

O‘ZBEKISTON DAVLAT JAHON TILLARI UNIVERSITETI
HUZURIDAGI ILMIY DARAJALAR BERUVCHI
DSc.03/2026.31.01.Fil.45.01 RAQAMLI ILMIY KENGASH ASOSIDAGI
BIR MARTALIK ILMIY KENGASH

O‘ZBEKISTON DAVLAT JAHON TILLARI UNIVERSITETI

ASROROVA NARGIZA ISOMITDINOVNA

SUN’IY INTELLEKT CHATBOTLARI BILAN OG‘ZAKI
SUHBATLARNING KORPUSGA ASOSLANGAN DISKURS TAHLILI

10.00.11 – Amaliy va kompyuter lingvistikasi.

FILOLOGIYA fanlari bo‘yicha falsafa doktori (PhD) dissertatsiyasi
AVTOREFERATI

Toshkent – 2026

Falsafa doktori (PhD) dissertatsiyasi avtoreferati mundarijasi
Content of dissertation abstract of the doctor of philosophy (PhD)
Оглавление автореферата диссертации доктора философии (PhD)

Asrorova Nargiza Isomitdinovna

Sun’iy intellekt chatbotlari bilan og‘zaki suhbatlarning korpusga asoslangan diskurs tahlili.....3

Asrorova Nargiza Isomitdinovna

Corpus-Based Discourse Analysis of Spoken Conversations with AI Chatbots.....25

Асророва Наргиза Исомитдиновна

Корпусный дискурс-анализ устных разговоров с ИИ-чатботами.....45

E’lon qilingan ishlar ro‘yxati

List of published works

Список опубликованных работ.....50

O‘ZBEKISTON DAVLAT JAHON TILLARI UNIVERSITETI
HUZURIDAGI ILMIY DARAJALAR BERUVCHI
DSc.03/2026.31.01.Fil.45.01 RAQAMLI ILMIY KENGASH ASOSIDAGI
BIR MARTALIK ILMIY KENGASH

O‘ZBEKISTON DAVLAT JAHON TILLARI UNIVERSITETI

ASROROVA NARGIZA ISOMITDINOVNA

SUN’IY INTELLEKT CHATBOTLARI BILAN OG‘ZAKI
SUHBATLARNING KORPUSGA ASOSLANGAN DISKURS TAHLILI

10.00.11 – Amaliy va kompyuter lingvistikasi.

FILOLOGIYA fanlari bo‘yicha falsafa doktori (phd) dissertatsiyasi
AVTOREFERATI

Toshkent – 2026

Filologiya fanlari bo'yicha falsafa doktori (PhD) dissertatsiyasi mavzusi O'zbekiston Respublikasi Fanlar akademiyasi huzuridagi Oliy attestatsiya komissiyasida B2025.2.PhD/Fil6395 raqam bilan ro'yxatga olingan.

Falsafa doktori (PhD) dissertatsiyasi O'zbekiston davlat jahon tillari universitetida bajarilgan.

Dissertatsiya avtoreferati uch tilda (o'zbek, ingliz, rus (rezyume)) Ilmiy kengashning veb-sahifasida (www.uzswlu.uz) hamda "ZiyoNet" axborot-ta'lim portalida (www.ziynet.uz) joylashtirilgan.

Ilmiy rahbar:

Hakimov Hamidulla Inomovich

filologiya fanlari bo'yicha falsafa doktori, dotsent

Rasmiy opponentlar:

Rahimov G'anisher Xudoyqulovich

filologiya fanlari doktori, professor

Djumabayeva Jamilya Sharipovna

filologiya fanlari doktori, professor

Yetakchi tashkilot:

Alisher Navoiy nomidagi Toshkent

davlat o'zbek tili va adabiyoti universiteti

Dissertatsiya himoyasi O'zbekiston davlat jahon tillar universiteti huzuridagi DSc.03/2026.31.01.Fil.45.01 raqamli Ilmiy kengash asosidagi bir martalik ilmiy kengashning 2026-yil "29" oktabr soat 9:00 dagi majlisida bo'lib o'tadi (Manzil: 100138, Toshkent shahri, Uchtepa tumani, Kichik xalqa yo'li ko'chasi, 21-A uy. Tel.: (99871) 230-12-91. Faks: (99871) 230-12-92. Email: uzswlu_info@mail.ru).

Dissertatsiya bilan O'zbekiston davlat jahon tillari universitetining Axborot-resurs markazida tanishish mumkin (_____ raqami bilan ro'yxatga olingan). Manzil: 100138, Toshkent shahri, Uchtepa tumani, Kichik Halqa yo'li ko'chasi, 21-A uy. Tel : (99871) 230-12-91. Faks: (99871) 230-12-92.

Dissertatsiya avtoreferati 2026-yil "___" _____ kuni tarqatildi.

(2026-yil "___" _____ dagi ___ - raqamli reyestr bayonnomasi).

A.A. Nasirov

Ilmiy darajalar beruvchi bir martalik
ilmiy kengash raisi, filol.f.d., professor

X.B. Samigova

Ilmiy darajalar beruvchi bir martalik
ilmiy kengash ilmiy kotibi, filol.f.d.,
professor

J.A. Yoqubov

Ilmiy darajalar beruvchi bir martalik
ilmiy kengash qoshidagi bir martalik
Ilmiy seminar raisi, filol.f.d., professor

KIRISH

Dissertatsiya mavzusining dolzarbligi va zarurati. XXI asr dunyoda har bir sohada sun'iy intellekt davri bo'lishi kutilmoqda. Jahon bo'ylab xorijiy til o'rganuvchilar sun'iy intellektidan foydalangan holda amaliyot olib bormoqdalar, ayniqsa, xorijiy tilda so'zlashuv ko'nikmasini rivojlantirish uchun sun'iy intellektidan suhbatdosh sifatida foydalanish keng tarqalmoqda. Til o'rganuvchilarning sun'iy intellekt chatbotlari bilan og'zaki tarzda olib boriladigan suhbatlari til bilish ko'nikmasini rivojlantirish, pragmatik ko'nikmalarni shakllantirish va diskurs xususiyatlarini namoyon qilish jihatlari bilan o'zgacha bir til o'rganish muhitiga aylanib ulgurgan. Mazkur suhbatlar til o'rganish maqsadida amalga oshirilsada, ular allaqachon diskursning suhbat janridagi o'ziga xos, yangi va texnologiya vositachiligida amalga oshiriladigan janriga aylandi. Natijada bugungi kunda zamonaviy tilshunoslik fani oldida turgan eng dolzarb vazifalardan biri – bu mazkur suhbatlarning diskurs xususiyatlarini tadqiq qilishdir. Garchi sun'iy intellektidan suhbatdosh sifatida til amaliyotida keng foydalanilsada, mazkur suhbat diskursining lingvistik tavsifi noma'lumligicha qolmoqda. Bugungi kunda sun'iy intellekt chatbotlarining ish samaradorligi, ulardan pedagogik maqsadlarda foydalanish va sun'iy intellektning til o'rganuvchilarga ta'siri bo'yicha keng ko'lamda izlanishlar olib borilmoqda, ammo mazkur tadqiqotlarning aksariyati ta'limiy maqsadlarga yo'naltirilgan bo'lib, ularda sun'iy intellekt bilan suhbatlar diskursi tadqiqi qamrab olinmagan. Korpus asosida sun'iy intellekt chatbotlari bilan suhbatlar diskursi esa faqat ona tili misolida yoki yozma muloqotga asoslangan suhbatlar asosida olib borilgan. Mazkur suhbatlarda til o'rganuvchilar nutqining syntaktik-leksik, prozodiyaga oid va pragmatik jihatlari yetarlicha tahlil qilinmagan. Mazkur bo'shliqni to'ldirishning ahamiyati shundaki, bugungi kunda sun'iy intellekt chatbotlari jadal rivojlanmoqda, ovoz orqali suhbatlashmoqda, hattoki, insonga xos bo'lgan jihatlardan urg'u, intonatsiya va turli lahjalardan samarali foydalanmoqda. Mazkur suhbatlar insonlar o'rtasidagi suhbatlardan tubdan farq qilishi aniq, ammo bu yetarlicha isbotlanmagan. Mazkur o'ziga xos suhbat diskursining lingvistik tavsifini yaratishga katta ehtiyoj bor.

Insonlar o'rtasida sodir bo'ladigan og'zaki suhbatlar diskursini tadqiq qilish uchun yaratilgan mezonlar, mavjud nazariyalar sun'iy intellekt bilan suhbatlar tahlilini yoritib berish uchun ishlamaganligi sababli, mazkur yo'nalishni nazariy va metodologik jihatlardan empirik usulda tadqiq etish zarur. Bunday empirik asosga ega bo'lmas ekanmiz, sun'iy intellekt chatbotlari va inson o'rtasidagi suhbatlar diskursi tavsifi aniq va ishonchli bo'la olmaydi. Mazkur dissertatsiyada til o'rganuvchilar va sun'iy intellekt o'rtasidagi og'zaki suhbatlarning empirik dalillangan, korpusga asoslangan tadqiqi mamlakatimizda hali olib borilmaganligiga to'xtalib o'tilgan. Mazkur suhbatlarning diskurs tuzilishini aniqlash va tahlil qilish vazifa sifatidan belgilab olingan.

Mamlakatimizda axborot-kommunikatsiya texnologiyalari va sun'iy intellekt infratuzilmasini rivojlantirish davlat siyosatining ustuvor yo'nalishlaridan biri sifatida belgilangan bo'lib, bu lingvistik izlanishlar va korpus tadqiqotlari uchun yangi imkoniyatlarni yaratmoqda. Zero, "Sun'iy intellekt va raqamli

texnologiyalarni rivojlantirish bugungi kunda barcha sohalar qatori ta'lim tizimini ham tubdan yangilashni taqozo etadi, bunda til resurslari va elektron korpuslar muhim strategik ahamiyat kasb etadi"¹. Mazkur ilmiy strategik yo'nalish, xususan, ikkinchi til sifatida ingliz tilini o'rganuvchilarning sun'iy intellekt chatbotlari bilan og'zaki diskursini korpusga asoslangan usullar orqali tahlil qilish kabi yangi izlanishlarni amaliyot bilan uzviy bog'lash imkonini beradi, bu esa milliy ta'lim amaliyotini yuqori texnologik bosqichga ko'tarish va til ta'limi sohasida innovatsion standartlarni belgilash uchun zarur shartlarni yaratadi.

Dissertatsiyaning tadqiqot maqsadlari O'zbekiston Respublikasining qonuniy va strategik hujjatlarida belgilangan ustuvor vazifalar bilan muvofiq keladi. Bular qatoriga "Ta'lim to'g'risida"gi O'zbekiston Respublikasi Qonuni, Kadrlar tayyorlash milliy dasturi, O'zbekiston Respublikasi Prezidentining 2019-yil 8-oktyabrdagi PF-5847-son "O'zbekiston Respublikasi oliy ta'lim tizimini 2030-yilgacha rivojlantirish konsepsiyasini tasdiqlash to'g'risida"gi Farmoni, O'zbekiston Respublikasi Vazirlar Mahkamasining 2019-yil 19-iyuldagi 606-son "Oliy malakali ilmiy-tadqiqot va ilmiy-pedagog kadrlarni maqsadli tayyorlash tizimini takomillashtirish to'g'risida"gi qarori, O'zbekiston Respublikasi Prezidentining 2020-yil 29-oktyabrdagi PF-6097-son "Ilm-fanni 2030-yilgacha rivojlantirish konsepsiyasini tasdiqlash to'g'risida"gi Farmoni, 2021-yil 19-maydagi PQ-5117-son "O'zbekiston Respublikasida xorijiy tillarni o'rganishni ommalashtirish faoliyatini sifat jihatidan yangi bosqichga olib chiqish chora-tadbirlari to'g'risida"gi qarori, 2022-yil 28-yanvardagi PF-60-son "2022-2026-yillarga mo'ljallangan Yangi O'zbekistonning taraqqiyot strategiyasi to'g'risida"gi Farmoni, "Raqamli O'zbekiston – 2030" Strategiyasi, 2022-yil 14-oktyabrdagi PQ-358-son "Sun'iy intellekt texnologiyalarini 2030-yilgacha rivojlantirish strategiyasini tasdiqlash to'g'risida"gi qarori kiradi. Tadqiqot natijalari ushbu qonuniy hujjatlarda belgilangan maqsadlarga xizmat qiladi.

Tadqiqotning respublika fan va texnologiyalari rivojlanishi ustuvor yo'nalishlariga mosligi. Mazkur tadqiqot respublika fan va texnologiyalari rivojlanishining I. "Axborotlashgan jamiyat va demokratik davlatni ijtimoiy, huquqiy, iqtisodiy, madaniy, ma'naviy-ma'rifiy rivojlantirishda innovatsion g'oyalar tizimini shakllantirish va ularni amalga oshirish yo'llari" ustuvor yo'nalishiga muvofiq bajarilgan.

Muammoning o'rganilganlik darajasi. Xorijiy mamlakatlarda korpus natijalarini o'z ichiga olgan dastlabki diskurs tadqiqotlariga L.Flaurdyu, N.Ferklo, A.Payper va A.Partingtonning siyosiy matnlar tahlili; M.Stabs va P.Uikensning o'quv materiallari tahlili; T.A.van Deyk, A.Morrison, C.Kaldas-Koulthardning ilmiy matnlar tahlili; C.Shalom va T.Makenerining² diskursda identifikatsiya va

¹ <https://president.uz/en/lists/view/8582> [Murojaat etilgan sana: 05.12.2025].

² Flowerdew L. Corpus-Based Discourse Analysis. Second Edition. – London: Routledge, 2023. – P.10-44.; Fairclough N. Discourse and social change. – Malden: Polity Press, 1992. – 135 p.; Piper A. Some people have credit cards and others have giro cheques: "Individuals" and "people" as lifelong learners in late modernity. Discourse and Society. – SAGE Journals, 2000. Vol.11. – P. 515-42.; Partington A. Patterns and Meanings. – Amsterdam: Benjamins, 1998. – 162 p.; Stubbs M. Discourse Analysis: The Sociolinguistic Analysis of Natural Language. – Chicago: University of Chicago Press, 1983. – 272 p.; Wickens P. Comparative analysis of the use of projecting

gender farqlari tahlili misol bo'la oladi. Korpus lingvistikasi nuqtai nazaridan diskursni keng ko'lamda o'rganishga misol sifatida T.MakEneri, J.Rey, D.Biber va S.Conrad kabi olimlarning izlanishlari va tadqiqotlarini keltirishimiz mumkin³.

So'nggi yillarda bir qator xorij olimlari, jumladan, P.Beyker, R.Graymz, A.Chavez, B.Li, H.Vang, Y.Fan, F.Jiang va P.Li⁴ kabilar korpusga asoslangan diskurs tahlilining uslublari, yondashuvlari va amaliy ahamiyatini tadqiq etdilar. X.Lu⁵ esa korpusda sintaktik va leksik murakkablik darajalarini kompyuter vositalari orqali tahlil qilish uchun dastur yaratdi. Shuningdek, sun'iy intellekt chatbotlarining matn va ovozli so'zlashmalarini o'rganish bo'yicha bir qancha meta-tahlil va meta-sintez tadqiqotlari o'tkazildi. F.Koch va P.Savash, C.Zai va S.Vibovo, V.Huang, K.Hyu, L.Frayer⁶ kabilar sun'iy intellekt chatbotlarining ovozli rejimda ishlash samaradorligi, afzalliklari, cheklovlari hamda bunday muloqotda lingvistik nazariyalarning qo'llanilishini o'rgandilar. A.Iram, G.Tudino va Y.Qin⁷ sun'iy intellekt chatbotlari bilan yozma muloqot jarayonida inson nutqini tadqiq etdilar.

O'zbekistonda korpus lingvistikasi va korpusga asoslangan diskurs tadqiqotlari asosan til o'qitish metodikasi hamda matn lingvistikasi doirasida rivojlantirilgan bo'lib, korpus asosida diskurs tahlili yangi yo'nalish bo'lib

clauses in pedagogical legal genres. Paper read at the 25th International Systemic Functional Congress. – University of Wales, Cardiff, July 1998. – 13 p.; Van Dijk T. A. Discourse and power. – London: Palgrave Macmillan, 2008. – P. 17-19.; Morrison A., Love A. A discourse of disillusionment: Letters to the Editor in two Zimbabwean magazines 10 years after independence. Discourse and Society. – SAGE Journals, 1996. Vol.7. – P. 39-76.; Caldas-Coulthard C. R. and Moon, R. Curvy, hunky, kinky: Using corpora as tools in critical analysis. Paper read at the Critical Discourse Analysis Meeting. – University of Birmingham, 1999. – P. 17-21.; Shalom C. That great supermarket of desire: attributes of the desired other in personal advertisements. – London: Routledge, 1997. – P. 186-203.; McEnery T., Hardie A. Corpus Linguistics: Method, Theory and Practice. – Cambridge: Cambridge University Press, 2012. – 312 p.

³McEnery T., Xiao R., Tono Y. Corpus-Based Language Studies. – London: Routledge, 2006. – 406 p.; Rey J. M. Changing gender roles in popular culture: Dialogue in Star Trek episodes from 1966 to 1993. – London: Longman, 2001. – P. 138-56.; Biber D., Conrad S., Cortes V. If you look at. . . : Lexical bundles in university teaching and textbooks. Journal of Applied Linguistics. – Oxford: Oxford University Press, 2004. Vol.25 No.3. – P. 371-405.

⁴Baker P. Using corpora in discourse analysis. – London: Continuum, 2021. – 207 p.; Grimes R.E., Grimes J.E. The thread of discourse. – The Hague: Mouton, 1975. – 420 p.; Chaves A.P., Doerry E., Egbert J., and Gerosa M. It's How You Say it: Identifying Appropriate Register for Chatbot Language Design. / In Proceedings of the 7th International Conference on Human-Agent Interaction (HAI '19). – Association for Computing Machinery, New York, NY, USA, 2019. – P.102–109.; Li, B. and Wang, H. TACOMORE: Leveraging the Potential of LLMs in Corpus-based Discourse Analysis with Prompt Engineering. – Cornell University Archive, 2024.; Fan, Y. Jiang, F. and Li P. Uncovering the Potential of ChatGPT for Discourse Analysis in Dialogue: An Empirical Study Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation. – Cornell University Archive, 2024. – P.16998-17010.

⁵Lu X. Computational methods for corpus annotation and analysis. – Dordrecht: Springer, 2014. – P. 4-158

⁶Koç FŞ, Savaş P. The use of artificially intelligent chatbots in English language learning: A systematic meta-synthesis study of articles published between 2010 and 2024. ReCALL. – Elsevier, 2025. Vol.37, No.1. – P. 4-21.; Zhai C., Wibowo S. A systematic review on cross-culture, humor and empathy dimensions in conversational chatbots: The case of second language acquisition. Heliyon. – Elsevier, 2022. Vol.8, No.12. – P. 5.; Huang W., Hew K.F., Fryer L.K. Chatbots for language learning-Are they really useful? A systematic review of chatbot-supported language learning. Journal of Computer Assisted Learning. – Online Journal, 2015. Vol. 38, No.1. – P. 237–257.

