

**ҚОЛДАНБАЛЫ ТІЛ БІЛІМІ
ПРИКЛАДНОЕ ЯЗЫКОЗНАНИЕ
APPLIED LINGUISTICS**

Мақала/Статья/Article
FTAMP 16.21.27

<https://doi.org/10.55491/2411-6076-2026-3-121-132>

Назира Серікқызы Әміржанова^{1*}, Анар Мұратқызы Фазылжан²

^{1*}автор-корреспондент, филология ғылымдарының кандидаты, Назарбаев Университеті, Қазақстан, Астана қ., ORCID: 0000-0002-3068-5907 E-mail: nazira.amirzhanova@mail.ru

²филология ғылымдарының кандидаты, қауымдастырылған профессор, Ахмет Байтұрсынұлы атындағы Тіл білімі институты, Қазақстан, Алматы қ., ORCID: 0000-0002-0562-4846 E-mail: nar20@bk.ru

**ҚАЗАҚ ТІЛІН ЕКІНШІ ТІЛ РЕТІНДЕ МЕНГЕРУШІЛЕРДІҢ
КОЛЛОКАЦИЯЛЫҚ ҚАТЕЛЕРІ: ҰЛТТЫҚ КОРПУС ДЕРЕКТЕРІ НЕГІЗІНДЕ**

Аңдатпа. Қазақ тілін екінші тіл (L2) ретінде меңгерушілердің жазба жұмыстары бүгінге дейін корпусдық тәсілмен зерттелген жоқ. Осымен байланысты зерттеудің мақсаты – қазақ тілін екінші тіл ретінде меңгеретін оқушылардың жазба мәтіндеріндегі коллокациялық қателерді оқушылар корпусы негізінде анықтау, жіктеу және себептерін интерпретациялау. Осы зерттеуде оқушылар корпусы арқылы коллокациялық қателерді анықтауға, жіктеуге және интерпретациялауға назар аударылады. Жұмыс екінші тілді меңгеру теориясына (SLA, Second Language Acquisition) жалғамалы түркі тілдері, оның ішінде қазақ тілі материалы бойынша үлес қосады. Алдыңғы зерттеулерге сүйене отырып, мақалада «мәдени коллокация» деген теориялық категория ұсынылады. Зерттеу материалы ретінде орта мектеп деңгейіндегі 25 оқушының «Наурыз мейрамы» тақырыбына жазған 65 эссесінен құралған корпус (жалпы көлемі 12500 – 14000 сөз) алынды. Эссе мәтіндері қолмен аннотацияланып, қателерді талдау (Error Analysis) әдістемесі қолданылды. Корпустан 94 коллокациялық қате анықталып, олар лексикалық алмастыру (47%), етістік пен зат есім тіркестеріндегі ауытқулар (24%), бірінші тілдің ықпалына байланысты қателер (18%) және сын есім мен зат есім тіркестеріндегі қателер (11%) деп төрт санатқа жіктелді. Талдау грамматикалық дұрыстық пен коллокациялық құзыреттіліктің бір-бірінен ажырайтынын көрсетеді. Нәтижелер қазақ тілін L2 ретінде оқыту бағдарламаларын, лексикографиялық ресурстарды және мұғалімдерді даярлауды жетілдіру үшін эмпирикалық негіз ұсынады.

Тірек сөздер: коллокациялық қателер; оқушылар корпусы; лексикалық алмастыру; қазақ тілі; эмпирикалық негіз.

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» атты бағдарламалық-нысаналы қаржыландыру жобасы аясында дайындалды.

Сілтеме жасау үшін: Әміржанова Н.С., Фазылжан А.М. Қазақ тілін екінші тіл ретінде меңгерушілердің коллокациялық қателері: Ұлттық корпус деректері негізінде. *Tiltanyam*, 2026. №3 (103). 121-132-бб.

DOI: <https://doi.org/10.55491/2411-6076-2026-3-121-132>

Назира Сериковна Амиржанова^{1*}, Анар Муратовна Фазылжан²

^{1*}автор-корреспондент, кандидат филологических наук, Назарбаев Университет, Казахстан, г. Астана, ORCID: 0000-0002-3068-5907 E-mail: nazira.amirzhanova@mail.ru

²кандидат филологических наук, ассоциированный профессор, Институт языкознания имени Ахмета Байтурсынулы, Казахстан, г. Алматы, ORCID: 0000-0002-0562-4846 E-mail: nar20@bk.ru

**КОЛЛОКАЦИОННЫЕ ОШИБКИ ИЗУЧАЮЩИХ КАЗАХСКИЙ ЯЗЫК КАК ВТОРОЙ:
АНАЛИЗ ДАННЫХ НАЦИОНАЛЬНОГО КОРПУСА**

Аннотация. Письменные работы изучающих казахский язык как второй (L2) до сих пор не исследовались с применением корпусного подхода. Цель исследования – выявить и классифицировать коллокационные ошибки в письменных текстах учащихся, изучающих казахский язык как второй, а также интерпретировать причины их возникновения на основе данных учебного корпуса. В данном исследовании внимание сосредоточено на

выявлении, классификации и интерпретации коллокационных ошибок на материале корпуса учащихся. Работа вносит вклад в теорию усвоения второго языка (SLA, Second Language Acquisition) на материале агглютинативных тюркских языков, в том числе казахского. Опираясь на предшествующие исследования, автор предлагает теоретическую категорию «культурной коллокации». Материалом послужил корпус из 65 эссе на тему «Праздник Наурыз», написанных 25 учащимися средней школы (общий объём 12500 – 14000 слов). Тексты эссе были размечены вручную, применялась методика анализа ошибок (Error Analysis). В корпусе выявлено 94 коллокационные ошибки, классифицированные по четырём категориям: лексические замены (47%), отклонения в глагольно-именных сочетаниях (24%), ошибки, связанные с влиянием первого языка (18%), и ошибки в сочетаниях прилагательного с существительным (11%). Анализ показывает, что грамматическая правильность и коллокационная компетенция расходятся между собой. Результаты дают эмпирическую основу для совершенствования программ обучения казахскому языку как L2, лексикографических ресурсов и подготовки учителей.

Ключевые слова: коллокационные ошибки; корпус учащихся; лексическая замена; казахский язык; эмпирическая база.

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 казахоязычного искусственного интеллекта, разработка его подкорпусов».

Для цитирования: Амиржанова Н.С., Фазылжан А.М. Коллокационные ошибки изучающих казахский язык как второй: анализ данных Национального корпуса. *Tiltany*, 2026. №3 (103). С. 121-132. (на каз. яз.)

DOI: <https://doi.org/10.55491/2411-6076-2026-3-121-132>

Nazira Amirzhanova^{1*}, Anar Fazylzhan²

^{1*}Corresponding Author, Candidate of Philological Sciences, Nazarbayev University,

Kazakhstan, Astana, ORCID: 0000-0002-3068-5907 E-mail: nazira.amirzhanova@mail.ru

²Candidate of Philological Sciences, Associate Professor, Akhmet Baitursynuly Institute of Linguistics,

Kazakhstan, Almaty, ORCID: 0000-0002-0562-4846 E-mail: nar20@bk.ru

COLLOCATIONAL ERRORS OF LEARNERS OF KAZAKH AS A SECOND LANGUAGE: BASED ON THE NATIONAL CORPUS DATA

Abstract. The written production of learners of Kazakh as a second language (L2) has not yet been examined with corpus-based methods. The aim of this study is to identify and classify collocational errors in the written texts of learners of Kazakh as a second language and to interpret their causes using learner corpus data. This study focuses on identifying, classifying and interpreting collocational errors in a learner corpus. It contributes to second language acquisition (SLA) theory with evidence from agglutinative Turkic languages, Kazakh in particular. Building on previous research, the article proposes the theoretical category of “cultural collocation”. The data consist of a corpus of 65 essays on the topic “The Nauryz Holiday” written by 25 secondary school students (12,500-14,000 words in total). The essays were annotated manually and analysed using the Error Analysis approach. A total of 94 collocational errors were identified and classified into four categories: lexical substitution (47%), deviations in verb-noun combinations (24%), errors related to first language influence (18%) and errors in adjective-noun combinations (11%). The analysis shows that grammatical accuracy and collocational competence diverge. The findings provide an empirical basis for improving L2 Kazakh curricula, lexicographic resources and teacher training.

