

**ЎЗБЕКИСТОН РЕСПУБЛИКАСИ ОЛИЙ ВА ЎРТА МАХСУС
ТАЪЛИМ ВАЗИРЛИГИ
ЖИЗЗАХ ДАВЛАТ ПЕДАГОГИКА ИНСТИТУТИ**

Қўлёзма ҳуқуқида
УДК _____

ХИДИРОВ ОТАБЕК ЖЎРАБОЕВИЧ

**МИЛЛИЙ КОРПУС УЧУН ПАРСИНГ ДАСТУРИ ЯРАТИШНИНГ
ЛИНГВИСТИК АСОСЛАРИ**

10.00.01 – Ўзбек тили

Филология фанлари бўйича фалсафа доктори (PhD)
илмий даражасини олиш учун тақдим этилган
ДИССЕРТАЦИЯ

**Илмий раҳбар: Б.Р.Менглиев,
филология фанлари доктори, профессор**

Жиззах - 2021

МУНДАРИЖА

КИРИШ (фалсафа доктори диссертацияси аннотацияси)	3
I БОБ. СИНТАКТИК БИРЛИКЛАРНИ ТЕГЛАШНИНГ НАЗАРИЙ АСОСЛАРИ.....	12
1.1. Жаҳон тилшунослигида синтактик теглаш назарияси ва амалиёти хусусида	12
1.2. Синтактик разметка ва унинг турли корпуслардаги имкониятлари	21
1.3. Синтактик разметкаланган корпуслар ва уларнинг дастурий таъминоти борасида айрим мулоҳазалар	28
Боб бўйича қисқа хулосалар	37
II БОБ. СЎЗ БИРИКМАСИ ВА СОДДА ГАПЛАРНИ СИНТАКТИК ТЕГЛАШНИНГ НАЗАРИЙ АСОСЛАРИ.....	39
2.1. Ўзбек тили корпусида сўз бирикмаларини синтактик теглаш хусусида .	39
2.2. Ўзбек тилида содда йиғиқ гапларни синтактик теглаш: лисоний-синтактик қолип ва лингвистик модел.....	52
2.3. Ўзбек тилидаги содда ёйиқ гапларни синтактик теглаш масалалари	63
Боб бўйича қисқа хулосалар	68
III БОБ. ЎЗБЕК ТИЛИДА ҚЎШМА ГАПЛАРНИ СИНТАКТИК ТЕГЛАШНИНГ НАЗАРИЙ АСОСЛАРИ.....	71
3.1. Ўзбек тилида боғланган ва боғловчисиз қўшма гапларни синтактик моделлаштириш масалалари	71
3.2. Ўзбек тилида эргашган қўшма гапларни синтактик теглаш ва моделлаштириш.....	82
3.3. Ўзбек тилида мураккаб қўшма гапларни синтактик теглаш хусусида...	102
Боб бўйича қисқа хулосалар	107
УМУМИЙ ХУЛОСА	109
Фойдаланилган адабиётлар рўйхати	113
ИЛОВАЛАР.....	123

КИРИШ (фалсафа доктори (PhD) диссертацияси аннотацияси)

Диссертация мавзусининг долзарблиги ва зарурати. Жаҳон тилшунослигида компьютер ва корпус лингвистикаси муаммоларига эътибор XX асрнинг иккинчи ярмидан бошлаб жадаллашди, XXI аср бошларида катта ҳажмли тил корпуслари пайдо бўлмоқда. Автоматик таржима, электрон луғат, лингвистик корпуслардан фойдаланиш имконияти янада кенгаймоқда. Мазкур янгиланишлар ахборот технологияларини тилшуносликка татбиқ этиш билан боғлиқ истиқболли илмий йўналишлар пайдо бўлишига йўл очди. Бу эса тил бирликларини корпус материали сифатида теглаш тамойилларига бўлган эҳтиёжни кун тартибига қўймоқда.

Дунё тилшунослигида XXI асрга келиб, корпус лингвистикасини ўрганиш ҳаракати жадал тус олмақда. Компьютер лингвистикаси йўналишида автоматик таржима сифатини яхшилаш, тил бирликларини теглаш назарияси, алгоритми ҳамда лингвистик таъминотини яратиш, матн таҳлил қилиш (таггинг, парсинг, спелккер) дастурлари тузиш жаҳон компьютер лингвистикасида долзарб масалага айланиб бормоқда. Тилшуносликда, хусусан, компьютер лингвистикаси соҳасида корпус бирликларини синтактик теглаш, тег моделлари, теглаш алгоритминини ишлаб чиқиш, шу асосда автоматик синтактик теглаш (парсинг) дастуринини ишлаб чиқишга зарурат сезилмоқда.

Истиқлол йилларида компьютер лингвистикасида автоматик таржима, сунъий интеллектнинг ўзбек тилини тушуниш ва қайта ишлашига эришиш борасида қатор ишлар амалга оширилмоқда. Бинобарин, “...давлат тилининг софлигини сақлаш, уни бойитиб бориш ва аҳолининг нутқ маданиятини ошириш; давлат тилининг замонавий ахборот технологиялари ва коммуникацияларига фаол интеграциялашувини таъминлаш”¹ бугунги кунда

¹Ўзбекистон Республикаси Президенти Шавкат Мирзиёевнинг 2020 йил 20 октябрдаги “Мамлакатимизда ўзбек тилини янада ривожлантириш ва тил сиёсатини такомиллаштириш чора-тадбирлари тўғрисида”ги ПФ-6084-сон фармони // <https://lex.uz/docs/5058351>

Ўзбек тилшунослиги олдида турган долзарб вазифадир. Мамлакатимизда давлат тилига эътибор давлат сиёсатининг устувор йўналишларидан бири даражасига кўтарилди. Корпус лингвистикаси истиқболли илмий йўналиш эканлигини инобатга олган ҳолда ўзбек тили миллий корпусини яратиш, лингвистик моделларни тузиш сингари масалаларни замонавий илмий тамойиллар асосида тадқиқ этиш фанимиз олдида турган долзарб вазифалардан биридир.

Ўзбекистон Республикаси Президентининг 2016 йил 13 майдаги ПФ-4997-сон “Алишер Навоий номидаги Тошкент давлат ўзбек тили ва адабиёти университетини ташкил этиш тўғрисида”, 2017 йил 7 февралдаги ПФ-4947-сон “Ўзбекистон Республикасини янада ривожлантириш бўйича Ҳаракатлар стратегияси тўғрисида”, 2019 йил 21 октябрдаги ПФ-5850-сон “Ўзбек тилининг давлат тили сифатидаги нуфузи ва мавқеини тубдан ошириш чора-тадбирлари тўғрисида”, 2020 йил 20 октябрдаги ПФ-6084-сон “Мамлакатимизда ўзбек тилини янада ривожлантириш ва тил сиёсатини такомиллаштириш чора-тадбирлари тўғрисида”ги фармонлари, 2019 йил 4 октябрдаги ПҚ-4479-сон “Ўзбекистон Республикасининг “Давлат тили ҳақида”ги Қонуни қабул қилинганлигининг ўттиз йиллигини кенг нишонлаш тўғрисида” қарорлари ҳамда бошқа меъёрий-ҳуқуқий ҳужжатларда белгиланган вазифаларни амалга оширишда ушбу тадқиқот муайян даражада хизмат қилади.

Тадқиқотнинг республика фан ва технологиялари ривожланишининг устувор йўналишларига мослиги. Диссертация республика фан ва технологиялари ривожланишининг I. «Ахборотлашган жамият ва демократик давлатни ижтимоий, ҳуқуқий, иқтисодий, маданий, маънавий-маърифий ривожлантиришда инновацион ғоялар тизимини шакллантириш ва уларни амалга ошириш йўллари» устувор йўналишига мувофиқ бажарилган.

Муаммонинг ўрганилганлик даражаси. Жаҳон тилшунослигида корпус соҳасидаги мақсадли тадқиқотлар XX асрнинг 40-йилларида

Блумфильд, Фрайс ва Бонджерслар томонидан бошланган²; Н.Френсис ва Г.Кучера илк марта корпус тузиш принципларини ишлаб чиққан³. Рус тилшунослигида В.П.Захаров, А.Б.Кутузов, Е.В.Недошивина, В.В.Риков, В.А.Плунгянлар корпус, унинг турлари, корпус тузиш ва теглаш тамойиллари борасида тадқиқот олиб боришган⁴. Корпус бирликларини синтактик теглаш муаммолари, парсинг дастурлари яратиш Д.Бибер, С.Конрад, Р.Реппен⁵, Э.Финеган⁶, М.В.Копотев, Г.Б.Гурин⁷, И.М.Ножов⁸, Ю.Д.Апресян, И.М.Богуславский, Л.Л.Иомдин⁹ тадқиқотлари предмети бўлган.

Туркий тилларнинг автоматик синтактик разметкаси масаласи Т.Френсис, Ж.Вашингтон, Ч.Чўлтекин, А.Макажанов¹⁰, В.П.Желтов, П.В.Желтов¹¹ ларнинг тадқиқотида кун тартибига қўйилган.

Ўзбек тилшунослигида компьютер лингвистикаси соҳасида бир канча тадқиқотлар амалга оширилган. Жумладан, А.Қ.Пўлатов, С.М.Муҳаммедов, С.Муҳаммедова, Д.Б.Ўринбоева, Н.З.Абдурахмонова, А.М.Норов ва

²Блумфильд Л. Язык. – Москва: Прогресс, 1968. – 608 с.; Fries Ch.C. The structure of English. An introduction to the construction of English sentences. – New York: Harcourt, 1952. – 304 p.; Bongers H. The history and principles of Vocabulary control. – Woerden: WOCOPI, 1947. – 442 p.

³Френсис Н., Кучера Г. Вычислительный анализ современного американского варианта английского языка. – Москва, 1967. – 150 с.; Синклер Д. Предисловие к книге “Как использовать корпуса в преподавании иностранного языка”/ Д. Синклер [Электронный ресурс]. – Режим доступа: <http://www.ruscorgora.ru/corgora-info.html>, свободный.

⁴Захаров В.П. Корпусная лингвистика. Учебно-методическое пособие. – Санкт-Петербург, 2005. – 69 с. – С. 27.; Кутузов А.Б. Корпусная лингвистика. – (Электрон ресурс): Лицензия Creative commons Attribution Share-Alike 3.0 Unported (Электрон ресурс) - //lab314.brsu.by/kmp-lite/kmp-video/CL/CorporeLingva.pdf; Недошивина Е.В. Программы для работы с корпусами текстов: обзор основных корпусных менеджеров. Учебно-методическое пособие. – Санкт-Петербург, 2006. – 126 с.; Риков В.В. Курс лекций по корпусной лингвистике. URL: <http://rykov-cl.narod.ru/c.html>; Плунгян В. Зачем мы делаем Национальный корпус русского языка? “Отечественные записки” 2005, №2. http://magazines.russ.ru/oz/2005/2/2005_2_20-pr.html

⁵Biber D., Conrad S., Reppen R. Corpus linguistics. Investigating language structure and use. – Cambridge University Press, 1998. – 189 p.

⁶Finegan E. Language: its structure and use. – N.Y: Harcourt Brace College Publishers, 2004. – 152 p.

⁷Копотев М.В., Гурин Г.Б. Принципы синтаксической разметки хельсинского аннотированного корпуса русских текстов ХАНКО. (электрон ресурс): www.dialog-21.ru.

⁸Ножов И.М. Морфологическая и синтаксическая обработка текста (модели и программы) сегментации русского предложения. – М.: АКД, 2003. – 152 с.

⁹Апресян Ю.Д., Богуславский И.М., Иомдин Л.Л. и др. Лингвистическое обеспечение системы ЭТАП-2. – М.: Наука, 1989. – 190 с.

¹⁰Francis M. Tyers, Jonathan Washington, Çağrı Çöltekin, Aibek Makazhanov. Оценка критериев морфо-синтаксической разметки для тюркских языков в проекте «UNIVERSAL DEPENDENCIES» // Пятая Международная конференция по компьютерной обработке тюркских языков «TurkLang 2017». – Труды конференции. В 2-х томах. Т 1. – Казань: Издательство Академии наук Республики Татарстан, 2017. – 380 с. – Б. 356-362.

¹¹Желтов В.П., Желтов П.В. Синтаксический анализатор национального корпуса чувашского языка // Пятая Международная конференция по компьютерной обработке тюркских языков «TurkLang 2017». – Труды конференции. В 2-х томах. Т 1. – Казань: Издательство Академии наук Республики Татарстан, 2017. – 380 с. – Б. 304-315.

бошқаларнинг изланишлари лингвостатистик, лексикографик муаммоларни компьютер ёрдамида ечиш масалаларига бағишланган¹².

Ўзбек тили матнларини синтактик теглаш муаммолари махсус тадқиқ предмети бўлмаган, ammo айрим изланишларда масаланинг муайян қирралари ҳақида фикр билдирилган. Чунончи, ҳаракат феъллари асосида компьютер дастурлари учун лингвистик таъмин яратиш¹³, ўзбек-инглиз тили машина таржимасининг лингвистик таъминоти¹⁴, ўзбек тили муаллифлик корпусини тузиш тамойиллари¹⁵, ўзбек тили бирликларини графематик таҳлил қилиш муаммолари¹⁶, тил корпуси лингвистик базасини тузиш тамойиллари¹⁷, ўзбек тили атов бирликларини семантик теглашнинг лингвистик асослари¹⁸, ўзбек тили морфологик анализаторининг лингвистик таъминоти муаммолари¹⁹ монографик планда ўрганилган. Шунингдек, ўзбек-

¹²Муҳаммедов С.М. Статистический анализ лексико-морфологической структуры узбекских газетных текстов: Автореф. дисс... канд. фил. наук.–Ташкент, 1980.-180 с.; Бабанаров А. Разработка принципов построения словарного обеспечения турецко-русского машинного перевода: Автореф. дисс... канд. фил. наук. – Ленинград, 1981.-221 с.; Айымбетов Н.К. Опыт лингвостатистического анализа лексики и морфологии каракалпакского публицистического текста: Автореф. дисс... канд. фил. наук.– Ташкент, 1987.-180 с.; Муҳаммедова С. Ҳаракат феъллари асосида компьютер дастурлари учун лингвистик таъмин яратиш. Методик қўлланма.–Тошкент, 2006.–45 б.; Ўринбоева Д.Б. Ўзбек фольклори матнларининг лингвостатистик тадқиқи. – Тошкент: Фан, 2010. – 121 б.; Жуманазарова Г.У. Фозил Йўлдош ўғли дostonлари тилининг лингвопоэтикаси: Фил. фан. док. дис...автореф.–Тошкент, 2017.-52 б; Абдурахмонова Н.З. Инглизча матнларни ўзбек тилига таржима қилиш дастурининг лингвистик таъминоти (содда гаплар мисолида): Фил.фан.бўйича фалсафа доктори (PhD)...дис. автореф. – Тошкент, 2018.-52 б; Пўлатов А. Компьютер лингвистикаси. – Тошкент: Akademnashr, 2011.-175 б.; Норов А. Компьютер лингвистикаси асослари. – Қарши, 2017. - 136 б.

¹³Муҳаммедова С. Ҳаракат феъллари асосида компьютер дастурлари учун лингвистик таъмин яратиш. Методик қўлланма. –Тошкент, 2006.-98 б.;

¹⁴Абдурахмонова Н.З. Инглизча матнларни ўзбек тилига таржима қилиш дастурининг лингвистик таъминоти (Содда гаплар мисолида). Филол.фан.бўйича фалсафа доктори (PhD)...дис. афтореф. – Тошкент, 2018. – 52 б.

¹⁵Хамроева Ш. Ўзбек тили муаллифлик корпусини тузишнинг лингвистик асослари: Филол.фан.бўйича фалсафа доктори (PhD)...диссер. – Қарши, 2018. – 250 б.