⁷Iram A. Hussain M., Shahbaz M. Evaluating Human and AI-Crafted Texts: A Comparative Analysis Using the 6+1 Trait Writing Model and Discourse Analysis. International Research Journal of Arts, Humanities and Social Sciences. – Elsevier, 2025. Vol.2, No.2. – P. 706–737.; Tudino G, Qin, Y. A corpus-driven comparative analysis of AI in academic discourse: Investigating ChatGPT-generated academic texts in social sciences. Lingua. – Elsevier, 2024. Vol.312. P. 45-46.

qolmoqda. Mamlakatimizda korpus tushunchasi va diskurs tahlilining nazariy asoslari Sh.Safarov, R.Qongurov, M. Tursunov, S.Karimov, Sh.Hamroyeva, B.Mengliyev, N.Ataboyev and A.Akhmedova⁸ kabi olimlar tomonidan yaratilgan va chuqur nazariy tahlil qilingan. Mazkur olimlarning ilmiy ishlari ushbu dissertatsiyaning korpusga yo'naltirilgan tadqiqotlari uchun metodologik asos bo'lib xizmat qiladi hamda "diskurs" va "korpus" tushunchalarining asosiy terminologik mezonlarini belgilashga yordam berdi.

Amaliy korpus tadqiqotlari ichida M.Abjalova⁹ olib borgan izlanishlarni alohida ta'kidlab o'tish zarur. Mazkur olim so'z turkumlarini teglash, lemmatizatsiya va korpus annotatsiyasi uchun dasturlasr yaratish bo'yicha izlanishlar olib borganlar. Shuningdek, M.Abjalova Alisher Navoiy mualliflik korpusi hamda O'zbekistondagi ilk ta'limiy korpusni yaratib, yurtimizda korpus lingvistikasi sohasiga ulkan hissa qo'shganlar. S.Tayirova va M.Usmanov¹⁰ esa korpusga asoslangan til o'qitishning til o'rganuvchilar lingvistik kompetensiyasini rivojlantirishdagi samaradorligini asoslab berganlar. K.Rizayeva tomonidan olib borilgan bir tadqiqotda media matnlari diskursi korpus lingvistikasi hamda regressiya metodlariga tayangan holda tahlil qilingan¹¹. So'ngi yillarda lingvistik tahlilda sun'iy intellektning roli oshishi bilan mazkur sohada izlanishlar ham ko'plab sodir bo'ldi. N.Ataboyev kompyuter texnologiyalari va sun'iy intellektning zamonaviy lingvistik tadqiqotlarda o'rnini tadqiq qildilar, shuningdek, diaxron korpus yo'nalishida ham korpusshunoslik rivojiga katta hissa qo'shmoqdalar.¹² Garchi mamlakatimizda korpus lingvistikasi faol rivojlanayotgan bo'lsa-da, til o'rganuvchilarning sun'iy intellekt chatbotlar bilan muloqot jarayoni tizimli ravishda korpus asosida tadqiq qilinmagan va ularning diskursi o'rganilmagan.

⁸ Safarov Sh. Pragmalingvistika. – Toshkent: O'zbekiston Milliy Ensiklopediyasi Davlat ilmiy nashriyoti, 2008. – 432 p.; Qo'ng'urov, R. Diskurs nazariyasi asoslari. // Til va adabiyot ta'limi jurnali. – Toshkent, 2017. Vol.4. – P. 12–18.; Tursunov M., Karimov S., Karshiiev A. Development of a Modern Corpus of Computational Linguistics. 2020 International Conference on Information Science and Communications Technologies – Tashkent, 2020. <https://doi.org/10.1109/ICISCT50599.2020.9351376> [Date of application: 05.10.2025]; Hamroyeva Sh., Mengliyev B., Abdullayeva O. Scopus-based bibliometric analysis on corpus linguistics for the period of 2017-2021.– E3S Web of Conferences, 2023. – 5 p. <https://doi.org/10.1051/e3sconf/202341303008> [Date of application: 08.11.2024]; Mengliyev, B. O'zbekistonning zamonaviy siyosiy diskursi: kognitiv-diskursiv jihat. – Toshkent: O'ZR FA "Fan" nashriyoti, 2021. – 192 b; Nuritdinov, A. Korpus lingvistikasida lingvostatistik tahlil metodi. // Maktabgacha va maktab ta'limi jurnali. – Toshkent, 2025. Vol.3(5). – B. 5.; Nurmatova, G. Corpus-based study of engineering terms: Linguistic and pedagogical implication. Filologicheskie nauki. – Navoiy, 2014. Vol. 81. – 7 p.; Ataboyev N. Corpus-Based Research on the Language Features of Corpus Linguistics: In the Example of ECOCL – International Journal of Scientific and Technology Research. – IJSTR Online Journal, 2020. – Vol.9, Issue 02. – pp. 4176-4179.; Akhmedova A. Korpusga asoslangan tarjimashunoslikni o'qitishning o'ziga xos xususiyatlari – Переводоведение: проблемы, решения и перспективы, 2022. – 1-son – bb. 34-36.

⁹ Abjalova M. Tagging and Annotation of Corpus Units. – International Journal of Language Learning and Applied Linguistics, 2023. – 5 p.

¹⁰ Tayirova S., Usmonov M. Corpus-Based Activities Use Real Language Data to Engage Students in Active Vocabulary Learning, Analysis, and Application. International Journal of Recently Scientific Researcher's Theory. – Tashkent, 2025. Vol.2, No.5. – P. 1-7.

¹¹ Rizayeva K. Cross-Linguistic Features of English and Uzbek Media Discourse. Innovations in Science and Technologies. – Tashkent, 2025. Vol.2, No.6 (Special issue).

¹² Ataboyev N. Mediamatnlar diaxronik korpusida til rivojining empirik tahlil tamoyillari (gazeta materiallari misolida). – Filologiya fanlari doktori (DSc) dissertatsiyasi, Farg'ona davlat universiteti, 2024. – 236 b.; Abdurahmonova N. Corpus Based Teaching Uzbek as a Foreign Language. Journal of Foreign Language Teaching and Applied Linguistics. – Samarkand, 2019. Volume 6 – No. 1. – P.131-137.

MDH mamlakatlarida kompyuter lingvistikasi mustahkam nazariya va kompyuterlashgan tahlil asosida rivojlanib kelmoqda. Bunday tahlillarning metodologik asoslari V.P.Zakharov, A.K.Kutuzov, E.V.Nedoshivina, V.Plungyan hamda M.Kopotev tomonidan ishlab chiqilib, ular korpus tuzish tamoyillari va statistik lingvistik tahlil yondashuvlarini takomillashtirdilar. S.Kiseleva chatbotlar bilan muloqot jarayonida o'quvchilar yo'l qo'yadigan xatolar va muloqotning tavsifiy tadqiqini, jumladan, ingliz-rus korpuslari doirasidagi muloqotlarni tahlil qildi. P.Sysoyev va E.Filatov til o'rganuvchilar va Yandex GPT chatboti o'rtasidagi yozma suhbatlarni o'rgangan bo'lsa, A.Avramenko, A.Akhmedova va E.Bulanova HelloTalk chatbotining pedagogik salohiyatini tadqiq etganlar¹³. Mazkur tadqiqotlarning birortasi til o'rganuvchilarning suhbat diskursini qamrab olmagan.

Bugungi kunda xalqaro miqyosda korpus lingvistikasi va diskurs tahlili izchil rivojlanib bormoqda. O'zbekistonda ham bu borada tadqiqotlar soni oshmoqda. Ammo bugungi kunga qadar til o'rganuvchilarning sun'iy intellekt chatbotlari bilan o'zaro suhbat diskursining lingvistik va pragmatik xususiyatlarini, ayniqsa O'zbekiston kontekstida, korpus asosida tizimli tadqiq etish hali amalga oshirilmagan. Mazkur bo'shliq ushbu tadqiqotning ilmiy dolzarbligi va yangiligini belgilab beradi.

Tadqiqotning dissertatsiya bajarilgan oliy ta'lim muassasasi ilmiy-tadqiqot ishlari rejalari bilan bog'liqligi. Dissertatsiya O'zbekiston davlat jahon tillari universiteti ilmiy tadqiqot ishlari rejasining "Lingvistik ekspertiza va nazariy tilshunoslik bo'yicha til muammolarini yechishda ilg'or tajribalarni tatbiq etish orqali ingliz tilining kommunikativ imkoniyatlarini oshirish" mavzusidagi ustuvor yo'nalishi doirasida bajarilgan.

Tadqiqotning maqsadi ingliz tilida sun'iy intellekt chatbotlari bilan og'zaki suhbatlarning korpusga asoslangan diskurs tahlilini olib borishdan iborat.

Tadqiqotning vazifalari quyidagilardan iborat:

sun'iy intellekt chatbotlari bilan suhbat korpusida til o'rganuvchilar nutqida eng ko'p uchraydigan sintaktik-leksik, prosodik va pragmatik qonuniyatlar (language patterns)ni tahlil qilish;

sun'iy intellekt chatbotlari bilan og'zaki muloqotda qo'llanadigan diskurs harakatlarini (discourse moves) aniqlash va bu diskurs harakatlarining lug'atga

¹³ Захаров В.П. Корпусная Лингвистика. – Санкт Петербург: СПбГУ Пресс, 2005. – 178 п.; Недошивина, Е.В. Современный политический дискурс: стратегии и тактики речевого воздействия. – Москва, 2017. – Р. 13.; Плунгян, В.А., Сичинава, Д.В. (Ред.). Национальный корпус русского языка. – Москва, 2003—; Копотев М.В. Введение в корпусную лингвистику: Учебное пособие для студентов филологических и лингвистических специальностей университетов. – Прага: Animedia Company, 2014. – 285 с.; Kiseleva S.V., Smirnova A.A., and Trofimova N.A. "Chatbot Communication" as an Object of Linguistic Research in the System of Digital Communications. Diskurs. – Moscow, 2022. Vol.8, No.3. – P.128-146.; Sysoyev P.V., Filatov E.M. Artificial intelligence in teaching Russian as a foreign language. Journal of Russian Language Studies. – Saint Petersburg, 2024. Vol.22, No 2. – P. 300-317.; Авраменко А., Ахмедова А., Буланова Е. Технология чат-ботов как средства формирования иноязычной грамматической компетенции при самостоятельном обучении. – Вестник Тамбовского университета, 2023. – С. 387-393. DOI: <https://doi.org/10.20310/1810-0201-2023-28-2-386-394> [Date of application: 13.04.2025].

asoslangan diskurs birliklari (Vocabulary-Based Units) bo'ylab tarqalishini o'rganish;

chatbot suhbatlari davomida til o'rganuvchilar tomonidan keng qo'llaniladigan muloqot strategiyalarini tadqiq qilish va ularning til o'rganuvchilar tomonidan diskursda qay darajada samara keltirishini tekshirish;

sun'iy intellekt chatbotlari bilan olib boriladigan og'zaki suhbatlarda diskurs markerlari, diskurs kogeziyasi, xushmuomalalik strategiyalari va nutq uslublarini tadqiq etish.

Tadqiqot obyekti sifatida ingliz tili o'rganuvchilarning sun'iy intellekt chatbotlari bilan og'zaki suhbatlardagi diskurs xususiyatlari tanlangan.

Tadqiqot predmetini sun'iy intellektga asoslangan chatbotlar bilan o'zaro suhbat diskursida til o'rganuvchilar nutqiy faoliyatida namoyon bo'ladigan lisoniy qonuniyatlar, diskurs harakatlari hamda muloqot strategiyalari tashkil etadi.

Tadqiqotning usullari. Tadqiqot jarayonida kontent tahlil, tasniflash, tavsiflash, qiyoslash, statistik, korpus va empirik tahlil metodlaridan foydalanilgan.

Tadqiqotning ilmiy yangiligi quyidagilardan iborat:

til o'rganuvchilarning sun'iy intellekt chatbotlari bilan og'zaki muloqotidagi sintaktik, leksik, pragmatik va prosodik birliklarning qo'llanishi o'rganilib, ularning funksional modellari aniqlangan;

chatbotlar suhbatida keng tarqalgan lisoniy qonuniyatlar (language patterns), so'z birikmalari (n-grammalar), konkordanslar, leksik va sintaktik murakkablik darajalari, diskurs markerlari hamda nutq aktlarining taqsimlanishi dalillangan;

sun'iy intellekt chatbotlari bilan og'zaki muloqot jarayonida qo'llanadigan muloqot strategiyalarining (interactional strategies) samaradorligi lug'atga asoslangan nutq birliklari (Vocabulary-Based Units) bo'yicha tahlil qilinib, ularning samaradorligi chastota, funksional xususiyatlar va diskurs bo'ylab takrorlanish asosida isbotlangan;

sun'iy intellekt bilan og'zaki muloqotlarda ijtimoiy munosabatlarning mavjud emasligi isbotlandi, bu holat "sun'iy intellekt xushmuomalalik paradoksi" sifatida ta'riflanib, mavqe (P) ko'rsatkichining pastligi va masofa (D) ko'rsatkichining yuqoriligi orqali dalillangan.

Tadqiqotning amaliy natijalari quyidagilardan iborat:

Sun'iy intellekt chatbotlari bilan suhbat diskursini korpus asosida tahlil qilishning nazariyasi takomillashtirilib, metodologik mezonlari ishlab chiqildi. Uning asosida O'zbekistonlik talabalarning nutqi tavsifiy va miqdoriy tahlil qilindi.

Ingliz tili misolida til o'rganuvchilarning sun'iy intellekt chatbotlari bilan og'zaki suhbatlari asosida kichik hajmli korpus to'planib, annotatsiyalandi va mazkur korpusdan asosiy diskurs tadqiqida usul sifatida foydalanildi.

Tadqiqot natijasida aniqlangan yangiliklar, sun'iy intellekt nutqidagi insonga xos bo'lmagan jihatlarni yaxshilash orqali sun'iy intellektning suhbatdosh sifatidagi rolini oshirish yuzasidan tavsiyalar yaratildi.

Tadqiqot natijalarining ishonchliligi. Tadqiqot davomida ilgari surilgan xulosalar tadqiqotchi tomonidan e'lon qilingan ilmiy maqola va tezislarda o'z ifodasini topgani hamda ushbu xulosalarning ilmiy jamoatchilik tomonidan ijobiy

baholanganligi, ikkinchi til o'rganuvchilarning sun'iy intellekt chatbotlari bilan og'zaki muloqotiga oid korpusga asoslangan diskurs tahlili doirasida qo'yilgan muammolarning miqdoriy va sifat jihatdan asoslangan metodlar – korpus tahlili, diskurs tahlili, chastotaviy-statistik hisob-kitoblar, sintaktik-leksik murakkablik tahlili, pragmatik va prosodik tavsiflash, qiyosiy va empirik tahlil usullari asosida aniqlab berilgani, ishlab chiqilgan xulosa va tavsiyalarning ta'lim jarayonida tajriba-sinov asosida qo'llangani hamda olingan natijalarning vakolatli ilmiy muassasalar tomonidan tasdiqlanganligi bilan izohlanadi.

Tadqiqot natijalarining ilmiy va amaliy ahamiyati. Mazkur tadqiqot natijalarining ilmiy ahamiyati unda belgilab berilgan nazariy va metodologik ko'rsatmalar kelajakda ushbu yangi yo'nalishda olib boriladigan tadqiqotlar uchun mezon bo'lib xizmat qiladi. Shuningdek, tavsifiy va miqdoriy tadqiqot usullarini uyg'unlashtirgan holda aralash tarzda korpus asosida og'zaki diskurs tadqiqotlari olib borishda manba bo'lib xizmat qiladi.

Tadqiqot natijalarining amaliy ahamiyati oliy ta'lim muassasalarida "Amaliy tilshunoslik", "Ikkinchi tilni o'zlashtirish", "Korpus lingvistikasi", "Matn lingvistikasi", "Zamonaviy tilshunoslik", "Lingvistik tahlil usullari" fanlarini o'qitishda nazariy va amaliy manba sifatida foydalanish, sun'iy intellektga asoslangan til o'rganish platformalari uchun maqsadli o'quv materiallari ishlab chiqish, til o'rganuvchilarning kommunikativ va pragmatik kompetensiyalarini rivojlantirishga xizmat qiluvchi chatbot dizaynini takomillashtirish hamda sun'iy intellekt chatbotlari orqali haqiqiy hayotdagi muloqotga o'tishni yengillashtiruvchi samarali diskurs modellarini yaratishda qo'llash mumkinligi bilan izohlanadi.

Tadqiqot natijalarining joriy qilinishi. Ingliz tilida sun'iy intellekt chatbotlari bilan og'zaki suhbatlarning korpusga asoslangan diskurs tahlili jarayonida olingan natijalar va amaliy tavsiyalar asosida:

til o'rganuvchilarning sun'iy intellekt chatbotlari bilan og'zaki muloqotidagi diskursiv xususiyatlar korpus asosida o'rganilib, sun'iy intellekt suhbatdosh sifatida ishtirok etadigan muloqotda sintaktik-leksik, pragmatik va prosodik birliklarning funksional qo'llanish modellari aniqlanganligiga oid ilmiy xulosalaridan Maktabgacha va maktab ta'limi vazirining 2025-yil 20-oktabrdagi 394-sonli buyrug'iga asosan amalga oshirilayotgan "Umumta'lim maktablarining informatika fani o'qituvchilarini Massachusetts texnologiya institutining sun'iy intellekt ta'lim dasturi asosida kasbiy rivojlantirish va mentor sifatida tayyorlash loyihasi" doirasida foydalanilgan (O'zbekiston Respublikasi Maktabgacha va maktab ta'limi vazirligi 2025-yil 23-dekabrda 02-33-4733-sonli ma'lumotnomasi). Natijada sun'iy intellekt asosida ishlaydigan til o'rgatish platformalari diskursini takomillashtirish borasida amaliy tavsiyalar ishlab chiqilgan;

takrorlanuvchi til qoliplari, n-grammalar va konkordanslar, leksik hamda sintaktik murakkablik darajalari, diskurs markerlari va suhbat aktlarining til o'rganuvchi nutqida taqsimlanishi sun'iy intellekt chatbotlarini interaktiv suhbatdosh sifatida qo'llash orqali dalillanganligiga oid ilmiy natijalaridan 19-12/147-sonli Oliy Majlis huzuridagi Fuqarolik jamiyati institutlarini qo'llab-quvvatlash jamoat fondi tomonidan moliyalashtirilgan O'zbekiston "Tadbirkor va

ishbilarmon ayollar” birlashmasi tomonidan amalga oshirilgan “Mamlakatimizda xotin-qizlar manfaati doimiy e’tiborda” nomli loyiha doirasida foydalanilgan (O‘zbekiston “Tadbirkorlar va ishbilarmon ayollar” birlashmasining 2025-yil 20-sentabrda 27-sonli ma’lumotnomasi). Natijada loyiha tinglovchilarning sun’iy intellekt bo’yicha kompetensiyalarini rivojlantirishda “Axborot xavfsizligi asoslari” va “Kiberxavfsizlik” nomli o‘quv qo‘llanmalari, hamda, maxsus ishlab chiqilgan sun’iy intellekt asosidagi o‘quv platformasini yaratishda foydalanildi;

sun’iy intellekt chatbotlari bilan og‘zaki muloqot jarayonida qo‘llanadigan muloqot strategiyalarining samaradorligi lug‘atga asoslangan nutq birliklari bo’yicha chastotaviy taqsimoti, funksional qo‘llanilishi va diskurs bo‘ylab barqaror takrorlanishi asosida isbotlanganligiga oid ilmiy natijalar asosida filologiya va tillarni o‘qitish bakalavriat ta’lim yo‘nalishlari uchun “Axborot xavfsizligi asoslari” nomli o‘quv qo‘llanmasi ishlab chiqilgan va amaliyotga joriy etilgan (O‘zbekiston Respublikasi Oliy ta’lim, fan va innovatsiyalar vazirligining 2026-yil 10-yanvardagi 02/17-4/4-son ma’lumotnomasi). Natijada ushbu o‘quv qo‘llanma metodik ta’minot vazifasini bajarib, ta’limni samarali tashkil etishda qo‘llanilgan;

shuningdek, sun’iy intellekt chatbotlari bilan og‘zaki muloqot jarayonida qo‘llanadigan muloqot strategiyalarining samaradorligi lug‘atga asoslangan nutq birliklari bo’yicha chastotaviy va funksional tahlil qilinib olingan ilmiy natijalardan O‘zbekiston Respublikasi Vazirlar Mahkamasining 2024-yil 27-martdagi “Davlat ta’lim standartlari va davlat ta’lim talablarini ishlab chiqish hamda joriy etish tartibini takomillashtirish chora-tadbirlari to‘g‘risida”gi 157-sonli qarori asosida Buyuk Britaniyaning Cambridge University Press & Assessment tashkiloti bilan hamkorlikda milliy o‘quv dasturini takomillashtirish loyihasi doirasida, chet tillari (L2 – ikkinchi yoki xorijiy til sifatida o‘qitish) yo‘nalishidagi fan dasturi mazmunini boyitishda foydalanilgan (O‘zbekiston Respublikasi Maktabgacha va maktab ta’limi vazirligi huzuridagi Respublika ta’lim markazining 2026-yil 8-apreldagi 01/11-6-219-son ma’lumotnomasi). Ushbu ilmiy natijalar sun’iy intellekt vositachiligidagi muloqotning o‘ziga xos diskursiv xususiyatlarini ochib berish hamda ularni til o‘qitish jarayoniga samarali integratsiya qilish imkonini berdi.