Keywords: collocational errors; Learner corpus; lexical substitution; Kazakh language; empirical basis.

Funding: 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 “Improving the National Corpus of the Kazakh Language as a Mega-project of Smart Texts and a Basis for Kazakh-language Artificial Intelligence, and Developing its Subcorpora”.

For citation: Amirzhanova, N., Fazylzhan, A. Collocational Errors of Learners of Kazakh as a Second Language: Based on the National Corpus Data. *Tiltany*, 2026. No.3 (103). Pp. 121-132. (in Kazakh)

DOI: <https://doi.org/10.55491/2411-6076-2026-3-121-132>

Кіріспе

Екінші тілді (L2) меңгеру (Second Language Acquisition, SLA) теориясында грамматикалық дұрыстық пен лексикалық ұқыптылық бірін-бірі алмастырмайтын дербес құзыреттілік салалары ретінде ажыратылады (Chomsky, 1965; Hymes, 1972). Қолданбалы лингвистика мен лексикология зерттеулері сөздерді оқшау меңгерудің коммуникативтік дискурс үшін жеткіліксіз екенін көрсетті: жатық әрі аутентикалық жазба тіл үшін сөздерді тілдің тарихи қалыптасқан тіркесу заңдылықтарына сай біріктіру, яғни коллокациялық құзыреттілік қажет (Nation, 2001; Lewis,

2000; Hoesy, 2005). Firth (1957: 11) ұсынған «сөзді оның серіктестерінен таны» (you shall know a word by the company it keeps) деген тұжырым лексиканы меңгерудің мәнін білдіретін теориялық позиция ретінде бүгінге дейін маңызын жойған жоқ. Грамматикалық құрылым мен мағынаның арақатынасы туралы генеративтік грамматика ішіндегі пікірталас та синтаксистік дұрыстықтың мағыналық табиғилықты толық қамтамасыз ете алмайтынын көрсетті (Harris, 2021). Sinclair (1991) корпуслық зерттеулер негізінде тілдік дискурста еркін тіркестерден гөрі алдын ала қалыптасқан формулалық бірліктердің, соның ішінде коллокациялардың басым болатынын көрсетті.

Коллокация лексикалық бірліктердің бірге жүру заңдылығын білдіреді, ал бұл заңдылық кездейсоқтыққа емес, тілдің ішкі нормасына негізделеді. Коллокацияның еркін сөз тіркесінен айырмашылығы сол, екі немесе одан да көп сөздің тіркесу таңдауын тілдің ішкі логикасы реттейді, оны аударма немесе ұқсастық арқылы болжау мүмкін емес (Granger, Meunier, 2008). Қазақ тілінде *мейрам ұйымдастыру*, *бақыт тілеу*, *дастарқан жаю*, *той өткізу* тіркестері нормаға сай, ал *мейрам жасау*, *бақыт сыйлау*, *дастарқан ашу* тіркестері грамматикалық тұрғыдан мүмкін болғанымен, тілдің табиғи жүйесінен алшақ.

Зерттеуде коллокациялық қате деп грамматикалық тұрғыдан мүмкін, бірақ нысаналы тілдің нормативтік лексикалық жүйесіне сәйкес келмейтін сөз тіркесі түсініледі (Howarth, 1998; Nesselhauf, 2003).

Алайда қазақ тілін L2 ретінде меңгерушілердің жазба жұмыстарындағы коллокациялық қателер корпуслық тәсіл арқылы әлі жүйелі зерттелмеген (толығырақ «Әдебиетке шолу» бөлімін қараңыз).

Зерттеудің мақсаты: қазақ тілін L2 ретінде меңгеретін оқушылардың жазба мәтіндеріндегі коллокациялық қателерді оқушылар корпусы арқылы анықтау, жіктеу және интерпретациялау.

Зерттеу мынадай төрт сұраққа жауап іздейді:

- 1) L2 қазақ тілінде жазатын оқушылардың мәтіндерінде қандай коллокациялық қате санаттары кездеседі және олардың сандық үлесі қандай?
- 2) Анықталған қателерді қандай себептермен түсіндіруге болады?
- 3) «Мәдени коллокация» категориясын L2 қазақ тілі контекстінде дербес ішкі жүйе ретінде негіздеуге бола ма?
- 4) Зерттеу нәтижелері қазақ тілін екінші тіл ретінде оқыту әдістемесіне қандай үлес қоса алады?

Материал және әдістер

Жұмыс Learner Corpus Research (LCR) парадигмасына негізделген сипаттамалы корпуслық зерттеу болып табылады (Granger, 1996, 2003). Талдау әдісі ретінде қателерді талдау (Error Analysis) қолданылды. Корпуслық лингвистика деректерге негізделген талдауға мүмкіндік беретіндіктен, қате үлгілерін сандық тұрғыдан анықтауға және оларды нормативтік дереккөздермен салыстыруға жағдай жасайды (McEnery, Wilson, 2001; Hunston, 2022).

Зерттеу материалы орта мектеп деңгейіндегі 25 оқушының «Наурыз мейрамы» тақырыбына жазған 65 эссесінен тұрады. Корпусың жалпы көлемі 12500 – 14000 сөз, бір эссенің орташа көлемі шамамен 200 сөз. Оқушылардың басым бөлігі үшін бірінші тіл орыс тілі, қазақ тілі екінші тіл болып табылады. Барлық мәтін бірыңғай тапсырма бойынша бірдей жағдайда жазылды, бұл тақырыптық лексиканың салыстырмалы болуын қамтамасыз етеді. Мәтіндер Қазақ тілінің ұлттық корпусының (<https://qazcorpus.kz/>) оқушылар ішкорпусына енгізілді. Ұлттық корпус деректері соңғы жылдары қазақ тілінің лексикалық-грамматикалық ерекшеліктерін зерттеуде белсенді қолданылып келеді (Omarova et al., 2025).

Материалдың бірыңғай тақырыпқа негізделуі әдейі таңдалған шешім. Ол тақырыптық лексика мен мәдени фондық білімді талап ететін тіркестерді бір ізбен салыстыруға мүмкіндік береді, алайда бұл шешім нәтижелердің жалпылану аясын шектейді, себебі басқа тақырыптарда қате үлестірімі өзгеше болуы ықтимал.

Мәтіндер қолмен аннотацияланды. Аннотациялау мынадай кезеңдермен жүргізілді. Алдымен, әрбір мәтіннен тіркес деңгейіндегі күмәнді қолданыстар бөліп алынды. Содан кейін әр қолданыс нормативтік дереккөздермен салыстырылды: «Қазақ әдеби тілінің сөздігі» (2011),

«Қазақ тілінің түсіндірме сөздігі» (2008) және Ұлттық корпустың негізгі ішкорпусындағы тіркесу деректері. Тіркес нормативтік дереккөздерде тіркелмеген немесе олардағы қолданыс аясына сәйкес келмеген жағдайда ғана коллокациялық қате деп белгіленді.

Грамматикалық қателер (септік, жіктік, шақ формаларының қателері) талдаудан шығарылды, өйткені зерттеу нысаны тіркесу нормасы болып табылады. Емле және пунктуация қателері де есепке алынбады.

Анықталған қателер тіркестің құрылымдық типі бойынша төрт санатқа бөлінді: лексикалық алмастыру, етістік пен зат есім тіркесі, сын есім мен зат есім тіркесі, бірінші тілдің ықпалы. Санаттар бір-бірін толық жоққа шығармайды: бір қатеде бірнеше фактор қатар байқалуы мүмкін, мұндай жағдайда контексте басым көрінген белгі бойынша шешім қабылданды. Осы шектеу нәтижелерді талқылау кезінде ескерілді.