¹⁶Абжалова М.А. Ўзбек тилидаги матнларни таҳрир ва таҳлил қилувчи дастурнинг лингвистик модуллари (Расмий ва илмий услубдаги матнлар таҳрири дастури учун): Филол.фан.бўйича фалсафа доктори (PhD)...диссер. – Фарғона, 2019. – 164 б.; Абжалова М. Матнларга автолингвистик ишлов бериш тизимлари // Шестая Международная конференция по компьютерной обработке тюркских языков «TurkLang-2018». (Труды конференции) – Ташкент, 2018. – 320 с.

¹⁷Эшмўминов А.А. Ўзбек тили миллий корпусининг синоним сўзлар базаси: Филол.фан.бўйича фалсафа доктори (PhD)...диссер. – Қарши, 2019. – 140 б.;

¹⁸Akhmedova D.B., Mengliev B.R. Semantic Tag Categories in Corpus Linguistics: Experience and Examination International Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277-3878, Volume-8, Issue-3S, October 2019. – P. 208-212.

¹⁹Хамроева Ш. Ўзбек тили морфологик анализаторининг лингвистик таъминоти. Филология фанлари доктори диссертацияси автореферати. – Фарғона, 2021. – 78 б.

инглиз параллел корпуси тузиш муаммолари²⁰, ўзбек тилидаги отларни автоматик таҳлил қилиш²¹, сўз ясалишининг формал моделлари²² хақида катор мақолалар нашр этилган. Бу тадқиқотлар компьютер лингвистикаси муаммоларини тадқиқ этганлиги билан долзарблик касб этган, аммо ўзбек тили парсинг дастури – матнни автоматик синтактик таҳлил қилиш муаммоси кун тартибига махсус қўйилмаган. Диссертацияни тайёрлаш жараёнида юқорида санаб ўтилган илмий изланишлар атрофлича ўрганилди, зарур ўринларда уларга муносабат билдирилди ва улардан тадқиқотда фойдаланилди.

Тадқиқотнинг диссертация бажарилган олий таълим муассасасининг илмий-тадқиқот ишлари режалари билан боғлиқлиги. Диссертация Жиззах давлат педагогика институти илмий-тадқиқот ишлари режасининг 1-сон «Филология фанларининг долзарб муаммолари ва уларни амалиётга жорий этишнинг янги педагогик технологиялари» (2018-2020 йй) мавзуси доирасида бажарилган.

Тадқиқотнинг мақсади ўзбек тилида синтактик бирликларни теглаш дастурини яратишнинг лингвистик асослари бўйича тавсиялар ишлаб чиқиш ҳамда ўзбек тилида сўз бирикмаси ва гапларни автоматик аннотациялашнинг лингвистик таъминотини яратишдан иборат.

Тадқиқотнинг вазифалари.

парсинг дастурларининг ўхшаш/фарқли томонлари ва имкониятларини тадқиқ этиш;

синтактик тег категориялари, тег моделларини ишлаб чиқиш, дастурлаш учун тавсиялар тайёрлаш;

²⁰Karimov R., Mengliev B. Theoretical fundamentals of uzbek-english parallel corpus / Journal of critical reviews. ISSN- 2394-5125. – VOL 7, ISSUE 17, 2020. – P. 73-76.; Karimov R.A., MenglievB.R. The Role of the Parallel Corpus in Linguistics, the Importance and the Possibilities of InterpretationInternational Journal of engineering and Advanced Technology (IJEAT). ISSN: 2249 – 8958, Volume-8, Issue-5S3 July 2019. – P. 388-391.

²¹Орхун М. Computational analysis of uzbek nouns / Шестая Международная конференция по компьютерной обработке тюркских языков «TurkLang-2018». (Труды конференции) – Ташкент, 2018. – 320 с.

²²Турсунов А. Вопросы словообразования в формальных моделях тюркских языков (на примере узбекского языка) // Шестая Международная конференция по компьютерной обработке тюркских языков «TurkLang-2018». Труды конференции) – Ташкент, 2018. – 320 с.

Ўзбек тилида синтактик муносабат ва синтактик бирлик ҳақидаги тадқиқотларнинг синтактик теглашдаги аҳамиятини ўрганиш;

содда ва қўшма гап синтаксиси борасидаги тадқиқотларни умумлаштириш ҳамда уларни парсинг дастури лингвистик базаси сифатида яратиш.

Тадқиқотнинг объекти сифатида ўзбек тили синтактик бирликлари танланган.

Тадқиқотнинг предметини синтактик теглашнинг лингвистик тег ва моделлари ташкил этади.

Тадқиқотнинг усуллари. Тадқиқот мавзусини ёритишда таснифлаш, тавсифлаш, статистик ҳамда компонент таҳлил усулларидан фойдаланилди.

Тадқиқотнинг илмий янгилиги қуйидагилардан иборат:

жаҳон корпус лингвистикасидаги Penn Treebank, SynTagger, Link Grammar Parser, ХАНКО каби синтактик таҳлил дастурларининг тобе-ҳоқимлик муносабати, тизимли схема грамматикаси, анъанавий синтактик таълимотлардан фойдаланишдаги ўхшаш, шунингдек, лингвистик жиҳатдан изоҳлаш ёки изоҳлай олмаслиги каби фарқли жиҳатлари аниқланган;

тилшунослар А.Ғуломов, М.Асқарова, Ғ.Абдурахмоновларнинг сўз бирикмаси бўйича, Н.Маҳмудов, А.Нурмонов, А.Бердиалиевларнинг семантик синтаксис ва валентлик ҳақидаги, Ҳ.Неъматов, М.Қодировларнинг йиғиқ гап ва ихчам гап тўғрисидаги қарашлари ўзбек тили корпуслари учун парсер дастурлари ишлаб чиқишда назарий асос вазифасини ўтаганлиги далилланган;

ўзбек тилида сўзнинг мослашув, битишув, бошқарув каби бирикиш усуллари, гапнинг мазмунига, тузилишига кўра турлари бўйича синтактик тег категориялари, сўз бирикмаси, содда ва қўшма гапларнинг якка ва жуфт теглари, ўзбек тили синтактик бирликларини анализ ва синтез қилувчи лингвистик моделлар ишлаб чиқилган;

ўзбек тилида боғланган, боғловчисиз, эргашган, мураккаб қўшма гаплар ҳамда кўчирма гапли қўшма гап конструкциясининг лингвистик синтактик

теглар тизими ва сўзнинг бирикиш усули, содда, қўшма, мураккаб гап турларининг даража-уядошлик ва тур-жинс асосидаги қидирув параметрлари ишлаб чиқилган.

Тадқиқотнинг амалий натижалари қуйидагилардан иборат:

жаҳон тилшунослигида синтактик теглаш назарияси ва амалиёти, синтактик разметка, унинг имкониятлари, синтактик разметкаланган корпуслар ва уларнинг дастурий таъминоти асослаб берилган;

ўзбек тили корпусида сўз бирикмасини синтактик теглаш тамойиллари ишлаб чиқилган;

ўзбек тилида содда йиғиқ гапларни синтактик теглаш ва лингвистик модел ишлаб чиқишда лисоний-синтактик қолипларнинг ўрни очиб берилган ҳамда ЛСҚлар асосида лингвистик моделлар тузилган;

ўзбек тилида содда ёйиқ гапларни синтактик теглашда асос бўлувчи назарий маълумотларга биноан содда ёйиқ гапнинг тег ва моделлари ишлаб чиқилган;

ўзбек тилида боғланган, боғловчисиз ва эргашган қўшма гапларнинг синтактик моделлари ишлаб чиқилган;

ўзбек тилида мураккаб қўшма гапларни синтактик теглашнинг назарий асослари белгиланган.

Тадқиқот натижаларининг ишончлилиги ўрганилган материалларнинг ўзбек тили табиатидан келиб чиққан ҳолда хулосалар қилишга ёрдам берганлиги, уларнинг асосли эканлиги, методологик мукамаллиги, ўзбек тили корпуси парсинг (синтактик таҳлил) дастурини ишлаб чиқиш тамойилларини белгилашда амалда исботланган манбаларга таянилганлиги билан изоҳланади.

Тадқиқот натижаларининг илмий ва амалий аҳамияти. Тадқиқот натижаларининг илмий аҳамияти ўзбек тили корпуслари учун парсинг дастурлари яратишнинг назарий асосларини ишлаб чиқишда, компьютер ва корпус лингвистикаси йўналишида тадқиқотлар яратишда, илмий-назарий манба сифатида хизмат қилиши билан белгиланади.

Тадқиқот натижаларининг амалий аҳамияти амалий филология бўлимлари: компьютер лингвистикаси, корпус лингвистикаси, компьютер лексикографияси, табиий тилни қайта ишлаш (NLP) каби фанларни ўқитишда дастур, режалар тузиш, мавзуларни баён этишда манба вазифасини ўташи, ўзбек тили парсинг – синтактик таҳлил дастури учун лингвистик таъминот бўла олиши билан изоҳланади.

Тадқиқот натижаларининг жорий қилиниши. Миллий корпус учун парсинг дастури яратишнинг лингвистик асослари тадқиқи бўйича олинган илмий натижалар асосида:

жаҳон корпус лингвистикасидаги Penn Treebank, SynTagger, Link Grammar Parser, ХАНКО каби синтактик таҳлил дастурларининг тобе-ҳоқимлик муносабати, тизимли схема грамматикаси, анъанавий синтактик таълимотлардан фойдаланишдаги ўхшаш, шунингдек, лингвистик жиҳатдан изоҳлаш ёки изоҳлай олмаслиги каби аниқланган фарқли жиҳатлари, шунингдек, тилшунослар А.Фуломов, М.Асқарова, Ғ.Абдурахмоновларнинг сўз бирикмаси бўйича, Н.Маҳмудов, А.Нурмонов, А.Бердиалиевларнинг семантик синтаксис ва валентлик ҳақидаги, Ҳ.Неъматов, М.Қодировларнинг йиғиқ гап ва ихчам гап тўғрисидаги қарашлари ўзбек тили корпуслари учун парсер дастурлари ишлаб чиқишда назарий асос вазифасини ўташи ҳақидаги хулосаларидан PZ-20170927147 рақамли “Қадимий туркий ёзувлар ва XIII асргача бўлган фолклор тадқиқи” мавзусидаги фундаментал тадқиқот лойиҳасида фойдаланилган (Алишер Навоий номидаги Тошкент давлат ўзбек тили ва адабиёти университетининг 2021 йил 08 апрелдаги 04/1-1238-сон маълумотномаси). Натижада фундаментал тадқиқот лойиҳасининг қадимий туркий ёзувларни автоматик қайта ишлаш усулларига бағишланган бобининг бойитилишига хизмат қилган;

ўзбек тилида сўзнинг мослашув, битишув, бошқарув каби бирикиш усуллари, гапнинг мазмунига, тузилишига кўра турлари бўйича синтактик тег категориялари, сўз бирикмаси, содда ва кўшма гапларнинг якка ва жуфт теглари, ўзбек тили синтактик бирликларини анализ ва синтез қилувчи

лингвистик моделлар ҳақидаги назарий фикрлардан 63-11/41 рақамли “Фориш кадриятлари, анъаналари, урф-одатлари, удумларини сақлаш ва улар билан кенг жамоатчиликни таништириш” мавзусидаги амалий тадқиқот лойиҳасида фойдаланилган (Ўзбекистон Республикаси Тоғжик миллий-маданий маркази Жиззах вилояти бўлимининг 2021 йил 05 майдаги 39-сон маълумотномаси). Натижада амалий тадқиқот лойиҳасининг Фориш кадриятлари, анъаналари, урф-одатлари, удумларини автоматик қайта ишлаш ва теглаш усулларига бағишланган бобининг бойитилишига хизмат қилган;

Ўзбек тилида боғланган, боғловчисиз, эргашган, мураккаб қўшма гаплар ҳамда кўчирма гапли қўшма гап конструкциясининг лингвистик синтактик теглар тизими ва сўзнинг бирикиш усули, содда, қўшма, мураккаб гап турларининг даража-уядошлик ва тур-жинс асосидаги қидирув параметрларига доир хулосаларидан ФА-А1-Г007 “Қорақалпоқ нақлмақоллари лингвистик тадқиқот объекти сифатида” мавзусидаги амалий тадқиқот лойиҳасида фойдаланилган (ЎзРФА Қорақалпоғистон бўлимининг 2021 йил 17 январдаги 17.01/112-сон маълумотномаси). Натижада тадқиқотнинг содда ва қўшма гапларда сўз ясаиш типи бўйича фарқланувчи бирликларнинг қўлланилиши тадқиқи бўлимининг бойитилишига имкон берган.

Тадқиқот натижаларининг апробацияси. Мазкур тадқиқот натижалари 5 та халқаро, 5 та республика анжуманида муҳокамадан ўтказилган.

Тадқиқот натижаларнинг эълон қилинганлиги. Диссертация мавзуси бўйича 14 та илмий иш нашр эттирилган, жумладан, Ўзбекистон Республикаси Вазирлар Маҳкамаси ҳузуридаги Олий аттестация комиссиясининг докторлик диссертациялари асосий илмий натижаларини чоп этиш учун тавсия этилган илмий нашрларда 4 та илмий мақола, улардан 1 таси хорижий журналларда чоп қилинган.

Диссертациянинг тузилиши ва ҳажми. Диссертация кириш, уч боб, хулоса, фойдаланилган адабиётлар рўйхати ва иловадан иборат. Умумий ҳажми 128 саҳифани ташкил этади.

І БОБ. СИНТАКТИК БИРЛИКЛАРНИ ТЕГЛАШНИНГ НАЗАРИЙ АСОСЛАРИ

1.1. Жаҳон тилшунослигида синтактик теглаш назарияси ва амалиёти хусусида

Синтактик теглаш (синтактик разметка) – матннинг синтактик таҳлилига тегишли теглар мажмуи, морфологик таҳлил асосига қуриладиган парсинг натижаси. Разметканинг бу кўриниши лексик, бошқа турли синтактик қурилмалар (сода гап, қўшма гап, кўчирма гап), бирликлар орасидаги синтактик алоқани кўрсатади²³.

Илк корпусларни яратишда синтактик разметкаш ноавтоматик бўлса, кейинги авлод корпуслари синтактик разметкаси парсинг дастури асосида, автоматик/ярим автоматик тарзда амалга оширилган. Корпус тузиш тажрибасини кузатишларимиз шуни кўрсатяптики, синтактик разметканинг турли усуллари мавжуд бўлади. Улардан бири гапда сўз боғланишининг шажара усули бўлса, иккинчиси матн бирликларига синтактик тег бириктириш орқали амалга ошади.

Маълумки, оғзаки ва ёзма матнлар лингвистик тадқиқот ўтказишнинг асосий манбаси саналади. *Корпус* ва *корпус лингвистикаси* XX асрнинг 60-йилларида пайдо бўлди. *Корпус* атамаси кўп маъноли бўлиб, тилшуносликда турли объектларни атаб келади. Умуман олганда, корпус – маълум тилга оид материаллар йиғиндиси: у тўлиқ матн ёки унинг катта фрагментини қамраб олади. К.Бауш *корпус* атамасига қуйидагича таъриф беради: “Корпус фақат (ёзма ёки оғзаки) матнлар ёки уларнинг қисмларидан иборат чекланган миқдордаги тил материалидир”²⁴. Tuscan Word Centre ходими Джон Синклейрнинг таъкидлашича, корпус тил ёки лингвистик хилма-хилликни тадқиқ этиш, маълумот манбаи сифатида намоиш этиш учун кўринадиган

²³ Қаранг: Захаров В.П., Богданова С.Ю. Корпусная лингвистика. –Иркутск: ИГЛУ, 2011. -98 с.