Tadqiqot natijalarining aprobatsiyasi. Tadqiqot natijalari 6 ta xalqaro va 2 ta respublika miqyosidagi ilmiy-amaliy anjumanlarda muhokama qilinib, taqdim etilgan.

Tadqiqot natijalarining e’lon qilinishi. Dissertatsiya mavzusi bo’yicha jami 14 ta ilmiy ish, shu jumladan O‘zbekiston Respublikasi Oliy attestatsiya komissiyasining PhD dissertatsiyalari asosiy ilmiy natijalarini chop etish tavsiya etilgan ilmiy nashrlarda 5 ta maqola nashr etilgan.

Dissertatsiyaning hajmi va tuzilishi. Dissertatsiya kirish qismi, uch bob, xulosa va foydalanilgan adabiyotlar ro‘yxatidan tashkil topgan. Dissertatsiyaning umumiy hajmi 142 betni tashkil etadi.

DISSERTATSIYANING ASOSIY MAZMUNI

Kirish qismida mavzuning dolzarbligi va zarurati, tadqiqotning maqsad va vazifalari, obyekti va predmeti, fan va texnologiyalarni rivojlantirishning ustuvor yoʻnalishlariga mosligi, tadqiqotning ilmiy yangiligi, amaliy natijalari, ishonchlilik darajasi, nazariy va amaliy ahamiyati, amaliyotga tatbiq etilishi, aprotatsiyasi, chop etilgan ishlar hamda dissertatsiyaning tuzilishi bayon etilgan.

Birinchi bob **“Sunʼiy intellekt chatbotlari bilan ogʻzaki suhbatlar diskursini korpus asosida tahlil qilishning nazariy asoslari”** deb nomlangan. 1.1-paragrafda *“Diskurs va korpus: korpusga asoslangan diskurs tahlilining nazariy asoslari”* mavzusi yoritilgan boʻlib, unda tadqiqotchilar tomonidan ilgari surilgan korpusga asoslangan diskurs tahlilini olib borishning nazariy asoslari belgilab berilgan. Bunda diskurs tahlilini inson muloqoti natijasini oʻrganish sifatida, korpusni esa lingvistik tahlil uchun tizimli ravishda toʻplangan yozma yoki ogʻzaki matnlar majmuasi sifatida taʼriflaydilar. Diskurs tahlili amaliy tilshunoslikning yoʻnalishi boʻlib hisoblanadi. T.Makeneri, A.Xardi va J.Blommaert tomonidan taklif qilingan “qoʻllanishdagi til” va “harakatdagi til” tushunchalari diskursning umumiy qabul qilingan taʼriflari sanaladi¹⁴. M.Fukolt “diskurs”ning koʻplik shaklini taklif etadi, J.Ji esa katta “D” va kichik “d” bilan yoziladigan diskurslarni farqlaydi. U “diskurs”ni “qoʻllanilayotgan til” sifatida, “Diskurs”ni esa lingvistik amaliyotlarga mos keladigan semiotik elementlar tarzida taʼriflaydi¹⁵.

Korpus tilshunoslikning asosiy tarkibiy qismidir. U lisoniy qonuniyatlarni, tuzilmalari va tilning haqiqiy hayotdagi qoʻllanilishini oʻrganish uchun maʼlumotlar bazasi vazifasini bajaruvchi korpus tilshunosligining asosiy birligi hisoblanadi. Korpus tilshunosligi turli uslublar va janrlar boʻyicha soʻzlarning qoʻllanish chastotasi, birikmalari, matn tuzilmalari va maʼno oʻzgarishlarini tadqiq etish uchun korpusdan foydalanadi¹⁶. Paragrafda, shuningdek, “corpus-driven” nazariya va “corpus-based” usul atamalarini muhokama qiladi. Xulosa qilib aytganda, “korpus asosida” (“corpus-based”) atamasi korpus tahlilining turli shakllarini oʻz ichiga olgan umumlashgan tushuncha sifatida taʼriflangan.

Bobda ularning asosiy tafovutlari batafsil bayon etilgan. Diskurs tahlili tavsifiy boʻlib, bunda matn yaxlitligi saqlab qolgan holda, uning mazmuni va maʼnosiga eʼtibor qaratadi. Korpus tahlili esa asosan miqdoriy boʻlib, ajratib olingan namunalar ustida ishlaydi va “tilning bevosita ichki tuzilishi bilan shugʻullanadi”¹⁷. Korpus tahlilining diskurs tahlilidan asosiy epistemologik farqi shundaki, matnlar ijtimoiy harakatlar yoki dinamik jarayonlar emas, balki statik mahsulotlar sifatida qaraladi. U

¹⁴McEnergy T., Hardie A. Corpus Linguistics: Method, Theory and Practice. – Cambridge: Cambridge University Press, 2012. – 312 p.; Blommaert J. Discourse. A Critical Introduction. – Cambridge: Cambridge University Press, 2005. – 299 p.

¹⁵Baker P. Using corpora in discourse analysis. – London: Continuum, 2021. – 207 p.

¹⁶Gee J. An Introduction to Discourse Analysis: Theory and Method. – London: Routledge, 2014. – 248 p.

¹⁷Biber D., Conrad S., Reppen R. Corpus linguistics: Investigating language structure and use. – Cambridge: Cambridge University Press, 2004. – 300 p

shbu farqlarga qaramay, ikki sohaning o'zaro aloqasi natijasida korpusga asoslangan diskurs tahlili (KADT) vujudga keldi. Bu "to'liq ishlab chiqilgan va nazariy jihatdan asoslangan metodologiya" hisoblanadi¹⁸.

Paragrafda korpusga asoslangan diskurs tahlilining uchta asosiy yo'nalishi – matnli, tanqidiy va kontekstual yondashuvlar haqida ma'lumot berilgan. Ushbu tadqiqot L.Flauerduning "muayyan ijtimoiy vaziyatlarda ishtirokchilar tomonidan turli ritorik va pragmatik vositalarning qo'llanilishi" tushunchasiga asoslanib, korpusga asoslangan kontekstual yondashuvni referent korpus va ikkinchi til diskursi o'rtasidagi ko'priq sifatida tanlaydi¹⁹.

Diskurs tahliliga korpusga asoslangan yondashuvning cheklovlari ham mazkur paragrafda o'rganib chiqilgan. KADT bilan bog'liq asosiy muammo shundaki, kompyuterlar til qo'llanishining etnografik tavsiflarini yarata olmaydi²⁰. Bundan tashqari, korpuslarning ko'p qirrali tabiatini inobatga olgan holda, o'rganuvchi va sun'iy intellekt o'rtasidagi muloqotlarni korpusga asoslangan diskurs tahlili tilshunoslikning endigina shakllanayotgan sohasi hisoblanadi va bu borada tadqiqotlarga ko'maklashadigan manbalar juda kam. Shunga qaramay, KADT zamonaviy va ishonchli tahlil usuli ekanligi isbotlangan.

Tadqiqotning 1.2-paragrafida "*Til o'rganuvchilar korpusini yig'ish va annotatsiyalashning nazariy jihatlari*" mavzusi ko'rib chiqilgan bo'lib, unda diskurs tahlili uchun ma'lumotlar manbaiga asosiy e'tibor qaratilgan. Paragrafda ona tili korpuslari va o'rganuvchi korpuslari o'rtasidagi farq ko'rsatilgan, bunda o'rganuvchi korpuslari til o'rganuvchilar tomonidan yaratilgan katta yoki kichik hajmdagi matnli yoki og'zaki ma'lumotlar to'plami sifatida ta'riflangan. BNC, ALICE, COCA, MICASE, TACOMORE, RST, WECCCL kabi mavjud korpuslar hamda ICLE, CANCODE va Cleverbotsingari o'rganuvchilar korpuslari o'rganib chiqilgan. Ushbu tadqiqot uchun yangi kichik hajmli o'quvchilar korpusini tuzish muhim deb topilgan, chunki mavjud korpuslar asosan pedagogik yoki o'qitishga yo'naltirilgan bo'lib, bu tadqiqot esa lingvistikasi va diskurs tadqiqotlariga xizmat qilish maqsadida til o'rganuvchilar va sun'iy intellekt chatbotlari o'rtasidagi og'zaki suhbatlarni tadqiq qilishga yo'naltirilgan alohida korpusni talab etadi.

Ushbu paragrafning nazariy asos sifatida o'rganuvchi va sun'iy intellekt muloqoti diskursining korpus jihatlarini o'rganadi. Mazkur jarayon tadqiqot savollari bilan yo'naltirilishi lozimligi alohida ahamiyatga molik. Tilshunoslar B.Bulte va A.Xausen o'rganuvchi diskursining korpus jihatlarini to'rtta asosiy qismga ajratib ko'rsatadilar: sintaksis, leksika, pragmatika va prosodiya²¹. Ular, shuningdek, diskurs tahlili uchun korpusning qaysi jihatlarini tanlab olish zarurligini ko'rsatib beradilar: sintaktik murakkablik (koordinatsiya, subordinatsiya, gap uzunligi, iboralarning murakkabligi), leksik murakkablik

¹⁸ Widdowson, H. The limitations of linguistics applied. Applied Linguistics. – Cambridge, 2000. Vol.21, No.1. – P. 3-25.

¹⁹ Flowerdew L. Corpus-Based Discourse Analysis. – London: Routledge, 2023. – P.10-44.

²⁰ Widdowson H. The limitations of linguistics applied. Applied Linguistics. – Cambridge, 2000. Vol.21, No.1. – P. 3-25.

²¹ Bulte B., Housen, A. Conceptualizing and measuring short-term changes in L2 writing complexity. Journal of Second Language Writing, 2014. Vol. 26. – P. 42-65.

(leksik xilma-xillik, zichlik, noziklik), pragmatik naqshlar (diskurs markerlar, lingvistik uslub, muloyimlik va tuzatish strategiyalari) hamda prosodik jihatlar (suprasegmental xususiyatlar, urg‘u, intonatsiya, ritm va nutqdagi uzilishlar).

Tadqiqotning yakuniy 1.3-paragrafida *“til o‘rganuvchi-sun‘iy intellekt og‘zaki suhbatlarining korpusga asoslangan diskurs tahlili borasida tadqiqot zarurati”* mavzusi ko‘tarilgan. Unda tadqiqot mavjud bo‘shliqni to‘ldirishga qaratilganligi ta‘kidlanadi. O‘zbekistonda korpusga asoslangan tadqiqotlar mavjud bo‘lsa-da, amaliy diskurs tadqiqotlari kamligi va korpusga asoslangan diskurs tadqiqotlari hali olib borilmaganligi qayd etilgan. Xalqaro miqyosda chatbot suhbatlarida jarayonida inson suhbat diskursini tahlil qiluvchi tadqiqotlar hamon yetarli emas. J.Jiyaning kalit so‘zlarga asoslangan o‘zaro suhbat tahlili yoki D.Koniamning chatbot samaradorligini baholash kabi tadqiqotlari ikkinchi til diskursining to‘liq, korpusga asoslangan tahlilini deb baholanmaydi²².

Paragraf KADT usulining afzalliklarini alohida ta‘kidlab, bildirilgan tanqidlarga munosib javob beradi. KADT usulida tadqiqotchi xolisligi ta‘minlanadi va nutq qoliplarini tizimli tadqiq qilish imkonini beradi. Tadqiqot predmeti sifatida til o‘rganuvchilarning korpusi belgilangan bo‘lib, u chatbot promptlari, ovozni tanib olish dasturida mavjud cheklovlar va foydalanuvchilarning sun‘iy intellekt chatbotlari bilan real vaqtdagi suhbatlarda qo‘llaydigan strategiyalari ta‘sirida shakllangan.

Xulosa qilib aytganda, I bob sun‘iy intellekt chatbotlari bilan og‘zaki suhbatlarning yangi, texnologik vositachilikdagi janrini tahlil qilish uchun asos yaratish maqsadida korpus lingvistikasi va diskurs tahlili nazariyalarini umumlashtiradi. Tadqiqot ikkinchi til o‘rganuvchilardan yig‘ib olinadigan kichik hajmli korpus asosida olib boriladi va KADTning kontekstual yondashuviga tayanadi. Nazariy xulosa shundan iboratki, til o‘rganuvchilarning sun‘iy intellekt chatbotlari bilan o‘zaro muloqoti alohida o‘rganishga loyiq bo‘lgan diskurs janrini tashkil etadi. Bu esa texnologik vositachilikdagi muhitda metodologik jihatdan puxta va kontekstual jihatdan ta‘sirchan bo‘lgan og‘zaki diskursini o‘rganish uchun asos yaratadi.

Tadqiqotning ikkinchi bobi **“Sun‘iy intellekt interlokutorligida til o‘rganuvchilar og‘zaki suhbat diskursining korpusga asoslangan tahlili yondashuvlari”** deb nomlangan. Bobda tadqiqot loyihasi va uslubiy asoslari uch paragrafda bayon etilgan. 2.1-paragrafda *“Sun‘iy intellekt chatbotlari bilan muloqotda til o‘rganuvchilar og‘zaki korpusini to‘plash va annotatsiyalash usullari hamda tadqiqot dizayni”* masalasi ko‘tarilib, til o‘rganuvchilar og‘zaki korpusini asosiy ma‘lumotlar bazasi sifatida tanlashni asoslashdan boshlanadi. Og‘zaki korpus texnologiyaning rivojlanishi bilan yangi paydo bo‘lgan tilshunoslik sohasi bo‘lib, o‘rganuvchilar korpuslaridan olingan natijalar sun‘iy intellekt davrida ingliz tilini o‘qitish va o‘rganish uchun yangi asos yaratishga yordam beradi. Mavjud keng ko‘lamli korpuslar sun‘iy intellekt vositasida yaratilmagani yoki faqat ta‘limiy

²² Jia J. The Study of the Application of a Keywords-based Chatbot System on the Teaching of Foreign Languages, Doctoral Thesis. – Cornell University, 2003. – 228 p.; Coniam D. Evaluating the Language Resources of Chatbots for Their Potential in English as a Second Language. ReCALL, 2008. Vol.20, No.1, Issues 98. – P.116.

maqsadlarga yo'naltirilgani sababli, ushbu tadqiqot yangi, kichik hajmli og'zaki o'rganuvchilar korpusini tuzishni talab qiladi.

Korpus tuzishning dastlabki bosqichi diskurs jamiyatini aniqlashdir. Diskurs jamiyati – mazkur diskursdan foydalanadigan kishilar guruhi bo'lib, u o'ziga xos janr, maxsus lug'at tarkibi va diskursdan umumiy maqsadlarda foydalanishi orqali belgilanadi.²³ Ushbu tadqiqot doirasida o'rganilayotgan diskurs jamiyati - bu til amaliyoti uchun sun'iy intellektga asoslangan suhbat chatbotlaridan foydalanadigan universitet darajasidagi ingliz tilini ikkinchi yoki xorijiy til sifatida o'rganuvchilardir. Korpusda reprezentativlikni oshirish uchun korpus yig'ish jarayonini boshlashdan oldin so'rovnoma o'tkazilib, unda ishtirokchilarning jinsi, ona tili, til bilish darajasi, suhbat chatbotlaridan foydalanish darajasi va foydalanish maqsadlari kabi ko'rsatkichlar orqali ularning mazkur tadqiqot maqsadiga mosligi aniqlandi.

Korpus tuzish jarayoni D.Biber mezonlariga asoslanadi: u tabiiy tildagi katta yoki kichik elektron to'plam bo'lishi, annotatsiyalangan bo'lishi va tilning maxsus ko'rinishini o'z ichiga olishi lozim²⁴. Reprezentativlikni ta'minlash maqsadida demografik namuna olish usuli qo'llaniladi. Og'zaki nutqni to'plash uchun mazkur bobda ASR (Avtomatik Nutqni Aniqlash) texnologiyalari imkoniyatlari muhokama qilingan va Flow Speechmatics dasturi tavsiya etilgan. Bunda uning annotatsiyalash dasturlari bilan yaxshi integratsilashishi hamda turli xil aksentlarni aniq taniy olishi sabab qilib ko'rsatilgan. Transkripsiya va audio yozuvlarni tahlil qilish uchun Praaline kabi qo'shimcha dasturdan foydalanish tavsiya etilgan.

Dastlabgi korpus ma'lumotlari to'planganidan so'ng, annotatsiya jarayoni boshlanadi. Annotatsiya rejasi har qanday korpus tadqiqotining eng muhim qismi bo'lib hisoblanadi. Annotatsiya rejasi I bobda belgilangan sintaktik-leksik, prosodik va pragmatik toifalarga asosida shakllantiriladi. Paragrafda korpus annotatsiyasi va tahlili uchun zarur bo'lgan quyidagi avtomatlashtirilgan jarayonlar batafsil yoritilgan: so'z turkumlarini teglash (POS-tagging), lemmatizatsiya, chastotalar ro'yxatlarini tuzish, n-grammalar tahlili, leksik birikma tahlili, leksik zichlik, leksik xilma-xillik, tur-token nisbati (TTR), leksik murakkablik, gap uzunligiga asoslangan sintaktik murakkablik tahlili, gaplarning murakkabligi, ergash gaplar chastotasi, T-birliklar, koordinatsiya, ot va fe'l birikmalarining murakkabligi, nutq tuzatishlarining prosodik tahlili, nutq sur'ati va nutqdagi uzilishlar.