Зерттеудің бірнеше шектеуін алдын ала атап өткен жөн. Корпус көлемі шағын және бір тақырыппен шектелген, сондықтан пайыздық үлестерді жалпы қазақ тілін L2 ретінде меңгерушілерге таратуға болмайды. Бақылау тобы, яғни ана тілі қазақ тілі болатын оқушылардың дәл сондай тапсырма бойынша жазған мәтіндері жоқ, сондықтан анықталған тіркестердің тек L2 меңгерушілерге тән екені салыстырмалы түрде тексерілген жоқ. Қате себептері мәтіндегі деректер мен нормативтік дереккөздерге сүйене отырып түсіндіріледі, алайда себеп туралы тұжырымдар интроспективтік деректермен, мысалы оқушылардан алынған сұхбатпен расталған жоқ. Сондықтан төменде келтірілетін себеп-салдарлық түсіндірмелер дәлелденген қорытынды емес, қолда бар деректерге негізделген интерпретация ретінде ұсынылады.

Әдебиетке шолу

Қазақ тіл білімінде бұл мәселенің өзіндік зерттелу тарихы бар. Құдайберген Жұбанов тілдегі тұрақты тіркестердің лексикалық жүйедегі орнын айқындай отырып, формулалық қабатты зерттеудің іргетасын қалады (Жұбанов, 1999). Әсет Болғанбаев (1970, 1988) синонимдік қатардағы сөздердің тіркесу мүмкіндіктері бірдей емес екенін, яғни мағынасы жақын сөздердің лексикалық серіктестері сәйкес келмейтінін нақты тілдік бірліктер арқылы көрсетті. Сейілбек Исаев (1998) мәселені семантика тұрғысынан жалғастырып, коллокациялық нормадан ауытқудың тіл ішіндегі себептерін қарастырды. Дегенмен аталған еңбектер негізінен қазақ тілін ана тілі ретінде қарастыру тұрғысынан жазылған. Олардың теориялық қорытындылары лексикалық үйлесімділік нормаларының зерттелгенін көрсетсе де, екінші тіл ретінде меңгеру контекстіне тікелей қатысты емес, әрі коллокациялар, оның ішінде мәдени коллокациялар жеке сараланбаған.

Халықаралық зерттеулерде коллокациялық қателер ағылшын (Nesselhauf, 2003, 2005; Howarth, 1998), неміс (Bahns, Eldaw, 1993) және испан (Alonso Ramos, 2006) тілдерін L2 ретінде үйренушілер арасында жан-жақты қарастырылған. Granger (1996, 2003) ICLE (International Corpus of Learner English) корпусына сүйене отырып, L2 үйренушілердің коллокациялық қателерінде тілдік трансфер мен шамадан тыс жалпылаудың жүйелі үлгілер құрайтынын анықтады. Nesselhauf (2005) ең жиі кездесетін коллокациялық қате мағынасы жақын, бірақ тіркесу нормасы ерекшеленетін сөздің таңдалуы екенін көрсетті. Кейінгі корпустық зерттеулер L2 үйренушілердің жиі кездесетін тіркестерді артық, ал сирек, бірақ өзара тығыз байланысқан коллокацияларды жеткіліксіз қолданатынын (Durrant, Schmitt, 2009), бұл үрдістің тілді меңгеру деңгейіне қарай өзгертетінін (Granger, Bestgen, 2014) және фразеологиялық құзыреттіліктің дамуын жазба мәтіндерде сандық тұрғыдан бақылауға болатынын (Bestgen, Granger, 2014, 2018) көрсетті. Соңғы жылдардағы зерттеулер бастауыш және орта деңгейдегі үйренушілердің коллокацияларды қолдануы жоғары деңгейдегілерден айтарлықтай ерекшеленетінін (Xia, Chen, Rae, 2023), көпсөзді құрылымдардың жазбада уақыт өте келе біртіндеп дамидынын (Candarlı, 2021) және оқушылар корпустары L2 фразеологиясын зерттеудің негізгі дереккөзіне айналғанын (Granger, 2021) көрсетеді. Wray (2002) формулалық тіл тұжырымдамасы бойынша L2 үйренушілер жеке лексемаларды меңгергенімен, олардың коллокациялық серіктестерін ішкі құрылымға айналдыра алмайды, ал бұл олқылық жазба тілде нормадан ауытқыған тіркестер арқылы көрінеді деген тұжырым жасайды (Wray, 1999; Wray, Perkins, 2000; Wray, 2012).

Johanson және Csató (1998) түркі тілдерінің типологиялық ерекшеліктерін сипаттай отырып, морфологиялық күрделілікті атап өткенімен, L2 меңгерудің лексикалық аспектілерін арнайы зерттеген жоқ.

Аталған зерттеулердің барлығы дерлік еуропалық тілдерге қатысты, ал агглютинативтік морфологиясы, синтетикалық сөзжасамы және мәдени лексикасы дамыған қазақ тілі бұл зерттеу өрісінен тыс қалып отыр. Осымен байланысты бірнеше олқылық байқалады. Біріншіден, қазақ тілін L2 ретінде меңгерушілердің жазба жұмыстарындағы коллокациялық қателер корпустық тәсіл арқылы жүйелі зерттелмеген. Екіншіден, қазақ тілі L2 бойынша коллокациялық қате түрлерінің типологиясы жасалмаған. Үшіншіден, бірінші тіл (L1) ретіндегі орыс тілінің қазақ тіліндегі коллокациялық таңдауға тигізетін ықпалы эмпирикалық материалда қаралмаған. Төртіншіден, мәдени-белгіленген тіркестер лексикалық құзыреттілік теориясында дербес санат ретінде айқындалмаған. «Қазақ әдеби тілінің сөздігі» (2011) нормативтік тіркестерді жүйелі беретін маңызды ресурс болғанымен, оның деректері оқушылардың нақты жазу практикасымен салыстырылмаған.

Нәтижелер және талқылау

Корпусты талдау барысында жалпы саны 94 коллокациялық қате анықталды. Қателер төрт негізгі санатқа топтастырылды (Кесте 1).

Кесте 1 – Коллокациялық қателердің санаттары мен сандық үлестірімі

Таблица 1 – Категории коллокационных ошибок и их количественное распределение

Table 1 – Categories and Quantitative Distribution of Collocational Errors

Санат	Жиілік (n)	Үлес (%)
Лексикалық алмастыру	44	47
Етістік пен зат есім тіркесі	23	24
Бірінші тілдің ықпалы	17	18
Сын есім мен зат есім тіркесі	10	11
Барлығы	94	100

Бұл үлестірім Nesselhauf (2005) пен Howarth (1998) ағылшын тілі материалында анықтаған жалпы бағытпен, атап айтқанда, лексикалық алмастыру санатының басым болуымен, үйлеседі. Дегенмен корпустардың көлемі мен құрамы әртүрлі болғандықтан, бұл сәйкестікті сандық салыстыру ретінде емес, бағыттас үрдіс ретінде қарастырған жөн.

Лексикалық алмастыру

Бұл санат корпустағы ең жиі кездесетін топ (n=44, 47%). Мұндай қателерде оқушы мағынасы жақын немесе категориялық жағынан ұқсас, бірақ нысаналы контексте нормативтік тұрғыдан орынсыз лексеманы таңдайды. Талданған мысалдар оқушының сөздің денотаттық мағынасын білгенімен, оның нақты контекстегі тіркесу мүмкіндіктерін толық меңгермегенін көрсетеді.

«Мереке» және «мейрам» лексемаларының алмасуы. Корпустағы ең жиі кездесетін жеке үлгі – осы. Қате 11 оқушы мәтінінде (№2, №6, №7, №11, №14, №18, №22, №24, №26, №27, №28) барлығы 18 рет тіркелді, бұл – оны жекелеген стилистикалық сәтсіздік емес, қайталанатын үлгі деп қарауға негіз береді.