²⁴ Махмутова М.К. Проблемы аннотирования (тагирования) текстов в корпусной лингвистике // Выпускная квалификационная работа ЮУрГУ. – Челябинск, 2018. – 94 с.

мезонларга мувофиқ танланган электрон кўринишдаги матнлар фрагментидан иборат²⁵.

Инглиз адабиётларида разметкаланган ва разметкаланмаган корпуслар фарқланади. Разметкаланмаган корпус – қайта ишланмаган, оддий, “хом” матн, лингвистик информацияга эга бўлмаган корпус. Масалан, present сўзи маълум контекстда от бўлиб келишини кўрсатувчи аниқ факт йўқ. Бу маълумот тушунилади, аммо юзага чиқарилмайди. Разметкаланган корпус ўзида лингвистик ахборот сақлайди: Масалан, present – noun/verb, яъни present сўзи ҳам от, ҳам феъл бўлиб кела олади. Ланкастер университети профессори Ричард Сяо бундай корпуснинг оддий матндан фарқлигини атрофлича шарҳлаб берган²⁶. Ўз навбатида, Ланкастер университети олимлари Тони Макенри, Эндрю Уилсон корпус тилшунослиги (корпус лингвистикаси) тўғрисида қуйидаги фикрларни айтишган: “Корпус лингвистикаси – бу тилни ўрганиш бўлиб, ёзма/оғзаки ўқиладиган матнни қайта ишлаш, фойдаланиш ва таҳлил қилиш билан боғлиқ барча жараёнларни ўз ичига олади. Корпус тилшунослиги нисбатан замонавий атама бўлиб, “жонли тил”дан фойдаланишга асосланган методологияга таянади. Ҳозирги вақтда корпус тилшунослигининг самарадорлиги, аҳамияти компьютер лингвистикасининг ривожланиши билан чамбарчас боғлиқ. Ҳар қандай матн таркибини яратиш тамойиллари унинг амалий мақсадига бевосита боғлиқ²⁷.

Мутахассислар изоҳланиш (тегланиш) даражасига кўра корпуснинг турли кўринишларини фарқлашади. Қуйида шундай корпус турларига қисқача тўхталамиз.

1. Ихтисослашган корпус – маълум бир турдаги матнлар тўплами: газета матни, илмий мақолалар.

²⁵Кўрсатилган манба.

²⁶XianyaoHu, RichardXiao. How do English translations differ from non-translated English writings? A multi-feature statistical model for linguistic variation analysis // https://www.researchgate.net/publication/294138492_How_do_English_translations_differ_from_non-translated_English_writings_A_multi-feature_statistical_model_for_linguistic_variation_analysis/references

²⁷ Махмутова М.К. Проблемы аннотирования (тагирования) текстов в корпусной лингвистике // Выпускная квалификационная работа ЮУрГУ. – Челябинск, 2018. – 94 с.

2. Умумий корпус турли хил матнларни ўз ичига олади, матн мазмуни ва жанрига алоҳида талаб қўйилмайди.
3. Қиёсий корпус – турли тилларнинг икки ёки ундан кўп кичик қисмлари, масалан, рус ҳамда немис тили ёхуд битта тилнинг вариантыдан иборат. Масалан, Австрия ва Швейцария немис тили версиялари.
4. Параллел корпус турли тиллардаги ўхшаш матнларни ўз ичига олган ички корпусдан иборат. Биринчи корпусдан аслиятдаги матн, иккинчисидан таржима матн ўрин олади.
5. Таълимий корпус – тил ўрганаётган шахс учун она тилида сўзлашувчилар томонидан ёзилган матнлар тўплами.
6. Дидактик корпус чет тилини ўқитиш жараёнида фойдаланиладиган тил маълумотларидан иборат.

А.Н.Баранов корпус технологияларига куйидагича таъриф беради: “Корпус – компьютер тилшунослигининг энг муҳим воситаларидан бири. Улар тилни таҳлил қилиш, матнни лингвистик корпус шаклида тақдим этишнинг амалий вазифаларини ҳал қилишга имкон беради”²⁸.

М.К.Махмутованинг фикрича, дастлаб, тилларни тадқиқ қилишда корпусдан фойдаланишнинг мақсади турли тил элементларининг частотасини ҳисоблаб чиқиш эди. Бундай элементлар сўз, сўзшакл, морфема ва иборалар бўлиши мумкин. Корпусдан тил ва нутқ бирликлари бўйича турли хил маълумот ва статистика олиш учун фойдаланиш мумкин²⁹. Мазкур технология, хусусан, лексикография, сўзни автоматик қайта ишлаш тизимлари соҳасида турли вазифаларни ҳал қилиш имконини беради. Нутқни аниқлаш, синтез қилиш, автоматлаштирилган машина таржимаси, имло ва грамматикани текшириш каби мураккаб лингвистик муаммоларни ҳал қилишда статистик усулдан ҳам фойдаланилади. Масалан, корпус материалида қайси сўзнинг турғун иборалар гуруҳига тегишли эканлигини

²⁸ Баранов А.Н. Введение в прикладную лингвистику [Текст]: учебное пособие / А.Н. Баранов. – М.: Изд-во Эдиториал УРСС, 2001. – 347 с.

²⁹ Чернякова, Т. А. Методика формирования навыков студентов на основе лингвистического корпуса [Текст] / Т. А. Чернякова. – Тамбов, 2012. – 149 с.

аниқлаш мумкин. Бунинг учун олинган маълумотда бирликлар ўзаро мунтазам равишда бирика олишини текшириш керак бўлади. Голландиянинг Нижмеген университетида ишлаб чиқиладиган грамматикалар корпус матни ҳолатларида синондан ўтказилади. Грамматика асосида корпусни қайта ишлайдиган таҳлил дастури тузилади. Олинган ишлов бериш натижалари грамматика маълумотларини қанчалик аниқ тасвирлашини кўрсатади. Шундан келиб чиққан ҳолда, корпус технологияси тилшуносликнинг янги назариялари, сўзни автоматик қайта ишлаш тизимларини синаб кўришга имкон беришини кўраимиз.

1993 йилда Ланкастер-Осло/Берген (ЛОБ) корпуси ва Британия миллий корпуси (BNC) яратувчиси Жеффри Леич томонидан 1993 йилда тузилган аннотациялаш постулатлардан бири тил белгиларининг аниқ ва тушунарли тавсифлаш принципи эътиборга молик. Шунингдек, унинг фикрига кўра, умумфойдаланишга мўлжалланган корпуснинг разметкаси учта принципга мувофиқ келиши керак.

1. Разметка (корпус аннотацияси) фойдаланувчи учун қўлланма ёки кўрсатма шаклида мавжуд бўлган таҳлил схемасига асосланган бўлиши, ҳар бир параметр ундан жой олиши керак.

2. Фойдаланувчи учун очиқ корпус разметкаси “назарий жиҳатдан нейтрал” бўлиши лозим: разметка параметрлари барча учун тушунарли бўлган тушунчалар тизимидан ташкил топган бўлиши талаб этилади. Агар корпус аниқ бир лойиҳа учун мўлжалланган бўлса, уни разметкалашда махсус, айнан муаллифга хос, умумқабул қилинган таснифдан фойдаланиш лозим: бунда ҳам тузувчидан у ёки бу тил назариясига таяниш талаб қилинади.

3. Корпус аннотацияси схемаси ким томонидан, қайси аудиторияга мўлжалланганлиги аниқ, равшан кўрсатилиши лозим, чунки корпусдан фойдаланишда турли юридик ва техник чегаралар мавжуд³⁰.

³⁰ Leech, G. Corpus annotation schemes / G. Leech Literary and Linguistic Computing, 1993. – 8/4. – P. 275-281.

Ж.Леичнинг биринчи постулати мукамал ишланган, аммо у ҳамма корпусда ҳам юзага чиқавермайди. Албатта, барча корпуслар у ёки бу даражада маълумот(тег)лар тизими билан таъминланади. Ҳар бир параметрнинг қандай маълумот ташиётганлигини аниқлашга доим ҳам эришиб бўлмайди. Бу борада “матнни автоматик қайта ишлаш” (<http://www.aot.ru>) гуруҳида фаолият олиб борган олимларнинг фаолиятини алоҳида таъкидлаш лозим. Улар асосан, рус тилидаги матнларга автоматик ишлов бериш жараёнларини ишлаб чиқишган; компьютер технологиялари ютуқлари назарий тилшунослик билан бирлаштирилиб, амалий натижаларга эришилган. Демак, ҳар қандай синтактик теглар тизимини ишлаб чиқиш учун компьютер технологиялари ютуқлари билан биргаликда ўзбек тилшунослигида синтаксис бўйича яратилган назариялар асосида корпус учун парсер дастури ишлаб чиқиш мумкин. Бу дастурдан синтаксис бўйича яратилган барча назарий материалларни тўплаш, ўрганиш, умумлаштириш ҳамда синтактик теглар яратишда фойдаланиш зарур. Шунингдек, ушбу гуруҳ рус, инглиз ва немис тилларида матн таҳлил қилишнинг қуйидаги модулларини ишлаб чиққан:

- 1) графематик (сўз чегарасини аниқлаш);
- 2) морфологик (сўз туркумини аниқлаш);
- 3) синтактик (гап бўлагини аниқлаш);
- 4) семантик (сўзлар орасидаги семантик алоқани аниқлаш).

Корпусни разметкалаш аннотациялаш мақсадидан келиб чиққан ҳолда турли даражада амалга оширилиши мумкин.

Морфологик разметка морфологик шакл ва унинг грамматик маъноси ҳақидаги ахборотдан (туркум, даража, майл, шахс, сон, келишик, замон, майл ва ҳ.) иборат бўлади. Бундай разметка фойдаланувчига керакли сўз, сўзшакл ҳамда конструкцияни топишга ёрдам беради.

Синтактик разметка. Маълумки, корпусни синтактик разметка маълумотлари билан таъминлаш жараёни автоматлаштирилса ҳам, морфологик разметкадан фарқли ўлароқ, инсон омилига кўпроқ эҳтиёж

сезилади. Шу сабабли катта ҳажмли матнни деталлаштирилган синтактик таҳлил қилиш талаб этилмайди. Синтактик разметкада қўлланиладиган асосий алгоритм синтактик *декомпозиция* саналади. *Кириш*да матн берилса, *натиж*да матннинг нечта гапдан ташкил топганлиги миқдори ҳамда матндаги гаплар рўйхати кўрсатилади. Бундай алгоритмни ишлаб чиқиш 1960-йилларда бошланган. Синтактик декомпозиция алгоритми белги асосида бўшлиқ, тиниш белгилари асосида жумлани танийди. Аммо бу ерда ҳам муаммога дуч келамиз: оддий матнни форматлаб бўлмаслиги мумкин, чунки ажратувчи сифатида ишлатиладиган белги нафақат жумланинг охирида, балки ўртада ҳам ишлатилиши мумкин. Сарлавҳа, бўлимни ажратиб кўрсатувчи бирлик, расм, жадвал номи, колонтитул – ушбу бирликларнинг барчаси гап бўлмаслиги мумкин, аммо гап кўринишида ифодаланган бўлади. Бундай ҳолатларда таҳлил тизими гапларни ажратиш/аниқлашда натижани хато кўрсатиши мумкин. Синтактик декомпозиция бир қатор фразавий жумлаларни аниқлаш алгоритмини бажариш учун асос бўлиб хизмат қилади. N-граммани ажратиш алгоритми жуда кенг тарқалган. Улар икки, уч ёки ундан ортиқ токен(биграмма, триграмма, тетраграмма ва б.)лардан таркиб топган иборалардан иборат³¹. Гапни жумлага бўлиш корпусда мавжуд токеннинг позициясини ҳисобга олган ҳолда амалга оширилади. N-граммани таниб олиш турли қоидаларга асосланади; N-грамм тақсимоти таҳлили статистик жиҳатдан муҳим ибораларни аниқлашга ёрдам беради ҳамда, кўпинча, нутқ парчаси билан стохастик аннотация алгоритмида қўлланилади. Жумланинг боши ва охири шартли теглар билан кўрсатилган бўлади, бу эса битта белгидан иборат бўлган жумлани триграмма сифатида кўриб чиқишга имкон беради. Бундан ташқари, у тег танлаш учун зарур бўлган тахминий параметрларни аниқлашга ёрдам беради. Масалан, Инглиз тилининг америка варианти корпуси(COCA)да *I like to* сўз бирикмаси 4810 марта; *I like women*

³¹ Betsch, M. Automatic Annotation of Russian texts: Evaluation of Different Tagging Methods / R.Meyer, P.Kosta, J.Blaszczak, J.Frasek, L.Geist, M.Zygis // Investigations into Formal Slavic Linguistics. Contributions of the Fourth European Conference on Formal Description of Slavic Languages (FDSLIV) held at Potsdam University, 2003. – 234.

бирикмаси эса 29 марта учрайди. Бундан кўринадики, тилда биринчи бирикиш усули эҳтимоли кўпроқ.

Морфологик жиҳатдан муҳим иборани таҳлил қилиш учун *чанкер*лар ишлатилади. Сўровга берилган жавобда улар маълум бир турдаги (номинал, феъл, сифат, равиш) иборалар рўйхатини чиқаради. Кўпинча, рўйхатга олинган қисм(от бирикмаси)лар чанкер сифатида фойдаланилади. Улар жумлани бошқарувчи от орқали танийди. Ушбу турдаги жумлаларни “таниб олиш” нутқ қисмлари теглари билан олдиндан изоҳлаш, нутқнинг айрим қисмларини грамматик қоидалар асосида ибораларга бирлаштириш асосида амалга оширилади.

Тубинген корпусидаги рус тилида синтактик белгилаш элементлари сўзнинг эквиваленти сифатида ишлайдиган барқарор синтактик бурилишлар изоҳини ўз ичига олади (масалан, *чунки, барибир, шунга кўра*). Бу корпусда кўшма гап ҳам тегланади, яъни ушбу корпус сўз бирикма, содда гап, кўшма гап каби бирликларини автоматик равишда фарқлай олади.

Хелсинки лойиҳаси муаллифлари грамматик маълумотни максимал қамраб олиш учун матнни чеклаб, корпусга синтактик тег (белги)ни батафсил кўшишни режалаштирганликлари мантиқан тўғри. Ҳозирда ушбу иш натижаси Хельсинки корпусининг онлайн версиясида мавжуд эмас, аммо яқин орада жойлаштирилиши кутилмоқда. Муаллифларнинг фикрига кўра, синтактик белги икки изоҳли схемани мураккаб шаклда бирлаштиради (аниқроғи, параллел равишда таклиф қилади): гап бўлаклари нуқтаи назаридан бизга кенг доирадаги фойдаланувчилар ишлатиши мумкин бўлган синтактик изоҳларни тасвирлаш имконини беради; шажара дарахтлари нуқтаи назаридан бизга, биринчи навбатда, синтактик бўғин турлари ҳамда иерархиясини аниқ кўрсатишга имкон беради³². Бундан ташқари, манбанинг мавжуд версиясида ноаниқ турғун бурилишлар қайд этилган: рўйхатлар

³² Мустайоки, А. Принципы создания Хельсинкского аннотированного корпуса русских текстов ХАНКО в сети Интернет / Научно-техническая информация. – 2003. – Сер. 2, № 6. – С. 33-37.

асосида Т.Ф.Ефремова³³, Р.П.Рогожникова³⁴ ишларида 2000 га яқин бундай бўлинмалар ажралиб туради.

Рус тили миллий корпуси таркибидаги матнларнинг барчасини синтактик разметкаш кўзда тутилмаган; корпусга бундай талаб ҳам қўйилмаган. Бутун корпусда юзага чиқариш талаб қилинадиган синтактик таҳлил элементларига турғун ибораларни теглаш (1), корпус воситасида уларнинг статистикасини чиқариш (2) тегишли. Шу билан бирга, рус тили миллий корпуси сайтида корпуснинг (ҳажми 30000 гапдан кўпроқ) синтактик разметкаланган фрагментини қўйиш режалаштирилган. Ушбу таҳлил шажара дарахти воситасида амалга оширилиши кутилмоқда.

