға үйрету үшін арнайы бейімделген әдістеме қажет. Мұндай әдістеме тілдердің айырмашылығын ескеріп, грамматикалық, фонетикалық және мәдени кедергілерді жеңуге бағытталуы тиіс.

Ағылшынтілді инофондарға арналған қазақ тілі курстарының бірнешеуін атап өтсек (мәліметтер сәйкес университеттердің ашық түрде қолжетімді сайттарынан алынды):

UCLA (Лос-Анджелес Калифорния университеті): славян, Шығыс Еуропа және Еуразия тілдері бөлімінде қазақ тілі мен мәдениеті Орталық Азияны тереңірек түсіну, тілдік және мәдениаралық құзыреттілікті дамыту мақсатында оқытылады. Бұл курстарды 2019 жылдан бері этнограф және тіл маманы Деннис Кин жүргізіп келеді.

University of Wisconsin-Madison: Орталық Еуропа институтының жазғы курстарында сегіз апталық интенсив қазақ тілі курсы өткізіледі. Сондай-ақ ұйғыр, тәжік, өзбек, әзірбайжан тілдері де оқытылады.

Indiana University Bloomington: қазақ тілі Central Eurasian Studies бөлімінде оқытылады. Кіріспе, орта және жоғары деңгейдегі курстар бар. Оқытушылар қатарында Ғайни Тыныштықова, Фатима Молдашова және Пиби-Кай Кивик бар. Курс төрт кредитке бағаланады және күзгі семестрде оқытылады.

Arizona State University: жаз айларында онлайн форматта ашық қазақ тілі курсы ұйымдастырылады. Курс жаңадан бастаушыларға арналған, грамматика мен лексикаға кіріспе беріп, студенттерді орта немесе жоғары деңгейге жеткізуді мақсат етеді.

University of Chicago: «түркі тілдеріне кіріспе» курсы аясында қазақ және өзбек тілдері аптасына бір рет оқытылады. Қазақ тілі бастапқы және орта деңгейде ұсынылады, ал жоғары деңгей сұраныс бойынша ұйымдастырылады. Оқытушы – доктор Каган Арик.

University of Cambridge: қазақ тіліне арналған толық курс жоқ, бірақ Джон Тримм тілдік ресурстары бөлімінде өзіндік оқу құралдары ұсынылған. Оларға З. Батаеваның *Colloquial Kazakh*, Т. Аяпова мен А. Томастың *Learn Kazakh in 70 steps*, Р. Мұхамедованың *Kazakh language: comprehensive grammar* және А. Тарановтың *Kazakh vocabulary for English speakers* атты еңбектері кіреді. Бұл материалдар тілдің фонетикасы, грамматикасы, сөздігі мен практикалық қолданылуына бағытталған.

Аталған оқу орындарындағы курстар мен ресурстар қазақ тілін үйренушілерге әртүрлі деңгейде қолдау көрсетеді, алайда олардың көбі қысқа мерзімді немесе қосымша сипатта ұсынылады. 8-10 апта ішінде қазақ тілі туралы жалпы түсінік беріліп, оқу, сөйлеу, жазу, тыңдау біліктері тек бастапқы деңгейге дейін ғана дамытылады. Мұндай қысқамерзімді курстардың бағдарламасы, бір жағынан, интенсивті болса, екінші жағынан, мәдени ерекшеліктерді, менталитетті, тарихи оқиғаларды есепке ала қоймайды. Демек, қазақ тілін Оксфорд университетінде жүйелі әрі сапалы оқыту мақсатында арнайы база жасап шығару, осы базаға барлық қажетті материалдарды енгізу қажеттілігі туып отыр.

Корпус базасына қажетті мәтіндерді іріктеу мен жазу барысында инофондарға арналған қазақ тілі оқулықтарында кездесетін мәтіндер жан-жақты талданды. Өлшемшарттар мынадай болды:

- мәтінде тілдік және логикалық қателердің болу-болмауы.
- мәтінде сюжеттік желінің, кейіпкердің, оқушыны қызықтыратын элементтердің болуы.
- мәтіннің сөз саны, мәтінде қамтылған лексика мен грамматиканың көрсетілген минимум шегінен аспауы.

Талданған мәтіндер негізінен мынадай оқулықтардан алынды:

1. Beginning Kazakh (Бастапқы деңгейдегі қазақ тілі), Аблахат Ибрахим. Аризона университетінің басып шығарған оқулығы, 1999. Курс 20 сабақтан тұрады. Қазір интерактивті курс ретінде қолданылып отыр. <https://www.languagecanvas.com/languages/kazakh/beginning-kazakh/>;

2. Мусабекова У.Е., Краева И.А., Фролова Г.М. Казахский язык для стран СНГ;

3. Нұржанова А.Б., Абнасырова Р.Ж., Нұрманова Г.Н., Дүйсебекова Ж.М. Шетелдіктерге

арналған қазақ тілі (қарапайым деңгей);

4. Күзекова, Ақбұзауова, қазақ тілі, бастапқы деңгей А1;

5. Ханкок Майкл. Сәлем әлем, қазақ грамматикасы;

6. Бейсенбаева, Балабекова, Жақыпжанова, Kazakh language: Elementary level (A1): Universal textbook for the kazakh diaspora abroad. – Nur-Sultan: 2021. – 202 p.;

7. Мұсаева Г. Қазақ тілі – шет тілі ретінде. Базалық деңгей: Оқу құралы. 1-кітап. – Астана, 2015. – 242 б.;

8. Хинят Бабақұмар, Такач Ибойа, Қапалбекұлы Бижомарт. Мажар тілділерге арналған қазақ тілі оқулығы. 1-кітап. Бастауыш және орта деңгей (A1, A2, B1). Мажар және қазақ тілінде. – 288 б.;

9. Нұржанова А.Б., Абнасырова Р.Ж., Нұрманова Г.Н., Дүйсебекова Ж.М. Шетелдіктерге арналған қазақ тілі (A1 қарапайым деңгей): жұмыс дәптері. / Алматы: Алматы-Болашақ, 2018;

10. Қамзебаева Р.Е., Камашева Г.Қ. Қазақ тілін үйренеміз. Тіл дамыту құралы: бастауыш және негізгі деңгей үшін. / Р.Е.Қамзебаева, Г.Қ.Камашева. – Алматы: Глобус, 2012. – 136 б.;

11. Қапалбеков Б., Жолшаева М. Қазақ тілінің этнограмматикасы. Тіл үйренушілер үшін: бастапқы деңгей. – Алматы: Сөз-Слово, 2005. – 208 б.

Мысалы:

Дәрісханада отыз парта, орындықтар және тақта бар. Тақтада бор жатыр. Дәрісханада екі терезе, бір есік бар. Терезелерде әдемі перде бар. Қабырғада карта, суреттер де бар. Парталарда кітаптар, оқулықтар, дәптерлер, қаламдар, қарындаштар жатыр. Дәрісханада жиырма бес студент отыр: он үш бала мен он екі қыз. Олар жақсы оқиды. Дәрісханада оқытушы да бар. Ол орындықта отыр (Нұржанова А.Б., Абнасырова Р.Ж., Нұрманова Г.Н., Дүйсебекова Ж.М., 2018, 12-бет). Бала – Түркістан, Шымкент төңірегінде ұл орнына қолданылатын диалект, болу керек: 13 ұл бала мен 12 қыз бала.