Pragmatik tahlil esa korpus natijalarini diskurs tahlili bilan bog'laydi hamda suhbat aktlari tahlili, diskurs belgilari tahlili va P.Braun va S.Levinsonning yuzni saqlovchi harakatlar (FTAs) nazariyasi asosida xushmuomalalik strategiyalari tahlili va nutq uslubi tahlilini o'z ichiga oladi.

2.2-paragrafda *“Og'zaki nutq tahlili uchun yondashuv va metodologiyani tanlash”* mavzusi yoritilgan bo'lib, unda diskurs tahlilining metodologiyasi belgilangan. Tadqiqot obyekti sifatida til o'rganuvchilarning suhbat diskursida namoyon bo'ladigan lisoniy qonuniyatlar, diskurs harakatlari va muloqot

²³ Swales J. M. Genre analysis: English in academic and research settings. – Cambridge: Cambridge University Press, 1990. – 260 p.

²⁴ Biber D., Connor U., Upton T.A. Discourse on the move: Using corpus analysis to describe discourse structure. – London: John Benjamins, 2007. – 304 p.

strategiyalari olingan. Tadqiqotda miqdoriy va sifat usullarini birlashtirgan aralash usul hamda korpus namunalaridan boshlab diskurs birliklarini aniqlashgacha bo'lgan eng kichik birlikdan diskursga qarab yo'naltirilgan tahlil tizimi qo'llanilgan. Tadqiqotchi tomonidan tadqiqot dizayni ishlab chiqilgan:



1-rasm. Til o'rganuvchilarning sun'iy intellekt chatbotlari bilan og'zaki suhbatlarining korpusga asoslangan diskurs tahlili dizayni

Tadqiqotning asosiy metodologik vazifasi korpus asosida diskurs birliklarini aniqlashdan iborat. Ushbu tadqiqotda lug'atga asoslangan diskurs birliklari (VBDUs) segmentatsiyasi usuli qo'llanilgan bo'lib, ular TextTiling dasturi singari vositalar yordamida leksik ko'lam va mavzudagi o'zgarishlarga ko'ra avtomatik ravishda aniqlanadigan qismlardir. Diskurs tahlilining bunday eng kichik birlikdan diskursga qarab yo'naltirilgan yondashuvi mazkur tadqiqot uchun tanlab olingan, chunki u aniq hisob-kitoblarga asoslangan bo'lib, korpusga asoslangan izlanishlar uchun maqbuldir²⁵. Jarayon quyidagilarni o'z ichiga oladi: segmentatsiya, har bir birlikning lingvistik tahlili, tasniflash, funksional guruhlash va nihoyat, umumiy diskurs harakatlarini aniqlash.

Diskurs tadqiqoti uchun tahliliy usul sifatida korpus lingvistikasi, kontekstual doiralash va tematik tahlilni o'z ichiga oluvchi Diskursni Kontekstlashtirish Texnikalari (DKT) tashkil etadi. Bu texnikalar diskursni o'rganishda keng qamrovli yondashuvni ta'minlaydi.

“Diskursda lisoniy qonuniyatlar, lingvistik harakatlar va o'zaro ta'sir strategiyalarini korpusga asoslangan tahlil qilish usullari va tamoyillari” nomli bo'limning yakuniy 2.3-paragrafida mazkur diskurs tahlilining uchta asosiy obyekti aniqlanib, to'rtta tadqiqot savollari orqali yo'naltirilgan. Bu tadqiqot savollari tadqiqot maqsadiga erishish uchun zarur bo'lgan vazifalarni tashkil etadi. Tadqiqotning asosiy maqsadi korpus lingvistikasi asosida sun'iy intellekt

²⁵ Biber D., Connor U., Upton T. Discourse on the move: Using corpus analysis to describe discourse structure. – London: John Benjamins, 2007. – 304 p.

chatbotlari va til o'rganuvchilar o'rtasidagi og'zaki muloqotlarning lisoniy qonuniyatlari, diskurs harakatlari va muloqot strategiyalarini tahlil qilishdan iborat.

Lisoniy qonuniyatlar tahlili sintaktik-leksik, pragmatik va prosodik jihatlar tahlilini o'z ichiga oladi. Mazkur tahlil til o'rganuvchilarining takrorlanuvchi til qoliplarini va diskursning lingvistik xususiyatlarini aniqlash imkonini beradi.

Diskurs harakatlari tahlili nutqdagi vazifaviy birliklarni aniqlaydi. J.Sveyls tomonidan taklif etilgan CARS modeli singari mavjud modellar yozma akademik janrlar uchun ishlab chiqilganligi sababli, ushbu tadqiqot V.Bhatiyaning o'n bosqichli yondashuviga asoslanib, sun'iy intellekt chatbotlari suhbatlariga moslashtirilgan yangi korpusga asoslangan diskurs harakatlari tahlili mezonini ishlab chiqdi²⁶. Tadqiqotchi tomonidan ishlab chiqilgan baholash mezonlari lug'atga asoslangan nutq birliklari bo'ylab qo'llaniladigan yetti harakatni va ularning tarkibiy qismlarini tasvirlaydi. Ushbu harakatlarning har biri o'z navbatida bir necha bosqichlardan tashkil topgan:

1-jadval

Sun'iy intellekt bilash suhbatlarda diskurs harakatlari tahlili mezonlari

Diskurs harakatlari	Harakat bosqichlari
1. Suhbatni boshlash	Salomlashish, shaxsni aniqlashtirish, hol-ahvol so'rash
2. Vazifani belgilash	Mavzuni belgilash, ma'lumot berish
3. Ma'lumot olish	Promptlar berish, mavzuni to'ldirish, tanlov imkonini berish
4. Savolga javob berish	Ma'lumot taqdim etish, fikr bildirish, funksional harakatlar (rozilik yoki norozilik bildirish, kechirim so'rash, iltimos qilish)
5. Fikr-mulohaza bildirish yoki tuzatish kiritish	O'z-o'zini tuzatish, SI tomonidan tuzatish, qayta ifodalash, aniqlashtirish
6. Mavzuni kengaytirish	Qo'shimcha savollar berish, mavzuni yoritish, yaqin mavzularga o'tish
7. Suhbatni yakunlash	Umumlashtirish, keyingi mashg'ulot rejasi

Muloqot strategiyalari tahlili so'zlovchilarning muloqotni qanday boshqarishini o'rganadi, bunda yuz va freymlash strategiyalaridan foydalaniladi. Ushbu tahlil suhbat aktlari, xushmuomalalik, nutq uslubi, diskurs markerlari va nutqdagi uzilishlarni tuzatish strategiyalarini pragmatik jihatdan o'rganish natijalarini o'z ichiga oladi. Nutqdagi uzilishlarni tuzatish strategiyalari uchun E.Shegloff, G.Jefferson va H.Saksning to'rt qismdan iborat nazariyasi qabul qilinadi²⁷.

Mazkur bob diskurs tahliliga korpusga asoslangan yondashuvning ahamiyatini asoslash bilan yakunlanadi. Bunda korpus ma'lumotlari diskurs tahlili uchun obyektiv va miqdoriy asos bo'lishi ta'kidlanadi. Bundan tashqari, bu yondashuv bosqichma-bosqich bo'lib, diskursdagi o'zgarishlarni kuzatish hamda tavsifiy va miqdoriy

²⁶ Bhatia V. Applied genre analysis: Analytical advances and pedagogical procedures. In A. M. Johns (Ed.), *Genre in the Classroom: Multiple Perspectives*. – London: Lawrence Erlbaum Associates, 2002. – P. 279-283.

²⁷ Schegloff E., Jefferson G., Sacks H. The preference for self-correction in the organization of repair in conversation. *Language*. – Cambridge, 1977. Vol. 53, No.2. – P. 361-382.

usullarni uyg'unlashtirish orqali natijalarni har tomonlama tekshirish imkonini beradi. Bob korpusga asoslangan tahlilning cheklovlarini, masalan, korpus ma'lumotlari faqat og'zaki tilning lingvistik jihatlari bilan cheklanganligi kabi jihatlarni tan olsa-da, qoliplar va strategiyalarning kontekstual tahlili uchun korpus diskursni o'rganishda ishonchli asos bo'la olishini ta'kidlab o'tadi.

Xulosa qilib aytganda, II bob tadqiqotning keng qamrovli, bosqichma-bosqich uslubiy rejasini taqdim etadi. U maxsus til o'rganuvchilar korpusini yaratish va tahlil qilish uchun texnik dasturlardan boshlab, kontekstual, eng kichik birlikdan diskursga qarab yo'naltirilgan, aralash usuldagi diskurs tahlili yondashuvini tanlashgacha bo'lgan jarayonni qamrab oladi. Bob tadqiqot maqsadlarini Flow, Stanford POS Tagger, LCA, TextTiling kabi aniq dasturiy vositalar; yetakchi diskurs olimlarining nazariy asoslari hamda VBDU segmentatsiyasi hamda yetti bosqichli rubrika kabi tahliliy usullar orqali amalga oshiradi. Bu usullar keyingi tahlilda til qoliplari, lingvistik harakatlar va o'zaro ta'sir strategiyalarini o'rganish uchun qo'llaniladi.

Dissertatsiyaning **“Sun'iy intellekt chatbotlari bilan og'zaki suhbatlarning korpusga asoslangan diskurs tahlili”** deb nomlangan yakuniy bobida og'zaki suhbatlarning keng qamrovli korpusga asoslangan diskurs tahlili keltirilgan. Tahlil bir necha bosqichlarni o'z ichiga oladi: korpus tuzish va demografik profillash, so'z turkumlarini belgilash va lemmatizatsiya, chastota va n-gramm tahlili, leksik va sintaktik murakkablikni baholash, prosodik va pragmatik izohlash hamda diskurs darajasida harakatlar, o'zaro ta'sir strategiyalari va nutq aktlarini o'rganish. Natijalar til o'rganuvchilarning sun'iy intellekt chatbotlari bilan til amaliyoti uchun xavfsiz va yuqori faollikka ega platforma sifatida qanday munosabatda bo'lishini ko'rsatadi, bu esa diskursning o'ziga xos lingvistik, pragmatik va interaktiv xususiyatlarini namoyon etadi.

3.1-paragrafda *“Sun'iy intellekt chatbotlari bilan og'zaki suhbatlarning korpusga asoslangan diskurs tahlili natijalari”* keltirilgan. Diskurs jamiyati demografik tanlov orqali aniqlandi, bunda ingliz tilini mashq qilish umumiy maqsadiga ega shaxslar tanlandi. Ishtirokchilarning aksariyati ayollar (89%), texnologiya ishqibozlari (93%) va o'zbek tilida so'zlashuvchilar (95%) bo'lib, ularning til bilimi darajalari B1 va C1 oralig'ida edi. Ko'pchilik (89%) til amaliyoti uchun sun'iy intellekt chatbotlari bilan tez-tez muloqot qilgan.

Flow va Chat GPT sun'iy intellekt chatbotlari bilan og'zaki suhbatlardan 13 450 so'zdan iborat 54 ta namunali korpus tuzildi. Korpus Stanford TreeTagger yordamida so'z turkumlari bo'yicha belgilandi va lemmatizatsiya qilindi. Chastotali tahlil otlar (17,1%) va fe'llarni (16,3%) eng ko'p uchraydigan toifalar sifatida ko'rsatdi. Eng ko'p qo'llangan so'zlar sirasiga *family, sound, good, important* va *talk* so'zlari kirdi. N-gramm tahlili korpusning 71% dan ortig'i ko'p so'zli iboralardan iborat ekanligini ko'rsatdi, bunda ikki va uch so'zli birikmalar eng ko'p uchraydi. Quyidagi jadvalda korpus har bir so'z turkumining uchrash holatlari miqdoriy tahlili qilingan:

2-jadval

Korpusda har bir so'z turkumining uchrash holatlari

Kategoriya	Uchrashi	Foiz
otlar	2426	17.1%
fe'llar	2317	16.3%
boshqa so'zlar	2144	14.7%
olmoshlar	1995	14.0%
sifatlar	1284	9.0%
predloglar	1157	8.1%
aniqlovchilar	697	5.5%
ravish	771	5.4%
maqolalar	559	3.9%
so'roq so'zlar	343	2.4%
Jami	14228	100.0%

Leksik murakkablik analizatori yordamida o'tkazilgan tahlil natijalariga ko'ra, til o'rganuvchi korpusining leksik zichligi 0,21, leksik murakkabligi o'rtacha 0,74 va leksik variatsiyasi cheklangan (o'rtacha 0,002) ekanligi aniqlandi. Sun'iy intellekt korpusi esa biroz yuqoriroq normallashtirilgan leksik murakkablikni namoyon etdi. Sintaktik tahlil shuni ko'rsatdiki, til o'rganuvchilari sun'iy intellektga nisbatan ko'proq gap tuzganlar, biroq ularning o'rtacha uzunligi qisqaroq (MLC=7,84 va 9,22 mos ravishda). Til o'rganuvchi korpusida o'rta darajadagi sintaktik murakkablik darajasi aniqlandi, ergashish va bog'lanishning boshlang'ich belgilarini namoyon bo'ldi.

3-jadval

Korpusida lug'aviy murakkablik tahlili

O'lchov	O'rtacha	Minimum	Maksimum
Leksik zichlik	0.85	0.00	32.00
Leksik mukammallik	0.74	0.62	0.86
Leksik xilma-xillik	0.002	0.00	0.02

4-jadval

Korpusining sintaktik murakkablik profili

O'lchov	O'rtacha	Minimum	Maksimum
Gap murakkabligi	5.20	1.11	12.30
Ergash gap murakkabligi	2.65	0.28	7.84
T-birlik murakkabligi	2.96	0.32	11.10
Iborali murakkablik	0.75	0.50	1.00
Umumiy sintaktik nazorat	3.39	0.28	12.30

Prosodik tahlil tez-tez uchraydigan nutq buzilishlarini aniqladi: 185 holatda to'ldiruvchi so'zlar, 70 holatda takrorlanishlar va 45 holatda pauzalar. O'z-o'zini tuzatish strategiyalari keng tarqalgan bo'lib, o'z-o'zini tuzatish 22 marta va qayta shakllantirish 18 marta kuzatilgan. Bu o'z-o'zini faol nazorat qilish va ravon nutqni saqlashga intilishni ko'rsatadi.

Pragmatik tahlil shuni ko'rsatdiki, suhbatlar asosan shaxsiy mavzularga qaratilgan bo'lib, ona shahri, oila, qiziqishlar va shaxsiy tajribalar haqida bo'lgan. Bog'lovchilar orqali tahlil qilingan bog'liqlik til o'rganuvchi nutqining o'rtacha ko'rsatkichidan biroz yuqori (har 1000 so'z kesimida 53,68) bo'lib, mantiqiy bog'lovchilar eng ko'p uchraydi. Diskurs belgilari asosan norasmiy (94,3%) bo'lib, "shunday", "yaxshi", "mayli" kabi so'zlar strukturaviy va xususiy vazifalarni bajaradi. O'rganuvchilar nutqida pragmatik nomuvofiqlik mavjud, norasmiy va rasmiy uslublarni aralashtirishgan holda ko'pincha to'g'ridan-to'g'ri xushmuomalalik strategiyalaridan foydalanishgan.

Dissertatsiyaning *"Diskurs xususiyatlari tahlili: lisoniy qonuniyatlar, diskurs harakatlari va strategiyalari"* deb nomlangan yakuniy qismida tanlangan diskurs xususiyatlari bo'yicha korpus natijalari tavsifiy tahlil etilgan. Diskurs 16 ta lug'atga asoslangan diskurs birliklariga (VBDUs) bo'linib, ular tavsiflovchi monologlar, hikoyalar, so'rov/ma'lumot olish birliklari, fikr/dalil birliklari va meta-diskurs belgilariga ajratilgan. Bu birliklar o'z navbatida uchta funksional toifaga birlashtirilgan: axborot beruvchi, interaktiv va ishontiruvchi. Diskursning mavzuga yo'naltirilganligi, qolipli tuzilmaga egaligi va xavfsizlik zonasi bilan ajralib turishi aniqlangan. Bu zonada o'quvchilar ravon so'zlashdagi cheklovlarni bartaraf etish uchun suhbat yo'nalishini boshqarganlar.

Diskurs harakatlari tahlili yetti turdagi harakatni aniqlagan bo'lib, savollarga javob berish (4-harakat) va vazifani belgilashda (2-harakat) yuqori chastota kuzatilgan. Til o'rganuvchilar faol ravishda mavzularni boshlab, ma'lumot so'ragan va o'z-o'zini tuzatgan holda, o'z amaliyotlarini yo'naltirishda tashabbuskorlik ko'rsatganlar.

5-jadval

Til o'rganuvchi korpusida diskurs harakatlari chastotasi

Harakat	Tavsif	Umumiy chastota	Eng ko'p ishlatilgan bosqichlari
Harakat 1	Suhbatni boshlash	17	Salomlashish (16) hol-ahvol so'rash (small talk) (1)
Harakat 2	Vazifani shakllantirish (topshiriq)	21	Mavzu boshlanishi (21)
Harakat 3	Ma'lumotlarni olish	13	Til yaratish uchun ko'rsatmalar (13)
Harakat 4	Savolga javob berish	80	Ma'lumotlarni taqdim etish (yuqori darajada)
Harakat 5	Fikr-mulohaza bildirish yoki tuzatish	10	O'z-o'zini tuzatish (3)
Harakat 6	Mavzuni kengaytirish	9	Yaqin mavzularga o'tish (9)
Harakat 7	Suhbatni tugatish	4	Faoliyatni yakunlash (3)

Muloqot strategiyalari xushmuomalalik nazariyasi va freyming yordamida tahlil etildi. Bevosita muloyimlik strategiyalari korpusda eng ko'p uchrab (har 1000 so'z kesimida 16,1), o'quvchilarning sun'iy intellektning noinsoniy tabiatiga moslashuvini aks ettirdi. Ijobiy xushmuomalalik munosabat o'rnatish uchun

qo'llanilib, salbiy xushmuomalalik kamroq kuzatildi. "Sun'iy intellekt xushmuomalaligi paradoksi" yuzaga keldi: FTAlarning hisoblangan vazni ($W_x=12$) vaziyatni yumshatish talab qilinsa-da, o'rganuvchilar to'g'ridan-to'g'ri strategiyalardan foydalanishni ma'qul ko'rishini bildirdi. Bu esa ularning sun'iy intellektni "yuzini yo'qotish" xavfi bo'lmagan tizim sifatida qabul qilishlarini ko'rsatadi.

Suhbat aktlari tahlili reprezentativlar (1000 so'zga 30,7) va direktivlar (13,9) eng ko'p uchraganini, komissivlar va deklarativlar esa kam qo'llanilganini aniqladi. Tuzatish tahlili o'z-o'zini tuzatish (SISR) ustunlik qilganini ko'rsatdi. Sun'iy intellekt kamdan-kam hollarda tuzatishni boshlagan, bu esa uzilishlarsiz, ravonlikka yo'naltirilgan, xavotirsiz muhitni shakllantirgan.