Аннотациялаш (теглаш) усуллари. Аннотациялаш(теглаш)нинг лингвистик ҳамда статистик усуллари мавжуд бўлиб, аннотациялаш (теглаш) усуллари маълум концепцияга асосланиб тузилади. Лингвистик аннотациялаш усули сифат қоидаларига, статистик усул эса микдорга асосланади. Статистик усул 1950-йилларда оммалашиб кетди. Афсуски, ушбу усулни ишлаб чиқиш ҳамда қўллаш жуда тез урфдан чиқди. Бу иккита омилга боғлиқ.

Биринчидан, маълумотнинг мавжудлиги муаммоси. Ўша даврда тил маълумотларига статистик усулни қўллашдаги муаммолардан бири маълумот тўпламининг ниҳоятда кичик эканлиги эди; кўп сонли тил ҳодисалари учун қизиқарли статистик умумлашма қилишнинг иложи йўқ эди.

Иккинчидан, ижтимоий фанларда умумий силжишлар бўлди: кейинчалик, тилшунос Ноам Хомский расмий методологияга ҳам, статистика усулини мутлақо рад этган тилшунослик назариясига ҳам алоҳида эътибор қаратди. Одатда, корпус фойдаланувчиси маълум матннинг мазмунига қизиқмайди, балки корпуснинг метаматн маълумотлари, тилнинг баъзи элементлари ва тузилмаларидан фойдаланиш мисоллари билан қизиқишади.

³³ Ефремова Т.Ф. Толковый словарь служебных частей речи русского языка / Т. Ф. Ефремова. – М.: Изд-во Астрель-АСТ, 2004. – 814 с.

³⁴ Рогожникова, Р.П. Толковый словарь сочетаний, эквивалентных слову / Р. П. Рогожникова. – М.: Изд-во Астрель-АСТ, 2003. – 416 с.

Бу фойдаланувчилар – биринчи навбатда, тилшунослар. Корпус ёрдамида ўтказилган дастлабки лингвистик тадқиқотлар турли тил элементларининг пайдо бўлиш частотасини ҳисоблашда камайтирилди.

Автоматик таржима, нутқни аниқлаш, синтез қилиш, имло ва грамматикани текшириш воситалари каби мураккаб тил муаммоларини ҳал қилиш учун статистик усулдан фойдаланилади. Шундай қилиб, барқарор иборалар семантик нуқтаи назардан ажралмас бирлик бўлиб, лексикографияда, матнни автоматик қайта ишлаш тизимларида ҳисобга олиш жуда муҳим. Статистик усул воситасида корпус материалида мунтазам равишда қандай сўзлар биргаликда ишлатилишини аниқлаш ҳамда шу тариқа барқарор ибораларга боғланиш мумкин. Шунингдек, корпус лексикография ва грамматик тадқиқотлар учун бой маълумот манбаи ҳам бўлади. Лексикография соҳасида амалга ошириладиган тадқиқотлар семантика соҳасидаги тадқиқотлар билан чамбарчас боғлиқ. Корпусдаги маълум бир лисоний бирлик куршовини кузатиш орқали ушбу бирликни тавсифловчи маълум семантик хусусиятларни аниқлаш мумкин.

Лингвистик аннотациянинг қўлланиш соҳалари. Таггинг таснифлаш, қариндошликни аниқлаш, чегарани белгилаш, шахс/сонни аниқлашга ёрдам берадиган восита сифатида маълум. Тегнинг қўлланилиш соҳаси кенг: тег сўз, расм ёки бирор-бир идентификатив белгини ифодалаши мумкин. Унинг аналог сифатида музей экспонатларининг илмий маълумотномасини кўрсатиш мумкин. Объект, маълумотни сўз ёрдамида аниқлаш ва таснифлаш компьютер пайдо бўлишидан анча олдин вужудга келган. Инсониятнинг ижтимоий тармоқдан кенг фойдаланиши (ўзаро маълумот ва тасвир алмашиш) теглаш жараёнининг оммалашишига олиб келди. Сайтлар фойдаланувчига тег яратиш, уни бошқариш имконини берди; ушбу теглар контентдан фойдаланишда оддий калит сўзни ишлатиш имконини берди. Бундан ташқари, маълумот олиш учун категориялар рўйхати (ёки тег, ёрлик, калит сўз деб аталадиган теглар) визуал тарзда тақдим этилади.

Одатда, веб-сайтларда калит сўз (тег)дан тасвирлаш ёки оддий матнни ифодалаш учун фойдаланилади. Теглар, кўпинча, алоҳида сўз бўлиб, ҳар бир тегнинг аҳамияти шрифт ҳажми ёки ранги билан белгиланади. Бундай тегишлилик энг машҳур, кўп қўлланилаётган атама/сўзни тез аниқлаш, бири-бирига нисбатан кўп қўлланилган шароитда тақсимлаш учун қулай.

Теглаш, одатда, парсинг учун ишлатилади. Масалан, Stanford Parser³⁵. Қидирув қисмига сўров берилади, натижада қуйидаги рўйхат чиқади: сўров, синовдан ўтган таклиф, таҳлил қилиш, универсал тобеланиш, сўровни кўриб чиқиш вақтини ўз ичига олган универсал боғлиқлик ва статистика.

<http://parts-of-speech.info/> сайти нутқ шакл(қисм)ларини теглаш ҳам намунали мисол бўлиши мумкин. Ушбу дастур Stanford таггери асосида ишлайди; у нутқнинг қисмига қараб, берилган жумладаги сўзларни турли хил рангларда белгилайди.

Genia Tagger Demo –бевосита таркибий қисмларни автоматик таҳлил қилиш дастури³⁶. Жумла киритилгандан сўнг ибора турлари, тегланган жумла, объектларни таниб олиш номи, таниб олиш тури умумий жадвалда акс эттирилади.

Аннотациялашнинг яна бир кўриниши – бу ресурсга бўлган изоҳни шакллантириш. Теглаш калит сўзлар тўплами, тушунчалар тўплами ва учлик тўплами ёрдамида бажарилади. Теглардан фойдаланиш изоҳни хужжат матнига кўшиш орқали амалга оширилади. Бу ишлов бериш дастури уларнинг ёрдами билан танланган қисмнинг анъанавий маъносини аниқлаш учун амалга оширилади.

1.2. Синтактик разметка ва унинг турли корпуслардаги имкониятлари

Тил корпусида синтактик аннотация (разметка) турлари. Тил корпусида синтактик разметка масалалари жаҳон, хусусан, рус тилшунослигида атрофлича тадқиқ этилган. Хусусан, М.В.Копотев,

³⁵<http://nlp.stanford.edu:8080/parser/index.jsp>.

³⁶Дастурни ушбу ҳавола орқали кўриш мумкин: <http://nactem7.mib.man.ac.uk/geniatagger/>.

Г.Б.Гурин³⁷, Ж.Лич³⁸ асарларидаги назарий маълумотлар ўзбек тили корпусида синтактик таҳлил дастурини ишлаб чиқиш учун назарий асос бўлиши билан аҳамиятли. Ж.Лич томонидан яратилган синтактик аннотациялаш принципи ҳақидаги назарий материалларни кузатишларимиз шуни кўрсатдики, синтактик аннотацияси мавжуд корпуснинг фойдаланувчилар аудиторияси кенг бўлади; бундай корпуснинг ахборот тизимлари билан алоқа ўрнатиш имконияти кенгроқ бўлади. Шу билан бирга, маркировканинг изчиллиги учун барча жавобгарликни муаллифлар зиммасига юкламайдиган, лекин мавжуд таснифларга асосланиб, корпусни яратишга ёндашув тил тавсифларидаги бўшлиқни аниқлашга, тилга бўлган турли ёндашувдаги нуқсон, қарама-қаршиликларни аниқлашга имкон беради. Ж.Личнинг сўнгги постулатида айнан шу масалага эътибор қаратилади: “There can be no claim that the annotation scheme represents ‘God’s truth’. Rather, the annotated corpus is made available to a research community on a caveat emptor principle. It is offered as a matter of convenience only, on the assumption that many users will find it useful to use a corpus with annotations already built in, rather than to devise and apply their own annotation schemes from scratch (a task which could take them years to accomplish)”³⁹.

М.В.Копотев, Г.Б.Гуриннинг фикрига кўра, бугунги кунда рус тили корпусини тўлиқ разметкалашда синтактик таҳлил дастури учун ишлатилиши мумкин бўлган учта назария мавжуд:

1. Тобе-ҳокимлик муносабати грамматикаси (И.А.Мелчук, И.М.Богуславский, Л.Л.Иомдин ва б.).
2. Тизимли схема грамматикаси (Н.Ю.Шведова, В.А.Белошапкина ва б.);
3. Анъанавий синтактик таълимотлар (А.А.Шахматов, В.В.Виноградов, Н.С.Валгина ва б.).

³⁷Копотев М.В., Гурин Г.Б. Принципы синтаксической разметки хельсинкского аннотированного корпуса русских текстов ХАНКО //Труды международной конференции «Диалог 2006» // Б.280-284

³⁸ Leech G. Corpus annotation schemes // Literary and Linguistic Computing, 1993, № 8/4. Pp. 275-81.

³⁹ Leech G. Corpus annotation schemes // Literary and Linguistic Computing, 1993, № 8/4. Pp. 275-81.

Кўриняптики, синтактик таҳлил дастури тузиш принципларини ишлаб чиқиш дастур учун назарий асос вазифасини ўтайдиган назариялар фарқланади; асосланиши кутилаётган назария танлаб олинади. Демак, ўзбек тили корпуси учун синтактик разметка яратишда ҳам ўзбек тили синтаксисига бўлган ёндашувларни фарқлаш талаб этилади. Шу сабабли ишнинг кейинги қисмларида ўзбек тили синтактик таҳлил (разметка) дастури учун назарий-методологик асос бўла оладиган таълимотларга алоҳида тўхталамиз.

Кўрсатилган ишда барча – уч назария, бошқа ёндашувлардан келиб чиққан ҳолда лингвистик материалнинг тўлиқ ва мустақил тавсифини бера олиши таъкидланади. Бундан ташқари, турли хил назарияларда тавсиф даражаси ҳар хил бўлади. Тобе-ҳокимлик грамматикаси синтактик муносабат турларига кўпроқ эътибор қаратади (масалан, ЭТАП-3 тизимида, у асосида яратилган корпусда номланган муносабатлар сони 80 га яқин), бошқа томондан, гап бўлаклари ҳақидаги анъанавий таълимот тадқиқотчига синтактик алоқаларни батафсил таснифлашни таклиф қилади⁴⁰.

Бундан хулоса қилиш мумкинки, ноаниқ синтактик назария билан муваффақиятли синтактик таҳлил дастури ҳамда разметка тизимини яратиб бўлмайди.

Шуни таъкидлаш керакки, рус олимлари барча ёндашувларни бир хил даражада қабул қилишмайди. “Назарий жиҳатдан энг нейтрал” синтаксис сифатида, шубҳасиз, гап бўлаклари таснифига асосланган анъанавий грамматика тан олинishi керак. Рус тили тиниш белгилари қоидалари мактаб ўқитувчиси учун ҳам, профессионал тилшунос учун ҳам бир хил терминологик тизимга эга. Бошқа томондан, узоқ вақт давомида Россиянинг кўплаб университетида асосий ҳисобланиб келган В.А.Белошапкина таҳриридаги дарсликнинг тарқалиши кўп ўқитувчиларнинг, асосан, таркибий схемалар синтаксисига таянишига олиб келди. Шу билан бирга, тобе-

⁴⁰Копотев М.В., Гурин Г.Б. Принципы синтаксической разметки хельсинкского аннотированного корпуса русских текстов ХАНКО //Труды международной конференции «Диалог 2006» // Б.280-284

хокимлик грамматикаси тилшуносларга нисбатан кам маълум, чунки у, одатда, амалий муаммоларни ҳал қилишда фаол ишлатилади. Шундай қилиб, “назарий нейтраллик постулати”ни қондирадиган, етарлича мукаммал синтактик назарияни танлаш энг масъулиятли вазифа саналади⁴¹.

Барча оммалашган ёндашувлар муҳокама қилингандан сўнг, **ХАНКО** яратувчилари синтактик аннотациялаш учун иккита алтернатив синтактик белгилаш схемадан фойдаланишга қарор қилишган: тобе-хокимлик муносабати ҳамда гап бўлаклари анъанавий синтаксиси. Ушбу ёндашувнинг аниқ эклектикаси билан, иккита схемани бирлаштириш қуйидаги муаммоларни ҳал қилади:

- 1) синтактик тузилишнинг икки томони, боғланишни батафсил тавсифлаш;
- 2) рус тили ўқитувчиси ҳамда профессионал тилшунос эҳтиёжларини қондириш;
- 3) фойдаланувчининг муқобил тартиб натижаларини мустақил равишда ёки биргаликда тақдим этиш талабини бажариш.

Маълумки, турли лингвистик топшириқлар матн структураси ҳақида ҳар хил ҳажм, характердаги ахборотни талаб қилади. Синтактик разметка борасидаги кузатишларимиз шуни кўрсатдики, корпус муфассал разметкаси билан фарқ қилувчи уч қисмдан иборат⁴².

1. Леммаланган матн – унда матндаги ҳар бир лексик сўзшаклга унинг луғатдаги шакли (леммаси) ҳамда сўз туркумига мансублиги тўғрисидаги маълумот кўрсатилади.
2. Морфологик разметкаланган матн – матндаги ҳар бир сўзшакл лемма, сўз туркуми ва сўз ўзгартирувчи қўшимчалар ҳақидаги морфологик ахборот (изоҳ) билан таъминланади.
3. Синтактик разметкаланган матн – унда морфологик разметка билан бирга ҳар бир гапга унинг синтактик структураси тўғрисидаги изоҳ ёзилади.

⁴¹Кўрсатилган манба.

⁴²Богуславский И.М., Григорьев Н.В., Григорьева С.А., Иомдин Л.Л., Крейдлин Л.Г., Фрид Н.Е. Разработка синтаксически размеченного корпуса русского языка // http://cl.iitp.ru/bibitems/corpus_SPB.pdf

Қуйида охири турдаги разметка, яъни корпуснинг энг ривожланган қисми борасида сўз юритилади. Корпус ҳажмини кўрсатувчи рақамлар айнан шу қисмга тегишли. Рус тилидаги матнлар шажара занжири ёрдамида бирикма доирасида аннотацияланади. Ҳар бир гап шажараси синтактик структурасига қўйиб кўрилади: шажара аъзолари гапнинг бўлакларига тўғри келади; шажара шохалари эса турли хилдаги синтактик муносабатлар номи билан белгиланади (бугунги кунда 78 номдаги турли хил муносабатлар аниқланган бўлиб, уларнинг қарийб ярми И.А.Мельчукнинг анъанавий “Мазмун – Матн” назариясида таклиф қилинган синтактик муносабат саналади). Ушбу корпусда гапнинг синтактик структурасини кўрсатиш усули таркибидаги бирликларнинг синтактик алоқадорлигини аниқловчи бошқа корпуслардан тубдан фарқ қилади. Ишлаб чиқиладиган ушбу корпуснинг энг яқин аналоги Праганинг Карлово университетиде ишлаб чиқилган чех тилининг аннотацияланган корпуси Prague Dependency Treebank (PDT)⁴³ саналади. Ушбу корпусда ҳоким-тобелик грамматикаси ҳам ишлайди, синтактик алоқа бу корпусга нисбатан умумийлаштириб олинган. Масалан, чех тили корпусида синтактик муносабатнинг жами шакли моделлаштирилган. Рус тили корпуси синтактик конструкциянинг жуда кўп турларини, энг кам учрайдиган боғланиш ҳолатларини ҳам акс эттира олади. Чех корпусида эса ниҳоятда қизиқарли, тадқиқот мақсадида ишлатиладиган дискурс структура ҳақидаги маълумот ҳам киритилган⁴⁴. Грамматик алоқа аннотацияланган (тегланган) корпуслар сирасига немис тили корпуси NEGRA⁴⁵ ҳамда Alpino голланд тили корпуси⁴⁶ни қўшиш ўринли. Ушбу корпуслар синтактик разметка тажрибасини ўрганиш учун намуна бўлиши билан аҳамиятли.