Мәтіннің ағылшын тіліндегі *there is / there are* құрылымдарын үйретуге арналған жаттығуларға ұқсайтыны байқалады. Мұндай аутентивті емес, жасанды мәтіндер тіл меңгерудің ең алғашқы кезеңдерінде пайдалы болғанымен, уақыт өте келе тиімділігін жоғалтады.

Бұл мәтін тіл үйренушіге грамматикалық құрылымдарды немесе лексикалық бірліктерді таныстыруға арналған сияқты көрінгенімен, шын мәнінде, тіл үйрену процесіне елеулі үлес қоспайды. Себебі мәтінде өмірлік маңызы бар не жағдаят, не әрекет не ой тудыратын мазмұн жоқ, басқаша айтқанда, когнитивтік және коммуникативтік жағынан бос кеңістік.

Біріншіден, мәтінде оқиға жоқ – ешқандай даму, динамика немесе кейіпкердің мақсатқа бағытталған әрекеті байқалмайды. Бұл оқушының мәтінге қызығушылығын төмендетеді. Тек заттардың орналасуын сипаттау – тілдік ортаны механикалық суреттеу ғана, ол тілдің табиғи қолданыс аясын ашпайды. Екіншіден, мәтін ақпараттық тұрғыдан құнарсыз. Онда айтылған нәрселер – өздігінен болжанатын, стандартты тіркестер: «дәрісханада парта бар», «терезеде перде бар» сияқты сөйлемдер – тілді өмірде қолдануға үйретпейді. Тіл үйрену тек құрылымдарды жаттап алу емес, сол құрылымдарды мағынасы бар контексте пайдалану арқылы меңгерілетін үдеріс. Мұндай мәтін тілдің ой жеткізу құралы екенін сезінуге мүмкіндік бермейді. Үшіншіден, оқушы үшін мұндай мәтін мағыналық жағынан шектеулі, өйткені ол ой тудыратын мәселе көтермейді, талқылауға ынталандырмайды, пікір айтуға мүмкіндік бермейді, мұндағы тілдік бірліктер тек тізбектелген күйде қалып отыр.

Тағы бір мысал. Бұл жерде диалогтық сөйлеу түрі бұзылып, логикалық қателер жіберілген:

Сәлеметсіз бе, апай!

– Сәлеметсіңдер ме, балалар!

– Апай, қалаға американдық режиссердің жаңа туындысы келіпті. Демалыс күні жиналып соған барсақ қайтеді? Қай жерде жүріп жатыр екен? Мұны оқушылар өздері мұғалымге айтып тұрып кімнен сұрайды. Бірнеше жерде ... деп мұғалім жауап бере ме?

– Қала бойынша бірнеше жерде көрсетіп жатыр. Онда қасымыздағы «Алатауға» барайық, автобус жалдаудың да керегі жоқ. Ендеше, сағат үште мектептің алдына жиналыңдар, балалар! (Қапалбеков Б., Жолшаева М., 2005, 168-бет).

Бұл диалог шынайы өмір шындығына сәйкес келмейді, себебі оқушы мен мұғалім

арасындағы қатынас шынайы мектеп контексінде орынсыз құрылған. Біріншіден, оқушы бірден ресми амандасып алып, күрт тақырып ауыстырып, америкалық режиссердің жаңа фильмі туралы айтады – бұл табиғи әңгіме ағымына сай келмейді. Екіншіден, мұғалімнің сол сәтте-ақ барлық ақпаратты біліп тұруы өмірлік шындықтан алшақ: ол бір репликамен фильмнің қайда, қашан көрсетілетінін, автобус жалдаудың қажет еместігін айтып, нақты шешімді де қабылдайды. Мұндай жоспарлау мектеп жүйесінде ата-анамен және мектеп әкімшілігімен келісусіз жүзеге аспайды. Сонымен қатар, «қала бойынша бірнеше жерде көрсетіп жатыр» деген секілді тіркестер ауызекі сөйлеу үлгісіне жатпайды (меніңше ауызекі), тілдің формалдығы мен эмоционалды бейтараптығы оқушылар мен мұғалім арасындағы шынайы қарым-қатынасты көрсете алмайды. Бұл диалог өмірдегі дидактикалық және ұйымдастырушылық шындықтарға сай келмейді, сондықтан тіл үйренуші үшін пайдалы бола алмайды – ол тілдік жағдаятты бере алмайды, тілдік форма мен мазмұн арасында алшақтық бар.

Жалпы алғанда, 550 мәтінді талдау арқылы қазақ тілі оқулықтарындағы бірнеше жүйелі кемшілікті анықтадық. *Біріншіден*, мәтіндер деңгейге сәйкес келмейді – не тым жеңіл, не шамадан тыс күрделі. *Екіншіден*, оқулықтарда бірсарынды мәтіндер басым: «Менің атым...», «Мен студентпін», «Мен ... ұнатамын» сияқты сөйлемдерге негізделіп, үшінші жақ мүлде енгізілмейді. *Үшіншіден*, тілдік қателері бар мәтіндер жиі кездеседі, мысал керек бұл тіл үйренушінің жүйелі түсінігін бұзады. *Төртіншіден*, тарихи және мәдени мәліметтерде дәлдік жоқ, мысал керек бұл ұлттық танымды қате қалыптастыруға алып келеді. *Бесіншіден*, кейбір мәтіндер қызықсыз, тілді меңгеруге тілтанымдық тұрғыдан да, когнитивтік тұрғыдан да үлес қоспайды. *Алтыншыдан*, диалог түріндегі мәтіндерде логиканың болмауы жиі орын алады – кейіпкерлер ұзақ сөйлейді, ақпарат қайталанады, сөйтіп шынайы диалог табиғатынан алшақтайды. Мұның бәрі оқу материалын функционалды және тілдік жағынан қайта қарауды қажет етеді.

II. *Сюжеттік мәтіндер және коммуникативті жағдайлар: бірегей мәтін жазу тәжірибесі*

Бірегей, сюжетке негізделген мәтіндер жазу – тіл үйренушіге ұсынылатын тілдік материалдың мазмұндық және танымдық тартымдылығын арттырып қана қоймай, сонымен қатар Стивен Крашеннің «қиындығы сәл жоғарырақ кіріс» (i+1) қағидасымен де үйлеседі. Крашеннің пікірінше, тіл үйренушіге ұсынылатын мәтіндер оның қазіргі деңгейінен бір саты жоғары болуы тиіс – яғни түсіну үшін күш салуды қажет етеді, алайда толық түсініксіз емес. Осы қағида негізінде *Learner Corpus* базасында ұсынылатын мәтіндер тым қарапайым, мағынасыз үлгілерден гөрі, ой салатын, сюжеттік мазмұны бар, тілдік тұрғыдан байытылған нұсқаларға сүйенеді.

Сонымен қатар сюжеттік мәтіндер мотивацияны да арттырады: оқушы әр жаңа мәтінді оқи отырып, оқиғаның қалай өрбитінін білгісі келеді, соның нәтижесінде тілді үйренуге деген ішкі қажеттілік пайда болады. Мұндай мотивация – тіл меңгерудің ең маңызды алғышарттарының бірі.

