



O‘ZBEK TILI GAP BO‘LAKLARINI MODELLASHTIRISHDA CONSTITUENCY GRAMMAR

Hamroyeva Shahlo Mirdjonova,
Filologiya fanlari doktori, professor v.b.
shaxlo.xamrayeva@navoiy-uni.uz
ToshDO‘TAU

Nuriddinova Gulira’no Baxtiyor qizi,
1-kurs magistrant
guliranonuriddinova@gmail.com
ToshDO‘TAU

Annotatsiya. Ushbu maqolada o‘zbek tili gap bo‘laklarini zamonaviy tilshunoslikning tarkibiy grammatika (*Constituency Grammar*) modeli asosida modellashtirish masalalari tadqiq etiladi. Maqolada mazkur modelning nazariy asoslarini o‘zbek tili sintaktik qurilishiga tatbiq etilish imkoniyatlari tahlil qilingan. Tadqiqot davomida o‘zbek tili kabi agglyutinativ tillarda gap bo‘laklarini iyerarxik konstituentlarga ajratishning afzalliklari va so‘z tartibi erkinligi bilan bog‘liq tizimli muammolar yoritib berilgan. Xulosa qismida tarkibiy grammatikaning o‘zbek tili sintaktik tahlilidagi o‘rni hamda tabiiy tilni qayta ishlash tizimlari uchun uning yutuq va kamchiliklari qiyosiy baholangan.

Kalit so‘zlar: *o‘zbek tili, tarkibiy grammatika, constituency grammar, gap bo‘laklari, modellashtirish, iyerarxik struktura, NLP, sintaktik analizator.*

Abstract. This article explores the issues of modeling Uzbek sentence constituents based on the Constituency Grammar framework of modern linguistics. The study analyzes the possibilities of applying the theoretical foundations of this model to the syntactic structure of the Uzbek language. Throughout the research, the advantages of partitioning sentence components into hierarchical constituents in agglutinative languages, such as Uzbek, are highlighted alongside the systemic challenges arising from word order flexibility. In the concluding section, the role of constituency grammar in the syntactic analysis of the Uzbek language is evaluated,



and its strengths and weaknesses for Natural Language Processing (NLP) systems are comparatively assessed.

Keywords: Uzbek language, constituency grammar, sentence constituents, modeling, hierarchical structure, NLP, syntactic parser.

Tilshunoslik fani har bir davrda o'z rivojlanish bosqichlariga ega bo'lib, gapning sintaktik qurilishini o'rganish bu davrlar uchun markaziy masalalardan biri bo'lib kelgan. Klassik tilshunoslikda gap bo'laklari an'anaviy qonun-qoidalar, tasniflar asosida o'rganilgan bo'lsa, XXI asrga kelib kompyuter lingvistikasi va sun'iy intellekt texnologiyalarining taraqqiyoti sintaksis sohasida yangicha yondashuvlarni talab etmoqda. O'zbek tili sintaktik qurilishida so'z birikmalarining iyerarxik tizimi gapning mazmuniy butunligini ta'minlovchi asosiy omildir. Gapdagi so'zlar shunchaki tartib bilan emas, balki ma'lum bir grammatik va mantiqiy qonuniyatlar asosida birikib, butun bir strukturani hosil qiladi [7:25]. Gap bo'laklarini modellashtirish esa tilni qayta ishlash jarayonidagi eng asosiy va murakkab masalalardan biridir. Umuman olganda, modellashtirishda tadqiqotchi obyektning o'zini emas, balki uning modelini o'rganadi [6:21]. Bugungi kunda jahon tilshunosligida gapning strukturaviy va funksional mohiyatini ochib berishda uchta asosiy model – tobelik grammatikasi (*Dependency Grammar*), tarkibiy grammatika (*Constituency Grammar*) va ularni umumlashtiruvchi *Universal Dependencies* tizimi ustuvorlik qilmoqda [4:12-15].