⁴³ Hajicova E., Panevova J., Sgall P. Language Resources Need Annotations To Make Them Really Reusable: The Prague Dependency Treebank. // Proceedings of the First International Conference on Language Resources & Evaluation, pp. 713-718.

⁴⁴ Bemova, A., J. Hajic, B. Hladka, J. Panevova. Morphological and Syntactic Tagging of the Prague Dependency Treebank. URL: <http://citeseer.nj.nec.com/296119.html>

⁴⁵ Brants, Th., W. Skut, and H. Uszkoreit H. Syntactic annotation of a German newspaper corpus. // Proceedings of the ATALA Treebank Workshop, pp. 69-76, Paris, France.

⁴⁶ Van der Beek, L., G. Bouma, R. Malouf, G. van Noord. The Alpino Dependency Treebank. // Proceedings of LINC 2001.

И.М.Богуславский матн разметкаси махсус тег – маркерлар билан амалга оширилишини айтар экан, тегларни якка тег (1) ҳамда контейнер тег (2)га ажратади. Якка тег матн бирлиги (сўз) ҳақида ахборот беради, контейнер тег эса разметка тизимида сақланадиган матн структураси тўғрисидаги ахборотни ташиydi. Буларни кодлаштириш тўғрисида қуйида алоҳида тўхталамиз.

1. Матнни гапга ажратиш жуфт контейнер теглар воситасида амалга оширилади: <C> : </C>. Очилувчи тег яна бир параметрга эга бўлиши мумкин, бу гап идентификатори <C ИД=идентификатор>. Ушбу тег матн таркибидаги гаплар орасидаги муносабатни ифодаловчи изоҳ ҳисобланади.

2. Матнни лексик элементларга ажратиш жуфт контейнер теглар билан амалга оширилади: <W> : </W>. Сўз ҳам ўз идентификаторига эга бўлиши мумкин <W ИД=идентификатор>.

3. Сўзнинг морфологик характеристикаси якка тег билан ёзилади: <НОМ>; улар контейнер теглар ичида жойлашади. <НОМ> тегининг 4 та майдони мавжуд: ИД – идентификатор, ЛЕММА – сўзнинг луғатдаги шакли (лексема), POS – сўз туркуми, FEAT – морфологик характеристикалар.

4. Гапнинг синтактик структураси тўғрисидаги ахборот <НОМ> теги ичида жойлашувчи алоҳида белги – DOM билан ифодаланади: <НОМ ДОМ=идентификатор / алоқа типи>. Идентификатор синтактик тобе сўзга ишора қилса, алоқа типи ҳоким ва тобе сўз ўртасидаги синтактик муносабат типини акс эттиради.

Формализм етарли мослашувчанликка эга: у нафақат тайёр тузилмани, балки матннинг оралиқ ҳолатини қайд этиш имконини ҳам беради. Хусусан, битта контейнер <W> : </W> теги орасига бир неча <НОМ> тегларини киритиш орқали сўзшакл морфологик анализининг бир неча варианты ҳақидаги ахборотни битта разметка таркибида сақлашга эришиш мумкин. <НОМ> теги таркибида бир қанча DOM тегларини киритиш билан шажара тузилишини сақлаш мумкин.

Аннотацияланган корпусга маълумотни киритиш амалиёти ҳамда разметкаланиш воситалари яримавтоматик равишда амалга оширилади. Дастлаб, тизим автоматик морфологик, синтактик таҳлилдан ўтказилади; кейинчалик тилшунос томонидан таҳрирланади. Структурани аниқлаш ЭТАП-3 машина таржимаси тизими⁴⁷ билан амалга оширилади. Ушбу тизим махсус дастурий мажмуа асосида тузилган бўлиб, икки: разметкаланмаган матнни гапларга ажратиш (сегментация) – (“Chopper”) ҳамда синтактик структурани тузиш ва таҳрирлаш – “Structure Editor” дастурини камраб олади (Иловадаги 1-расмга қаранг).

Жараёнда тилшуноснинг иштирок этиш даражаси ишлов бериладиган матннинг мураккаблигига боғлиқ. Аксарият матнлар тизимдан тўғри структурани олади, бундай ҳолда тилшуносга умумий текшириш ҳамда структурани тасдиқлаш вазифаси қолади, холос. Агар анализатор томонидан тузилган структура нотўғри бўлса, тилшунос уни график муҳаррир ёрдамида таҳрирлаши мумкин (Иловадаги 2-расмга қаранг).

Фақат айрим ҳолларда, агар ибора жуда мураккаб, катта ҳажмли бўлса, “ЭТАП-3” тизими унинг структурасини туза олмайди. Бундай пайтда тилшуносга фавқулодда режим – “split and run” ёрдамга келади: бунда мутахассис томонидан ҳар бир гап икки ёки ундан ортиқ фрагментга ажратилади; ажратилган фрагмент эса кейинчалик таҳлилга берилади. Синтактик анализатор автоматик тарзда ҳар бир фрагментдан шажара дарахтини тузади, тилшуноздан фақатгина шу шажара (фрагмент)ни олдинги асосий шажарага қўшиш талаб этилади. Агар тилшунос батафсил кўриб чиқиш талаб қилинадиган ностандарт/зиддиятли синтактик тузилишга дуч келса, у бутун ибора ёки шажара дарахтининг тузилмадаги ўрни аниқ бўлмаган тугунни алоҳида белгилайди. Шундан сўнг тизимнинг ўзи ушбу

⁴⁷ Апресян, Ю.Д., Богуславский И.М., Иомдин Л.Л., Лазурский А.В., Перцов Н.В., Санников В.З., Цинман Л.Л. Лингвистическое обеспечение системы ЭТАП-2. М.: Наука, 1989.; Апресян, Ю.Д., Богуславский И.М., Иомдин Л.Л., Лазурский А.В., Митюшин Л.Г., Санников В.З., Цинман Л.Л. Лингвистический процессор для сложных информационных систем. М.: Наука, 1989.

тугунни ўз ичига олган жумланинг кўшимча тахрирга муҳтожлигини билдирадиган белги билан ажратади.

1.3.Синтактик разметкаланган корпуслар ва уларнинг дастурий таъминоти борасида айрим мулоҳазалар

1.3.1. *Разметкаланган матнда лингвистик ахборот типи.Морфологик ахборот.* Морфологик таҳлил ҳар бир сўзнинг туркуми, шу билан бирга гапдаги вазифасини аниқлайди; қуйидаги морфологик категориялар граммемасини кўрсатади (мас., рус тилида жонли/жонсиз предмет, род, келишик, шахс, сон, нисбат, даража, замон, майл, репрезентация ва ҳ.)

Синтактик ахборот. Гапнинг синтактик белгиси, юқорида айтиб ўтилганидек, шажара дарахти билан ифодаланади. Бу ерда ҳар бир ўқ “ҳоким”дан “тобе”га йўналтирилади ҳамда синтактик муносабатлардан бирининг номи билан белгиланади⁴⁸. Анъанавий ҳолда таркибда синтактик гуруҳ ҳамда таркибий қисмлар мавжуд эмас; аслида, шажара дарахтининг ҳар қандай “тупи” гуруҳ деб ҳисобланиши мумкин, унинг тепаси “бута”нинг ташқи алоқаларида унинг вакили сифатида ишлайди. Одатда, дарахт тузилишидаги тугунлар сони жумладаги сўзлар сонига тенг. Бир томондан, истисно ҳолатини келтириб чиқарувчи ҳодисалар ҳам мавжуд: айрим сўз шакллари занжири лексик бирликни билдирганда, юқоридаги қоидага амал қилинмайди, истисно ҳолати юзага келади. Бундай пайтда, сўз сонига қараганда структурада аъзо сони камаяди. Бошқа томондан, бирор бир сўз структурага “ёпиштирилиши лозим” бўлса ёки реал матнда мавжуд бўлмаган сўз структурага киритилиши керак бўлса ҳам, матндаги сўз шакллари занжир билан бир хил бўлмаслиги мумкин. (Отам врач бўлиб ишларди, онам эса ўқитувчи //Отам врач бўлиб ишларди, онам эса ўқитувчи БЎЛИБ ИШЛАРДИ). Таъкидлаш керакки, теглаш жараёнида ҳар қандай қайта ишланган матннинг лексик, синтактик омонимияси тўлиқ ечимини топади. Агар буни автоматик амалга ошириш имкони бўлмаса (тизим томонидан

⁴⁸Богуславский И.М., Григорьев Н.В., Григорьева С.А., Иомдин Л.Л., Крейдлин Л.Г., Фрид Н.Е. Разработка синтаксически размеченного корпуса русского языка // http://cl.iitp.ru/bibitems/corpus_SPB.pdf

тузилган, изоҳловчи ёки муҳаррирнинг фикрига кўра, жумлада учрайдиган тузилма мос келмаса), албатта қўл билан тузатилади.

Тегловчи мутахассис ҳам омонимияни ҳал қила олмайдиган алоҳида ҳолатларда (масалан, *қасддан қичқириқ бўлса, руҳ каби ажойиб импульс мавжуд*), жумла бир нечта тузатиш билан таъминланиши мумкин. Разметкаланган матнлар корпуси бирор-бир тилнинг алоҳида луғати билан узвий боғлиқ эмаслиги сабабли муаллиф сўзнинг аниқ лексик маъносини маълум тарзда, ноаниқлик ва лексик омонимия билан белгилаш ғоясидан воз кечади. Масалан, *механик*¹ (шахс оти) ва *механик*² (физик ҳолат) сўзи, *қалб*¹ (юрак) ҳамда *қалб*² (ногўғри, эгри) ҳеч қандай индекс билан таъминланмаган. Истиснолар баъзи “ишчи” сўзлар(хусусан, олд қўшимчалар)га оид бўлиб, уларнинг лексик маъноси корпус ҳужжатларида тасвирланган.

Демак, матнга автоматик ишлов бериш аннотацияланган корпусдан фойдаланганда аниқроқ, хатосиз амалга оширилиши ойдинлашади. Корпуснинг синтактик тег(изоҳ)ланган қисмини синаб кўришни бошлаган биринчи турдаги дастур – ЭТАП-3 тизимида рус тилидан инглиз тилига таржима қилинганда синтактик ноаниқликни автоматик ҳал қилиш аниқ натижа берган. Бундан келиб чиқадики, матнни автоматик қайта ишлаш учун лингвистик таъминот ва дастурий тизим ишлаб чиқилиши талаб этилади.

Синтактик таҳлил алгоритми ишлаб чиқилганда қўшимча филтр яратиш ҳам талаб этилган: 2-4 аъзодан ташкил топган ушбу восита таҳлил қилинаётган гапни потенциал тармоқлар воситасида таҳлилдан ўтказди. Бундай тажриба натижасини корпуснинг кейинги қисмини қуришда ҳам қўллаш мумкин, чунки янги, автоматик равишда қурилган гапларни таҳлил қилиш янада осонлашади.

О.И.Бабина, Н.Ю.Дюминларнинг таъкидлашича, ҳар бир корпус разметкаси асосида тил назарияси ётади, корпус асосидаги ҳар қандай хулоса шу концепцияга асосланган ҳолда чиқарилади⁴⁹. Ҳар бир тадқиқотчи тилни

⁴⁹Бабина О.И., Дюмин Н.Ю. Автоматизация лингвистической разметки корпуса текстов // <http://helling100.narod.ru/pubs/AutomationBabinaDyumin.pdf>

моделлаштиришда объектив/субъектив сабабларга кўра маълум нуктаи назарни қўллаб-қувватлаши мумкин: масалан, тил структурасидан келиб чиқиб, тилни формаллаштириш назарияси маълум бир тилга нисбатан қўлланади, бошқа тилга тўғри келмаслиги мумкин; тил модели, кўпинча, ҳатто бир тил доирасида ҳам, ундаги хилма-хиллик ва кўпмаънолиликни акс эттиролмайди. Тадқиқот мақсадидан келиб чиққан ҳолда, унча катта бўлмаган тадқиқий корпусларни лингвистик аннотациялаш (разметкалаш) чуқур синтактик ва семантик разметкани қамраб олиши, шунингдек, фақатгина морфологик изоҳ (разметка компоненти) билан чекланиб қолиши ҳам мумкин. Катта корпуслар разметкасида ифодаланган муфассал лингвистик ахборот ниҳоятда катта меҳнатни талаб қилади. Тадқиқот мақсади меҳнат сарфини камайтиришга қаратилганда, фақат зарурий изоҳлар мажмуини ўрганиш мақсадга мувофиқ. Демак, бу нуктаи назардан, корпус разметкасини чуқурлаштириш ўзини оқламайди, разметкани соддалаштириш йўлидан бориш ҳамда қидирув натижасини аниқлаштиришга эътибор қаратиш керак бўлади. Минималлаштириш концепциясига асосланадиган бўлсак, разметка тизимида фақат зарурий ахборотни киритиш мақсадга мувофиқ. Бошқа томондан, корпусда қўлланивчи воситанинг бошқа тадқиқий корпусда қўллаш мумкин бўлган автоматик разметка методологияси сифатида фойдаланиш самарали натижа беради. Шу концепциядан келиб чиққан ҳолда, О.И.Бабина, Н.Ю.Дюминлар лингвистик разметка воситаларини тузиш принциплари сифатида қуйидагиларни санаб ўтишади⁵⁰:

1. Дастурий таъминот воситаси турли хил тизимлар ўртасидаги мувофиқлик муаммосининг олдини олишга ёрдам берадиган Unicode белгилар кодлаш тизимини қўллаб-қувватлаши керак, шу билан кирилл ёки латин алифбоси бўлмаган алифболар диакритикаси ёки алифбодан фойдаланадиган тилларни тасвирлаш имкониятини бериши керак.

2. Матнлар тўплами ҳамда тегишли лингвистик маълумот ягона маълумотлар базасида сақланиши керак, бунда матн корпуси билан ишлаш

⁵⁰Ўша манба.

учун турли хил функцияларни амалга оширадиган тизимнинг дастурий компонентидан стандартлаштирилган кириш таъминланади.

3. Корпусга ишлов беришга мўлжалланган маълум воситалар ундан алоҳида бўлиши талаб қилинади. Шунинг баробарида, тизимнинг умумий универсаллигини таъминлашишга эришиш лозим. Маълумотнинг лингвистик базадан бошқа тизимларда ҳам такрорий қўлланишига эришиш мумкин.

4. Тизимнинг ҳар бир компоненти алоҳида лингвистик вазифа бажаради, тизимнинг модул ташкилотини таъминлайди.

5. Тил материали(матн)нинг репрезентацияси асосий омил; барча ҳосилавий лингвистик маълумотлар (хусусан, лексикон) матн корпусидаги позицияларга боғланган. Матн репрезентативлиги принципи корпусдаги турли сўзшакл, сўз бирикмаси лексик ва грамматик контекстига эркин киришга шароит яратади.

6. Лексик бирликни матнга бириктириш эса омоним сўзшакл ва сўз бирикмаларига ёзиладиган турли морфологик теглар мажмуини ажратиб олишга йўл очади.

Матннинг автоматик разметка мажмуи дастурий воситалари кўйидагилардан ташкил топади:

- 1) корпусни бошқариш модули (CorpusManager);
- 2) автоматик морфологик разметка модули (AutoPOSTagger);
- 3) морфологик разметканинг автоматик корректори (Corrector);
- 4) автоматик синтактик разметка модули (SynTagger).