Uslub tahlili til amaliyoti maqsadi, o'rganuvchi-vosita munosabatlari va og'zaki uslub asosida shakllantirilgan gibrid, nosimmetrik registrni aniqladi. Bu registr norasmiy va rasmiy xususiyatlarni o'zida mujassam etgan. O'quvchilar turli ijtimoiy rollarni (mutaxassis, tengdosh, rahbar, ishonchli shaxs) sinab ko'rdilar.

Ushbu bob sun'iy intellekt chatbotlari bilan o'zaro muloqotda til o'rganuvchilarning tavakkallchilik, yuqori faol ishtirokchi ekanligini ta'kidlaydi. Mazkur jarayonda til o'rganuvchilar o'z diskurslarini boshqaradilar, xatosiz gapirishdan ko'ra ko'ra ravonlik, o'zini ifodalay olish va pragmatik tajribaga ustuvorlik beradilar. Lingvistik profil so'z turkumlaridan muvozanatli foydalanish, cheklangan lug'at boyligi, o'rtacha sintaktik murakkablik va yuqori darajada o'z-o'zini nazorat qilish bilan xarakterlandi. Diskurs o'rganuvchi tomonidan boshqariladi, tranzaksion xususiyatga ega va inson-sun'iy intellekt o'zaro ta'siriga xos bo'lgan moslashuvchan xushmuomalalik hamda muloqot strategiyalari bilan ajralib turadi.

XULOSALAR

Mazkur dissertatsiyada sun'iy intellekt chatbotlari bilan og'zaki suhbatlarning diskursi empirik tahlil qilindi. Tahlilda tabiiy tildan yig'ib olingan korpus asos bo'lib xizmat qildi. Sun'iy intellekt bilan og'zaki suhbatlarni korpus asosida tahlil qilish natijasida quyidagi xulosalarga kelindi:

1. Ushbu tadqiqotda korpus asosida diskurs tahlil qilishning nazariyasi, metodologiyasi saralanib, ular asosida og'zaki suhbat korpusini yig'ish va annotatsiyalash mezonlarini belgilab berildi.

2. Universitet talabalari nutqi asosida O'zbekistondagi birinchi kichik hajmli og'zaki suhbatlar korpusi to'plandi. Mazkur korpus kompyuterlashtirilgan dasturlar yordamida annotatsiyalandi va elektron manba sifatida saqlandi.

3. Til o'rganuvchilar korpusining sintaktik-leksik, prosodik va pragmatik xususiyatlari tahlil qilindi. Bunda takrorlanuvchi lisoniy qonuniyatlar, asosiy so'z turkumlarining uchrashi, barcha so'zlarning chastotasi, kollokatsiyalar va n-grammalar, konkordanslar, sintaktik va leksik murakkablik darajalari, nutq buzilishlari va tuzatishlar chastotasi, suhbat aktlarining tarqalishi va turlari, diskurs

markerlarining tarqalishi va chastotasi hamda til o'rganuvchilar nutqi kogeziyasi tahlil qilindi.

4. Til o'rganuvchilar suhbat korpusida 16 turdagi lug'atga asoslangan birliklar avtomatik kompyuter tahlili vositasida aniqlandi va ma'lumot beruvchi, interaktiv va persuasiv funksional toifalarga ajratildi. Korpus natijalarining tavsifiy tahlilida diskursning asosiy tendensiyalari namoyon bo'ldi a) mazkur diskurs mavzuga yo'naltirilgan bo'lib, til o'rganuvchilarga tanish mavzular atrofida tashkil etilgan va mavzu doirasi hamda boshqa mavzularga o'tish aniq belgilab berilgan; b) diskurs lingvistik "xavfsiz zona"da bo'lib, nutqdagi kamchiliklarni to'ldirishga xizmat qiladi; c) diskursning ochilish va yopilish qismlari shablon jumlar ko'p uchraydi; d) diskurs birliklari bo'ylab turli darajadagi til bilish ko'rsatgichlari mavjud, jumladan, ma'lumot beruvchi funksional birliklarda gaplar soda tuzilgan va lug'at boyligi chegaralangan, persuasiv birliklarda esa leksin va sintaktik murakkablik darajasi nisbatan yuqori e) diskursda til qo'llanishi borasida xatolar ko'p uchraydi, shunga qaramay, so'zlovchilarning pragmatik xabardorligi ancha yuqori.

5. Diskurs harakatlarini belgilashning maxsus mezonini ishlab chiqildi; u yettita diskurs harakati va bir nechta bosqichlardan iborat etib belgilandi. Diskursda keng tarqalgan va ko'p qo'llanilgan harakatlar nisbati hisoblandi va tahlil qilindi. Diskursda mavzuga kirish harakati eng ko'p uchrab, til o'rganuvchilarning sun'iy intellekt bilan suhbat diskursining o'ziga xos jihati sifatida baholandi.

6. Diskurs birliklari bo'ylab talabalar qanday muloqot strategiyalaridan foydalanishi o'rganildi. Til o'rganuvchilar: a) ko'pincha (har 1000 so'zda 16,1 marta hisobida) to'g'ridan-to'g'ri xushmuomalalik strategiyalaridan foydalanadilar; b) "bilimdon shaxs" va "tengdosh" kabi ijobiy ijtimoiy yuz ramkalaridan, "rahbar" va "imtihon oluvchi" salbiy yuz ramkalaridan foydalanadilar, "ishonchli do'st" ramkasidan foydalanganda sun'iy intellektda ijtimoiy yuz mavjud emasligini yaxshi his etadilar; c) eng ko'p reprezentativ (30.7) va direktiv (13.9) suhbat aktlaridan foydalanadilar; d) nutqdagi uzilishlarni tuzatishning asosiy strategiyasi sifatida o'z-o'zini tashabbuskor ravishda tuzatishdan foydalanadilar; e) kundalik va rasmiy uslub xususiyatlariga ega bo'lgan gibriddan, aralash uslubdan so'zlashishlari aniqlandi.

7. Ijtimoiy munosabatlarda mavqe'ni belgilashning ikki o'zgaruvchisi P – mavqe' va D – masofa nisbatlari ijtimoiy yuzga tahdid kalkulyatori formulasi asosida hisoblanganda, diskurs past P va maksimal D miqdorini ko'rsatdi, bu insonlar orasidagi suhbatlarga xos bo'lmagan holat deb hisoblandi va uning asosida past darajadagi P va yuqori darajadagi D ko'rsatgichiga ega bo'lgan xushmuomalalikning o'ziga xos ko'rinishi aniqlandi, bu "sun'iy intellekt xushmuomalalik paradoksi" sifatida baholandi; bunda mazkur suhbatlar jarayonida ijtimoiy munosabatlarning yo'qligi vat il o'rganuvchilar interlokuturning inson emasligidan yaxshi xabardorligini isbotladi.

8. Tadqiqot natijalari asosida yangi avlod til o'rganish platformalarini loyihalash bo'yicha tavsiyalar ishlab chiqildi. Bu tavsiyalar sun'iy intellekt diskursini inson diskursiga yaqinlashtirish, suhbatdoshga moslashuvchanlik, to'xtalishlar va nutqdagi uzilishlarga javobni takomillashtirishni o'z ichiga oladi. Eng muhimi, til

o'rganuvchilarning sun'iy intellekt bilan suhbatlar orqali insonlar bilan suhbat diskursiga tayyorlana olishini ta'minlaydigan dasturlar yaratish tavsiya etildi.

9. Inson va sun'iy intellekt suhbat diskursining o'ziga xos lingvistik va pragmatik dinamikasi O'zbekistonda ilk marotaba tadqiq etildi. Bu tadqiqot bugungi kunda "sun'iy intellekt vositachiligidagi muloqot diskursi"ning paydo bo'lishi bilan yuzaga kelgan diskurs tahlili amaliyotidagi bo'shliqlarni to'ldirishga xizmat qiladi.

10. Tadqiqotda qo'llanilgan namunalarda gender jihatdan nomutanosiblik mavjud bo'lib, bu ma'lumotlar to'plangan muassasadagi ingliz filologiyasi fakulteti demografik tarkibini aks ettiradi. Tadqiqot natijalari shu jihatdan asosan ayol (89%) til o'rganuvchilarining diskursini ifodalaydi va erkaklarga qo'shimcha tadqiqotlarsiz umumlashtirilishi tavsiya etilmaydi. Kelgusida sun'iy intellekt vositachiligidagi og'zaki suhbat diskursini tadqiq qilish maqsadida gender jihatdan muvozanatlashgan tadqiqot o'tkazish tavsiya etiladi.

**ONE TIME SCIENTIFIC COUNCIL ON THE BASIS OF SCIENTIFIC
COUNCIL DSc.03/2026.31.01.Fil.45.01 ON AWARDING SCIENTIFIC
DEGREES AT UZBEKISTAN STATE WORLD
LANGUAGES UNIVERSITY**

UZBEKISTAN STATE WORLD LANGUAGES UNIVERSITY

ASROROVA NARGIZA ISOMITDINOVNA

**CORPUS-BASED DISCOURSE ANALYSIS OF SPOKEN
CONVERSATIONS WITH AI CHATBOTS**

10.00.11 Applied and Computational Linguistics

ABSTRACT

of the Dissertation for Doctor of Philosophy (PhD) in Philological Sciences

Tashkent – 2026

The theme of the dissertation of Doctor of Philosophy (PhD) was registered at the Supreme Attestation Commission of the Academy of Sciences of the Republic of Uzbekistan under number B2025.2.PhD/Fil6395.

The dissertation has been carried out at Uzbekistan State World Languages University.

The abstract of the dissertation is posted in three languages (Uzbek, English, Russian (abstract)) on the website of Scientific Council (www.uzswlu.uz) and on the website «Ziyonet» informational and educational portal (www.ziyonet.uz).

Scientific consultant:

Hakimov Hamidulla Inomovich

PhD in Philological Sciences,
Associate Professor

Official opponents:

Rahimov Ganisher Khudoykulovich

Doctor of Philological Sciences, Professor

Djumabayeva Jamilya Sharipovna

Doctor of Philological Sciences, Professor

Leading organization:

**Tashkent State University
of Uzbek Language and Literature
named after Alisher Navoiy**

The defence will take place on October 29, 2026 at 9:00 at the meeting of the One Time Scientific Council on the basis of the Scientific Council DSc.03/2026.31.01.Fil.45.01 at Uzbekistan State World Languages University (Address: 100138, Tashkent city, Uchtepa district, Kichik halqa yoli street, 21a. Tel.: (99871) 230-12-91, fax: (99871) 230-12-92, e-mail: uzswlu_info@mail.ru).

The dissertation can be reviewed at the Information Resource Centre of Uzbekistan State World Languages University (registered under the number_____) (Address: 100138, Tashkent city, Uchtepa district, Kichik halqa yoli street, 21a. Tel.: (99871) 230-12-91, fax: (99871) 230-12-92, e-mail: uzswlu_info@mail.ru).

Abstract of dissertation was sent out on “_____” _____2026

(mailing report № _____ “_____” _____2026)

A.A. Nasirov

Chairman of One-time Scientific
Council Awarding Scientific Degrees,
Doctor of Philology, Professor

Kh.B. Samigova

Scientific Secretary of One-time
Scientific Council Awarding Scientific
Degrees, Doctor of Philology,
Professor

J.A. Yakubov

Chairman of the One-time Scientific
Seminar of the One-time Scientific
Council Awarding Scientific Degrees,
Doctor of Philology, Professor

INTRODUCTION (Annotation of the Doctor of Philosophy (PhD) dissertation)

Topicality and necessity of the research. The contemporary era is predicted to be the era of the artificial intelligence (AI) and technological reforms regardless of field. Throughout the globe, language learners are also turning to AI to improve their language skills and it has become a frequent speaking partner. When language learners interact with AI on voice mode, this crucial environment is shaping their language practice, pragmatic awareness and discursive aspects. These oral conversations with the end goal of language practice are giving rise to a new and AI-mediated form of spoken discourse. Consequently, one of the most urgent tasks of contemporary linguistics is to identify the discursive features of these conversations because despite the proliferation of AI chatbot use for speaking practice, the linguistic nature of these conversations remains largely unexplored. The effectiveness of AI chatbots, their pedagogical applications and learner perceptions of AI are popular research areas, however, these studies have primarily focused on learning outcomes rather than the discourse of AI-mediated conversations. The existing corpus-based studies have examined human-AI interactions either in native language contexts or through written, text-based communication. Therefore, the syntactic-lexical, prosodic and pragmatic features of AI-learner spoken conversation are still an uncharted territory of linguistic research. This gap is significant because AI chatbots are now capable of voice-based interactions, even using paralinguistic elements of stress and intonation, also, variety of language accents which accurately depict or imitate human-human conversations, yet entirely distinct from them. The linguistic portrait of this unique communicative environment needs to be painted.

These distinctive features necessitate systematic empirical investigation, as existing discourse theories, developed primarily for human-human interactions, may not adequately account for human-AI communication dynamics. The claims about AI-mediated communication remain speculative without an empirical grounding, and discourse descriptions will lack accuracy and reliability. The absence of empirically grounded, corpus-based analysis of learner spoken discourse in AI-mediated contexts is the central problem addressed by this dissertation. While language learners increasingly rely on AI chatbots for speaking practice, the linguistic characteristics of these interactions and their syntactic, lexical, prosodic and pragmatic features, and discourse organization remain undocumented.

In our country, the development of information and communication technologies and artificial intelligence has been designated as one of the priority areas of state policy, opening up new, advanced scientific horizons for linguistic and corpus research. Indeed, “The development of artificial intelligence and digital technologies today necessitates a fundamental renewal of the education system, alongside all other sectors, with language resources and electronic

corpora playing a crucial strategic role”¹. This scientific strategic direction, in particular, enables the organic integration of new research with practice, such as analyzing the oral discourse of English language learners with artificial intelligence chatbots using corpus-based methods. This creates the necessary conditions for elevating national educational practices to a high technological level and establishing innovative standards in the field of language education.

The goals of the research align with the national priority areas outlined in legal and strategic documents of the Republic of Uzbekistan, including the Law of the Republic of Uzbekistan “On Education”, the National program for personnel training, the Presidential Decree № PF-5847 “On Approval of the concept for the development of the higher education system of the Republic of Uzbekistan until 2030” of October 8, 2019, the resolution № 606 of the Cabinet of Ministers of the Republic of Uzbekistan “On Improving the system for targeted training of highly qualified research and academic-pedagogical staff” of July 19, 2019, the Presidential Decree № PF-6097 “On Approval of the concept for the development of science until 2030” of October 29, 2020, the resolution № PQ-5117 “On Measures to raise the activities for promoting the learning of foreign languages in the Republic of Uzbekistan to a new qualitative level” of May 19, 2021, the Presidential Decree № PF-60 “On the Development strategy of new Uzbekistan for 2022 – 2026” of January 28, 2022, the “Digital Uzbekistan – 2030” Strategy, the Presidential Resolution № PQ-358 “On Approval of the strategy for the development of artificial intelligence technologies until 2030” of October 14, 2024 and the research findings will serve the objectives set by these legislative acts.

Alignment of the research with priority directions of developing science and technology in Uzbekistan. This research was conducted in accordance with the priority direction I of the science and technology development of the Republic: “Formation of a system of innovative ideas and ways of implementing them in the social, legal, economic, cultural, spiritual, and educational advancement of an information society and a democratic state”.

Problem development status. Internationally, early discourse studies incorporating corpus findings include the analysis of political texts by L.Flowerdew, N.Fairclough, A.Piper and A.Partington; the analysis of educational materials by M.Stubbs and P.Wickens; the analysis of scientific texts by T.A.van Dijk, A.Morrison and C.Caldas-Coulthard; and the analysis of identification and gender differences by

¹ <https://president.uz/en/lists/view/8582> [Date of application: 05.12.2025].

C.Shalom and T.McEnery². T. McEnery, J.Rey, D.Biber, and S.Conrad³ conducted large-scale discourse studies from the perspective of corpus linguistics.

In recent years, several scholars have studied the methods, approaches, and practical significance of corpus-based discourse analysis, including P.Baker, R.Grimes, A.Chavez, B.Li and H.Wang, Y.Fan, F.Jiang, and P.Li⁴. Additionally, H.Lu⁵ established a theoretical foundation for analyzing the syntactic and lexical complexity of corpora using computer tools. Several meta-analyses and meta-synthesis studies were conducted to examine the textual and voice outputs of artificial intelligence chatbots, as an example, F.Koc and P.Savash, C.Zhai and S.Wibowo, V.Huang, K.Hew and L.Fryer⁶ investigated the effectiveness, advantages, and limitations of the voice mode in AI chatbots, as well as the application of pedagogical and linguistic theories in AI chatbot communication. A.Iram, G.Tudino and Y.Qin, examined human speech patterns during written interactions with AI chatbots⁷.

² Flowerdew L. *Corpus-Based Discourse Analysis*. Second Edition. – London: Routledge, 2023. – P.10-44.; Fairclough N. *Discourse and social change*. – Malden: Polity Press, 1992. – 135 p.; Piper A. Some people have credit cards and others have giro cheques: “Individuals” and “people” as lifelong learners in late modernity. *Discourse and Society*. – SAGE Journals, 2000. Vol.11. – P. 515-42.; Partington A. *Patterns and Meanings*. – Amsterdam: Benjamins, 1998. – 162 p.; Stubbs M. *Discourse Analysis: The Sociolinguistic Analysis of Natural Language*. – Chicago: University of Chicago Press, 1983. – 272 p.; Wickens P. Comparative analysis of the use of projecting clauses in pedagogical legal genres. Paper read at the 25th International Systemic Functional Congress. – University of Wales, Cardiff, July 1998. – 13 p.; Van Dijk T. A. *Discourse and power*. – London: Palgrave Macmillan, 2008. – P. 17-19.; Morrison A., Love A. A discourse of disillusionment: Letters to the Editor in two Zimbabwean magazines 10 years after independence. *Discourse and Society*. – SAGE Journals, 1996. Vol.7. – P. 39-76.; Caldas-Coulthard C. R. and Moon, R. Curvy, hunky, kinky: Using corpora as tools in critical analysis. Paper read at the Critical Discourse Analysis Meeting. – University of Birmingham, 1999. – P. 17-21.; Shalom C. That great supermarket of desire: attributes of the desired other in personal advertisements. – London: Routledge, 1997. – P. 186-203.; McEnery T., Hardie A. *Corpus Linguistics: Method, Theory and Practice*. – Cambridge: Cambridge University Press, 2012. – 312 p.

³ McEnery T., Xiao R., Tono Y. *Corpus-Based Language Studies*. – London: Routledge, 2006. – 406 p.; Rey J. M. Changing gender roles in popular culture: Dialogue in Star Trek episodes from 1966 to 1993. – London: Longman, 2001. – P. 138-56.; Biber D., Conrad S., Cortes V. If you look at. . . : Lexical bundles in university teaching and textbooks. *Journal of Applied Linguistics*. – Oxford: Oxford University Press, 2004. Vol.25 No.3. – P. 371-405. McEnery T., Xiao R., Tono Y. *Corpus-Based Language Studies*. – London: Routledge, 2006. – 406 p.; Rey J. M. Changing gender roles in popular culture: Dialogue in Star Trek episodes from 1966 to 1993. – London: Longman, 2001. – P. 138-56.; Biber D., Conrad S., Cortes V. If you look at. . . : Lexical bundles in university teaching and textbooks. *Journal of Applied Linguistics*. – Oxford: Oxford University Press, 2004. Vol.25 No.3. – P. 371-405.