Мысалы: *Біздің үйде мысық бар. Оның аты – Тәтті. Ол өте кішкентай, бірақ сүйкімді және ақылды. Қазақ тілінде мысықтың баласын марғау деп атайды. Тәтті ойнағанды, жүгіргенді, ұйықтағанды ұнатады. Мен оны тамақтандырамын. Тәтті қант, бал сияқты тәтті нәрселерді жемейді. Сүтті, балықты, етті жақсы көреді. Ал мен Тәттіні де, тәтті нәрселерді де жақсы көремін.*

Ғылыми-әдістемелік тұрғыдан талдасақ, мәтін бірнеше маңызды тілдік және когнитивтік дағдыларды қалыптастыруға ықпал етеді. Ең алдымен, мәтін коммуникативтік құндылыққа ие. Ол нақты бір кейіпкер – автордың мысығы туралы әңгіме арқылы тілдік бірліктерді шынайы өмірлік контексте ұсынады. Бұл оқушыға тілдің қалай қолданылатынын интуитивті сезінуге көмектеседі. Яғни, мұнда тіл статикалық емес, функционалды түрде, нақты эмоция мен іс-әрекетке байланған күйде ұсынылған. Мәтіннің құрылымы да қарапайым, бірақ мағыналық жағынан айтарлықтай күрделі. Сөйлемдердің көбісі жай сөйлем болғанымен, олардың ішіндегі мазмұн когнитивтік деңгейде маңызды. Мысалы, «Қазақ тілінде мысықтың баласын марғау деп атайды» деген сөйлем арқылы оқушы тек лексикалық бірлік емес, сонымен қатар мәдени ақпарат

алады – бұл лингвомәдени құзыреттілік үшін маңызды.

Тағы бір мысал.

Қыс келді. Маржанға жаңа күрте қажет. Астананың қысы суық, сондықтан Маржан жылы, ұзын, қалың күрте іздеді. Бірнеше киім дүкеніне барды. Бірде күртенің үлгісі ұнамайды, бірде бағасы тым қымбат, ал бірде Маржанның өлшеміне сай келмейді. Ақырында бір кішігірім дүкеннен өте жақсы, орта бағадағы күрте табылды. Маржан қатты қуанды. Күртемен бірге жаңа бөкебай, бас киім мен қолғап сатып алды.

Бұл мәтін – тіл үйретуде ситуативтілік қағидатына толық сай келетін үлгі. Ол оқырманды бірден нақты жағдаятқа енгізеді: жыл мезгілі – қыс, кейіпкер – Маржан, мәселе – жылы күртенің қажеттігі. Мұндай кіріспе оқушыларға абстрактілі емес, өмірлік жағдаятта тілдік құралдарды қалай қолдануға болатынын көрсетуі арқылы олардың түсіну, байланыстыру қабілеттерін дамытады. *Біріншіден*, мәтін табиғи тілдік ахуалдан туындаған. Кейіпкердің іс-әрекеті (дүкен аралау, өлшем мен бағаға сай келетін күрте іздеу) оқушының күнделікті өмірімен үндеседі, сондықтан оқырман оқиғаға қатыса алады, кейіпкердің уәждерін, көңіл-күйін түсіне алады. Бұл – тіл үйренуші үшін контексті қабылдауды және мазмұнды талдауды жеңілдетеді. *Екіншіден*, мәтін тиянақты, оқиғалар логикалық тұрғыдан жүйелі дамиды: мәселенің қойылуы (қыс келді – күрте қажет), шешімді іздеу үдерісі (бірнеше дүкенге баруы), кедергілердің болуы (үлгісі, бағасы, өлшемі сәйкес келмеу) және шешімнің табылуы (жақсы күрте сатып алу). Бұл тіл үйренушіге әрекеттер мен себеп-салдарды түсінуге көмектеседі. *Үшіншіден*, бұл мәтінде оқушы өзін «жоғалтып» алмайды – мәтін тым жеңіл де, тым күрделі де емес. Ол тілдік тұрғыдан бай, алайда артық ғылыми термин немесе жасанды диалогтан ада. Бұл – Стивен Крашен ұсынған « $i + 1$ » қағидасына сәйкес келеді: оқушының қазіргі деңгейінен сәл жоғары деңгейдегі тілдік үлгі ұсынылған.

Жалпы алғанда, осындай сюжеттік-коммуникативтік тәсілмен 100 тақырып бойынша 250 мәтін жазылып шықты. Бірегей мәтін жазу жұмысы әлі де жалғастырылуда.

III. *Тілдік интерференция және оны жеңу жолдары: ағылшынтілді инофондарға арналған қазақ тілінің салғастырмалы грамматикасын әзірлеу тәжірибесі*

Тілдік интерференция ұғымын алғаш жүйелі түрде зерттеген ғалым – америкалық лингвист Уриэль Уэйнрих (Weinreich, 1953). Ол 1953 жылы жарық көрген *Languages in Contact: Findings and Problems* атты еңбегінде тілдер арасындағы өзара ықпалдастықты, соның ішінде интерференцияны жан-жақты сипаттап, бұл құбылысты тілдік байланыс нәтижесінде пайда болатын заңды үдеріс ретінде танытты. Уэйнрих интерференцияны тек қана грамматикалық немесе фонетикалық деңгейдегі ауытқу емес, сонымен қатар әлеуметтік және психолингвистикалық факторлармен байланысты кешенді құбылыс деп түсіндіреді.

Кейін бұл тақырыпты Леонард Блумфилд, Эрих Хаупт, сондай-ақ кеңестік дәуірдегі А.И. Смирницкий, Е.М. Верещагин, Ю.Н. Караулов сияқты ғалымдар да дамытты. Кеңестік лингвистикада интерференция әсіресе екінші тілді меңгерудің психологиялық және лингвистикалық аспектісі ретінде қарастырылды. XX ғасырдың екінші жартысында тілдік интерференцияны екінші тілді оқыту әдістемесінде маңызды фактор ретінде қарастыра бастады. Сондықтан бұл құбылысты қазіргі қолданбалы лингвистика, социолингвистика және психолингвистика салаларындағы зерттеулердің өзегі деуге болады.

Ағылшын тілі мен қазақ тілінің арасындағы айырмашылықтар тілдік интерференцияны одан әрі күшейтеді. Ағылшын тілі – Батыс Герман тілдік тобына жататын аналитикалық-флективті тіл. Отто Йеспersenнің айтуынша, ағылшын тілі өзінің тарихи даму барысында флективті жүйеден біртіндеп алшақтап, қазіргі аналитикалық құрылымға ие болды (Jespersen, 1905). Бұл үдеріс негізінен Орта ағылшын тілі кезеңінде, шамамен XI ғасырдан XIV ғасырға дейін белсенді түрде жүрді. Ерте ағылшын тілі (Old English) кезеңінде ағылшын тілі германдық флективті тілдердің қатарына жататын – зат есімдерде, сын есімдерде, есімдіктерде және етістіктерде күрделі септелу мен жіктелу жүйесі болған. Алайда норман шапқыншылығы (1066 жылдан кейін), латын және француз тілдерінің ықпалы, сондай-ақ фонетикалық өзгерістер мен дыбыстардың редукциясы нәтижесінде көптеген флексиялық жалғаулар жойылып, сөздердің формалық түрленуі қысқарды. Осылайша, грамматикалық мағыналар негізінен сөз тәртібі мен

көмекші сөздер арқылы беріле бастады.