Айтиш жоизки, ушбу тизим матнга ҳар томонлама ишлов беришга мўлжалланган. Шу билан бирга, турли вазифаларни бажарувчи воситалар алоҳида ишлаб чиқилади.

SynTagger синтактик таҳлил дастури хусусида. О.И.Бабина, Н.Ю.Дюминлар томонидан таклиф этилган *автоматик синтактик разметка модули* (SynTagger) матннинг синтактик жиҳатдан бир-бирига бўйсунувчи, тобе-ҳоким бўлак бўлиб келган лексик бирликни қавслар билан

бириктирилган кўшилма сифатида ўз ичига олади (Қаранг: иловадаги 3-расм).

Фойдалунувчи синтактик блокнинг боши ва охирини белгилаши, унинг типи (отли бирикма, феълли бирикма, сонни ифодаловчи бирикма ва х.)ни аниқлаши тавсия этилади. SynTagger модули морфологик разметка мавжуд бўлган тақдирда автоматик равишда турли хилдаги синтактик структура гуруҳларини ажратишга имкон беради. Турли функционал услуб ёки лаҳжанинг ўзига хослигини кўрсатувчи синтактик тадқиқотларда фойдаланиш мумкин бўлади. Тавсифланаётган тизим махсус қурилишга эга: ундан бир вақтнинг ўзида бир неча тадқиқотчи фойдаланиши мумкин (онлайн/офлайн форматда бирдек ишлайди). Лингвистик базанинг маълумотлар омборига киритилиши ёки ундан фойдаланиш учун очиқлик тамойили сервер сўровларини қайта ишлаш, сўровга жавоб бериш учун лингвистик таъминот вазифасини бажаради.

Олдин айтиб ўтилганидек, жаҳон корпус лингвистикасида разметкаланган ва разметкаламаган корпуслар фарқланади. Разметкаланган (тегланган) корпус лингвистик маълумотга эгаллиги билан ажралиб туради. Бундай корпуслар сирасига Рус тили миллий корпуси⁵¹ таркибидаги синтактик корпус⁵² ҳамда Британия миллий корпуси (BNC) таркибидаги Treebank⁵³ни киритиш мумкин. Рус тили миллий корпуси таркибидаги синтактик корпус РФА тилшунослик институти лойиҳаси асосида яратилган бўлиб, ундан www.ruscorpora.ru сайти орқали фойдаланиш мумкин. 2500000 гапдан ташкил топган синтактик разметкаланган корпус дастурчи Сергей Протасов томонидан ишлаб чиқилган Link Grammar Parser⁵⁴ воситасида таҳлил қилинади. Шунингдек, EVALITA⁵⁵, AMALGAM⁵⁶ каби дастурий

⁵¹www.ruscorpora.ru

⁵²<http://otipl.philol.msu.ru/~soiza/testsynt/files/info-parse.htm>

⁵³<http://www.evalita.it/>

⁵⁴<http://www.link.cs.cmu.edu/link/index.html>

⁵⁵<http://www.scs.leeds.ac.uk/amalgam/amalgam/amalghome.htm>

⁵⁶<http://www ldc.upenn.edu/ldc/online/treebank/>

воситалар ҳам синтактик таҳлил дастури – парсинг сифатида яхши натижалар кўрсатгани билан эътиборга молик.

Жаҳон корпуслари, хусусан, инглиз тили корпуслари орасида ҳам синтактик разметкаланган корпуслар мавжуд; улар ҳам ўзига хос парсинг дастурларига эга. Улар орасида Penn Treebank⁵⁷ таркибидаги воситалари бошқа парсерлар учун намуна бўла оладиган, синтактик таҳлил натижалари аниқ чиқадиган энг мукамал парсер саналади. Сайтларда синтактик аннотацияланган турли тилдаги куйидаги, хусусан, инглиз тилининг тарихий корпуси ҳам мавжуд:

-
1. Penn Parsed Corpus of Middle English (PPCME).
 2. Penn Chinese Treebank.
 3. Penn Korean Treebank.
 4. Prague Dependency Treebank.
 5. Arabic Syntactic/Predicate-Argument annotation.

Кузатишларимиз шуни кўрсатдики, ушбу синтактик таҳлил дастурлари – парсерлар бошқа турли лойиҳалар учун “олтин стандарт” намунаси сифатида хизмат қила олади, чунки ушбу корпусларда синтактик таҳлил методларига тўғри ёндашилган.

Prague Dependency Treebank⁵⁸ – словян тиллар шажара дарахти шакли жиҳатдан синтактик разметкаланган корпуслар орасида энг намунали корпус. Праганинг Калов университети (Институт формальной и прикладной лингвистики, физико-математический факультет, Карлов университет, Прага, Чехия (ÚFAL MFF UK) лойиҳаси асосида амалга оширилган. Лойиҳа ҳақида бир неча илмий ишлар⁵⁹ нашр қилинган бўлиб, унда синтактик таҳлил дастури – парсер тузиш принциплари таҳлил қилинган. Турли тиллар(дат, француз, португал, эстон, эсперанто)ни синтактик таҳлил қилиш бўйича лойиҳа ва дастурларнинг маълумотлар омбори ҳақидаги ахборот интернет

⁵⁷ https://www.ling.upenn.edu/courses/Fall_2003/ling001/penn_treebank_pos.html

⁵⁸ <http://ufal.mff.cuni.cz/pdt2.0/>

⁵⁹ Недолужко А., Гаич Я. и др. *Синтаксически аннотированный корпус чешского языка*. The Prague Dependency Treebank. // <http://www.dialog-21.ru/dialog2008/materials/html/62.htm>

сайти⁶⁰ да очик бўлиб, улардан айрим тиллар базасини яратишда фойдаланиш мумкин. Treebank – парсерлар рўйхати, номенклатураси Википедия электрон энциклопедияси⁶¹ да ҳам мавжуд.

Рус тилининг on-line синтактик таҳлил дастурлари имкониятлари сирасига Aot.ru ишчи гуруҳи томонидан синтактик таҳлил дастурининг ишлаб чиқилганлиги⁶²; табиий тилдаги матнлар таҳлили технологияларини яратиш билан шуғулланадиган “Диктум” компанияси томонидан рус тили шажара дарахтини тузиш принципларининг ишлаб чиқилган⁶³; рус тил учун Link Grammar Parser дастурининг мавжудлигини алоҳида таъкидлаш жоиз.

Инглиз тилини синтактик таҳлил қилувчи бир қанча тизимлар мавжуд бўлиб, қуйида уларнинг айримларига ҳаволаларни келтириб ўтамыз:

1. Link Grammar Parser: <http://www.link.cs.cmu.edu/link/dict/introduction.html> - ушбу сайтда дастурий тизимнинг тўлиқ изоҳи келтирилган.
2. <http://www.link.cs.cmu.edu/link/submit-sentence-4.html>.
3. Шажара дарахти парсери: <http://www.connexor.com/demo/>
4. VISL лойиҳаси.
5. <http://faculty.washington.edu/dillon/GramResources/GramResources.html> сайти морфологик ва синтактик разметка муаммолари билан қизиқувчи мутахассислар учун қўлланма вазифасини бажара олади.

ХАНКОда анъанавий синтаксис. Маълумки, анъанавий ёндошув асослари XIX асрда рус тилшунослари тадқиқотларида мукамал ишланган. Рус тили синтаксисининг энг тўлиқ тавсифини 1960 йилларнинг академик грамматикаси деб ҳисоблаш мумкин. Замонавий таснифлар ҳозирги рус тили университет дарсликларида бироз ўзгариш, фарқ билан акс эттирилган⁶⁴. Ушбу ёндашувнинг афзалликлари қуйидагилар:

- 1) умумийлик ва соддалик;

⁶⁰<http://corp.hum.sdu.dk/>

⁶¹<http://en.wikipedia.org/wiki/Treebank>

⁶²<http://aot.ru/demo/synt.html>

⁶³<http://www.dictum.ru/ru/syntax/blog>

⁶⁴ Валгина Н.С., Современный русский язык. Синтаксис. М.: Высшая школа, 2003.; Кустова Г.И., Мишина К.И., Федосеев В.А. Синтаксис современного русского языка. М., 2005.

2) бошқа синтактик ёндашувлар асосида (биринчи навбатда, структур синтаксис), тадқиқот учун материални билвосита излаш имкониятининг мавжудлиги;

Анъанавий синтаксиснинг камчиликлари қуйидагиларда кўринади:

- 1) синтактик тузилмалар табиати тўғрисидаги замонавий ғоялар билан номувофиклик;
- 2) синтактик бирликлар тавсифи ҳамда синтактик алоқани эътиборсиз қолдириш;
- 3) тавсифдаги номувофиклик, муқаррар қарама-қаршиликлар (предпозицион гуруҳнинг йўқлиги, турли бўлак ичидаги бўлакни аниқ ажратиб олмаслик);
- 4) автоматик ишлов беришнинг мураккаблиги.

Бироқ ёндашувнинг кўрсатилган афзаллик/камчиликлари иш натижасини оқламайди; аксинча, улар потенциал фойдаланувчининг тахминларини янада чигаллаштиради.

ХАНКО (Хельсинский аннотированный корпус қисқартмаси) яратувчилари оддий фойдаланувчига тушунарли бўлган разметка даражасини сақлаб қолиш учун умумқабул қилинган назариялардан фойдаланишни маъқул кўришган. Қуйида ХАНКОни ишлаб чиқилган тамойилларни келтиришни маъқул кўрдик, зеро, ўзбек тили корпуси синтактик теглар тизимини ишлаб чиқишда ҳамда шунга ўхшаш ҳолатларда ушбу тамойилларга таяниш мумкин.

1. Муайян муаммони ҳал қилишда ХАНКО яратувчилари доим бу ёки бошқа синтактик маълумот ва унинг ахборот қайта ишлашни автоматлаштиришда қанчалик муҳимлигига эътибор беришган. Қўлда бажариладиган ишларнинг кутилган ҳажми ва натижа қиймати кўпинча бири-бирига зид келади: масалан, детерминантни бирлик сифатида ажратиб кўрсатиш қўлда қилинадиган ишларнинг сезиларли даражада кўпайишига олиб келади (детерминантни қидиришни автоматлаштириш мумкин эмас), ammo бу ишни кетма-кет бажариш қийин бўлади, чунки концепциянинг кўлами “детерминант”ни турли матнда турлича аниқлайди.

2. Интерфейснинг қулайлигини ҳисобга олиш ҳам диққат-эътибордаги масала саналади. Синтактик маълумотлар турли хил бирликка, жумладан, морфологик маълумотни ўз ичига олган матн шаклига, икки марта белги қўйилган ҳолларда киритилади.

Миллий корпусларда (мас., Рус тили миллий корпуси (НКРЯ) ҳар бир сўзшаклга морфологик изоҳ бириктириш билан бирга сўз(гап бўлаги)га унинг қайси гап бўлаклиги эканлигини билдирувчи изоҳ (тег) ҳам ёзилган; ҳар бир гап учун унинг синтактик структураси кўрсатиб берилган. Ушбу корпуснинг разметкаси яримавтомат режимда ишлайди. Дастлаб, ҳар бир матн ЭТАП-3 дастури асосида морфологик ва синтактик таҳлилдан ўтади, олинган натижа мутахассис томонидан таҳрирланади. Ушбу корпус разметкаси шундай аниқ ишланганки, корпус материали морфологик, синтактик омонимиядан холи. Морфосинтактик разметка билан бирга, лексик тавсиф ҳам мавжуд бўлиб, ушбу функция идиоматик birlikларни изоҳлаш (теглаш)да қўл келади, чунки турғун бирикмалар ҳам семантик муносабатлар асосида таҳлилдан ўтказилади. Рус тили миллий корпусининг СинТагРус қисмида 100дан ортиқ лексик функциялар аниқланган; натижада маълумотлар базасида 21 000 сўз бирикмаси занжири аниқланган. СинТагРус, асосан, икки турдаги матнлар йиғиндисидан иборат:

- 1) 1980 йилдан бошлаб ҳозирги кунга қадар нашр этилган журнал ва интернетдаги илмий-оммабоп, сиёсий-ижтимоий ҳамда ахборот мақолалари;
- 2) XXасрнинг иккинчи ярми рус бадиий насри намуналари.

Ушбу синтактик таҳлил тизимлари ўзбек тили синтактик таҳлил тизимини яратиш учун зарурий тажриба майдони бўлиб хизмат қилади. Юқорида санаб ўтилган парсер(синтактик таҳлил тизимлари)ни ўрганар эканмиз, синтактик таҳлил тизими қандай таркибий қисмлардан ташкил топиши, синтактик таҳлил тегларини ишлаб чиқиш учун қандай лингвистик билимлар керак бўлишини кузатдик. Хулоса сифатида айтиш мумкинки, ҳар бир тилдаги синтактик разметка тизимини ишлаб чиқиш учун ўша тилнинг синтактик қурилиши моделлаштирилиши талаб этилади. Моделлаштиришдан

кейинги босқич синтактик теглар тизимини тузиш, сўнгги қадам эса матн тил бирликларига синтактик тегларни бириктиришдир. Ишнинг кейинги бобларида ушбу масалалар батафсил ёритилади.

Боб бўйича қисқа хулосалар

1. Синтактик теглаш матннинг синтактик таҳлилига тегишли теглар мажмуи, морфологик таҳлил асосига қуриладиган парсинг натижаси бўлиб, у синтактик қурилмалар орасидаги синтактик алоқани кўрсатади. Илк корпусларни яратишда синтактик разметкаланиш қўлда амалга оширилган бўлса, кейинги авлод корпуслари синтактик разметкаси парсинг дастури асосида, автоматик/ярим автоматик тарзда амалга оширилган.

2. Синтактик разметканинг турли усуллари мавжуд: биринчиси гапда сўз боғланишининг шажара усули бўлса, иккинчиси матн бирликларига синтактик тег бириктириш саналади.

3. Синтактик разметкада қўлланиладиган асосий алгоритм – синтактик декомпозиция. Синтактик декомпозиция алгоритми белги асосида, яъни бўшлиқ (пробел), тиниш белгилари асосида жумлани фарқлайди. Аммо бу усул мукамал эмас, чунки ажратувчи сифатида ишлатиладиган белги нафақат жумланинг охирида, балки ўртасида ҳам ишлатилиши мумкин. Сарлавҳа, бўлимни ажратиб кўрсатувчи бирлик, расм, жадвал номи, колонтитул каби бирликлар гап бўлмаса-да, гап кўринишида ифодаланади. Бундай ҳолатларда таҳлил тизими гапларни аниқлашда натижани хато кўрсатиши мумкин.

4. Автоматик таржима, нутқни аниқлаш, синтез қилиш, имло ва грамматикани текшириш воситалари каби мураккаб тил муаммоларини ҳал қилиш учун статистик усулдан фойдаланилади.

5. Тагнинг таснифлаш, қариндошликни аниқлаш, чегарани белгилаш, шахс/сонни аниқлашга ёрдам берадиган восита сифатида маълум. Тегнинг қўлланилиш соҳаси кенг: тег сўз, расм ёки бирор-бир идентификатив белгини ифодаланиши мумкин.

6. Корпуснинг муфассал разметкаси уч қисмдан иборат. Леммаланган матнда ҳар бир сўзшаклга сўз туркумига мансублик тўғрисида маълумот кўрсатилса, морфологик разметкаланган матнда сўзшакл лемма, сўз туркуми ва сўз ўзгартирувчи қўшимчалар ҳақидаги морфологик ахборот (изоҳ) билан таъминланади. Синтактик разметкаланган матнда морфологик разметка билан бирга ҳар бир гапга унинг синтактик структураси тўғрисидаги изоҳ ёзилади.