⁴ Baker P. *Using corpora in discourse analysis*. – London: Continuum, 2021. – 207 p.; Grimes R.E., Grimes J.E. The thread of discourse. – The Hague: Mouton, 1975. – 420 p.; Chaves A.P., Doerry E., Egbert J., and Gerosa M. It's How You Say it: Identifying Appropriate Register for Chatbot Language Design. / In Proceedings of the 7th International Conference on Human-Agent Interaction (HAI '19). – Association for Computing Machinery, New York, NY, USA, 2019. – P.102–109.; Li, B. and Wang, H. TACOMORE: Leveraging the Potential of LLMs in Corpus-based Discourse Analysis with Prompt Engineering. – Cornell University Archive, 2024.; Fan, Y. Jiang, F. and Li P. Uncovering the Potential of ChatGPT for Discourse Analysis in Dialogue: An Empirical Study Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation. – Cornell University Archive, 2024. – P.16998-17010.

⁵ Lu X. *Computational methods for corpus annotation and analysis*. – Dordrecht: Springer, 2014. – P. 4-158.

⁶ Koç FŞ, Savaş P. The use of artificially intelligent chatbots in English language learning: A systematic meta-synthesis study of articles published between 2010 and 2024. *ReCALL*. – Elsevier, 2025. Vol.37, No.1. – P. 4-21.; Zhai C., Wibowo S. A systematic review on cross-culture, humor and empathy dimensions in conversational chatbots: The case of second language acquisition. *Heliyon*. – Elsevier, 2022. Vol.8, No.12. – P. 5.; Huang W., Hew K.F., Fryer L.K. Chatbots for language learning-Are they really useful? A systematic review of chatbot-supported language learning. *Journal of Computer Assisted Learning*. – Online Journal, 2015. Vol. 38, No.1. – P. 237–257.;

⁷ Iram A. Hussain M., Shahbaz M. Evaluating Human and AI-Crafted Texts: A Comparative Analysis Using the 6+1 Trait Writing Model and Discourse Analysis. *International Research Journal of Arts, Humanities and Social Sciences*. – Elsevier, 2025. Vol.2, No.2. – P. 706–737.; Tudino G, Qin, Y. A corpus-driven comparative analysis of AI in

In Uzbekistan, corpus linguistics and corpus-based discourse studies have developed primarily within the theory of discourse, language pedagogy and text linguistics. The theoretical foundations for corpus analysis and discourse analysis were established by scholars such as Sh.Safarov, R.Qongurov, M. Tursunov, S.Karimov, Sh.Hamroyeva, B.Mengliyev, N.Ataboyev and A.Akhmedova⁸. Their research provided the methodological basis for corpus-oriented investigations of this dissertation and helped define the key terms of discourse and corpus.

Applied corpus studies were advanced by M.Abjalova⁹, who examined methods of tagging parts of speech, lemmatization and corpus analysis programs. M.Abjalova is the author of Alisher Navoiy author corpus and the first educational corpus of Uzbek language. S.Tayirova and M.Usmanov¹⁰ further demonstrated the effectiveness of corpus-based activities in developing learners' linguistic competence and K.Rizayeva applied corpus and regression-based methods to discourse and media text analysis¹¹. More recently, attention has shifted toward artificial intelligence in education. N.Ataboyev studied the role of ICT and AI in modern linguistic research and longitudinal diachronic corpora. N.Abdurahmonova of National University of Uzbekistan examined learner corpora and conducted annotation of error patterns, providing recommendations for pedagogical applications of corpus studies in Uzbek language¹². Although corpus linguistics is actively developing in the country, systematic corpus-based discourse analysis of learner interactions with AI chatbots is still emerging.

Corpus linguistics has been supported by strong theoretical and computational traditions in CIS countries. V.P.Zakharov, A.K.Kutuzov, E.V.Nedoshivina,

academic discourse: Investigating ChatGPT-generated academic texts in social sciences. *Lingua*. – Elsevier, 2024. Vol.312. P. 45-46.

⁸Safarov Sh. Pragmalingvistika. – Toshkent: O'zbekiston Milliy Ensiklopediyasi Davlat ilmiy nashriyoti, 2008. – 432 p.; Qo'ng'urov, R. Diskurs nazariyasi asoslari. // Til va adabiyot ta'limi jurnali. – Toshkent, 2017. Vol.4. – P. 12–18.; Tursunov M., Karimov S., Karshiyev A. Development of a Modern Corpus of Computational Linguistics. 2020 International Conference on Information Science and Communications Technologies – Tashkent, 2020. <https://doi.org/10.1109/ICISCT50599.2020.9351376> [Date of application: 05.10.2025]; Hamroyeva Sh., Mengliyev B., Abdullayeva O. Scopus-based bibliometric analysis on corpus linguistics for the period of 2017-2021.– E3S Web of Conferences, 2023. – 5 p. <https://doi.org/10.1051/e3sconf/202341303008> [Date of application: 08.11.2024]; Mengliyev, B. O'zbekistonning zamonaviy siyosiy diskursi: kognitiv-diskursiv jihat. – Toshkent: O'zR FA "Fan" nashriyoti, 2021. – 192 b; Nuritdinov, A. Korpus lingvistikasida lingvostatistik tahlil metodi. // Maktabgacha va maktab ta'limi jurnali. – Toshkent, 2025. Vol.3(5). – B. 5.; Nurmatova, G. Corpus-based study of engineering terms: Linguistic and pedagogical implication. *Filologicheskie nauki*. – Navoiy, 2014. Vol. 81. – 7 p.; Ataboyev N. Corpus-Based Research on the Language Features of Corpus Linguistics: In the Example of ECOCL – International Journal of Scientific and Technology Research. – IJSTR Online Journal, 2020. – Vol.9, Issue 02. – pp. 4176-4179.; Akhmedova A. Korpusga asoslangan tarjimashunoslikni o'qitishning o'ziga xos xususiyatlari – Переводоведение: проблемы, решения и перспективы, 2022. – 1-son – bb. 34-36.

⁹ Abjalova M. Tagging and Annotation of Corpus Units. – International Journal of Language Learning and Applied Linguistics, 2023. – 5 p.;

¹⁰ Tayirova S., Usmonov M. Corpus-Based Activities Use Real Language Data to Engage Students in Active Vocabulary Learning, Analysis, and Application. *International Journal of Recently Scientific Researcher's Theory*. – Tashkent, 2025. Vol.2, No.5. – P. 1-7.

¹¹ Rizayeva K. Cross-Linguistic Features of English and Uzbek Media Discourse. *Innovations in Science and Technologies*. – Tashkent, 2025. Vol.2, No.6 (Special issue).

¹²Ataboyev N. Mediamatnlar diaxronik korpusida til rivojining empirik tahlil tamoyillari (gazeta materiallari misolida). – Filologiya fanlari doktori (DSc) dissertatsiyasi, Farg'ona davlat universiteti, 2024. – 236 b.; Abdurahmonova N. Corpus Based Teaching Uzbek as a Foreign Language. *Journal of Foreign Language Teaching and Applied Linguistics*. – Samarkand, 2019. Volume 6 – No. 1. – P.131-137.

V.Plungyan, and M.Kopotev made important methodological contributions by designing corpus compilation principles and quantitative linguistic approaches. S.Kiseleva analyzed learners' errors and communicative patterns in chatbot interaction, including English - Russian corpora. P.Sysoyev and E.Filatov examined conversations between learners and Yandex GPT, while A.Avramenko, A.Akhmedova and E.Bulanova studied the pedagogical potential of the HelloTalk chatbot¹³. A large-scale corpus-based discourse analysis of AI chatbot communication, especially, focusing on linguistic, interactional and discourse features, remains insufficiently explored in the CIS region despite growing popularity of this research trend.

While corpus linguistics and discourse analysis are strongly developed internationally, and are steadily expanding in Uzbekistan, systematic corpus-based investigation of linguistic and pragmatic features of learner interactions with AI chatbots, particularly, in Uzbek educational context, remains under-researched. This gap determines the scientific relevance and novelty of the present study.

Alignment of the research topic with scientific-research goals of the higher educational institution of its implementation. The dissertation aligns with the framework of the priority research direction of "Enhancing the communicative potential of English language by applying advanced practices in solving linguistic problems in linguistic analysis and theoretical linguistics" in the research work plan of the Uzbekistan State World Languages University.

The aim of the research is to conduct a corpus-based discourse analysis of oral conversations with artificial intelligence chatbots in a second language.

The research objectives are as follows:

to analyze recurrent syntactic-lexical, prosodic, and pragmatic language patterns in the corpus of language learners when communicating with artificial intelligence chatbots;

to identify discourse moves used by language learners in oral communication with artificial intelligence chatbots and study the structure of these moves across various vocabulary-based discourse units;

to examine the interactional strategies widely employed by language learners during chatbot conversations and assess how well these strategies align with the learners' discourse management;

¹³ Захаров В.П. Корпусная Лингвистика. – Санкт Петербург: СПбГУ Пресс, 2005. – 178 п.; Недошивина, Е.В. Современный политический дискурс: стратегии и тактики речевого воздействия. – Москва, 2017. – Р. 13.; Плунгян, В.А., Сичинава, Д.В. (Ред.). Национальный корпус русского языка. – Москва, 2003–.; Копотев М.В. Введение в корпусную лингвистику: Учебное пособие для студентов филологических и лингвистических специальностей университетов. – Прага: Animedia Company, 2014. – 285 с.; Kiseleva S.V., Smirnova A.A., and Trofimova N.A. "Chatbot Communication" as an Object of Linguistic Research in the System of Digital Communications. Diskurs. – Moscow, 2022. Vol.8, No.3. – P.128-146.; Sysoyev P.V., Filatov E.M. Artificial intelligence in teaching Russian as a foreign language. Journal of Russian Language Studies. – Saint Petersburg, 2024. Vol.22, No 2. – P. 300-317.; Авраменко А., Ахмедова А., Буланова Е. Технология чат-ботов как средства формирования иноязычной грамматической компетенции при самостоятельном обучении. – Вестник Тамбовского университета, 2023. – С. 387-393. DOI: <https://doi.org/10.20310/1810-0201-2023-28-2-386-394> [Date of application: 13.04.2025].

to investigate discourse markers, conversation acts, discourse cohesiveness, politeness strategies and register types used by language learners in conversations conducted with artificial intelligence chatbots.

The object of the research is the discourse characteristics of language learners in oral communication with artificial intelligence chatbots.

The subject of the research are the linguistic patterns, discourse moves and interactional strategies manifested in the spoken conversational discourse of language learners during interactions with AI chatbots.

Research methods. The research process employed methods of content analysis, classification, description, comparison, statistical analysis, corpus analysis, and empirical analysis.

Scientific novelty of the research:

The scientific novelty of the research comprises the following:

the discursive features of language learners' verbal communication with AI chatbots were examined using a corpus-based approach, and patterns of functional usage of syntactic-lexical, pragmatic, and prosodic units in communication with artificial intelligence as an interlocutor were identified;

the distribution of recurrent language patterns, n-grams, concordances, dysfluencies, linguistic repair, discourse markers and conversational acts, levels of lexical and syntactic complexity, cohesiveness of discourse was documented in the process of language communication with AI as an interlocutor;

the effectiveness of interactional strategies and types of linguistic moves employed by language learners during oral communication with AI chatbots have been substantiated based on their frequency distribution in Vocabulary-Based Units throughout the discourse;

within the framework of corpus-based discourse analysis, the "*AI politeness paradox*" was revealed using the model for calculating the weight of actions, which indicates a low P (power) indicator and a high D (distance) indicator, reflecting the absence of social relations between the speaker and the interlocutor.

Practical results of research include:

development of theoretical framework on procedures of the corpus-based analysis of AI-mediated spoken conversational discourse including the approach, methodology and techniques;

compilation of the first small-scale reference corpus of language learners of one university in Uzbekistan when interacting with AI chatbots through voice-mode conversations;

development of practical recommendations on the design plans for improving the role of AI as a conversation partner.

Reliability of the research results. The reliability of the research results is substantiated by the following factors: the conclusions drawn during the research are reflected in scientific articles and theses published by the researcher and have been positively evaluated by the scientific community; the issues raised within the framework of corpus-based discourse analysis of second language learners' oral communication with artificial intelligence chatbots have been determined using

quantitative and qualitative methods, including corpus analysis, discourse analysis, frequency-statistical calculations, syntactic-lexical complexity analysis, pragmatic and prosodic description, and comparative and empirical analysis methods; the developed conclusions and recommendations have been applied in the educational process through experimental testing; and the obtained results have been validated by authorized scientific institutions.

Scientific and practical significance of research results. The scientific significance of the research results lies in their potential to serve as a scientific and theoretical resource for developing theoretical and methodological foundations for corpus-based discourse analysis of oral communication with artificial intelligence chatbots. They also contribute to establishing a mixed research model that combines quantitative corpus analysis results with qualitative discourse analysis.

The research has practical scientific significance in teaching “Applied Linguistics”, “Second Language Acquisition”, “Corpus Linguistics”, “Text Linguistics”, “Modern Linguistics”, and Methods of Linguistic Analysis” in higher education institutions. Additionally, the research findings help improve the designs of AI-based language learning platforms by providing recommendations on enhancing AI’s discursive features to facilitate the transition from AI chatbot conversations to real-life communication.

The implementation of research results. The research findings have been applied in practice and are being utilized by several organizations.

Scientific findings regarding the study of discursive features in oral communication between language learners and AI chatbots, based on corpus analysis, and the identification of functional usage models for syntactic-lexical, pragmatic, and prosodic units in conversations where artificial intelligence participates as an interlocutor, were used within the framework of the “Professional development and mentor training project for computer science teachers in general education schools based on the Massachusetts Institute of Technology’s artificial intelligence education program” implemented according to Order № 394 of the Minister of Preschool and School Education dated October 20, 2025 (Reference № 02-33-4733 of the Ministry of Preschool and School Education of the Republic of Uzbekistan dated December 23, 2025). As a result, practical recommendations were developed for second language learning and improving AI-based language learning platforms.

Scientific results demonstrating the use of AI chatbots as interactive interlocutors in second language communication, as evidenced by recurring language patterns, n-grams and concordances, levels of lexical and syntactic complexity, and distribution of discourse markers and conversational acts, were utilized in the project “The interests of women in our country are under constant attention,” implemented by the Association of Entrepreneurs and Businesswomen of Uzbekistan, funded by the Public Foundation for Support of Civil Society Institutions under the Supreme Assembly № 19-12/147 (Reference № 27 of the Association of Entrepreneurs and Businesswomen of Uzbekistan dated September 20, 2025). The project applied the research findings as practical guidelines for

developing communicative modules in the textbooks titled “Fundamentals of Information Security” and “Cybersecurity”, as well as a specifically developed AI-based learning platform, to enhance participants’ competencies in communicating with artificial intelligence.

Based on scientific results proving the effectiveness of linguistic strategies used by language learners in oral communication with artificial intelligence chatbots, determined by the frequency distribution, functional load, and consistent repetition of Vocabulary-Based Units throughout the discourse, a textbook titled “Fundamentals of Information Security” was developed and implemented for bachelor’s degree programs in philology and language teaching (Reference № 02/17-4/4 of the Ministry of Higher Education, Science and Innovation of the Republic of Uzbekistan dated January 10, 2026). The module “AI agents and voice-mode communication” was structured according to the seven-move framework identified in the dissertation. Additionally, the design of scaffolded speaking activities and model dialogues that help learners develop communicative competence in AI-mediated settings were informed by the frequency-based profiles of linguistic strategies present in the research. The task-based learning modules of the textbook applied the VBDUs classification. Thus, the textbook provided practical discourse templates for students to engage effectively with AI in voice mode.

Based on the first comprehensive study of the corpus-based discourse analysis of language learners’ oral communication with AI chatbots, particularly, identified linguistic strategies, discourse units, and their frequency and functional characteristics were used in the project of the Subject Curriculum Reform in the field of foreign languages of the national educational program in cooperation with Cambridge University Press & Assessment of the United Kingdom, based on Resolution № 157 of the Cabinet of Ministers of the Republic of Uzbekistan dated March 27, 2024, “On measures to improve the procedure for developing and implementing state educational standards and state educational requirements.” (Reference № 01/11-6/219 of the Republican Center for Education under the Ministry of Preschool and School Education of the Republic of Uzbekistan dated April 8, 2026). The implementation reflects the research results on the discursive features of AI-mediated communication and provides ground for their successful integration into the language teaching process.

Approbation of research results. The research results have been discussed and presented at 6 international and 2 republican scientific-practical conferences.

Publication of research results. A total of 14 scientific works has been published on the dissertation topic, including 5 articles in scientific publications recommended by the Higher Attestation Commission of the Republic of Uzbekistan for publishing the main scientific results of PhD dissertations.

The structure and the size of the dissertation. The dissertation consists of an introduction, three chapters, conclusion and references. Total page count of the dissertation is 142 pages.

THE MAIN CONTENT OF THE DISSERTATION

The introduction part outlines the relevance and necessity of the topic, the goals and objectives of the research, its object and subject, the alignment with the priorities in the development of science and technology, the scientific novelty of the research, its practical results, reliability, theoretical and practical significance, implementation in practice, approbation, published works, and the structure of the research.

The first chapter is titled **“Theoretical Framework for Corpus-Based Discourse Analysis of Spoken Conversations with AI Chatbots”** and in paragraph 1.1., titled *“1.1. Theoretical foundations of corpus-based discourse analysis”*, the study established theoretical foundations of conducting corpus-based discourse analysis. It defines discourse analysis as the analysis of *the result of human communication*, and the corpus as *systematically collected body of written or spoken texts used for linguistic analysis*. Discourse analysis is a field within applied linguistics. Generally agreed definitions of the discourse are “language-in-use” and “language-in-action” suggested by T.McEnery, A.Hardie and J.Blommaert¹⁴. M.Foucolt suggests the plural form of “discourses” and J.Gee distinguishes between discourse with capital “D” and with little “d”; he defines “discourse” as “language-in-use” and “Discourse” as semiotic elements in line with linguistic practices¹⁵.

Corpus is a building block of corpus linguistics. Corpus is the basic unit of corpus linguistics which serves as a database for studying language patterns, structures and usage in real-world contexts. Corpus linguistics relies on corpus to investigate frequency, collocations, discourse structures and semantic variations across different registers and genres¹⁶. The paragraph also discusses terms “corpus-driven”, which is a theory and “corpus-based”, which stands as a method. As a summary, “corpus-based” was defined as an umbrella term that includes various forms of corpus analysis.

The chapter details their fundamental differences. Discourse analysis is qualitative and preserves text integrity, focusing on content and meaning, while corpus analysis is largely quantitative, works with extracted samples, and is “inherently concerned with language structure itself”¹⁷. A key epistemological distinction is that for corpus analysis, “texts are static products rather than as social actions or dynamic processes”. Despite these distinctions, the synergy of the two fields yields corpus-based discourse analysis (CBDA), a “fully developed and theoretically grounded methodology”¹⁸.