Бір жағынан, ағылшын тілінің толық флективті болмауы оны қазақ тілінің агглютинативті жүйесіне жақындатады. Ал екінші жағынан, қазақ тіліндегі етістік парадигмалары мен ағылшын тіліндегі сәйкес етістік парадигмаларын салыстырып көрсек, алшақтықтың өте үлкен екенін байқаймыз. Мысалы, «бару» етістігінің парадигмасы (Кесте 1).

Кесте 1 – Қазақ және ағылшын тілдеріндегі «бару» етістігінің парадигмасы

Таблица 1 – Парадигма глагола «бару» в казахском и английском языках

Table 1 – Paradigm of the Verb “baru” in Kazakh and English

Қазақ	Ағылшын
Бар (барыңыз, барындар, барыңыздар) Барсын Бару Барды (бардым, бардың, бардыңыз) Баратын (баратынымын, баратынсың, баратынсыз, баратынбыз, баратынсыздар, баратынсыңдар) Барған (барғанмын, барғанбыз, барғансыз, барғансыздар, барғансыңдар) Барады (барамын, барасың, барасыз, барасындар, барасыздар, барамыз) Бара Барып Бармақ (бармақпыз, бармақсындар, бармақсыздар, бармақпын) Бармақшы (бармақшымын, бармақшысың, бармақшысыз, бармақшысыздар, бармақшымыз, бармақшысыңдар) Барушы (еді) Барса (барсақ, барсаң, барсам, барсаңдар, барсаңыздар) Барғысы (барғым, барғың, барғыңыз, барғымыз, барғыларың) келеді Болымсыз түр: бармайды, бармаса, барғысы келмейді, бармайтын, бармасын, бармаған т.б.	Go To go Went Gone Бұл етістіктің өзі одан әрі түрленбейді. Басқа шақ және етіс формалары көмекші сөздер арқылы жасалады: <i>will go, would go, would have gone</i> т.б.

Осындай жағдайда лексика-грамматикалық білімдер жүйесін берумен қатар, осы жүйені салғастырмалы түрде үйретуге қажеттілік туады.

Ағылшынтілді инофондар қазақ тілін үйренгенде кездесетін интерференциялық қателер мен қиындықтарды әліпби мен дыбыстық жүйе деңгейінде қарастырып көрейік.

Қазақ тіл білімінде фонема теориясын зерттеудің негізгі екі бағыты қалыптасқан. Бірінші бағыт – дәстүрлі акцентті-фонологиялық мектеп (басты өкілдері: Қ. Жұбанов, І. Кеңесбаев, Ж. Аралбаев, Т. Тәліпов және тағы басқалар), ал екінші бағыт – сингармониялық тәсілге негізделген көзқарас (басты өкілдері: Ә. Жүнісбек, Ж. Әбуов, М. Жүсіпов және т.б.). Сингармониялық бағытты ұстанатын зерттеушілер Ахмет Байтұрсынұлының дыбыстық жүйелеу қағидаларын басшылыққа алып, тек қазақ тілінің төл дыбыстарымен шектеледі және әліпбидің 26-28 әріптен тұратын ықшам нұсқасын пайдаланады (Рсалиева, 2008). Learner Corpus жобасын әзірлеу барысында және Оксфорд университетіндегі англофондарға қазақ тілін үйрету кезінде қазіргі жазба және ауызша қолданысқа кеңінен енген, өзге тілдік элементтермен табиғи үндескен дәстүрлі акцентті-фонологиялық бағытқа негізделген, 42 әріптен тұратын әліпби қолданылуда.

Қазақ тілі – фонемалық тұрғыда бай тіл. Онда дауысты және дауыссыз дыбыстардың үндесуі мен ерекшеліктері бар.

Ағылшын тілі мен қазақ тілінің фонемалық құрамында айырмашылық көп, салыстыруды кестеге түсіріп көрелік (Кесте 2).

Кесте 2 – Ағылшын және қазақ тілдерінің фонемалық құрамы

Таблица 2 – Фонематический состав английского и казахского языков

Table 2 – Phonemic Composition of English and Kazakh languages

Категория	Қазақ тілі	Ағылшын тілі	Айырмашылықтың сипаттамасы
Дауысты дыбыстар саны	12 (негізгі фонемалар: а, ә, е, и, о, ө, ұ, ү, ы, і, у, э)	20-дан астам (monophthongs және diphthongs)	Ағылшын тілінде дифтонгтар кең таралған; қазақ тілінде дифтонг аз (Аралбаев, 1988).
Дауыссыз дыбыстар саны	25 фонема шамасында	24-26 фонема шамасында	Дауыссыздардың түрі мен позициялық өзгерістері әртүрлі
Фонетикалық жүйенің түрі	Сингармониялық (буын үндестігіне негізделген)	Үндестік жоқ, бейтарап жүйе	Қазақ тілінде буын үндестігі негізгі заң ретінде қызмет етеді (Аралбаев, 1988).
Дауысты үндестік	Бар (жуан-жіңішке, ерін үндестігі)	Жоқ	Қазақ тілінде фонемалар сөздің барлық морфемаларында үйлесімде болады
Ұяң және қатаң дауыссыздар	Қатаң/ұяң айқын ажыратылады	Көптеген дауыссыздар позициялық түрде алмасады	Ағылшын тілінде аспирация мен позициялық дауыс ерекшеліктері бар (Roach, 2009).
Аспирация (demarcation through breath)	Жоқ	Бар (мысалы, <i>pat</i> [p ^h æt])	Қазақ тілінде [п], [к] сияқты қатаң дауыссыздар аспирациясыз айтылады
Мұрын дыбыстары	[м], [н], [ң]	[m], [n], [ŋ]	Ұқсастық көп, әсіресе «ң» дыбысы
Дауысты дыбыстардың позициялық өзгерісі	Үндестік арқылы реттеледі	Жиі, редукция мен әлсіреу бар	Ағылшын тілінде дыбыстар екпін мен буынға байланысты жиі өзгеріске түседі (Roach, 2009).
Редукция	Жоқ	Бар (мысалы, <i>banana</i> → [bə'neɪnə])	Қазақ тілінде барлық буындар анық айтылады
Экспираторлық ерекшелік	Экспираторлық, бірақ экстра қысымсыз	Күшті экспирация мен екпінге байланысты	Ағылшын тілінде екпін астындағы дыбыс қаттырақ айтылады (Roach, 2009).

Өліпби және дыбыстық жүйе деңгейінен кейінгі деңгейлерде кездесетін интерференциялық құбылыстарды атай кетейік (Кесте 3, Кесте 4).