Корпуста мынадай қолданыстар кездесті: *22 наурыз шығыс өңірінің мерекесі; Наурыз ол қазақтың ұлттық мерекесі; Наурыз, жалпы халықтардың мерекесі*. Аталған контекстерде нормативтік нұсқа «мейрам» лексемасын талап етеді: *22 наурыз, шығыс халықтарының мейрамы; Наурыз, қазақтың ұлттық мейрамы; Наурыз, жалпыхалықтық мейрам*.

«Қазақ әдеби тілінің сөздігі» (2011) мен Уәли (2007) бойынша бұл екі лексеманың арасында стилистикалық және семантикалық айырмашылық бар. «Мереке» жалпы бейтарап атау ретінде күнделікті қуанышты оқиғаны білдіреді (туған күн мерекесі, отбасылық мереке). «Мейрам» сөзі – ресми, ұлттық немесе діни сипаттағы, тарихи тамыры бар мерекелерге қатысты қолданылатын жоғары регистрдегі лексема. Демек, Наурыз сияқты ұлттық мейрамды сипаттауда «мереке» сөзін қолдану – коллокациялық қана емес, прагматикалық-мәдени тұрғыдан да орынсыз. Nation (2001)

ұсынған лексикалық тереңдік ұғымы тұрғысынан бұл қатені оқушының екі сөздің денотаттық мағынасын білгенімен, олардың стилистикалық коннотациясын ажырата алмауымен түсіндіруге болады.

«Азық» және «дәм» лексемаларының алмасуы. Екінші жиі кездесетін үлгі – «азық», «азық-түлік» сөздерінің «дәм» орнына жұмсалуды. Қате №3, №4, №5, №10, №15, №19, №23 мәтіндерінде барлығы 9 рет тіркелді. Корпустағы қолданыстар: *Наурыз көжеге 7 қоспа қосылады; Наурыз көже жеті түрлі азықпен жасалады; Наурыз көже жеті азық-түліктен дайындалғаннан кейін.* Нормативтік нұсқа: *Наурыз көжеге жеті түрлі дән қосылады; Наурыз көже жеті түрлі дәмнен жасалады; Наурыз көже жеті дәмнен дайындалады.*

Уәли (2007) көрсеткендей, «дәм» сөзі – қазақтың дәстүрлі тағам мәдениетінде жай ғана ас атауы емес, дәстүрлік мазмұны бар лексема. «Жеті түрлі дәм» тіркесінде жеті саны киелі ұғым ретінде қабылданады және әр дәмнің өзіндік мәні бар. Осы тұрғыдан алғанда, мұндай алмастыруды тек сөз таңдау қатесі емес, мәдени-танымдық кодты меңгермеудің көрінісі деп қарастыруға болады. Бұл интерпретация төменде ұсынылатын мәдени коллокация категориясының эмпирикалық негіздерінің бірі болып табылады.

«Жалғыз рет» және «бір-ақ рет» тіркестерінің алмасуы. Үшінші үлгі – «жалғыз рет» тіркесінің «бір-ақ рет» орнына қолданылуы. Қате №38, №43, №49, №51, №52 мәтіндерінде 7 рет кездесті. Мысалы: *Бізге жалғыз рет қана берлетін өмірді жақсы өткізу керек.* Норма бойынша: *Бізге бір-ақ рет берлетін өмірді мәнді өткізу керек.*

Қазақ тілінде «жалғыз» сөзі «жеке, дара, серіксіз» мағынасын береді және кеңістіктік немесе сандық жекеленуді білдіретін контекстерде қолданылады: *жалғыз бала, жалғыз үй.* Өмірдің қайталанбайтынын білдіруде «бір-ақ рет» немесе «бір рет» тіркестері нормативтік болып табылады (Уәли, 2007). «Жалғыз рет» тіркесі сандық мөлшерді емес, даралықты меңзейді, сондықтан бұл контексте орынсыз.

Регистрлік және морфологиялық сипаттағы басқа қателер. Лексикалық алмастыру санатына жататын тағы үш үлгі корпуста жалпы 15 рет кездесті (Кесте 2).

Кесте 2 – Регистрлік және морфологиялық сипаттағы лексикалық алмастыру мысалдары (n=15)

Таблица 2 – Примеры лексических замен регистравого и морфологического характера (n=15)

Table 2 – Examples of Register-related and Morphological Lexical Substitutions (n=15)

Қате нұсқа	Нормативтік нұсқа	Мәтін №	Жиілік	Түсіндірме
Наурыз тұжырымдамасы бірлік пен жаңарудың рәмізін білдіреді.	Наурыз бірлік пен жаңарудың мағынасын білдіреді.	№9, №12, №16, №21	5	Регистрлік сәйкессіздік: ғылыми-ресми стильге тән лексема мәдени сипаттамада қолданылған.
Наурыз қазақ халқының ең ірі мейрамы болып келеді.	Наурыз – қазақ халқының ең ірі мейрамы.	№13, №27, №33, №41, №44, №58	7	Аспекттік сәйкессіздік: динамикалық форма тұрақты сипаттамада қолданылған.
Мұндай мінез-құлық кешірулі.	Мұндай мінез-құлық кешірілуі мүмкін.	№9	3	Морфологиялық ауытқу: «-улі» жұрнағы пассивтік рұқсат мағынасын бермейді.

«Тұжырымдама» сөзінің нормативтік коллокаторлары – «жасау», «әзірлеу», «бекіту», «ұсыну»; ол мәдени мейрамның мазмұнын білдіру үшін қолданылмайды. Бұл контексте нормативтік нұсқалар: *мағынасын білдіреді, мәнін ашады, рухын жеткізеді, мазмұнын қамтиды.*

«Болып келеді» аналитикалық формасы үдерістің біртіндеп жүзеге асуын білдіреді, сондықтан тұрақты, уақытқа тәуелсіз сипаттамаларды беруге жарамайды. Орыс тіліндегі «является» сөзінің қазақша баламасы «болып табылады» немесе «болады». Бұл қатені Jarvis және Pavlenko (2008) сипаттаған семантикалық калькалау механизмімен түсіндіруге болады, дегенмен

интроспективтік дерек болмағандықтан, оқушының нақты қандай үлгіге сүйенгенін біржақты айту мүмкін емес.

«Кешірулі» тұлғасы қазақ тілінің морфологиялық жүйесінде тіркелмеген. Пассивтік рұқсат мағынасы аналитикалық конструкциямен беріледі: *кешірілуі мүмкін, жол берілуі мүмкін*.

Етістік пен зат есім тіркестеріндегі қателер

Бұл санаттағы қателер (n=23, 24%) мағына жағынан күрделі топ құрайды. Мұндай тіркестерді Howarth (1998) рестриктивті коллокациялар деп атаған. Қазақ тілінде *киім киеміз*, бірақ *орамал тағамыз*: айырмашылық тілдік дәстүрмен қалыптасқан және оны семантикалық ұқсастықтан болжау мүмкін емес. Етістік пен зат есім коллокацияларының L2 жазбада ерекше қиындық тудыратыны кейінгі зерттеулерде де көрсетілген (Garner, 2022).

«Жаңалық сыйлау» (№42, №47). *Әлемге жаңалық сыйлауды армандайды* тіркесінде коллокациялық ауытқу бар; нормативтік нұсқалар: *жаңа идея ұсыну, жаңалық ашу*. «Сыйлау» етістігінің нормативтік коллокаторлары – «сый», «сыйлық», «құрмет» сияқты нақты немесе әлеуметтік құндылығы бар зат есімдер. «Қазақ тілінің түсіндірме сөздігіне» (2008) сәйкес «жаңалық» сөзімен тіркесетін етістіктер: *ашу, жасау, ұсыну*.