7. Синтактик разметкада теглар якка ҳамда контейнер тегга ажратилади. Якка тег матн бирлиги (сўз) ҳақида ахборот беради, контейнер тег эса разметка тизимида сақланадиган матн структураси тўғрисидаги ахборотни ташийди. Матнни гапга ажратиш жуфт контейнер теглар воситасида амалга оширилади. Идентификатор синтактик тобе сўзга ишора қилса, алоқа типи ҳоким ва тобе сўз ўртасидаги синтактик муносабат типини акс эттиради.

8. Ҳар бир тилдаги синтактик разметка тизимини ишлаб чиқиш учун ўша тилнинг синтактик қурилиши моделлаштирилиши талаб этилади. Моделлаштиришдан кейинги босқич синтактик теглар тизимини тузиш, сўнгги қадам эса матн тил бирликларига синтактик тегларни бириктиришдир.

II БОБ. СЎЗ БИРИКМАСИ ВА СОДДА ГАПЛАРНИ СИНТАКТИК ТЕГЛАШНИНГ НАЗАРИЙ АСОСЛАРИ

2.1. Ўзбек тили корпусида сўз бирикмаларни синтактик теглаш хусусида

2.1.1. *Синтактик теглаш принциплари ҳақида.* Синтактик теглаш – тил корпусларида синтактик бирликлар қидирувини юзага чиқарувчи омил. Тил корпусида қидирувнинг содда кўриниши сифатида лемма, сўзшакл, ибора ва коллокация қидирувини санаш мумкин. Ушбу белгилар қарийб барча аннотацияланган (разметкаланган) тил корпусларида мавжуд. Шунингдек, бирликларига лингвистик тег бириктирилган корпусда морфологик белги асосидаги қидирув ҳам етакчилик қилади. Синтактик ва семантик белги асосидаги қидирувни шакллантириш эса анча мураккаб жараён бўлганлиги сабабли ҳамма корпусларда ҳам қидирувнинг ушбу кўриниши бўлавермайди.

1993 йилда Ланкастер-Осло/Берген (ЛОБ) корпуси ва Британия миллий корпуси (BNC) яратувчиси Жеффри Леич томонидан 1993 йилда тузилган аннотациялаш постулатлардан бири бўлган “тил белгиларининг аниқ ва тушунарли тавсифлаш принципи” эътиборга молик. Унинг фикрига кўра, умумфойдаланишга мўлжалланган корпус теглари уч принципга мувофиқ келиши керак⁶⁵.

1. Тег фойдаланувчи учун қўлланма/кўрсатма шаклида мавжуд бўлган таҳлил схемасига асосланган бўлиши ҳамда ҳар бир параметр ундан жой олиши керак.

2. Фойдаланувчи учун очиқ корпус теглари “назарий жиҳатдан нейтрал” бўлиши лозим: разметка параметрлари барча учун тушунарли бўлган тушунчалар тизимидан ташкил топган бўлиши талаб этилади. Агар корпус аниқ бир лойиҳа учун мўлжалланган бўлса, уни аннотациялаш(теглаш)да махсус, айнан муаллифга хос, умумқабул қилинган таснифдан фойдаланиш лозим: бунда ҳам тузувчидан у ёки бу тил назариясига таяниш талаб қилинади.

⁶⁵ Leech G. Corpus annotation schemes / G.Leech. Literary and Linguistic Computing, 1993. – 8/4. – P. 275-281.

3. Корпус аннотацияси схемаси ким томонидан, қайси аудиторияга мўлжалланганлиги аниқ, равшан кўрсатилиши лозим, чунки корпусдан фойдаланишда юридик ва техник жиҳатдан турли чегаралар мавжуд.

Демак, ҳар қандай аннотация турида бўлгани каби синтактик теглар тизимини ишлаб чиқиш учун ҳам компьютер технологиялари ютуқлари билан биргаликда ўзбек тилшунослигида синтаксис бўйича яратилган назария(қараш)лар асосида корпус учун парсер дастурини ишлаб чиқиш мумкин. Синтаксис бўйича яратилган барча назарий материалларни ўрганиш, умумлаштириш, қиёслаш ҳамда синтактик теглар яратишда ёндашувни тўғри танлаш муҳим. О.И.Бабина, Н.Ю.Дюминларнинг таъкидлашича, ҳар бир корпус разметкаси асосида тил назарияси ётади, корпус асосидаги ҳар қандай хулоса шу концепцияга асосланган ҳолда хулосаланади⁶⁶. Ҳар бир тадқиқотчи тилни моделлаштиришда объектив/субъектив сабабларга кўра маълум нуқтаи назарни қўллаб-қувватлаши мумкин: тил структурасидан келиб чиқиб, тилни формаллаштириш назарияси маълум бир тилга қўлланса, бошқа тилга тўғри келмаслиги мумкин. Олдинги қисмларда айтилганидек, минималлаштириш концепциясига асосланадиган бўлсак, теглар тизимига фақат зарурий ахборотни киритиш мақсадга мувофиқ бўлади.

Кўриняптики, синтактик таҳлил дастурини тузиш принциплари ёки синтактик теглар тизимини ишлаб чиқиш учун назарий асос вазифасини ўтайдиган назариялар фарқланади; асосланиши кутилаётган назария танлаб олинади. Шу сабабли, қуйида ўзбек тили синтактик таҳлил (разметка) дастури ёки синтактик теглар тизими учун назарий-методологик асос бўла оладиган таълимотларга алоҳида тўхталамиз.

2.1.2. *Ўзбек тилида сўз бирикмаси ва унинг тадқиқи.* Ўзбек формал грамматик талқинлари асосларини А.Ғуломов, М.Асқароваларнинг (ўзбек формал синтактик таълимотининг мукамал шарҳи бўлган, яқингача олий таълим тизими учун энг мукамал дарслик саналиб, академик

⁶⁶Бабина О.И., Дюмин Н.Ю. Автоматизация лингвистической разметки корпуса текстов // <http://helling100.narod.ru/pubs/AutomationBabinaDyumin.pdf>

Ғ.Абдурахмонов етакчилигида тузилган қўлланма билан бир қаторда ҳозиргача ҳам амалиётдан тушмаган “Ҳозирги замон ўзбек тили. Синтаксис (–Т.: Ўрта ва олий мактаб, 1960)) ишлари⁶⁷ орқали кенг оммалашди. Ўтган асрнинг 70-йилларига келиб, формал таҳлил усулига таянган ўзбек синтактик назарияси ва ўзбек тили синтактик қурилишининг талқини тугал шаклланди, 1976-йилда нашр этилган илмий грамматика⁶⁸ да умумлаштирилди, формал таҳлил ютуқлари асосида ўзбек тили синтактик қурилишининг бошқа тадқиқ методлари, жумладан, систем-структур таҳлил усуллари асосида чуқурлаштириш, формал тавсиф асосида қатор устқурма таҳлилларда тилимиз синтактик қурилишидаги турли-туман парадигма ва парадигматик муносабатларни очиш ишларига замин ҳозирланди.

Формал синтактик тавсиф натижаларига таяниб, ўзбек тили синтаксиси бўйича амалга оширилган ҳамда катта илмий самара берган йўналишлар сифатида систем-структурал таҳлил усулларига таяниб олиб борилган семантик синтаксис ва валентлик бўйича тадқиқотларни алоҳида санаб ўтиш лозим. Бу соҳада Н.Маҳмудов, А.Нурмонов, А.Бердиалиев, Д.Лутфуллаева, А.Аҳмедов, И.Қўчқортоев, Р.Расулов, М.Қурбонова, А.А.Цалкаламанидзе, С.Султонсаидова, С.Муҳамедова, Ҳ.Усмонова каби олимлар эришган ютуқларнинг илмий салоҳияти ўзбек тилшунослигидагина эмас, славистика ва германистика – умуман тил қурилиши таҳлил назариясида ўзига хос аҳамиятга моликлиги, мустақиллиги билан ажралиб туради⁶⁹.

Ф.Ф.Фортунатов ва унинг издошлари тобе алоқага киришган сўзлар қўшилмасини, жумладан, эга билан кесимнинг ўзаро алоқаси асосида ҳосил бўлган конструкцияни ҳам, тенг алоқага киришган сўзлар қўшилмасини ҳам сўз бирикмаси санашади. А.А.Шахматов сўз бирикмасини фақат бир сўзнинг бошқасига тобеланиб бирикуви асосида ҳосил бўлган грамматик бутунлик

⁶⁷ Ғуломов А. Синтаксис ва пунктуация бўйича машқлар тўплами. – Т.: Ўздавнашр, 1938; 1939; 1947); Ўзбек тили грамматикаси. 2-қисм. Синтаксис. –Т.:Ўздавнашр, 1940; Ғуломов А. Ўзбек тилида аниқловчилар. –Т.: Ўздавнашр, 1941;Ўзбек тили грамматикаси. 2-қисм. Синтаксис. –Т.: Ўздавнашр, 1944-1960; Ғуломов А. Содда гап. Ҳозирги замон ўзбек тили курси бўйича материаллар. –Т.: ЎзФАнашр., 1955.

⁶⁸Ўзбек тили грамматикаси. 2-том. Синтаксис. – Т.: Фан, 1976.-180 б.

⁶⁹Ўзбек тили грамматикаси. 2-том. Синтаксис. – Т.: Фан, 1976.-180 б.

сифатида таърифлайди⁷⁰. Шунинг учун олим икки ёки ундан ортиқ сўздан ташкил топган гапларни ҳам тугалланган фикр бирлигига мувофиқ келадиган “тугалланган сўз бирикмаси” деб билади. Бироқ А.А.Шахматовнинг фикрича, гап фақат сўз бирикмасининг оддий бир тури, хусусий кўриниши эмас; гап алоҳида синтактик birlikдир: у сўз бирикмаси тарзидагина эмас, балки бир сўз формаси тарзида ҳам намоён бўлиши мумкин⁷¹.

Сўз бирикмасини принципиал бошқача тушунишни В.В.Виноградов асарларида учратамиз. Унинг бу соҳадаги концепциясига А.А.Шахматовнинг сўз бирикмаси тўғрисидаги таълимотининг ўзига хос назарий йўналиши бевосита асос бўлди. Бунда В.В.Виноградов тилда иккита синтактик birlikнинг – сўз бирикмаси ва гапнинг мавжудлиги ҳақидаги А.А.Шахматов томонидан олдинга сурилган тезисдан келиб чиқди. В.В.Виноградов тенг алоқа асосида бириккан сўз қаторини – уюшиқ бўлакларни, ажратилган бўлакларда учрайдиган изоҳланувчи билан изоҳловчининг бирикувини сўз бирикмаси ҳисобламайди. В.В.Виноградовнинг нуқтаи назарича, сўз бирикмаси тўғрисидаги таълимот фақат баъзи бир бирикувларни тасвирлаши лозим. Сўз бирикмаси синтаксисининг вазифаси – турли сўзларнинг у ёки бу шаклдаги тобе сўзлар билан кенгайишини – ёйилишини ва кенгайиш-ёйилиш натижасида ҳосил бўлган мураккаб номлар қандай маънолар ифодалашини аниқлаш⁷².

Кўриниб турибдики, синтактик birlikларга қарашлар турлича; корпус birlikларини теглашда birlikлар тизимини аниқлаш ҳам асосий иш саналади. Манбаларда сўз бирикмасининг ташқи томондан гапга мос келиши ҳолатлари ҳақида айтилади. Мисол учун *ойдин кеча* сўз бирикмаси билан кўйидаги контекстдаги шу лексик-грамматик таркибдан ташкил топган гапни қиёслайлик: *1944 йилнинг баҳори. Ойдин кеча. Осмон тўла юлдуз. Ойдинда ва лампа ёруғида гуллаган ўриклар кўринади.* Мисолдаги “сифат+от” қўшилмаси

⁷⁰Шахматов А.А. Синтаксис русского языка. Л., 1941, С. 274.

⁷¹Нурмонов А., Маҳмудов Н., Аҳмедов А., Солихўжаев С. Ўзбек тилининг мазмуний синтаксиси. Тошкент: Фан, 1992. – 9-10-б.

⁷²Нурмонов А., Маҳмудов Н., Аҳмедов А., Солихўжаев С. Ўзбек тилининг мазмуний синтаксиси. Тошкент: Фан, 1992. – 11-б.

предметнинг ҳозирги замонда реал мавжудлиги маъносини билдиради, яъни у предикативликни ифодалайди. Шунинг учун бу қўшилма коммуникатив бирлик функциясини бажаради⁷³. Маълум бўладики, *ойдин кеча* каби бирикув корпусда гап ёки бирикма эканлиги фарқланиши керак. Бундай бирикувларни фарқлаш учун алоҳида алгоритм, филтер ишлаб чиқилиши лозим.

Гап билан сўз бирикмасини фарқловчи бошқа белгилар ҳам бор. Лекин улар ҳар иккала синтактик бирлик учун умумий эмас. Жумладан, гап бир компонентдан, яъни потенциал бир сўздан ташкил топиши мумкин: *Баҳор! Йўқол! Курсинлар!* каби. Сўз бирикмаси минимум икки компонентли бўлади. Демак, гапнинг бир компонентли ҳам бўла олиши, сўз бирикмасининг эса камида икки компонентдан тузилиши гап билан сўз бирикмасини фарқловчи белгилардан. Шундан келиб чиқиб, корпусда гап ва сўз бирикмасини аниқлашнинг яна бир алгоритми сифатида бир компонентли ва икки компонентли эканлигига асосланиш мумкин. Бироқ олимлар бу белгининг иккала синтактик бирлик учун умумий эмаслигини таъкидлашади. Шунинг учун ҳамма вақт ҳам бу белгига асосланиб, сўз бирикмасини гапдан фарқлаб бўлмайди. Шундай гаплар борки, улар учун икки компонентлилик зарурий шарт. Бу компонентлар гапнинг предикатив бирлик сифатида тузилишида, предикативликни ифодалашда муҳим роль ўйнайди. Масалан: *Болалар ёш. Кетмон чопмоқ – олмоқ-солмоқ, қўш ҳайдамоқ – бормоқ-келмоқ* (Мақол). *Сен талабасан* каби⁷⁴.

Эга билан кесимдан ташкил топган икки компонентли гапларда компонентлар орасидаги икки томонлама тобелик муносабатини бундай гаплар билан сўз бирикмасини фарқловчи белги сифатида ажратиш мумкин. Бунда эга кесимнинг ўзига мос келадиган шахс ва сонда қўлланишини тақозо этади: *Улар келдилар. Биз ўқидик* каби. Шу билан бирга, кесим эганинг бош

⁷³ Нурмонов А., Маҳмудов Н., Аҳмедов А., Солихўжаев С. Ўзбек тилининг мазмуний синтаксиси. Тошкент: Фан, 1992. – 9-10-б.

⁷⁴ Нурмонов А., Маҳмудов Н., Аҳмедов А., Солихўжаев С. Ўзбек тилининг мазмуний синтаксиси. Тошкент: Фан, 1992. – 15-16-б.

келишик формасида қўлланишини талаб қилади. Эга билан кесим алоқасида эга кесимни ўзига мувофиқ келадиган шаклга киришини талаб қилса ҳам, гапликни кўрсатишда, предикативликни ифодалашда кесим муҳим роль ўйнайди. Эга фақат предметни атаб кўрсатади, мавзуни билдиради, у ҳақда нима дейишини кесим кўрсатади, яъни кесим эганинг предикатив белгисини, у ҳақдаги ҳукмни, баённи англатади⁷⁵. Демак, “эга+кесим” типидagi компонентлар муносабати гапга хос. Бундай муносабат сўз бирикмасига мутлақо хос эмас. Бунга сабаб “эга+кесим” типидagi муносабатда предикативликнинг ифодаланишидир. Бунда кесим – предикативликни ифодаловчи восита. Бу умумлашма кесим гапнинг асосий конструктив бўлаги эканлигидан келиб чиқади⁷⁶.