Кесте 3 – Лексикалық деңгейдегі интерференция (калькалау)

Таблица 3 – Интерференция на лексическом уровне (калькирование)

Table 3 – Lexical-level Interference (Calquing)

Ағылшынша үлгі	Қазақша қате нұсқа	Дұрыс нұсқа
Thank you very much	үлкен рақмет	көп рақмет
Make a photo / take a photo	сурет жасау / сурет алу	суретке түсу / түсіру
Have a rest	демалу алу	демалу
Take a shower	душ алу	жуыну / душқа түсу

Кесте 4 – Сөз тіркестері мен сөйлем деңгейіндегі интерференция

Таблица 4 – Интерференция на уровне словосочетаний и предложений

Table 4 – Interference at the Phrase and Sentence Level

Ағылшын тілі	Қазақ тіліндегі қате нұсқа	Дұрыс нұсқа
I think	Мен ойлаймын	Меніңше, Менің ойымша
A group of students	Бір топ студенттер	Бір топ студент (ағылшын тілінде заттың көпше түрін -s жалғау арқылы жасаудың әсері)

Ағылшын тіліндегі жалпылама ұғымдарға қазақ тіліндегі нақты мәндегі сөздерді дұрыс сәйкестендірмеу де кездеседі. Мысалы:

“*Friend*” сөзі қазақшада дос, жолдас, сыныптас, әріптес, т.б. бөлінеді. Инофондар барлығын «дос» деп атай береді.

“*Love*” сөзі ағылшыншада кең ұғым (ата-анаға, досқа, тамаққа т.б.). Қазақ тілінде сүйу, жақсы көру, ұнату, ғашық болу – мағыналық реңкі әртүрлі. Англофондар бәріне бірдей «мен сені сүйемін» деп айтуы мүмкін, бұл ситуациялық сәйкессіздікке ұшыратады.

«Оксфорд Birsoz» бастамасы аясында Оксфорд университетінде қазақ тілінің академиялық курсы әзірленуде. Бұл курс 24 аптаны қамтымақ әрі интенсивті болмақ. Сәйкесінше, А.Байтұрсынұлы институтында әзірленіп жатқан *Learner Corpus* интерактивті базасының құрамындағы лексика-грамматикалық білімдер жүйесі, сабақтар, салғастырмалы грамматика, бір жағынан, жоғарыдағы интерференциялық «бөгеттермен» күресуге бағытталса, екінші жағынан Оксфорд университетіндегі түркі тілдерін оқыту бағдарламаларына негізделеді.

Англофондарға академиялық қазақ тілін үйрету әдістемесі жалпы және дәстүрлі тіл оқыту әдістемелерінен ерекшеленеді. Алдымен, оқытушыға қойылатын талап жоғары, кез келген екінші тілді академиялық мақсатта үйрету үшін оқытушының мета тілі (біздің жағдайда бұл – ағылшын тілі) академиялық деңгейде болуды көздейді. Демек, оқытушы қазақ және ағылшын тілдерін жай ғана еркін меңгеріп қоймай, екі тілде ғылыми салғастырмалы зерттеулер жасауға күзіретті маман болуы қажет. Тілдерді меңгерудің жалпыеуропалық құзыреттер шеңбері (А) бойынша қазақ тілін үйрету «даярлық» кезеңінен басталады. Бұл CEFR жүйесі бойынша «Pre A1», қазақша айтқанда, «A1 даярлық» деп аталады. Бұл – A1 деңгейін бастауға мүмкіндік беретін, кемінде 2 академиялық сағатты қажет ететін тілдік кезең. Бұл деңгейде студенттер:

- қазақ әліпбиіндегі әріптерді тануды;
- қолжазуды (handwriting);
- дыбыстарды (фонемаларды) ажыратуды;
- әріп, дыбыс, буын сияқты графикалық және фонематикалық элементтерді үйренеді.

Демек, жалпылай айтқанда, «A1 даярлық» – қазақ тілі туралы жалпы және тілдік-құрылымдық ақпарат беретін фонетикалық-графикалық дайындық кезеңі. Оны тіл деңгейлерінің алғышарты ретінде қарастыруға болады. «A1 даярлық» кезеңінің негізгі мақсаты – қазақ әліпбиін студенттер санасында когнитивті деңгейде автоматтандыру, буынға бөліп оқуға машықтандыру, қарапайым сөздер мен сөйлемдерді дұрыс дыбыстату.

Нәтижесінде осы айтылған мәселелерді бірнеше деңгейде былайша шешу көзделеді:

- әліпби және дыбыстық деңгей: ұқсас фонема арқылы үйрету тактикасы, дыбыстың айтылымына баса назар аудару, практикалық жаттығулар, салғастыру;
- лексикалық деңгей: нақты бірегей мәтіндерге негізделген лексикалық минимум ұсыну;
- грамматикалық деңгей: қазақ-ағылшын салғастырмалы грамматикасын әзірлеу. Бұл еңбекте ағылшын және қазақ тілдеріндегі түрлі грамматикалық құрылымдар ғана емес, ағылшынтілділердің қазақ тілін үйренгенде жіберетін негізгі қателері сипатталады;
- мәтіндік деңгей: сюжеттік-коммуникативтік бірегей мәтіндер даярлау;
- лингвомәдени деңгей: мәтіндерге қазақ мәдениеті, менталитеті, тарихынан ақпарат беретін мәліметтер енгізу.

Алынған нәтижелер *Learner Corpus* жобасының қажеттілігін және оның ғылыми-әдістемелік тұрғыдан жаңа деңгей ұсынатынын көрсетеді. Қазақ тілін шет тілі ретінде оқыту

саласында бұрын-соңды қолжетімді болған базалар фрагментарлық сипатта болды. Олардың көбінде жүйелі сюжеттік мәтіндер, когнитивтік жүктемесі дұрыс есептелген контент, салғастырмалы грамматикалық түсіндірмелер немесе инофондарға бейімделген лингвомәдени материал болмады. Сондықтан қазіргі әзірленіп жатқан корпус өзінің әмбебаптығымен және толық құрылымымен ерекшеленеді.

Халықаралық Learner Corpus үлгілерімен салыстыру осы базаның өз алдына дербес ғылыми-әдістемелік бағыт қалыптастырып жатқанын көрсетеді. British National Corpus немесе Cambridge Learner Corpus сияқты ресурстардың басты артықшылығы – кең ауқымды деректер және жалпылама ағылшын тілін оқытуға бейімделу. Ал қазақ тілін ағылшын тілді аудиторияға үйрету – типологиялық қиындығы жоғары, тілдік алшақтыққа негізделген ерекше сала. Сондықтан әзірленіп жатқан корпус әлемдік аналогтардан мазмұны, құрылымы, салғастырмалы талдауы және әдістемелік бағдарлауымен ерекшеленеді.

Қазақстандағы корпустық лингвистиканың қазіргі деңгейінде зерттеулер негізінен жалпы тілдік корпус құру, сөз түбірі мен түбірлес формаларды белгілеу, мәтін жиілігі сияқты бағыттарда жасалған. Ал арнайы *инофондарға арналған*, деңгейлік бөлінісі, лингвомәдени түсіндірмесі, арнайы сюжеттік мәтіндері және ағылшын тілімен салғастырмалы грамматикасы бар корпус бұрын жасалған емес. Демек, бұл база – мазмұны мен мақсаттық аудиториясына қарай Қазақстанда бірінші рет дайындалып жатқан бірегей жоба.