Ushbu tadqiqotning dolzarbligi shundaki, an'anaviy grammatik tahlil usullari inson nutqini tushunish uchun yetarli bo'lsa-da, mashinali tarjima va matn tahlili kabi raqamli texnologiyalar uchun matematik aniqlikdagi sintaktik modellar zarurdir. N. Chomsky tomonidan asos solingan tarkibiy grammatika gapni iyerarxik frazalar yig'indisi sifatida ko'rib, qat'iy so'z tartibiga ega tillarni tahlil qilishda samarali natija beradi [5:44-48]. Aksincha, L. Tesniere asarlarida poydevor qo'yilgan tobelik grammatikasi so'zlar orasidagi bevosita bog'liqlikni asos qilib

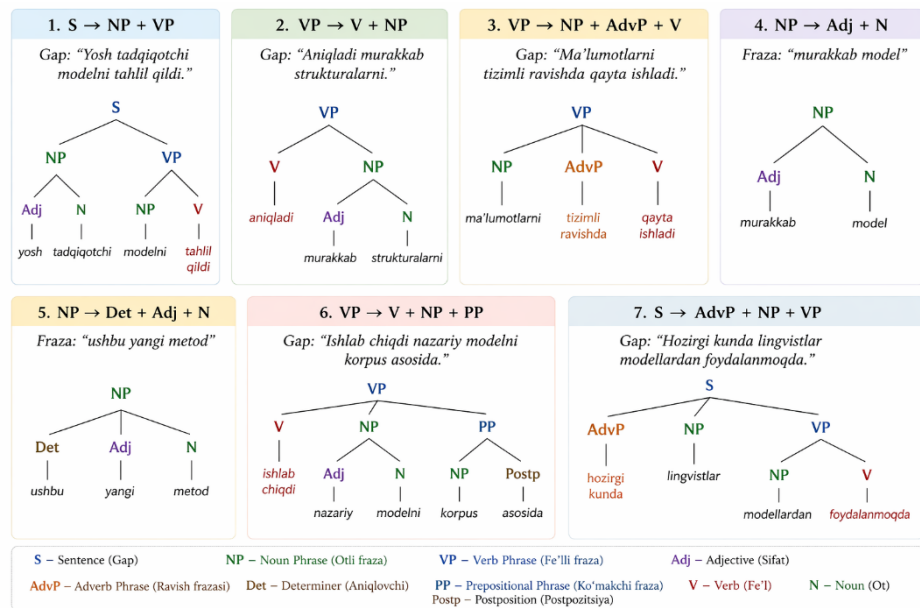
oladi va o'zbek tili kabi erkin so'z tartibiga ega hamda agglyutinativ xususiyatli tillarni modellashtirishda keng imkoniyatlar yaratadi [8:102-105]. Ushbu maqolaning maqsadi jahon tilshunosligidagi sintaktik modellarning o'ziga xos xususiyatlarini qiyosiy tahlil qilish va ularning o'zbek tili misolidagi tatbiqini ko'rib chiqishdan iborat. Ushbu tahlillar o'zbek tili korpus lingvistikasini rivojlantirish va milliy tilimizning raqamli muhitdagi mavqeyini mustahkamlashda muhim nazariy asos bo'lib xizmat qiladi.

Constituency grammar (tarkibiy grammatika). Tarkibiy grammatika sintaktik tahlilning muhim yo'nalishlaridan biri hisoblanadi. Gapni iyerarxik tarzda tashkil topgan tarkibiy qismlarga ajratish g'oyasiga asoslanadi. Ushbu yondashuv XX asr o'rtalarida Noam Chomsky tomonidan ishlab chiqilgan generativ grammatika nazariyasiga borib taqaladi [5:45–46]. Tarkibiy grammatika doirasida gap fraza tuzilmasi asosida tahlil qilinadi, ya'ni gap alohida so'zlar yig'indisi emas, balki o'zaro bog'langan frazalardan tashkil topgan tizim sifatida qaraladi [5:112]. Tarkibiy grammatikada asosiy birliklar sifatida otli fraza (NP), fe'lli fraza (VP), ko'makchili fraza (PP) qaraladi. Bu yondashuv fraza tuzilma qoidalari orqali tushuntiriladi [5:67]. Tarkibli grammatikada asosiy qoida $S=NP \setminus VP$ (Gap = Otli iboralar + Fe'lli ibora). Ishlatilgan belgilari (Teglar) esa quyidagicha:

1-jadval. Constituency grammardagi belgilar tavsifi

Belgilar	Tavsifi
S (Sentence)	Gap
NP (Noun Phrase)	Otli ibora
VP (Verb Phrase)	Fe'lli ibora
DP (Determiner Phrase/Possessive)	Egalik/Aniqllovchi iborasi
AP (Adjective Phrase)	Sifatli ibora

Qoidalarga asoslangan yondashuvga ko'ra Constituency grammarning modellari quyidagicha:



1-rasm. Tarkibli grammatikada modellashtirish

Constituency grammaring o'zbek tili gap bo'laklarini modellashtirishdagi ahamiyati. Yuqoridagi rasmdan ko'rinib turibdiki, Constituency grammaring o'zbek tilidagi gap bo'laklarini aniqlashdagi bir qancha yutuq va kamchiliklari ko'zga tashlanadi. Modelning ijobiy tomonlari sifatida dastlab uning iyerarxik aniqligini aytishimiz mumkin. Ushbu model gapni shunchaki so'zlar zanjiri emas, balki so'zlar yig'indisi sifatida ko'radi. Bu, ayniqsa, o'zbek tilidagi murakkab aniqlovchi va aniqlanmish birikmalarini (masalan, [qizil matodan tikilgan]ko'ylak) yaxlit bitta sintaktik birlik (NP) sifatida tahlil qilishda juda qo'l keladi. Bundan tashqari, u gapdagi sintaktik noaniqliklarni daraxtsimon struktura orqali aniq ajratib beradi. Masalan, "Guruhdagi bolalar va qizlar" birikmasida "guruhdagi" sifati faqat "bolalar"ga yoki ikkala guruhga ham tegishli ekanligini tarkibiy qismlar orqali ko'rsatadi. Generativ imkoniyati ham borki, model tildagi cheksiz miqdordagi gaplarni chekli miqdordagi qoidalar (masalan, S=NP+VP) orqali hosil qilish imkonini beradi, bu esa kompyuter lingvistikasi uchun muhim algoritmik asosdir. Modelning kamchiliklari esa o'zbek tili agglyutinativ va so'z tartibi nisbatan erkin til bo'lganligi uchun, tarkibiy qismlar ko'pincha gap ichida bir-biridan uzoqda joylashishi mumkin. Masalan, "Men kitobni bugun o'qishni boshladim" gapida



obyekt va fe'l orasiga boshqa bo'laklar tushishi qat'iy qoidalarga asoslangan ibora strukturasi (CFG) tahlilini murakkablashtiradi. Bundan tashqari, tarkibiy grammatika ko'proq so'zlarning ketma-ketlik tartibiga tayanadi. O'zbek tilida esa gap bo'laklari orasidagi munosabat ko'pincha so'z tartibi bilan emas, balki qo'shimchalar orqali ifodalanadi. Bu modelda morfologik ko'rsatkichlarni ifodalash uchun qo'shimcha qoidalar talab etadi. O'zbek tili uchun tarkibiy daraxtlarni qurishda juda ko'p miqdordagi “mavhum” tugunlar (NP, VP, AP) hosil bo'lishi ham tabiiy tilni qayta ishlashda modelning ishlash tezligini pasaytiradi.

Xulosa qilib aytganda, *Constituency Grammar* o'zbek tili sintaktik birliklarini iyerarxik guruhlashda muhim o'ringa ega. Biroq, amaliyotda –xususan, o'zbek tili uchun sintaktik analizatorlar va Treebanklar yaratishda ushbu modelni *Universal Dependencies* prinsiplari bilan uyg'unlashtirish yoki tobelik grammatikasiga ustuvorlik berish maqsadga muvofiqdir. Bu yondashuv o'zbek tili grammatikasining ham morfologik, ham strukturaviy xususiyatlarini to'laqonli aks ettirish imkonini beradi.

Foydalanilgan adabiyotlar ro'yxati

1. Carnie A. *Syntax: A Generative Introduction*. Oxford, UK: Wiley-Blackwell, 2013.
2. G'ulomov A., Asqarova M. *Hozirgi o'zbek adabiy tili. Sintaksis*. – Toshkent: O'qituvchi, 1987.
3. Jurafsky D., Martin J. *Speech and Language Processing*. Stanford, USA: Stanford University, 2023.
4. Nivre J. *Dependency Grammar and Dependency Parsing*. MSI report, Vaxjo University, 2005.
5. Noam Chomsky. *Syntactic Structures*. The Hague, Netherlands: Mouton Publishers, 1957.
6. Po'latov A. *Kompyuter lingvistikasi*. – Toshkent, 2009.



7. Tesniere L. Elements de syntaxe structurale. – Klincksieck, 1959.
8. Zeman D. Universal Dependencies 2.6: Guidelines for Annotation.
LINDAT/CLARIAH-CZ, 2020