«Бақыт жаяды» (№11). *Мол бақыт тілеп жаяды* тіркесіндегі «жаяды» етістігі нормативтік тұрғыдан нақты физикалық нысандармен тіркеседі: *дастарқан жаяды, кілем жаяды*. Абстрактілі «бақыт» сөзімен тіркесуі метафоралық кеңейтудің сәтсіз үлгісі болып табылады.

«Апарады» (№8). *Оның аяғы ойын-сауық, ән-күйге апарады* сөйлеміндегі «апарады» етістігі негізінен жағымсыз нәтижені білдіретін контекстерде қолданылады: *апатқа апарады, сотқа апарады*. Бейтарап немесе жағымды үдерісті білдіру үшін «ұласады» етістігі сәйкес келеді.

Осы санатта тағы үш үлгі тіркелді. *Ойын-сауықтар болады* тіркесінде «болады» етістігі оқиғаның өздігінен жүзеге асуын білдіреді, ал ойын-сауық ұйымдастыруды талап етеді, сондықтан нормативтік нұсқа: *ойын-сауық ұйымдастырылады. Деп атайды* тіркесі тілек айту контекстінде қолданылған, ал «атау» етістігі ат қою немесе санатқа жатқызу мағынасын береді; нормативтік нұсқа: *деген тілек айтылады. Жақсылық әкелуде* тіркесі орыс тіліндегі «приносить добро» құрылымына сәйкес келеді; қазақ тілінде «жақсылық» сөзі «жасау» және «тигізу» етістіктерімен тіркеседі.

Бұл санаттағы қателердің ортақ белгісі: нақты әрекетті білдіретін етістіктер абстрактілі ұғымдармен тіркестірілген. Мұны оқушылардың етістіктің семантикалық аясын кеңейтуге бейімділігімен түсіндіруге болады.

Сын есім мен зат есім тіркестеріндегі қателер

Бұл санат сандық жағынан ең аз (n=10, 11%), дегенмен сын есімнің семантикалық аясын меңгеру мәселесін айқын көрсетеді.

«Асқарлы наурыз көже» (№17). «Асқарлы» сын есімінің нормативтік қолданыс аясы биіктік пен тау жоталарын сипаттау: *асқарлы тау, асқарлы шың*. Тағам атауымен тіркесуі семантикалық аяның бұзылуын білдіреді. Мәтін мазмұнынан оқушының «ерекше, керемет» мағынасын беруді көздегені байқалады. Нормативтік нұсқа: *дәмді наурыз көже дайындап*.

«Ұлы наурыз көже» (№8). «Ұлы» сын есімі мейрам атауымен тіркескенде нормативтік болып табылады: *Ұлы Наурыз, ұлы мейрам*. Тағам атауымен тіркесуі семантикалық үйлесімсіздік тудырады, өйткені тағам «дәмді», «қою», «хош иісті» сияқты сипаттамаларды қабылдайды. Бұл жағдайда сын есімді алып тастау жеткілікті.

«Халықаралық күн» (№22). *Наурыз ол Қазақстанда халықаралық күн* деген қолданыс лексикалық-семантикалық деңгейден бөлек, энциклопедиялық білім деңгейіндегі ауытқуды да көрсетеді. «Халықаралық» сын есімі халықаралық ұйымдар белгілеген күндерге қатысты қолданылады. Наурыздың бірнеше елде тойлануы оны бұл санатқа жатқызуға негіз болмайды.

Бірінші тілдің ықпалына байланысты қателер

Бұл санаттағы қателер (n=17, 18%) тілдік трансфердің үлгілерін көрсетеді. Jarvis және Pavlenko (2008) тілдік трансферді үш механизмге бөліп қарастырады: тікелей лексикалық ауысу, калькалау және категориялық шатасу. Талдауда осы жіктеме негізге алынды.

Тікелей кодтық ауысу. №28 мәтінінде мейрамханаға или қыдырады тіркесі тіркелді; мұнда

орыс тіліндегі «или» жалғаулығы қазақша мәтінге тікелей енгізілген. Норма бойынша: *мейрамханаға барады немесе қыдырады*. Odlin (1989) мұндай ауыстыруды *loanshift* деп атаған.

Калькалау. №18 мәтінде *жалпы халықтардың мерекесі* тіркесі қолданылған; бұл орыс тіліндегі «праздник всего народа» құрылымына сәйкес келеді, ал қазақша нормативтік нұсқа *жалпыхалықтық мейрам* болып табылады. Қасым (2013) көрсеткендей, қазақ тіліндегі күрделі сын есімдердің өзіндік сөзжасамдық моделі бар, ал оның аналитикалық үлгімен алмастырылуы тілдік табиғилықты жояды.

Мәдени-энциклопедиялық білім тапшылығымен астасқан қателер. №2 және №7 мәтіндерінде *күнмен ай теңесіп, күнмен ай теңеседі* тіркестері кездесті. Мұнда астрономиялық-мәдени білімнің жеткіліксіздігі байқалады: теңелетін күн мен түн, күн мен ай емес.

«Теңесу» және «теңелу» етістіктерінің алмасуы. №7 және №27 мәтіндерінде *күн мен түн теңесіп* тіркесі кездесті, норма бойынша *күн мен түн теңеліп*. «Теңесу» етістігі екі нәрсені белсенді тең ету мағынасында жұмсалады (*есепті теңестірді*), ал «теңелу» уақыт өте өздігінен тең болу үдерісін білдіреді. Астрономиялық контексте пассивтік, өздік мәндегі форма талап етіледі. Бұл жағдайда тіл ішіндегі семантикалық ажырату дағдысының жетіспеушілігі де, бірінші тілдің ықпалы да қатар әсер еткен болуы ықтимал; қолда бар деректер екеуін ажыратуға мүмкіндік бермейді.

Талдау нәтижелері L2 қазақ тілінде жазатын оқушылардың грамматикалық дұрыстығы мен colloкациялық құзыреттілігі арасында алшақтық барын көрсетеді. Мәтіндердің басым бөлігінде морфологиялық және синтаксистік ережелер сақталған, ал лексикалық тіркестерде нормадан ауытқу қайталанып отырады. Бұл байқау Nesselhauf (2005) ағылшын тілі материалында жасаған, colloкациялық қателер грамматикалық қателерден бөлек когнитивтік деңгейде орналасады деген тұжырымымен үйлеседі.

Зерттеуші Nation лексикалық білімнің екі өлшемін ажыратады: кеңдік, яғни оқушы қанша сөздің мағынасын білетіні және тереңдік, яғни сол сөздерді нақты контексте, colloкациялық нормаға сай, стилистикалық регистрді ескере отырып қолдана алу дағдысы (Nation, 2001: 27; Read, 2000; Qian, 2002). Анықталған қателердің сипаты оқушылардың лексикалық кеңдігі салыстырмалы түрде жеткілікті болғанымен, тереңдік деңгейінің төмен екенін көрсетеді.

Корпустық деректер қателердің төрт себебін көрсетеді. Бұл себептер бір-бірін жоққа шығармайды және бір қателеде қатар әрекет етуі мүмкін.

Регистрлік шатасу. Оқушылар ауызекі сөйлеуде қолданатын лексиканы жазба мәтінде де сол күйінде пайдаланады (*мереке, тамақ, жасап берді*). Мұның бір ықтимал себебі, тілдік ортада мұндай қолданыстардың норма ретінде қабылдануы, алайда бұл болжамды тексеру үшін оқушылардың тілдік ортасы туралы қосымша дерек қажет.

Бірінші тілдің ықпалы. Интерференция негізінен екі жолмен көрінеді: тікелей лексикалық ауысу және семантикалық калькалау. Екінші жағдайда сөздер қазақ тілінде болғанымен, олардың орналасу үлгісі өзге тілдің құрылымына сәйкес келеді.