Three key approaches to corpus-based discourse analysis – textual approach, critical approach and contextual approach have been outlined in the paragraph. The

¹⁴ McEnery T., Hardie A. Corpus Linguistics: Method, Theory and Practice. – Cambridge: Cambridge University Press, 2012. – 312 p.; Blommaert, J. Discourse. A Critical Introduction. – Cambridge: Cambridge University Press, 2005. – 299 p.

¹⁵ Baker P. Using corpora in discourse analysis. – London: Continuum, 2021. – P.19.; Gee J. An Introduction to Discourse Analysis: Theory and Method. – London: Routledge, 2014. – 248 p.

¹⁶ Gee J.P. An Introduction to Discourse Analysis: Theory and Method. – London: Routledge, 2014. – 248 p.

¹⁷ Biber D., Conrad S., Reppen, R. Corpus linguistics: Investigating language structure and use. – Cambridge: Cambridge University Press, 2004. – 300 p.

¹⁸ Widdowson, H. The limitations of linguistics applied. Applied Linguistics. – Cambridge, 2000. Vol.21, No.1. – P. 3–25.

study opts for corpus-based contextual approach as a bridge between the reference corpus and L2 discourse based on Flowerdew's concept of "largely concerned with how various rhetorical and pragmatic devices are operationalized by participants in specific social situations"¹⁹.

Limitations of corpus-based approach to discourse analysis have also been analyzed. The main problem with CBDA is that "computers cannot produce ethnographic descriptions of language use"²⁰. Furthermore, given the multimodal nature of corpora, corpus-based discourse analysis of learner-AI conversations is a fledgling area of linguistics and few works exist to aid the study.

In the paragraph 1.2. of the study, titled "*Theoretical aspects of learner corpus collection and annotation*", the central focus is the data source for discourse analysis. The paragraph distinguishes between *native corpora* and *learner corpora* and defines the latter as large or small-scale compilations of textual or spoken data generated by language learners. Existing corpora such as BNC, ALICE, COCA, MICASE, TACOMORE, RST, WECCL and learner corpora as ICLE, CANCODE and Cleverbot have been investigated and compiling a new small-scale learner corpus is deemed imperative for this study because existing corpora are largely pedagogic or teaching-oriented, whereas this study requires a learning-oriented corpus of practice conversations between language learners and AI chatbots to serve Second Language Acquisition (SLA) and discourse studies.

The theoretical framework of this paragraph examines corpus aspects of learner-AI conversational discourse, which must be guided by research questions. Linguists B.Bulte and A.Housen identify four key parts of corpus aspects as syntax, lexis, pragmatics, and prosody of learner speech²¹. It also outlines what corpus aspects need to be identified for serving discourse analysis: syntactic complexity (coordination, subordination, sentence length, phrasal sophistication), lexical complexity (diversity, density, sophistication), pragmatic patterns (discourse markers, linguistic register, politeness theory and repair strategies) and prosodic aspects (suprasegmental features, stress, intonation, rhythm and dysfluencies).

The final 1.3. paragraph of the study, titled "*The need for research in corpus-based discourse analysis of learner-AI conversations*", positions the study within a research gap, emphasizing that while corpus-based studies exist in Uzbekistan, practical discourse studies do not exist and corpus-based discourse studies have not been undertaken. Internationally, research analyzing the human interlocutor's language in chatbot conversations remains limited, with studies like J.Jia's analysis of keyword-based interactions or D.Coniam's assessment of chatbot suitability not providing a comprehensive, corpus-based discourse analysis of learner speech²².

¹⁹ Flowerdew L. Corpus-Based Discourse Analysis. 2nd Edition. – London: Routledge, 2023. – P.10-44.

²⁰ Widdowson, H. The limitations of linguistics applied. Applied Linguistics. – Cambridge, 2000. Vol.21, No.1. – P. 3–25.

²¹ Bulte B., Housen A. Conceptualizing and measuring short-term changes in L2 writing complexity. Journal of Second Language Writing, 2014. Vol. 26. – P.42-65.

²² Jia J. The Study of the Application of a Keywords-based Chatbot System on the Teaching of Foreign Languages, Doctoral Thesis. – Cornell University, 2003. – 228 p.; Coniam, D. Evaluating the Language Resources of Chatbots for Their Potential in English as a Second Language. ReCALL, 2008. Vol.20, No.1, Issues 98. – P.116.

The paragraph argues for the advantages of CBDA, countering criticisms by the advantages it provides, such as reducing researcher bias and allowing for tracking discourse patterns. The subject of the study is thus defined as the corpus of L2 learner which is influenced by chatbot prompts, voice recognition constraints and user adaptation strategies in real-time conversations with an AI chatbot.

In summary, Chapter I synthesizes theories from corpus linguistics and discourse analysis to construct a framework for analyzing a new, technologically mediated genre of L2 spoken discourse with AI chatbots. It establishes that the investigation will focus on a specially compiled learner corpus, examining it through the contextual CBDA approach to analyze syntactic, lexical, pragmatic, and prosodic features. The theoretical conclusion is that L2 learner interactions with AI chatbots constitute a distinct genre of discourse meriting dedicated investigation, thereby providing the basis for a methodologically rigorous and contextually sensitive exploration of L2 spoken discourse in a technologically mediated environment.

The second chapter of the research is titled **“An Approach to Corpus-Based Discourse Analysis of Learner Corpora When AI is an Interlocutor”**.

The chapter outlines the research design and methodological framework for the study and it has three paragraphs. In the paragraph 2.1., the issue of *“Research design and methods of spoken learner corpus collection and annotation”* begins by justifying the choice of a spoken learner corpus as the main dataset. Spoken corpus is a new linguistic field facilitated by advances in technology and corpus findings from learner corpora can help create new framework for teaching and learning English language in the age of AI. Since existing large-scale corpora are either not AI-mediated or are teaching-oriented, the study requires the compilation of a new, small-scale spoken learner corpus.

The initial step in corpus compilation is identifying the discourse community, defined as “a group of people who must use their own genre, specialized vocabulary and shared goals”²³. For this study, this community is university-level English language learners who use AI chatbots for language practice. A pre-task survey is prescribed to establish participant identities with indicators of gender, L1 background, proficiency, chatbot use frequency, and goals to ensure the corpus represents the target variety.

The corpus compilation process follows D.Biber’s criteria: it must be a large or small electronic collection of naturally occurring language, be annotated, and represent a variety of language²⁴. Demographic sampling is adopted to ensure representativeness. For data collection, the chapter selects ASR (Automatic Speech Recognition) technology, specifically the Flow platform by Speechmatics, due to its accuracy with diverse accents and its integration with annotation workflow tools. To obtain both transcriptions and audio, an additional program like Praaline is recommended.

Following raw data collection, the annotation process begins. The annotation plan is the most critical part of any corpus study. The annotation scheme is guided by

²³ Swales J. M. *Genre analysis: English in academic and research settings*. – Cambridge: Cambridge University Press, 1990. – 260 p.

²⁴ Biber D., Connor, U., Upton, T. A. *Discourse on the move: Using corpus analysis to describe discourse structure*. – London: John Benjamins, 2007. – 304 p.

the linguistic categories established in Chapter I which are syntactic-lexical, prosodic, and pragmatic. The paragraph details the automated processes of POS-Tagging, lemmatization, frequency lists, n-gram analysis, lexical bundles analysis, lexical density, lexical diversity, TTR, lexical sophistication, syntactic complexity analysis based on sentence length, clause sophistication, frequency of dependent clauses, T-units, coordination, complexity of nominals and verb phrases, prosodic analysis of repairs, speech rates and dysfluencies for corpus annotation and analysis.

The pragmatic analysis bridges corpus findings to discourse analysis and investigates the categories of conversational acts, discourse markers, politeness strategies (based on Brown and Levinson's FTAs framework) and register.

In the 2.2. paragraph titled "*Selection of approach and methodology for spoken discourse analysis*" and defines the methodology for the discourse analysis. The object of the study is language learners' language use beyond the sentence level. The study employs a mixed-method of quantitative and qualitative methods and a bottom-up analytical framework, starting from corpus patterns to identifying discourse units. The following research design was created by researcher:

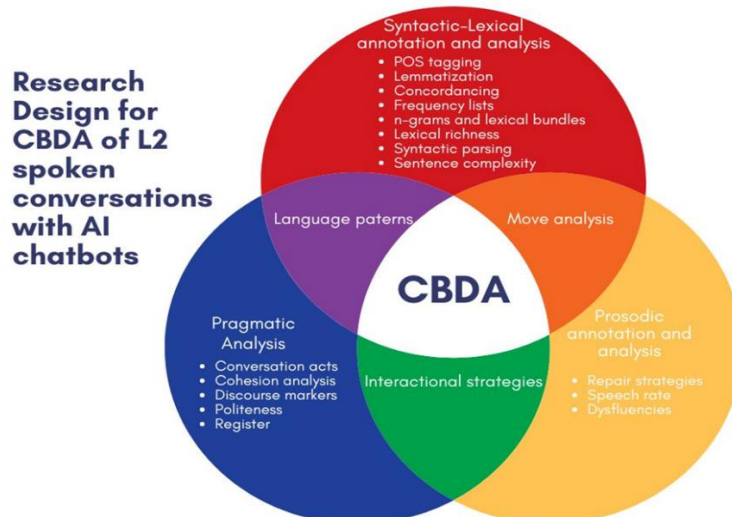


Figure-1. Research design for corpus-based discourse analysis of learner spoken conversations with AI chatbots.

A central methodological task is defining discourse units within the spoken corpus. The study adopts Vocabulary-Based Discourse Units (VBDUs), which are segments automatically identified by shifts in lexical range and topic using tools like the TextTiling program. This bottom-up approach of discourse analysis is adopted for this study as it is computationally driven and suitable for corpus-based studies²⁵. The process involves segmentation, linguistic analysis of each unit, classification, functional grouping, and finally, the identification of general discourse patterns.

The analytical framework for the discourse investigation is the Discourse Contextualization Techniques (DCTs), which applies the lenses of corpus linguistics, contextual framing and thematic analysis techniques.

²⁵Biber D., Connor, U., Upton, T. A. Discourse on the move: Using corpus analysis to describe discourse structure. – London: John Benjamins, 2007. – 304 p.

The final 2.3. paragraph of “*Methods and principles for corpus-based analysis of language patterns, linguistic moves and interactional strategies in discourse*” identifies three core domains of discourse analysis, guided by seven research questions. The research questions form the tasks which serve the goal of the research. The goal of the research is to analyze language patterns, linguistic moves and interactional strategies of spoken conversations between AI chatbots and language learners based on the corpus linguistics.

Language patterns analysis synthesizes corpus findings into a three-dimensional contextual frame: syntactic-lexical, pragmatic and prosodic dimensions. The calculations of the above frames reveal frequencies of recurrent language patterns of learners and profile the discourse community’s linguistic profile.

Linguistic move analysis identifies functional units in the discourse. Since existing models like J.Swales’ CARS model are designed for written academic genres, the study designs a new corpus-based move analysis rubric tailored to AI chatbot conversations, following V.Bhatia’s ten-step approach²⁶. The researcher-developed rubric outlines 7 moves of initiate the conversation, frame the task, elicit information, respond to question, give feedback or repair, expand the topic and end the conversation, their constituent steps, which will be applied to the VBDUs:

Table 1

Move analysis rubric designed by the researcher for AI chatbot conversation analysis

Move categories	Steps
Move 1. Initiating the conversation	Greetings, identity establishment, small talk
Move 2. Framing the task	Initiating the topic, providing background for the topic
Move 3. Eliciting information	Prompts for producing language, elaborating the topic, Offering choices
Move 4. Responding to the question	Providing requested information, giving opinions, performing functional acts of agreeing, disagreeing, apologizing, requesting
Move 5. Giving feedback or providing repair	Self-repair, other-initiated repair, rephrasing, requesting clarification
Move 6. Expanding the topic	Follow-up questions, elaboration on the topic, shifting to related topics
Move 7. Ending the conversation	Summarizing the performance, initiating next practice

Interactional strategies analysis examines how speakers manage communication, drawing on face strategies and framing strategies. The analysis integrates findings from the pragmatic investigation of conversational acts, politeness,

²⁶Bhatia V. Applied genre analysis: Analytical advances and pedagogical procedures. In A. M. Johns (Ed.), *Genre in the Classroom: Multiple Perspectives*. – London: Lawrence Erlbaum Associates, 2002. – P. 279-283.

register, discourse markers and repair strategies. For repair, the four-part framework by Schegloff, Jefferson and Sacks is adopted²⁷.

The chapter concludes by justifying the corpus-based approach to discourse analysis, highlighting that corpus data provides objective, quantitative grounding for discourse analysis. Also, it is incremental, which allows for tracking changes in evolving discourse genres and enables triangulation of results by combining qualitative and quantitative methods. While the paragraph acknowledges limitations of corpus-based analysis, such as corpus data being restricted to verbal language, it successfully asserts that for contextual analysis of patterns and strategies, corpus can provide reliable ground for discourse investigation.

In summary, Chapter II provides a comprehensive, step-by-step methodological blueprint. It moves from the technical design for building and analyzing a specialized L2 learner corpus, to the selection of a contextual, bottom-up, mixed-method discourse analysis approach. The chapter operationalizes the study's goals by specifying precise tools such as Flow, Stanford POS Tagger, LCA, TextTiling; frameworks by leading discourse scholars and analytical procedures of VBDU segmentation and the seven-move rubric that is used to investigate language patterns, linguistic moves and interactional strategies in the subsequent analysis.

In the final chapter of the dissertation, titled **“Corpus-Based Discourse Analysis of Spoken Conversations with AI Chatbots”**, the study presents a comprehensive corpus-based discourse analysis of spoken conversations with AI chatbots. The analysis is conducted through multiple stages: corpus compilation and demographic profiling, part-of-speech tagging and lemmatization, frequency and n-gram analysis, lexical and syntactic complexity assessment, prosodic and pragmatic annotation, and finally, discourse-level examination of moves, interactional strategies and speech acts. The findings illustrate how language learners engage with AI chatbots as a low-risk, high-agency platform for language practice, revealing distinct linguistic, pragmatic, and interactional patterns of the discourse.

In the paragraph 3.1. *“Research results of the corpus-based discourse analysis of spoken conversations with AI chatbots”*, the discourse community was identified using demographic sampling, selecting individuals with the common goal of practicing English language. Participants were predominantly female (89%), technophiles (93%), Uzbek-speaking (95%), with proficiency levels between B1 and C1. Most communicated with AI chatbots very frequently for language practice (89%).

A corpus of 54 samples constituting of 13,450 words was compiled from spoken conversations with AI chatbots Flow and Chat GPT. The corpus was POS-tagged and lemmatized using Stanford TreeTagger. Frequency analysis showed nouns (17.1%) and verbs (16.3%) as the most frequent categories. The top content words included family, sound, good, important, and talk. N-gram analysis indicated that over 71% of the corpus consisted of multi-word expressions, with bigrams and trigrams being most frequent.

²⁷Schegloff E.A., Jefferson G., H.Sacks. The preference for self-correction in the organization of repair in conversation. *Language*. – Cambridge, 1977. Vol.53, No.2. – P. 361-382.

Table 2**Occurrences of each POS in the corpus.**

Category	Occurrences	Percentage
nouns	2426	17.1%
verbs	2317	16.3%
other	2144	14.7%
pronouns	1995	14.0%
adjectives	1284	9.0%
prepositions	1157	8.1%
modifiers	697	5.5%
adverbs	771	5.4%
articles	559	3.9%
question words	343	2.4%
Total	14228	100.0%

Lexical complexity analysis using the Lexical Complexity Analyzer showed that the L2 corpus had a lexical density of 0.21, lexical sophistication of 0.74 (mean), and limited lexical variation (mean 0.002). The AI corpus showed slightly higher normalized lexical sophistication. Syntactic analysis revealed that L2 learners produced more clauses but with shorter mean length (MLC=7.84) compared to AI (MLC=9.22). The L2 corpus demonstrated moderate syntactic complexity with signs of developing subordination and coordination.

Table 3**Lexical complexity profile of L2 speech corpora.**

Dimension	Mean	Minimum	Maximum
Lexical Density	0.85	0.00	32.00
Lexical Sophistication	0.74	0.62	0.86
Lexical Variation	0.002	0.00	0.02

Table 4**Syntactic complexity profile of L2 speech corpora.**

Dimension	Mean	Minimum	Maximum
Sentence Complexity	5.20	1.11	12.30
Clause Complexity	2.65	0.28	7.84
T-Unit Complexity	2.96	0.32	11.10
Phrasal Complexity	0.75	0.50	1.00
Overall Syntactic Control	3.39	0.28	12.30

Prosodic analysis highlighted frequent dysfluencies: fillers in 185 instances, repetitions in 70 instances and pauses in 45. Repair strategies were common, with self-initiated repair occurring 22 times and reformulation 18 times, indicating active self-monitoring and attempts to maintain fluency.

Pragmatic analysis revealed that conversations were predominantly personal, focusing on topics like hometown, family, hobbies, and personal experiences.

Cohesiveness, analyzed via connectives, was slightly higher than average for learner speech (53.68 per 1000 words), with logical connectives being most frequent. Discourse markers were mostly informal (94.3%), with *so*, *well*, *OK*, and *first, then* serving structural and interpersonal functions. Learners showed pragmatic inconsistency, blending informal and formal registers, and frequently used bald-on-record politeness strategies.

In the last paragraph of the dissertation, titled “*Analysis of discursive features: patterns, moves, and strategies*”, corpus findings were qualitatively analyzed for selected discourse features. The discourse was segmented into 16 Vocabulary-Based Discourse Units (VBDUs), categorized into descriptive monologues, narratives, inquiry/elicitation units, opinion/argument units, and meta-discourse markers. These were further grouped into three functional categories: informative, interactive, and persuasive. The discourse was found to be topic-centered, formulaically framed, and characterized by a security zone, in which learners controlled conversation flow to compensate for fluency limitations.

Move analysis identified seven move types, with high frequency in responding to questions (Move 4) and framing the task (Move 2). Learners proactively initiated topics, elicited information, and engaged in self-repair, demonstrating agency in directing their practice.

Table 4

Frequency of moves in L2 corpus.

Move	Description	Total frequency	Most frequent steps
Move 1	Initiate the conversation	17	Greetings (16) Small talk (1)
Move 2	Frame the task (prompt)	21	Topic Initiation (21)
Move 3	Elicit information	13	Prompts for producing language (13)
Move 4	Respond to the question	80	Providing required information / Giving opinions (high)
Move 5	Give feedback or repair	10	Self-repair (3)
Move 6	Expand the topic	9	Shifting to related topics (9)
Move 7	End the conversation	4	Summarizing the performance (3)

Interactional strategies were analyzed using politeness theory and framing. Bald-on-record strategies were most frequent (16.1 per 1000 words), reflecting learners’ adaptation to the AI’s non-human status. Positive politeness was used to build rapport, while negative politeness was less common. The “AI politeness paradox” emerged: despite a calculated Weight of FTAs ($W_x=12$) suggesting need for mitigation, learners used direct strategies, indicating they perceive AI as having no “face” to threaten.