Зерттеудің белгілі шектеулері де жоқ емес. Біріншіден, Learner Corpus базасын әзірлеу әлі толық аяқталмағандықтан, талдау жасалған материалдар көлемі шектеулі және англофон үйренушілердің барлық деңгейін толық қамти алмайды. Екіншіден, ұсынылған құрылым мен лексика-грамматикалық прогрессия бастапқы кезеңдегі деректерге негізделгендіктен, болашақта корпус кеңейген сайын әдістемелік үлгілерді қайта қарау қажеттілігі туындауы мүмкін. Үшіншіден, ағылшын және қазақ тілдері арасындағы интерференциялық құбылыстар іріктелген мысалдар негізінде сипатталды, сондықтан үйренушілердің әртүрлі тілдік орталардан келуіне байланысты интерференцияның барлық типтері толық көрініс таппайды. Сонымен қатар зерттеу шетелдік оқу орындарындағы ашық деректерге сүйенетіндіктен, бағдарламалар мен оқыту әдістеріндегі жасырын немесе қолжетімсіз компоненттер есепке алынбаған болуы ықтимал. Болашақта корпус деректерінің ұлғаюы мен үйренушілердің кең ауқымын қамту бұл шектеулерді едәуір азайтады.

Қорыта айтқанда, болашақта Learner Corpus жобасын одан әрі жетілдіру үшін бірнеше бағытты назарда ұстау маңызды. Ең әуелі, англофондар жасаған тілдік қателерді жүйелі тіркеп отыратын қосымша модуль енгізу корпус деректерінің қолданбалы құндылығын арттыра алады. Сонымен қатар платформаға фонетикалық тренажерлер, сөйлем құрау алгоритмдері және сөйлеу әрекетіне негізделген интерактивті сценарийлер қосылса, үйренушінің тілдік интуициясы тереңірек қалыптасады. Оқу материалдары мәдени-танымдық мазмұнмен толықтырылып, диалогтік жағдаяттар нақты өмірлік контекстерге негізделген жағдайда, корпустың когнитивтік тиімділігі күшейе түседі. Сондай-ақ жүйені болашақта көптілді аудиторияға бейімдеп кеңейту, автоматты тексеру және жеке оқу траекториясын құратын жасанды интеллект элементтерін енгізу жобаның әдістемелік әлеуетін анағұрлым арттыра түсер еді. Толықтай жетілдіріліп, дамытылған жағдайда бұл бастама тек қазақ тілін оқытудың әдістемелік негіздерін нығайтып қана қоймай, қазақ тілінің халықаралық дәрежеде оқытылуына серпін береді, сондай-ақ оны заманауи білім беру кеңістігіне интеграциялай алады деп толық сеніммен айтуға негіз бар.

Қорытынды

Жүргізілген зерттеу англофондарға қазақ тілін оқыту саласында бұрын жүйеленбеген бірқатар маңызды мәселелерді айқындап, оларды шешуге бағытталған кешенді негіз ұсынды. Алдымен, қазіргі уақытта қолданылып жүрген 550-ден астам мәтінді қамтыған оқулық материалдары жан-жақты талданып, олардың мазмұндық, когнитивтік және лингвистикалық сапасына қатысты жүйелі кемшіліктер анықталды. Бұл кемшіліктер – мәтіндердің деңгейге сай келмеуі, оқиғасыз әрі мағынасыз үлгілердің жиілігі, логикалық және тілдік қателердің болуы, жасанды диалогтар, мәдени және тарихи ақпараттың дәл берілмеуі, сондай-ақ тілдік прогрессияның жүйесіздігі. Мұндай проблемалар оқулықтардың инофондардың когнитивтік

және коммуникативтік қажеттіліктеріне сай келмейтінін дәлелдеді.

Таңдалған материалдарды талдау нәтижесінде үйренушілерге қиындық туғызатын грамматикалық және құрылымдық ерекшеліктер де анықталды: *-ын/-ін* көсемше тұлғасының артық қолданылуы, *-атын* және *-ған* есімше формаларының ерте кезеңде берілуі, «барғысы келеді» үлгісіндегі қалау рай формаларының күрделілігі және олардың ағылшын тілінде тікелей баламасының болмауы. Бұл тұжырымдар Стивен Крашеннің грамматиканы меңгеру реттілігі туралы теориясымен үйлеседі және қазақ тілін оқытудың қазіргі моделінде когнитивтік шамадан тыс жүктеме бар екенін көрсетеді. Сонымен қатар талдау қазақ және ағылшын тілдерінің типологиялық айырмашылықтары салдарынан туындайтын интерференциялық қателердің – фонемалық, грамматикалық, лексикалық және прагматикалық деңгейлерде көрінетінін көрсетті. Осыған байланысты лексика-грамматикалық білімдер жүйесін салғастырмалы принциптерге сүйене отырып құру қажеттілігі дәлелденді.

Бірегей сюжетке негізделген 250 мәтіннің жазылып шығуы зерттеудің практикалық маңызын күшейтті. Бұл мәтіндер тілдік материалды механикалық үлгілер арқылы емес, өмірлік жағдаят пен оқиғалық дамуды қамтитын мазмұн арқылы ұсынады. Олардың когнитивтік және коммуникативтік құндылығы жоғары, өйткені студентті жағдаятқа енгізеді, тілдік формаларды нақты контексте қолдануға мүмкіндік береді және мотивацияны арттырады. Мұндай тәсіл *i+1* қағидасына сәйкес келетін тілдік прогрессияны қамтамасыз етеді.

Зерттеу нәтижесінде Learner Corpus базасының архитектурасы қалыптаса бастады: лексикалық минимум, деңгейлік мәтіндер, салғастырмалы грамматикаға негізделетін грамматикалық материалдар, лингвомәдени түсіндірмелер және интерференциялық қателер картасы жасалды. Осы құрылым болашақта корпустың тиімділігін арттырып, оны оқыту бағдарламаларына енгізуге мүмкіндік береді. Қазіргі кезеңде база оқу үдерісіне толық енгенімен, 2026-2027 оқу жылынан бастап Оксфорд университетінің қазақ тілі курсына факультативті құрал ретінде қолданылуы жоспарланып отыр. Бұл корпус мазмұнының шынайы студенттік аудиторияда қолданылуына, әрі қарай толықтырылып, жетілдірілуіне жағдай жасайды.

Талқылау бөлімінде көрсетілгендей, зерттеу белгілі методологиялық шектеулерге ие: материалдар көлемінің әлі толық болмауы, интерференциялық мысалдардың іріктемелік сипаты, халықаралық бағдарламалар туралы қолжетімді деректердің шектеулігі, сондай-ақ корпус әзірге эмпирикалық сынақтан өтпегені. Соған қарамастан, алынған нәтижелер қазақ тілін шет тілі ретінде оқыту саласына жаңа ғылыми-әдістемелік бағыт ұсынады, ал корпус базасының құрылымы болашақта тілдік қателерді тіркеу, сөйлеу әрекетін модельдеу, интерактивті сценарийлер құру сияқты практикалық мүмкіндіктерді кеңейтеді.

Қорытындылай келе, зерттеу қазақ тілін англофондарға оқытудағы басты мәселелерді анықтап қана қоймай, оларды жүйелі түрде шешуге бағытталған нақты әдістемелік және лингвистикалық негіз ұсынады. Learner Corpus базасының қалыптасуы – тек оқыту материалдарының жаңартылуы емес, қазақ тілін халықаралық академиялық кеңістікке интеграциялаудың жаңа кезеңі. Бұл жоба тіл үйрету тәсілдерін жетілдіруге, тілдердің типологиялық айырмашылығын ескеретін әдістемелер жасауға, интерференциялық құбылыстарды азайтуға және қазақ тілінің жаһандық білім беру жүйесінде лайықты орын алуына ықпал етеді. Болашақта корпус кеңейіп, эмпирикалық деректермен толыққан кезде, ол қазақ тілін үйренушілердің нақты тілдік қажеттіліктеріне бағытталған, дәлелді, заманауи, тиімді платформаға айналары сөзсіз.