Шамадан тыс жалтылау. Барлық қатені бірінші тілдің ықпалымен түсіндіруге болмайды. *Теңесу және теңелу* (№2, №7, №27), *қайталанбай және қайталанбас* (№40), *талпыныс және ұмтылыс* (№52) жұптарында оқушылар жалпы мағынаны түсінеді, бірақ қай сөздің қай контексте қолданылатынын ажырата алмайды. Уәли (2007) синонимдердің айырмашылығы тек мағыналық өріске емес, стильдік, прагматикалық және мәдени қабаттарға да тірелетінін көрсеткен болатын; корпус материалы осы тұжырымға сәйкес мысалдар береді.

Мәдени colloкациялар. Аталған үш себептен бөлек, қолданыстағы жіктемелерге қиын орналасатын қателер тобы анықталды. *Күн мен ай теңеледі* (№2, №7), *жеті азық* (№3-№19), *сырға киеді* (№24) сияқты мысалдарда тіркестің дұрыс не бұрыс болуы тілдік нормамен қатар мәдени-танымдық білімге тәуелді. *Жеті дәм* тіркесінде жеті саны сандық мөлшер ғана емес, мәдени мазмұны бар белгі; «дәм» сөзін «азықпен» алмастырғанда тіркестің мәдени мазмұны жоғалады.

Осы негізде мәдени colloкация деп тіркесу нормасы тек тіл ішіндегі заңдылықпен емес, тілдік ұжымның мәдени-танымдық білімімен айқындалатын colloкацияны атауды ұсынамыз. Мұндай тіркестерді меңгеру лексикалық білімнен бөлек, этномәдени білімді талап етеді (Schmitt,

2000). Бұл категорияның дербестігін толық негіздеу үшін оны басқа тақырыптық материалда және басқа мәдени салаларда тексеру қажет; осы зерттеу оны ұсынуға жеткілікті эмпирикалық негіз береді, бірақ түпкілікті бекітуге жеткіліксіз.

Зерттеу нәтижелері оқыту практикасына қатысты бірнеше бағыт ұсынады.

Біріншіден, оқу бағдарламасының мазмұнын қайта қарау қажет. Қазіргі бағдарламаларда лексика көбіне жеке сөздер тізімі түрінде беріледі, ал коллокациялық құзыреттілік оқу мақсаттарының қатарынан орын алуы тиіс. Бұл оқытудың бірлігі жеке сөз емес, тіркес болуын білдіреді.

Екіншіден, зерттеу барысында жинақталған қателер базасы (қате нұсқа, нормативтік нұсқа, түсіндірме) сыныпта тікелей қолданылатын дидактикалық материалға айналдырылуы мүмкін.

Үшіншіден, мұғалімдердің лексикологиялық және лексикографиялық дайындығын тереңдету қажет, өйткені коллокациялық норманы оқыту нормативтік дереккөздермен жүйелі жұмыс істеуді талап етеді.

Қорытынды

Зерттеу – L2 қазақ тілінде жазатын оқушылардың мәтіндеріндегі коллокациялық қателерді корпустық тәсілмен жүйелі анықтауға бағытталған алғашқы эмпирикалық жұмыстардың бірі. 65 эсседегі тұратын корпустан 94 коллокациялық қате анықталып, төрт санатқа жіктелді.

Бірінші зерттеу сұрағы бойынша: лексикалық алмастыру қателері (47%) жиілік бойынша басым, одан кейін етістік пен зат есім тіркестері (24%), бірінші тілдің ықпалы (18%) және сын есім мен зат есім тіркестері (11%) орналасады.

Екінші сұрақ бойынша: қателерді төрт себеппен түсіндіруге болады, атап айтқанда, регистрлік шатасу, бірінші тілдің ықпалы, нысаналы тілдің ішкі жүйесін жеткілікті ажыратпаудан туындайтын шамадан тыс жалпылау және мәдени-танымдық білімнің тапшылығы. Бұл түсіндірмелер корпустық деректерге негізделген интерпретация болып табылады және интроспективтік әдістермен қосымша тексеруді қажет етеді.

Үшінші сұрақ бойынша: мәдени коллокация категориясын дербес ішкі жүйе ретінде қараудың эмпирикалық негізі бар, өйткені бұл топтағы тіркестердің нормативтілігі тілдік заңдылықпен ғана емес, мәдени-танымдық білімге тәуелділікпен айқындалады. Категорияны түпкілікті бекіту үшін оны басқа тақырыптық және мәдени материалда тексеру қажет.

Төртінші сұрақ бойынша: нәтижелер L2 қазақ тілін оқытудың бағдарламалық мазмұнын, оқушыларға арналған коллокациялық сөздікті, қателерге бағытталған жазу нұсқаулығын және мұғалімдерді даярлауды жаңарту үшін эмпирикалық негіз ұсынады.

Зерттеу SLA саласындағы лексикалық құзыреттілік зерттеулерінің типологиялық өрісін кеңейтіп, Nation (2001), Wray (2002) және Schmitt (2000) тұжырымдамаларының агглютинативтік морфологиялы түркі тілдерінде де қолданылатынын көрсетеді. Алдағы зерттеулерде корпус көлемін ұлғайту, тақырыптық аясын кеңейту және ана тілі қазақ тілі болатын оқушылардың мәтіндерінен бақылау корпусын құру қажет.

Әдебиеттер

- Болғанбаев Ә. Қазақ тіліндегі синонимдер. – Алматы: Ғылым, 1970. – 214 б.
 Болғанбаев Ә. Лексикология. – Алматы: Мектеп, 1988. – 184 б.
 Жұбанов Қ. Қазақ тілі жөніндегі зерттеулер. – Алматы: Ғылым, 1999. – 581 б.
 Исаев С. Қазақ тіліндегі сөздердің грамматикалық сипаты. – Алматы: Рауан, 1998. – 304 б.
 Қазақ әдеби тілінің сөздігі: он бес томдық. – Алматы: Арыс, 2011.
 Қазақ тілінің түсіндірме сөздігі. – Алматы: Дайк-Пресс, 2008. – 968 б.
 Қасым Б. Қазақ тілінің лексикологиясы мен лексикографиясы. – Алматы: Қазақ университеті, 2013. – 272 б.
 Уәли Н. Қазақ сөз мәдениетінің теориялық негіздері. – Алматы: Дайк-Пресс, 2007. – 342 б.
 Alonso Ramos, M. Towards the synthesis of combinatorial information in the dictionary entry. *Phraseology in foreign language learning and teaching* / ed. by F. Meunier, S. Granger. Amsterdam: John Benjamins, 2006. Pp. 87–104.
 Bahns, J., Eldaw, M. Should we teach EFL students collocations? *System*. 1993. Vol. 21. No. 1. Pp. 101–114. DOI: 10.1016/0346-251X(93)90010-E.
 Bestgen, Y., Granger, S. Quantifying the development of phraseological competence in L2 English writing: an automated approach. *Journal of Second Language Writing*. 2014. Vol. 26. Pp. 28–41. DOI: 10.1016/j.jslw.2014.09.004.
 Bestgen, Y., Granger, S. Tracking L2 writers' phraseological development using collgrams: evidence from a