Speech act analysis showed representatives (30.7 per 1000 words) and directives (13.9) as most common, while commissives and declaratives were rare. Repair analysis indicated that self-initiated self-repair (SISR) was dominant, with AI rarely initiating correction, fostering a fluency-focused, low-anxiety environment.

Register analysis revealed a hybrid, asymmetrical register blending casual and formal features, shaped by the purpose of language practice, the learner–tool

relationship, and oral mode. Learners experimented with social identities (expert, peer, director, confidant) without social risk.

This chapter concludes that AI chatbot interactions serve as a low-stakes, high-agency discourse space where L2 learners strategically navigate their interlanguage, prioritizing fluency, self-expression, and pragmatic experimentation over normative accuracy. The linguistic profile shows controlled, intermediate-level production with balanced POS use, limited lexical range, moderate syntactic complexity, and high self-monitoring. Discourse is learner-directed, transactional, and marked by adaptive politeness and register strategies typical for human–AI interaction.

CONCLUSION

The study conducted corpus-based discourse analysis of spoken conversations with artificial intelligence chatbots which was based on empirical investigation and naturally occurring data in contrast to traditional discourse analysis guided by researcher intuition. The findings establish that this approach enhances the validity and reliability of research results, analyzing the most subtle details which are impossible to notice in traditional discourse analysis. The study developed methodological principles and frameworks for spoken corpus collection and annotation along with corpus-based discourse analysis of learner conversations with artificial intelligence chatbots as interlocutors which added to the novelty and significance of the research. The results demonstrate timely and valuable response to growing use of artificial intelligence as speaking partners and challenges posed by emergence of new, technologically-mediated conversational discourse.

1. The theoretical framework and methodological guideline for collecting and compiling spoken corpora to analyze learner discourse have been established by this research. Approaches, methodologies, methods and techniques of corpus-based discourse analysis have been reviewed, most relevant ones selected and recommended.

2. A small-scale reference corpus containing spoken learner corpora was collected from the students of English Philology department of one university and compiled into the first small-scale spoken learner corpus in Uzbekistan. The corpus went through annotation using computational tools available to research community and stored electronically for future use.

3. Corpus analysis of syntactic-lexical, prosodic and pragmatic features of learner corpora was conducted and recurrent linguistic patterns, occurrences of main POS, frequency of all the words, collocations and n-grams, concordances, syntactic and lexical complexity levels, frequency of dysfluencies and repair, distribution of conversational acts and their types, distribution and frequency of discourse markers and cohesiveness of the speech of language learners were identified.

4. Vocabulary-based units of 16 types were identified through computational analysis in learner corpus and classified into informative, interactive and persuasive functional categories. Qualitative analysis of corpus findings revealed the key discourse tendencies: a) the discourse is topic-centered and organized around series of familiar

topics to language learners with shifts marked by verbal clues, b) the discourse is in linguistic “security zone” – it compensates for potential limitations in fluency, c) the discourse is formulaically framed, especially in opening and closing, d) the discourse units display wide spectrum of proficiency with informative functions being simple in structure and word range and persuasive functions containing higher lexical and syntactic complexity indexes and e) the discourse demonstrates strong pragmatic awareness despite frequent inaccuracies.

5. A specific rubric for identifying linguistic move types was developed; the frequency of seven moves in the discourse and the most predominant move steps initiated by learners were calculated and analyzed. High frequency of topic initiation moves was found to be the most distinctive feature of learner discourse.

6. Interactional strategies learners employed throughout the discourse units were investigated. The investigation has revealed that learners a) mostly use Bald-on-Record politeness strategies with the frequency of 16.1 per 1000 words; b) use positive face frames of a “knowledgeable individual” and a “peer”, negative face frames of a “director” and an “examiner”, not applying FTA for “confidant” frame and treating AI as non-human entity with no face to lose; c) employ four types of speech acts with representatives (30.7) and directives (13.9) being the most frequent; d) apply SISR (self-initiated self-repair) as the main repair strategy; e) use hybrid, blended register with features of casual and formal conversation.

7. Contextual power dynamics variables of power (P) and distance (D) were calculated using Weight of FTAs (Wx) formula. The discourse showed low P and maximum D which is not typical of human interaction patterns. A distinct pattern of politeness emerged combining low P and high D which created “AI politeness paradox” which accurately reflected the lack of a social relationship during the conversations and learners’ awareness of non-human status of the interlocutor.

8. Based on findings, the recommendations on design plans of next-generation language learning platforms, including accommodating natural discourse structures, providing adaptive scaffolding, incorporating greater tolerance for planning pauses and dysfluencies and, most importantly, designing AI tutors to facilitate learners’ transition from AI-optimized conversations to human-human interactions were developed.

9. The linguistic and pragmatic dynamics of learner-AI discourse was examined using a corpus-based approach in the Uzbek higher education context for the first time. The findings contribute to filling the gap in existing theoretical frameworks challenged by the emergence of AI-mediated communication as a distinct discourse genre. As the corpus is of small-scale, the findings should be understood as exploratory contribution. The study can be extended and validated through larger-scale and diverse corpora of learner conversations with AI chatbots.

10. The study is limited by less diverse gender sampling which is the reflection of demographic composition of the institution where data collection took place. As the balance of the female participants is high (89%), the findings may not be generalized to male language learners without further investigation. Future research should aim for a more balanced gender distribution to explore potential gender-based differences in AI-mediated learner conversational discourse.

**РАЗОВЫЙ НАУЧНЫЙ СОВЕТ ПРИ НАУЧНОМ СОВЕТЕ
DSc.03/2026.31.01.Fil.45.01 ПО ПРИСУЖДЕНИЮ УЧЕНЫХ
СТЕПЕНЕЙ ПРИ УЗБЕКСКОМ ГОСУДАРСТВЕННОМ
УНИВЕРСИТЕТЕ МИРОВЫХ ЯЗЫКОВ**

**УЗБЕКСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
МИРОВЫХ ЯЗЫКОВ**

АСРОРОВА НАРГИЗА ИСОМИТДИНОВНА

**КОРПУСНЫЙ ДИСКУРС-АНАЛИЗ УСТНЫХ РАЗГОВОРОВ С ИИ-
ЧАТБОТАМИ**

10.00.11 – Прикладная и компьютерная лингвистика.

АВТОРЕФЕРАТ
диссертации доктора философии (PhD) по филологическим наукам

Ташкент – 2026

Тема диссертации доктора философии (PhD) зарегистрирована в Высшей аттестационной комиссии при Академии наук Республики Узбекистан за № B2025.2.PhD/Fil6395.

Диссертация выполнена в Узбекском государственном университете мировых языков.

Автореферат диссертации размещен на трех языках (узбекский, английский, русский (резюме)) на веб-странице Научного совета (www.uzswlu.uz) и на Информационно-образовательном портале «ZiyoNet» (www.ziynet.uz).

Научный руководитель:

Хакимов Хамидулла Иномович
кандидат филологических наук, доцент

Официальные оппоненты:

Рахимов Ганишер Худойкулович
доктор филологических наук, профессор

Джумабаева Джамиля Шариповна
доктор филологических наук, профессор

Ведущая организация:

**Ташкентский государственный
университет узбекского языка и
литературы им. А.Навои**

Защита диссертации состоится «29» октября 2026 года в 9:00 часов на заседании Разового научного совета при научном совете DSc.03/2026.31.01.Fil.45.01 по присуждению ученых степеней при Узбекском государственном университете мировых языков. (Адрес: 100138, г. Ташкент, Учтепинский район, улица Кичик халка йули, дом 21 А. Тел.: (99871) 230-12-91, факс: (99871) 230-12-92, e-mail: uzswlu_info@mail.ru).

С диссертацией можно ознакомиться в Информационно-ресурсном центре Узбекского государственного университета мировых языков (зарегистрирована за № _____). Адрес: 100138, г. Ташкент, Учтепинский район, улица Кичик халка йули, дом 21 А. Тел.: (99871) 230-12-91, факс: (99871) 230-12-92.

Автореферат диссертации разослан «___» _____ 2026 года.

(Протокол реестра рассылки № _____ от «___» _____ 2026 года).

А.А. Насиров

Председатель разового Научного совета по присуждению ученых степеней, д. филол. н., профессор

Х.Б. Самигова

Секретарь разового Научного совета по присуждению ученых степеней, д. филол. н., профессор

Ж.А. Якубов

Председатель разового научного семинара при разовом Научном совете по присуждению ученых степеней, д. филол. н., проф.

ВВЕДЕНИЕ (аннотация диссертации доктора философии (PhD))

Целью исследования является проведение корпусного дискурсивного анализа устных разговоров с чат-ботами искусственного интеллекта.

Объектом исследования выбраны речевые особенности изучающих иностранный язык в устном общении с чат-ботами искусственного интеллекта.

Предметом исследования являются языковые модели, лингвистические действия и коммуникативные стратегии разговорного дискурса, проявляющиеся в речевой деятельности изучающих иностранный язык в процессе взаимодействия с чат-ботами на основе искусственного интеллекта.

Методы исследования. В процессе исследования использовались методы контент-анализа, классификации, описания, сравнения, статистического, корпусного и эмпирического анализа.

Научная новизна исследования заключается в следующем:

на основе корпуса изучены дискурсивные особенности устного общения изучающих иностранный язык с ИИ-чатботами, выявлены модели функционального использования синтактико-лексических, прагматических и просодических единиц в общении, в котором искусственный интеллект выступает в качестве собеседника;

выявлено распределение повторяющихся языковых шаблонов, n-грамм и конкордансов, уровней лексической и синтаксической сложности, дискурсивных маркеров и речевых актов посредством использования ИИ-чатботов в качестве интерактивного собеседника в процессе общения на втором языке;

эффективность лингвистических стратегий, используемых изучающими иностранный язык в процессе устного общения с чат-ботами искусственного интеллекта, доказана на основе частотного распределения, функциональной нагрузки и устойчивого повторения лексических единиц речи (Vocabulary-Based Units) в дискурсе;

в рамках корпусного анализа дискурса «парадокс вежливости искусственного интеллекта» выявлен на основе модели расчета веса действий: низкий показатель Р (сила), отражающий отсутствие социальных отношений между говорящим и собеседником, и высокий показатель D (дистанция).

Внедрение результатов исследования. Результаты исследования внедрены в практику и используются рядом организаций:

научные выводы о выявленных на основе корпусного анализа дискурсивных особенностях устного общения изучающих иностранный язык с ИИ-чатботами, а также об установленных моделях функционального использования синтаксико-лексических, прагматических и просодических единиц в общении с искусственным интеллектом в качестве собеседника, были использованы в рамках «Проекта профессионального развития учителей информатики общеобразовательных школ и их подготовки в качестве

менторов на основе образовательной программы по искусственному интеллекту Массачусетского технологического института», реализуемого согласно приказу Министра дошкольного и школьного образования № 394 от 20 октября 2025 года (Справка Министерства дошкольного и школьного образования Республики Узбекистан № 02-33-4733 от 23 декабря 2025 года). В результате были разработаны практические рекомендации по изучению второго языка и совершенствованию языковых платформ, работающих на основе ИИ;

научные результаты, подтверждающие использование повторяющихся языковых шаблонов, n-грамм и конкордансов, уровней лексической и синтаксической сложности, маркеров дискурса и распределения речевых актов в процессе общения на втором языке с использованием ИИ-чатботов в качестве интерактивных собеседников, были применены в рамках проекта «Интересы женщин в нашей стране находятся под постоянным вниманием», реализованного Объединением «Предприниматели и деловые женщины» Узбекистана при финансировании Общественного фонда поддержки институтов гражданского общества при Олий Мажлисе № 19-12/147 (Справка № 27 от 20 сентября 2025 года Объединения «Предприниматели и деловые женщины» Узбекистана). В результате эти научные выводы были использованы при разработке учебных пособий «Основы информационной безопасности» и «Кибербезопасность», а также специально созданной учебной платформы на основе искусственного интеллекта для развития компетенций слушателей в области ИИ;

на основе научных результатов, доказывающих эффективность лингвистических стратегий, используемых изучающими язык иностранный язык в процессе устного общения с чат-ботами искусственного интеллекта, через частотное распределение, функциональную нагрузку и устойчивое повторение на протяжении всего дискурса речевых единиц, основанных на словаре (Vocabulary-Based Units), разработано и внедрено в практику учебное пособие «Основы информационной безопасности» для направлений бакалавриата филологии и преподавания языков (Справка Министерства высшего образования, науки и инноваций Республики Узбекистан № 02/17-4/4 от 10 января 2026 г.). В результате данное учебное пособие выполняло функцию методического обеспечения и использовалось для эффективной организации обучения;

в рамках проекта по совершенствованию содержания учебной программы по направлению «Иностранные языки (L2 – преподавание в качестве второго или иностранного языка)» национальной образовательной программы, осуществляемого на основе Постановления Кабинета Министров Республики Узбекистан от 27 марта 2024 года № 157 «О мерах по совершенствованию порядка разработки и внедрения государственных образовательных стандартов и государственных образовательных требований» в сотрудничестве с организацией Cambridge University Press & Assessment (Великобритания), были использованы результаты первого

комплексного исследования, основанного на корпусном дискурс-анализе устной коммуникации изучающих L2 с диалоговыми агентами на основе искусственного интеллекта. В частности, были применены выявленные в ходе данного исследования лингвистические стратегии, дискурсивные единицы, а также их частотные и функциональные характеристики (Справка Республиканского центра образования при Министерстве дошкольного и школьного образования Республики Узбекистан № 01/11-6/219 от 8 апреля 2026 года). Полученные научные результаты позволили раскрыть уникальные дискурсивные особенности коммуникации, опосредованной искусственным интеллектом, и эффективно интегрировать их в процесс обучения языку.

Объем и структура диссертации. Диссертация состоит из введения, трех глав, заключения и списка использованной литературы. Общий объем диссертации составляет 142 страницы.

E'OLON QILINGAN ISHLAR RO'YXATI
LIST OF PUBLISHED WORKS
СПИСОК ОПУБЛИКОВАННЫХ РАБОТ
I bo'lim (I part; I chast)

1. Asrorova N.I. Corpus-based discourse analysis: Automatic Speech Recognition (ASR) technologies and spoken corpus collection // Xorijiy lingvistika va lingvodidaktika. – Samarkand, 13-may, 2025. № -4(3). ISSN 2181-3701. DOI: 10.47689/2181-3701. – P. 1-8. (10.00.00; OAK rayosati 30-sentabr 2023 yil 343-son qarori). <https://inscience.uz/index.php/foreign-linguistics/article/view/6918>

2. Asrorova N.I. Corpus-Based Discourse Analysis Tools: A Comparative Study of AntConc and Praaline // The Lingua Spectrum. – Tashkent, 8-iyul 2025. №7(1). – B. 41–48. ISSN 3060-4958. (10.00.00; OAK Rayosatining 2024 yil 30-noyabrdagi 364-son qarori). <https://lingvospektr.uz/index.php/lingsp/article/view/969>

3. Asrorova N.I. Selecting an Approach and Methods for Corpus-Based Discourse Analysis // The Lingua Spectrum. – Tashkent, 5-oktabr 2025. №10(1). – B. 33-39. ISSN 3060-4958. (10.00.00; OAK Rayosatining 2024 yil 30-noyabrdagi 364-son qarori). <https://lingvospektr.uz/index.php/lingsp/article/view/1085>

4. Asrorova N.I., Asrorov X.Q. Ta'lim natijalari (learning outcomes)ga asoslangan xalqaro baholash dasturlarining didaktik imkoniyatlari. // Filologiya va pedagogika, 2025. – №6(10)-son. – B. 99-101. (10.00.00; OAK Rayosatining 2024 yil 30-noyabrdagi 364/5-son qarori). https://tilvaadabiyot.uz/f/6-sonnn_031992.pdf

5. Asrorova N.I. Advantages and Limitations of Corpus-Based Discourse Analysis: A Case Study of Learner Conversations with AI Chatbots // Journal of Advanced Linguistic Studies (JALS). – Delhi, India 2025. №12(2). – P. 951-962. ISSN 2231-4075. (10.00.00; № 13). <https://bahripublications.co.in/journal-of-advanced-linguistic-studies/>
(<https://bahripublications.co.in/pdf/JALS/JALS%2C%20No.%202%2C%20Jul-Dec%202025/70%20-%20Advantages%20and%20Limitations%20-%20FULLY%20FINAL.pdf?t=1763229034>)

6. Asrorova N.I. Combining Corpus Linguistics with Critical Discourse Analysis // “Kompyuter lingvistikasining zamonaviy texnologiyalari – CTCL.2025” nomli xalqaro ilmiy-amaliy konferensiyasining ilmiy maqolalari va tezislari to'plami. – Toshkent, O'ZMU, 2025. – B. 209-212. <https://ctcl.uz/index.php/uzmu/article/view/53>

7. Asrorova N.I. The Analytical Lenses in Corpus-Based Discourse Analysis // Materials of ISOC International Online Conference, Turkey, November 2025. – P. 109-112. <https://interoncof.com/index.php/turkey/article/view/15928>

8. Asrorova N.I. Analysis of AI-Based Tools for Enhancing Second Language Speaking Skills // “Ilm-fan, innovatsiyalar, ta'lim: xorijiy tillarni o'qitishning dolzarb masalalari” mavzusidagi Respublika ilmiy amaliy anjumani materiallari. TIFT Universiteti, Toshkent, 17-may 2025. – B. 379-382. <https://zenodo.org/records/22824688>

9. Asrorova N.I. The Basics of Corpus Collection for Research: Principles, Methods and Pitfalls // “Ma’lumot asrida ta’lim tizimlarini muvofiqlashtirish: Amaliyot, innovatsiya va madaniy yondashuvlar” mavzusidagi Respublika ilmiy amaliy anjumani materiallari. Toshkent xalqaro moliyaviy boshqaruv va texnologiyalar universiteti, Toshkent, 25-aprel 2026. – B. 138-141.
<https://doi.org/10.5281/zenodo.19825255>

II bo‘lim (II part; II chast)

10. Asrorova N.I. Corpus-Based Analysis of Linguistic Moves in Discourse // Materials of the “Studies in Economics and Education in the Modern World” International Scientific Conference, Canada, November 2025. – B. 57-63.
<https://doi.org/10.5281/zenodo.22824897>

11. Asrorova N.I. Compiling and Utilizing L2 Speech Corpora: A Framework for Research and Pedagogy in Second Language Acquisition // Научный импульс. – Moskva, Sentabr 2025. №37 (100). – B. 272-276.
<https://nauchnyimpuls.org/index.php/impuls/article/view/1703>

12. Asrorova N.I. A Comparative Analysis of Methodological Approaches in Discourse Studies: Corpus-Based and Traditional Paradigms // Научный импульс. – Moskva, Sentabr 2025. №37 (100). – B. 277-281.
<https://nauchnyimpuls.org/index.php/impuls/article/view/1704>

13. Asrorova N.I. Spoken Discourse Analysis // Новости образования: исследование в XXI веке. – Moskva, Oktabr 2025. №38 (100). – B. 119-122.
<https://nauchnyimpuls.org/index.php/obrazovaniya/article/view/1453>

14. Asrorova N.I. Advantages of a Corpus-Based Approach in Discourse Analysis // Materials of the “Studies in Economics and Education in the Modern World” International Scientific Conference, Canada, November 2025. – B. 64-70.
<https://doi.org/10.5281/zenodo.22824862>