Әдебиеттер

- Аралбаев Ж. Қазақ фонетикасы бойынша этюдтер. – Алма-Ата: Ғылым, 1988. – 176 б.
- Байтұрсынұлы А. Алты томдық шығармалар жинағы. Өліппелер мен мақалалар жинағы (IV т.). – Алматы: Ел-Шежіре, 2013. – 320 б.
- Қапалбеков Б., Жолшаева М. Қазақ тілінің этнограмматикасы. Тіл үйренушілер үшін: бастапқы деңгей. – Алматы: Сөз-Слово, 2005. – 208 б.
- Нұржанова А.Б., Абнасырова Р.Ж., Нұрманова Г.Н., Дүйсебекова Ж.М. Шетелдіктерге арналған қазақ тілі (A1 қарапайым деңгей): жұмыс дәптері. – Алматы: Алматы-Болашақ, 2018. – 96 б.

- Рсалиева Н.М. Қазақ онимдерін ағылшын тілінде транслитерациялаудың ғылыми-теориялық негіздері. 10.02.20 – салыстырмалы-тарихи, типологиялық және салғастырмалы тіл білімі. Филология ғылымдарының кандидаты ғылыми дәрежесін алу үшін дайындалған диссертация қолжазбасы. – Алматы, 2008. – 238 б.
- Asselborn, T., Johal, W., Tleubayev, B., et al. (2021) The transferability of handwriting skills: From the Cyrillic to the Latin alphabet. *NPJ Science of Learning*, 6(6). <https://doi.org/10.1038/s41539-021-00084-w>
- Ayapova, T. (1991) *Learn the Kazakh language in 70 steps*. Dunwoody Press, 432 p.
- Basal, A. (2024) Corpus-informed language teaching: A review of recent trends. *Journal of English Language Teaching and Applied Linguistics*, 6(1), 45–62.
- Bissenova, A. (2017) The “Trilingual” Language Policy in Kazakhstan: Tracing its Origins and Predicting its Future. *Central Asian Survey*, 36(4), 123–145.
- Boulton, A., Vyatkina, N. (2024) Corpora in language learning and teaching: Developments 2020–2024. *Language Learning & Technology*, 28(2), 1–24.
- Cambridge University Press & Cambridge English Language Assessment (2025) (n.d.). *Cambridge Learner Corpus*. Retrieved April 25, from https://www.cambridge.org/elt/corpus/learner_corpus2.htm.
- Cohen, S.A., Duncan, T., Thulemark, M. (2015) *Lifestyle Mobilities: The Crossroads of Travel, Leisure and Migration*. London: Routledge, 256 p.
- Crosthwaite, P. (2024) Data-driven learning for secondary school learners: A longitudinal study of passive constructions. *System*, 117, 102989.
- Crystal, D. (2000) *Language Death*. Cambridge University Press, 198 p.
- Curry, N., McEnery, T. (2025) A research agenda for corpus linguistics and language pedagogy. *Language Teaching Research*, 29(1), 5–28.
- Dave, B. (2007) *Kazakhstan: Ethnicity, Language and Power*. London: Routledge, 320 p.
- Emir, B. (2024) Corpus literacy and classroom integration: Implications for EFL teachers. *TESOL Quarterly*, 58(3), 755–780.
- Fierman, W. (2006) Language and Education in Post-Soviet Kazakhstan: Kazakh-Medium Instruction in Urban Schools. *The Russian Review*, 65(1), 98–120.
- Fulbright U.S. Student Program. (2022) *Annual Reports*. Retrieved from <https://us.fulbrightonline.org/>
- Gilquin, G., De Cock, S., Granger, S. (Eds.). (2010) *The Louvain International Database of Spoken English Interlanguage: Handbook and CD-ROM*. Louvain-la-Neuve: Presses universitaires de Louvain, 210 pp.
- Granger, S., Dupont, M., Meunier, F., Naets, H., Paquot, M. (2020) *The International Corpus of Learner English (Version 3)*. Louvain-la-Neuve: Presses universitaires de Louvain, 420 p.
- He, L., Xie, J. (2024) Using learner corpora to enhance collocation learning: Evidence from Chinese EFL learners. *ReCALL*, 36(1), 112–130.
- Johanson, L., Csató, É.Á. (1998) *The Turkic Languages*. London: Routledge, 474 p.
- Krashen, S.D. (1982) *Principles and Practice in Second Language Acquisition*. Oxford: Pergamon Press, 250 pp.
- Leppänen, S., Elo, A. (2016) Language in Social Media: Communication, Identity, and Power. In *The Routledge Handbook of Language and Digital Communication*. London: Routledge.
- Lewis, M. (1993) *The Lexical Approach: The State of ELT and a Way Forward*. Hove: Language Teaching Publications, 200 p.
- Lusta, D., Demirel, R., Mohammadzadeh, M. (2023) A meta-analysis of corpus-based instruction in EFL contexts. *Computer Assisted Language Learning*, 36(7), 1372–1398.
- Ma, Q., Mei, Y. (2021) Corpus tools in vocabulary teaching and learning: Pedagogical implications. *Language Teaching Research*, 25(4), 587–604.
- Ma, Q., Tang, R., Lin, A. (2022) Training language teachers in corpus-based pedagogy: A two-stage intervention model. *International Journal of Corpus Linguistics*, 27(3), 420–442.
- McEnery, T., Xiao, R. (2006) *Corpus-based Language Studies: An Advanced Resource Book*. London: Routledge, 386 pp.
- Polinsky, M., Kagan, O. (2007) Heritage Languages: In the “Wild” and in the Classroom. *Language and Linguistics Compass*, 1(5), 368–395.
- Roach, P. (2009) *English Phonetics and Phonology: A Practical Course (4th ed.)*. Cambridge University Press, 421 p.
- Spivey, H. (2023) Young learners’ perceptions of corpus-based EFL tasks: A qualitative study. *Journal of Language and Education*, 9(2), 63–77.
- Sun, Y. (2023) Data-driven learning and collocation acquisition: An experimental study. *System*, 113, 102888.
- Warschauer, M. (2000) The Changing Global Economy and the Future of Distance Language Learning. *Language Learning & Technology*, 4(2), 3–12.
- Weinreich, U. (1953) *Languages in Contact: Findings and Problems*. New York: Linguistic Circle of New York, 192 p.