- longitudinal EFL corpus. *Corpora and lexis* / ed. by S. Hoffmann, A. Sand, S. Arndt-Lappe, L.M. Dillmann. Leiden; Boston: Brill, 2018. Pp. 277–301.
- Candarli, D. A longitudinal study of multi-word constructions in L2 academic writing: the effects of frequency and dispersion. *Reading and Writing*. 2021. Vol. 34. No. 5. Pp. 1191–1223. DOI: 10.1007/s11145-020-10108-3.
- Chomsky, N. *Aspects of the theory of syntax*. Cambridge, MA: MIT Press, 1965. 261 p.
- Durrant, P., Schmitt, N. To what extent do native and non-native writers make use of collocations? *International Review of Applied Linguistics in Language Teaching*. 2009. Vol. 47. No. 2. Pp. 157–177. DOI: 10.1515/iral.2009.007.
- Firth, J.R. *A synopsis of linguistic theory, 1930–1955*. *Studies in linguistic analysis*. Oxford: Blackwell, 1957. Pp. 1–31.
- Garner, J. The cross-sectional development of verb-noun collocations as constructions in L2 writing. *International Review of Applied Linguistics in Language Teaching*. 2022. Vol. 60. No. 3. Pp. 909–935. DOI: 10.1515/iral-2019-0169.
- Granger, S. From CA to CIA and back: an integrated approach to computerized bilingual and learner corpora. *Languages in contrast* / ed. by K. Aijmer, B. Altenberg, M. Johansson. Lund: Lund University Press, 1996. Pp. 37–51.
- Granger, S. Error-tagged learner corpora and CALL: a promising synergy. *CALICO Journal*. 2003. Vol. 20. No. 3. Pp. 465–480.
- Granger, S. (ed.) *Perspectives on the L2 phrasicon: the view from learner corpora*. Bristol: Multilingual Matters, 2021.
- Granger, S., Meunier, F. (eds.) *Phraseology: an interdisciplinary perspective*. Amsterdam; Philadelphia: John Benjamins, 2008. 422 p. DOI: 10.1075/z.139.
- Granger, S., Bestgen, Y. The use of collocations by intermediate vs. advanced non-native writers: a bigram-based study. *International Review of Applied Linguistics in Language Teaching*. 2014. Vol. 52. No. 3. Pp. 229–252. DOI: 10.1515/iral-2014-0011.
- Harris, R.A. *The linguistics wars: Chomsky, Lakoff, and the battle over deep structure*. 2nd ed. Oxford: Oxford University Press, 2021. 568 p.
- Hoey, M. *Lexical priming: a new theory of words and language*. London: Routledge, 2005. 202 p.
- Howarth, P. Phraseology and second language proficiency. *Applied Linguistics*. 1998. Vol. 19. No. 1. Pp. 24–44. DOI: 10.1093/applin/19.1.24.
- Hunston, S. *Corpora in applied linguistics*. 2nd ed. Cambridge: Cambridge University Press, 2022. DOI: 10.1017/9781108616218.
- Hymes, D. On communicative competence. *Sociolinguistics: selected readings* / ed. by J.B. Pride, J. Holmes. Harmondsworth: Penguin, 1972. Pp. 269–293.
- Jarvis, S., Pavlenko, A. *Crosslinguistic influence in language and cognition*. New York: Routledge, 2008. 264 p. DOI: 10.4324/9780203935927.
- Johanson, L., Csátó, É.Á. (eds.) *The Turkic languages*. London; New York: Routledge, 1998. 480 p. DOI: 10.4324/9780203066102.
- Lewis, M. *Teaching collocation: further developments in the lexical approach*. Hove: Language Teaching Publications, 2000. 256 p.
- McEnery, T., Wilson, A. *Corpus linguistics: an introduction*. 2nd ed. Edinburgh: Edinburgh University Press, 2001. 235 p.
- Nation I.S.P. *Learning vocabulary in another language*. Cambridge: Cambridge University Press, 2001. 477 p. DOI: 10.1017/CBO9781139524759.
- Nesselhauf, N. The use of collocations by advanced learners of English and some implications for teaching // *Applied Linguistics*. 2003. Vol. 24. No. 2. Pp. 223–242. DOI: 10.1093/applin/24.2.223.
- Nesselhauf, N. *Collocations in a learner corpus*. Amsterdam: John Benjamins, 2005. 331 p. DOI: 10.1075/scl.14.
- Odlin, T. *Language transfer: cross-linguistic influence in language learning*. Cambridge: Cambridge University Press, 1989. 210 p.
- Omarova, S., Ospanova, D., Aitova, N., Tokenkyzy, G., Ormanova, A., Alshynbekova, M. A corpus approach in language discovery: a word frequency analysis based on the corpus outcomes in Kazakh. *Forum for Linguistic Studies*. 2025. Vol. 7. No. 2. Pp. 869–881. DOI: 10.30564/fls.v7i2.8317.
- Qian, D.D. Investigating the relationship between vocabulary knowledge and academic reading performance: an assessment perspective. *Language Learning*. 2002. Vol. 52. No. 3. Pp. 513–536. DOI: 10.1111/1467-9922.00193.
- Read, J. *Assessing vocabulary*. Cambridge: Cambridge University Press, 2000. 292 p. DOI: 10.1017/CBO9780511732942.
- Schmitt, N. *Vocabulary in language teaching*. Cambridge: Cambridge University Press, 2000. 224 p.
- Sinclair, J. *Corpus, concordance, collocation*. Oxford: Oxford University Press, 1991. 179 p. DOI: 10.1093/oso/9780194371445.001.0001.
- Wray, A. Formulaic language in learners and native speakers. *Language Teaching*. 1999. Vol. 32. No. 4. Pp. 213–231. DOI: 10.1017/S0261444800014154.
- Wray, A., Perkins, M.R. The functions of formulaic language: an integrated model. *Language & Communication*. 2000. Vol. 20. No. 1. Pp. 1–28. DOI: 10.1016/S0271-5309(99)00015-4.
- Wray, A. *Formulaic language and the lexicon*. Cambridge: Cambridge University Press, 2002. 330 p. DOI: 10.1017/CBO9780511519772.
- Wray, A. What do we (think we) know about formulaic language? An evaluation of the current state of play. *Annual*

Review of Applied Linguistics. 2012. Vol. 32. Pp. 231–254. DOI: 10.1017/S026719051200013X.

Xia, D., Chen, Y., Pae, H.K. Lexical and grammatical collocations in beginning and intermediate L2 argumentative essays: a bigram study. *International Review of Applied Linguistics in Language Teaching*. 2023. Vol. 61. No. 4. Pp. 1421–1453. DOI: 10.1515/iral-2021-0188.

References

Alonso Ramos, M. (2006) Towards the synthesis of combinatorial information in the dictionary entry. In: Meunier, F., Granger, S. (eds.) *Phraseology in Foreign Language Learning and Teaching*. Amsterdam: John Benjamins. Pp. 87–104. (in English)

Bahns, J., Eldaw, M. (1993) Should we teach EFL students collocations? *System*. Vol. 21. No. 1. Ppp. 101–114. DOI: 10.1016/0346-251X(93)90010-E (in English)

Bestgen, Y., Granger, S. (2014) Quantifying the development of phraseological competence in L2 English writing: an automated approach. *Journal of Second Language Writing*. Vol. 26. Pp. 28–41. DOI: 10.1016/j.jslw.2014.09.004 (in English)

Bestgen, Y., Granger, S. (2018) Tracking L2 writers' phraseological development using collgrams: evidence from a longitudinal EFL corpus. In: Hoffmann, S., Sand, A., Arndt-Lappe, S., Dillmann, L.M. (eds.) *Corpora and Lexis*. Leiden, Boston: Brill. Pp. 277–301. (in English)

Bolganbaev, A. (1970) Qazaq tilindegi sinonimder. Almaty: Gylym. 214 b. [Bolganbayev, A. (1970) Synonyms in the Kazakh Language. Almaty: Gylym. 214 p.] (in Kazakh)

Bolganbaev, A. (1988) *Leksikologija*. Almaty: Mektep. 184 b. [Bolganbayev, A. (1988) *Lexicology*. Almaty: Mektep. 184 p.] (in Kazakh)

Candarli, D. (2021) A longitudinal study of multi-word constructions in L2 academic writing: the effects of frequency and dispersion. *Reading and Writing*. Vol. 34. No. 5. Pp. 1191–1223. DOI: 10.1007/s11145-020-10108-3 (in English)

Chomsky, N. (1965) *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press. 261 p. (in English)

Durrant, P., Schmitt, N. (2009) To what extent do native and non-native writers make use of collocations? *International Review of Applied Linguistics in Language Teaching*. Vol. 47. No. 2. Pp. 157–177. DOI: 10.1515/iral.2009.007 (in English)

Firth, J.R. (1957) A synopsis of linguistic theory, 1930–1955. In: *Studies in Linguistic Analysis*. Oxford: Blackwell. Pp. 1–31. (in English)