References

- Aralbaev, Zh. (1988) *Qazaq fonetikasy boiynsha jetjudter*. Alma-Ata: Gylym, 176 b. [Aralbayev, Zh. (1988) *Etudes on Kazakh Phonetics*. Alma-Ata: Gylym, 176 p.] (in Kazakh)
- Asselborn, T., Johal, W., Tleubayev, B., et al. (2021) The transferability of handwriting skills: From the Cyrillic to

- the Latin alphabet. *NPJ Science of Learning*, 6(6). <https://doi.org/10.1038/s41539-021-00084-w> (in English)
- Ayapova, T. (1991) *Learn the Kazakh language in 70 steps*. Dunwoody Press, 432 p. (in English)
- Baitursynuly, A. (2013) *Alti tomdyq shygarmalar zhinagy. Alippeler men maqalalar zhinagy (IV t.)*. Almaty: El-Shezhire, 320 b. [Baitursynuly, A. (2013) *Collection of Works in Six Volumes. Collection of Alphabets and Articles (Vol. IV)*. Almaty: El-Shezhire, 320 p.] (in Kazakh)
- Basal, A. (2024) Corpus-informed language teaching: A review of recent trends. *Journal of English Language Teaching and Applied Linguistics*, 6(1), 45–62. (in English)
- Bissenova, A. (2017) The “Trilingual” Language Policy in Kazakhstan: Tracing its Origins and Predicting its Future. *Central Asian Survey*, 36(4), 123–145. (in English)
- Boulton, A., Vyatkina, N. (2024) Corpora in language learning and teaching: Developments 2020–2024. *Language Learning & Technology*, 28(2), 1–24. (in English)
- Cambridge University Press & Cambridge English Language Assessment (2025) (n.d.). *Cambridge Learner Corpus*. Retrieved April 25, from https://www.cambridge.org/elt/corpus/learner_corpus2.htm. (in English)
- Cohen, S.A., Duncan, T., Thulemark, M. (2015) *Lifestyle Mobilities: The Crossroads of Travel, Leisure and Migration*. London: Routledge, 256 p. (in English)
- Crosthwaite, P. (2024) Data-driven learning for secondary school learners: A longitudinal study of passive constructions. *System*, 117, 102989. (in English)
- Crystal, D. (2000) *Language Death*. Cambridge University Press, 198 p. (in English)
- Curry, N., McEnery, T. (2025) A research agenda for corpus linguistics and language pedagogy. *Language Teaching Research*, 29(1), 5–28. (in English)
- Dave, B. (2007) *Kazakhstan: Ethnicity, Language and Power*. London: Routledge, 320 p. (in English)
- Emir, B. (2024) Corpus literacy and classroom integration: Implications for EFL teachers. *TESOL Quarterly*, 58(3), 755–780. (in English)
- Fierman, W. (2006) Language and Education in Post-Soviet Kazakhstan: Kazakh-Medium Instruction in Urban Schools. *The Russian Review*, 65(1), 98–120. (in English)
- Fulbright, U.S. (2022) *Student Program. Annual Reports*. Retrieved from <https://us.fulbrightonline.org/> (in English)
- Gilquin, G., De Cock, S., Granger, S. (Eds.). (2010) *The Louvain International Database of Spoken English Interlanguage: Handbook and CD-ROM*. Louvain-la-Neuve: Presses universitaires de Louvain, 210 p. (in English)
- Granger, S., Dupont, M., Meunier, F., Naets, H., Paquot, M. (2020) *The International Corpus of Learner English (Version 3)*. Louvain-la-Neuve: Presses universitaires de Louvain, 420 p. (in English)
- He, L., Xie, J. (2024) Using learner corpora to enhance collocation learning: Evidence from Chinese EFL learners. *ReCALL*, 36(1), 112–130. (in English)
- Johanson, L., Csátó, É. Á. (1998) *The Turkic Languages*. London: Routledge, 474 p. (in English)
- Krashen, S.D. (1982). *Principles and Practice in Second Language Acquisition*. Oxford: Pergamon Press, 250 p. (in English)
- Leppänen, S., Elo, A. (2016) Language in Social Media: Communication, Identity, and Power. In *The Routledge Handbook of Language and Digital Communication*. London: Routledge. (in English)
- Lewis, M. (1993) *The Lexical Approach: The State of ELT and a Way Forward*. Hove: Language Teaching Publications, 200 p. (in English)
- Lusta, D., Demirel, R., Mohammadzadeh, M. (2023) A meta-analysis of corpus-based instruction in EFL contexts. *Computer Assisted Language Learning*, 36(7), 1372–1398. (in English)
- Ma, Q., Mei, Y. (2021) Corpus tools in vocabulary teaching and learning: Pedagogical implications. *Language Teaching Research*, 25(4), 587–604. (in English)
- Ma, Q., Tang, R., Lin, A. (2022) Training language teachers in corpus-based pedagogy: A two-stage intervention model. *International Journal of Corpus Linguistics*, 27(3), 420–442. (in English)
- McEnery, T., Xiao, R. (2006) *Corpus-based Language Studies: An Advanced Resource Book*. London: Routledge, 386 p. (in English)
- Nurzhanova, A.B., Abnasyrova, R.Zh., Nurmanova, G.N., Duisebekova, Zh.M (2018) *Sheteldikterge arnalgan qazaq tili (A1 qarapaiym denggej): zhumys dapteri*. Almaty: Almaty-Bolashaq, 96 b. [Nurzhanova, A.B., Abnasyrova, R.Zh., Nurmanova, G.N., Duisebekova, Zh.M. (2018) *Kazakh language for foreigners (simple level A1): workbook*. Almaty: Almaty-Bolashak, 96 p.] (in Kazakh)
- Polinsky, M., Kagan, O. (2007) Heritage Languages: In the “Wild” and in the Classroom. *Language and Linguistics Compass*, 1(5), 368–395. (in English)
- Qapalbekov, B., Zholshaeva M. (2005) *Qazaq tilining jetnogrammatikasy. Til uirenushiler ushin: Bastapqy denggej*. Almaty: Soz-Slovo, 208 b. [Kapalbekov B., Zholshayeva, M. (2005) *Ethnogrammatics of the Kazakh language. For language learners: Entry level*. Almaty: Soz-Slovo, 208 p.] (in Kazakh)
- Roach, P. (2009) *English Phonetics and Phonology: A Practical Course (4th ed.)*. Cambridge University Press, 421 p. (in English)
- Rsalieva, N.M. (2008) *Qazaq onimderin agylshyn tilinde transliteracijaladyng gylmi-teorijalyq negizderi*. 10.02.20 – salystyrmaly-tarihi, tipologijalyq zhane salgastyrmaly til bilimi. Filologija gylymdarynyng kandidati gylmi darezhesin alu ushin daiyndalghan dissertaciya qolzhazbasy. Almaty, 238 b. [Rsaliyeva, N.M. (2008) *Scientific and Theoretical Foundations of Transliteration of Kazakh onim in English*. 10.02.20 - comparative-historical, typological and comparative linguistics. Dissertation manuscript prepared for the degree of Candidate of Philological Sciences. Almaty, 238 p.] (in Kazakh)

- Spivey, H. (2023) Young learners' perceptions of corpus-based EFL tasks: A qualitative study. *Journal of Language and Education*, 9(2), 63–77. (in English)
- Sun, Y. (2023) Data-driven learning and collocation acquisition: An experimental study. *System*, 113, 102888. (in English)
- Warschauer, M. (2000) The Changing Global Economy and the Future of Distance Language Learning. *Language Learning & Technology*, 4(2), 3–12. (in English)
- Weinreich, U. (1953) *Languages in Contact: Findings and Problems*. New York: Linguistic Circle of New York, 192 p. (in English)

Мақала туралы ақпарат / Информация о статье / Information about the article
Редакцияға түсті / Поступила в редакцию / Entered the editorial office: 13.06.2025.

© Рсалиева Н.М., Чукаева Т.Қ., 2025

© А. Байтұрсынұлы атындағы Тіл білімі институты, 2025