Garner, J. (2022) The cross-sectional development of verb-noun collocations as constructions in L2 writing. *International Review of Applied Linguistics in Language Teaching*. Vol. 60. No. 3. Pp. 909–935. DOI: 10.1515/iral-2019-0169 (in English)

Granger, S. (1996) From CA to CIA and back: an integrated approach to computerized bilingual and learner corpora. In: Aijmer, K., Altenberg, B., Johansson, M. (eds.) *Languages in Contrast*. Lund: Lund University Press. Pp. 37–51. (in English)

Granger, S. (2003) Error-tagged learner corpora and CALL: a promising synergy. *CALICO Journal*. Vol. 20. No. 3. Pp. 465–480. (in English)

Granger, S. (ed.) (2021) *Perspectives on the L2 Phrasicon: The View from Learner Corpora*. Bristol: Multilingual Matters. (in English)

Granger, S., Meunier, F. (eds.) (2008) *Phraseology: An Interdisciplinary Perspective*. Amsterdam, Philadelphia: John Benjamins. 422 p. DOI: 10.1075/z.139 (in English)

Granger, S., Bestgen, Y. (2014) The use of collocations by intermediate vs. advanced non-native writers: a bigram-based study. *International Review of Applied Linguistics in Language Teaching*. Vol. 52. No. 3. Pp. 229–252. DOI: 10.1515/iral-2014-0011 (in English)

Harris, R.A. (2021) *The Linguistics Wars: Chomsky, Lakoff, and the Battle over Deep Structure*. 2nd ed. Oxford: Oxford University Press. 568 p. (in English)

Hoey, M. (2005) *Lexical Priming: A New Theory of Words and Language*. London: Routledge. 202 p. (in English)

Howarth, P. (1998) Phraseology and second language proficiency. *Applied Linguistics*. Vol. 19. No. 1. Pp. 24–44. DOI: 10.1093/applin/19.1.24 (in English)

Hunston, S. (2022) *Corpora in Applied Linguistics*. 2nd ed. Cambridge: Cambridge University Press. DOI: 10.1017/9781108616218 (in English)

Hymes, D. (1972) On communicative competence. In: Pride, J.B., Holmes, J. (eds.) *Sociolinguistics: Selected Readings*. Harmondsworth: Penguin. Pp. 269–293. (in English)

Isaev, S. (1998) Qazaq tilindegi sozderding grammatikalyq sipaty. Almaty: Rauan. 304 b. [Issayev, S. (1998) *Grammatical Characteristics of Words in the Kazakh Language*. Almaty: Rauan. 304 p.] (in Kazakh)

Jarvis, S., Pavlenko, A. (2008) *Crosslinguistic Influence in Language and Cognition*. New York: Routledge. 264 p. DOI: 10.4324/9780203935927 (in English)

Johanson, L., Csató, É.Á. (eds.) (1998) *The Turkic Languages*. London, New York: Routledge. 480 p. DOI: 10.4324/9780203066102 (in English)

Lewis, M. (2000) *Teaching Collocation: Further Developments in the Lexical Approach*. Hove: Language Teaching Publications. 256 p. (in English)

McEnery, T., Wilson, A. (2001) *Corpus Linguistics: An Introduction*. 2nd ed. Edinburgh: Edinburgh University Press. 235 p. (in English)

- Nation, I.S.P. (2001) *Learning Vocabulary in Another Language*. Cambridge: Cambridge University Press, 477 p. DOI: 10.1017/CBO9781139524759 (in English)
- Nesselhauf, N. (2003) The use of collocations by advanced learners of English and some implications for teaching. *Applied Linguistics*. Vol. 24. No. 2. Pp. 223-242. DOI: 10.1093/applin/24.2.223 (in English)
- Nesselhauf, N. (2005) *Collocations in a Learner Corpus*. Amsterdam: John Benjamins. 331 p. DOI: 10.1075/scl.14 (in English)
- Odlin, T. (1989) *Language Transfer: Cross-Linguistic Influence in Language Learning*. Cambridge: Cambridge University Press. 210 p. (in English)
- Omarova, S., Ospanova, D., Aitova, N., Tokenkyzy, G., Ormanova, A., Alshynbekova, M. (2025) A corpus approach in language discovery: a word frequency analysis based on the corpus outcomes in Kazakh. *Forum for Linguistic Studies*. Vol. 7. No. 2. Pp. 869-881. DOI: 10.30564/fls.v7i2.8317 (in English)
- Qasym, B. (2013) *Qazaq tilining leksikologijasy men leksikografijasy*. Almaty: Qazaq universiteti. 272 b. [Kassym, B. (2013) *Lexicology and Lexicography of the Kazakh Language*. Almaty: Kazakh University. 272 p.] (in Kazakh)
- Qazaq adebi tilining sozdigi: on bes tomдық (2011) Almaty: Arys. [Dictionary of the Kazakh Literary Language: in 15 volumes (2011) Almaty: Arys.] (in Kazakh)
- Qazaq tilining tusindirme sozdigi (2008) Almaty: Dajk-Press. 968 b. [Explanatory Dictionary of the Kazakh Language (2008) Almaty: Dyke-Press. 968 p.] (in Kazakh)
- Qian, D.D. (2002) Investigating the relationship between vocabulary knowledge and academic reading performance: an assessment perspective. *Language Learning*. Vol. 52. No. 3. Pp. 513-536. DOI: 10.1111/1467-9922.00193 (in English)
- Read, J. (2000) *Assessing Vocabulary*. Cambridge: Cambridge University Press. 292 p. DOI: 10.1017/CBO9780511732942 (in English)
- Schmitt, N. (2000) *Vocabulary in Language Teaching*. Cambridge: Cambridge University Press. 224 p. (in English)
- Sinclair, J. (1991) *Corpus, Concordance, Collocation*. Oxford: Oxford University Press. 179 p. DOI: 10.1093/oso/9780194371445.001.0001 (in English)
- Uali, N. (2007) *Qazaq soz madenietining teorijalyq negizderi*. Almaty: Dajk-Press. 342 b. [Uali, N. (2007) *Theoretical Foundations of Kazakh Speech Culture*. Almaty: Dyke-Press. 342 p.] (in Kazakh)
- Wray, A. (1999) Formulaic language in learners and native speakers. *Language Teaching*. Vol. 32. No. 4. Pp. 213-231. DOI: 10.1017/S0261444800014154 (in English)
- Wray, A., Perkins, M.R. (2000) The functions of formulaic language: an integrated model. *Language & Communication*. Vol. 20. No. 1. Pp. 1-28. DOI: 10.1016/S0271-5309(99)00015-4 (in English)
- Wray, A. (2002) *Formulaic Language and the Lexicon*. Cambridge: Cambridge University Press. 330 p. DOI: 10.1017/CBO9780511519772 (in English)
- Wray, A. (2012) What do we (think we) know about formulaic language? An evaluation of the current state of play. *Annual Review of Applied Linguistics*. Vol. 32. Pp. 231-254. DOI: 10.1017/S026719051200013X (in English)
- Xia, D., Chen, Y., Pae, H.K. (2023) Lexical and grammatical collocations in beginning and intermediate L2 argumentative essays: a bigram study. *International Review of Applied Linguistics in Language Teaching*. Vol. 61. No. 4. Pp. 1421-1453. DOI: 10.1515/iral-2021-0188 (in English)
- Zhubanov, Q. (1999) *Qazaq tili zhonindegi zertteuler*. Almaty: Gylym. 581 b. [Zhubanov, K. (1999) *Research on the Kazakh Language*. Almaty: Gylym. 581 p.] (in Kazakh)

Мақала туралы ақпарат / Информация о статье / Information about the article

Редакцияға түсті / Поступила в редакцию / Entered the editorial office: 29.05.2026.

Жариялауға қабылданды / Принята к публикации / Accepted for publication: 21.09.2026.

Интернетте қолжетімді / Доступно онлайн / Available online: 30.09.2026.