Юқоридаги ёндашувлардан келиб чиққан ҳолда, тил корпусларида синтактик бирликларни фарқлашнинг назарий асослари сифатида қуйидагиларни ажратиш мумкин:

биринчидан, корпусда қидирувнинг содда кўринишлари сифатида лемма, сўзшакл, ибора ва коллокация фарқланса-да, аннотацияли корпусларда морфологик белги асосидаги қидирув етакчилик қилади. Синтактик ва семантик белгилар асосидаги қидирув нисбатан мураккаб жараён бўлганлиги сабабли ҳамма корпусларни ҳам синтактик/семантик теглаш имкони, воситаси мавжуд эмас;

иккинчидан, синтактик теглар тизимини ишлаб чиқишда замонавий автоматик таҳлил воситалари/дастурлар билан бирга ўзбек тилшунослигида синтаксис бўйича мавжуд қарашлар асосида ўзбек тили корпуслари учун парсер дастурини ишлаб чиқиш мумкин. Бунда синтактик бирликлар ҳақида яратилган барча назарий материалларни ўрганиш, умумлаштириш ҳамда синтактик теглар яратишда ёндашувни тўғри танлаш муҳим;

⁷⁵ Нурмонов А., Маҳмудов Н., Аҳмедов А., Солихўжаев С. Ўзбек тилининг мазмуний синтаксиси. Тошкент: Фан, 1992. – 15-16-б.

⁷⁶ Нурмонов А., Маҳмудов Н., Аҳмедов А., Солихўжаев С. Ўзбек тилининг мазмуний синтаксиси. Тошкент: Фан, 1992. – 16-б.

учинчидан, синтактик бирликларга турлича қараш/ёндашувнинг мавжудлиги корпус бирликларини теглашда ўзбек тили синтактик бирликлар тизимини аниқлаш ҳам асосий вазифа саналади;

тўртинчидан, корпусда гап ёки бирикма эканлиги ташқи белгилари асосида фарқланмайдиган бирикувларни аниқлашнинг махсус алгоритми, фильтри ишлаб чиқилиши лозим. Бунда гап/сўз бирикмасини аниқлашнинг бир компонентли ёки икки компонентли эканлигига асосланиш натижа бера олади.

2.1.3. Синтактик теглашда синтактик алоқа, синтактик муносабат ва синтактик валентлик тушунчаси. Синтактик алоқа деганда, одатда, икки сўзнинг ўзаро боғланиши тушунилади. Аслида синтактик алоқа бундан кўра кенгроқ тушунчани ўз ичига олади. У гапда сўз ва сўз бирикмаларининг ўзаро боғланишини билдириш билан бирга, нутқда гапларнинг маълум усул ва воситалар ёрдамида бир-бири билан боғланишини ҳам билдиради. Маълумки, нутқ гаплардан ташкил топади. Муайян текстдаги гаплар бир-бири билан мантиқий муносабатда бўлади. Синтактик алоқа туфайли мантиқий муносабат грамматик шаклланади. Масалан: *Куз эрта келди. Ҳаво совий бошлади. Деҳқонлар мўл ҳосилни тезда йиғиштириб олишга киришдилар.* Бу мисолдаги гаплар орасида мантиқий муносабат бор. Бироқ унда ҳали синтактик алоқа йўқ. Бу гапларни ўзаро боғлаб, қуйидагича қайта тузайлик: *Куз эрта келиб, ҳаво совий бошлади, шунинг учун деҳқонлар мўл ҳосилни тезда йиғиштириб олишга киришдилар.* Қайта тузилган гапдаги компонентлар орасида мантиқий муносабат билан бирга синтактик алоқа ҳам бор. Олдинги мисол учта мустақил содда гапдан иборат эди, кейинги мисол ўша содда гаплардан ташкил топган битта қўшма гап. Демак, синтактик алоқа туфайли содда гаплар янги сифатга эга бўлди; учта содда гап битта қўшма гапга бирлашди. Синтактик алоқа қўшма гапни шакллантиради⁷⁷. Айтиш мумкинки, синтактик теглашда синтактик алоқанинг тури инобатга

⁷⁷ Ўзбек тили грамматикаси. II том. Синтаксис. – [Масъул муҳаррирлар: Ф. А. Абдурахмонов ва бошқ.]. – Т., «Фан», 1976. – 560 б. – Б. 7-8.

олинади ҳамда теглар тизимидан синтактик алоқани ифодалайдиган тег жой олади.

Манбаларда бу икки атаманинг фарқланмай, аралаш ишлатилиши айтилади. Синтактик алоқа сўз бирикмасидаги ҳамда гапдаги сўзлар, нутқдаги гапларнинг ўзаро грамматик боғланишини билдиради. Синтактик муносабат эса сўз бирикмаси, гапдаги сўзларнинг ҳамда нутқдаги гапларнинг бир-бирига нисбатан қандай грамматик маъно ва функцияда қўлланишини билдиради. Масалан: *уйга бормоқ, сенга олмоқ*. Бу мисолларда синтактик алоқа бир хил: ҳар иккала бирикмадаги элементлар бир хил грамматик восита – жўналиш келишиги қўшимчаси *-га* ёрдамида бошқарув алоқасига киришган. Бироқ бу бирикмалардаги элементларнинг синтактик муносабатини бир хил деб бўлмайди. Буларда муносабат икки хил: биринчи бирикмада ҳолли муносабат (ўрин муносабати) ифодаланган бўлса, иккинчи бирикмада объектли муносабат ифодаланган⁷⁸. Маълум бўладикки, тил корпуси материали(матн)ни синтактик теглашда синтактик алоқа ҳамда синтактик муносабат, албатта, фарқланиши лозим. Аммо шаклан фарқлашнинг имкони йўқ: боғловчи воситалар бир хил. Шу сабабли корпус разметкасида синтактик алоқа ва синтактик муносабатни фарқловчи тег киритилиши керак. Бизнингча, матнда синтактик алоқа ёки синтактик муносабатнинг тегланишини аниқлаш муҳим.

Синтактик алоқа. Мустақил сўзларнинг нутқ жараёнидаги эркин боғланиши *синтактик алоқа* дейилади: *Сувлар тиниқ, тоза ҳаво*. Синтактик бирлик бир неча аъзоли бўлади. Фақат гап бир аъзоли бўлиши мумкин. Нутқда сўзларнинг бир-бири билан боғланиши ҳар бир тилнинг лисоний-синтактик қолипи, лексеманинг бирикиш имконияти асосида содир бўлади. Сўзнинг синтактик муносабатга киришуви натижасида ҳосил бўлган синтактик қурилма *сўз қўшилмаси* дейилади. Бу сўзнинг бир-бирига эргашиши (*аълочи ўқувчи*) ёки тенглашиши (*олма ва анор*) шаклларида

⁷⁸Ўзбек тили грамматикаси. II том. Синтаксис. – [Масъул муҳаррирлар: Ғ. А. Абдурахмонов ва бошқ.]. – Т., «Фан», 1976. – 560 б.

бўлиши мумкин. Демак, сўзлар орасидаги алоқа икки хил: тенглашиш ва тобеланиш. Улар эса уч хил ҳосилани беради: сўз бирикмаси, сўз тизмаси ва гап. Тенглашиш алоқаси асосида сўз тизмаси ҳосил бўлса, тобеланиш алоқаси асосида сўз бирикмаси ва гап шаклланади⁷⁹.

Корпус синтактик разметка тизимида коллокатив қидирув жуда муҳим: қидирув натижасининг сифати қидирувнинг мукамаллиги билан тўғри пропорционал. Коллокатив қидирувни ташкил этиш учун корпус бирликлари сўз бирикмаси ва сўз қўшилмасини ажратувчи тегга эга бўлиши талаб қилинади.

Сўз бирикмаси бирикувчи сўзларнинг ҳоким-тобелик муносабати асосида ҳосил бўлади. Бунда изоҳланувчи аъзо кенгаювчи, изоҳловчи аъзо *кенгайтирувчи* аъзо деб юритилади: *китобни ўқимоқ* (китобни – изоҳловчи, кенгайтирувчи, ўқимоқ – изоҳланувчи, кенгаювчи).

Тобеланиш алоқасида ҳоким аъзо ҳокимлик мавқеи ва воситасига, тобе аъзо тобелик мавқеи ва уни таъминловчи воситага эга бўлади. *китобни ўқимоқ* бирикмасида *китобни* аъзоси олдин келиб, тушум келишиги кўрсаткичи билан шаклланган бўлса, *ўқимоқ* аъзоси ҳокимлик мавқеи ва шаклига эга. Аъзоларнинг ҳоким-тобелигини билдирадиган шаклий кўрсаткич ҳар доим ҳам бўлавермайди. Морфологик кўрсаткич бўлмаганда тобелик ва ҳокимлик белгиси бирикувчи сўзларнинг луғавий-грамматик табиатидан, синтактик ўрнидан ҳамда тартиб ва оҳангдан билиниб туради. Масалан: тиниқ сув (тобе+ҳоким), сув тиниқ (тобе+ҳоким)⁸⁰.

Морфологик кўрсаткичсиз тобе-ҳокимлик муносабатини аниқлаш учун лисоний-синтактик қолип ҳамда морфологик теглар тизимидан фойдаланиш ўринли. Масалан, *қизиқарли китоб* бирикмасида морфологик кўрсаткич йўқ.

⁷⁹Замонавий ўзбек тили: Синтаксис./ Муаллифлар жамоаси. Масъул муҳаррирлар Ҳ.Ғ.Неъматов, Р.Сайфуллаева. ЎзР Олий ва ўрта махсус таълим вазирлиги, Мирзо Улуғбек номидаги Ўзбекистон Миллий университети. – Тошкент.: Мумтоз сўз, 2011. – 312 б. – Б. 26.

⁸⁰Замонавий ўзбек тили: Синтаксис./ Муаллифлар жамоаси. Масъул муҳаррирлар Ҳ.Ғ.Неъматов, Р.Сайфуллаева. ЎзР Олий ва ўрта махсус таълим вазирлиги, Мирзо Улуғбек номидаги Ўзбекистон Миллий университети. – Тошкент.: Мумтоз сўз, 2011. – 312 б. – Б. 59-61.

Контекстда бу бирликларнинг сўз бирикма эканлигини лисоний-синтактик қолип билан аниқлаш мумкин. Лисоний-синтактик қолипда қайси сўз туркуми ўзаро бирикиши кўрсатиб берилган. Корпус бирликлари морфологик тегланган бўлса, тег ва лисоний-синтактик қолип асосида морфологик кўрсаткичсиз синтактик муносабатли бирикмаларни ҳам автоматик теглаш имкони пайдо бўлади. Ушбу ҳолатни қуйида батафсил изоҳлаймиз.

2.1.4. *Ўзбек тилида сўз бирикмаларининг лисоний-синтактик қолипларига асосланган синтактик теглаш тизими.* Бошқа лисоний ҳодиса каби сўз бирикмаларининг лисоний-синтактик қолипларини тиклаш ва унинг воқеланиши фарқланади. СБ ЛСҚнинг воқеланиши «лисон-нутқ» йўналишида бўлиб, бу йўналиш умумийликдан оралиқ кўриниш орқали хусусийликка – нутқий ҳодисага қараб боради⁸¹. Бу борада С.Назарова томонидан ишлаб чиқилган ЛСҚ⁸²ларга алоҳида эътибор қаратиш муҳим. Жумладан, С.Назарова ЛСҚларнинг $[W_{\text{морфологик восита}} - W_{\text{морфологик восита}}]$ кўринишдаги инвариант, $[W_{\text{қараткич келишиги}} - W_{\text{эғалик кўшимчаси}}]$, $[Исм_{\text{қараткич келишиги}} - Исм_{\text{эғалик кўшимчаси}}]$, $[От_{\text{қараткич келишиги}} - От_{\text{эғалик кўшимчаси}}]$, $[От_{\text{атокли келишиги}} - От_{\text{турдош эғалик кўшимчаси}}]$ каби вариантларини ажратади. Албатта, тил корпусида сўз бирикмаларни теглаш муаммоси умумий (инвариант) ЛСҚлар билан ҳал этилмайди, балки сўз бирикмаларни аниқлашда нисбатан аниқроқ қолиплар талаб этилади.

С.Назароваушбу ЛСҚнинг қуйи бўлинишлари унинг тегишли сўз туркуми томонидан тўлдирилиши асосида қуйидагича таснифлайди⁸³:

1) $[От_{\text{қараткич келишиги}} \rightarrow От_{\text{эғалик кўшимчаси}} = От_{\text{қаратувчи}} + От_{\text{қаралмиш}}]$: *китобнинг варағи*;

⁸¹ <https://uz.denemetr.com/docs/769/index-313952-1.html?page=30>

⁸² Назарова С. Бирикмаларда сўзларнинг эркин боғланиш омиллари: филол.фан. номз.... дисс. автореф. – Тошкент: 1997. – 26 б.

⁸³ Назарова С. Бирикмаларда сўзларнинг эркин боғланиш омиллари: филол.фан. номз.... дисс. автореф. – Тошкент: 1997. – 26 б.

- 2) [от^{қаратқич келишиги} → сифат^{эғалик қўшимчаси} = от^{қаратувчи} + сифат^{қаралмиш}]:
дарахтнинг мўрти;
- 3) [сифат^{қаратқич келишиги} → от^{эғалик қўшимчаси} = сифат^{қаратувчи} + от^{қаралмиш}]:
гулнинг/қизилининг ҳиди;
- 4) [сифат^{қаратқич келишиги} → сифат^{эғалик қўшимчаси} = сифат^{қаратувчи} + сифат^{қаралмиш}]:
олманинг/каттасининг чучуги;
- 5) [от турдош^{қаратқич келишиги} → сифат^{эғалик қўшимчаси} = от^{қаратувчи} + сон^{қаралмиш}]:
гулнинг биттаси;
- 6) [сон^{қаратқич келишиги} → сон^{эғалик қўшимчаси} = сон^{қаратувчи} + сон^{қаралмиш}]: *ўннинг ярми;*
- 7) [от^{қаратқич келишиги} → ҳаракат номи^{эғалик қўшимчаси} = от^{қаратувчи} + феъл^{қаралмиш}]:
Отабекнинг қайтиши;
- 8) [ҳаракат номи^{қаратқич келишиги} → от^{эғалик қўшимчаси} = феъл^{қаратувчи} + от^{қаралмиш}]:
уялишининг ўрни;
- 9) [ҳаракат номи^{қаратқич келишиги} → ҳаракат номи^{эғалик қўшимчаси} = феъл^{қаратувчи} + феъл^{қаратувчи}]: *олмоқнинг бермоғи;*
- 10) [от^{қаратқич келишиги} → сифатдош^{эғалик қўшимчаси} = от^{қаратувчи} + сифатдош^{қаралмиш}]: *юракнинг тўхтагани;*
- 11) [сифатдош^{қаратқич келишиги} → от^{эғалик қўшимчаси} = сифатдош^{қаратувчи} + от^{қаралмиш}]:
кўрққаннинг кўзи;
- 12) [от^{қаратқич келишиги} → равиш^{эғалик қўшимчаси} = от^{қаратувчи} + равиш^{қаралмиш}]:
меҳнатнинг кеча-кундузи;
- 13) [олмош^{қаратқич келишиги} → от^{эғалик қўшимчаси} = олмош^{қаратувчи} + от^{қаралмиш}]:
менинг ватаним;
- 14) [равиш^{қаратқич келишиги} → от^{эғалик қўшимчаси} = равиш^{қаратувчи} + от^{қаралмиш}]:
ҳозирнинг ҳузури.

Ушбу қолиплар С.Назарова томонидан тадқиқ этилган: юзлаб нутқий ҳосилаларда синаб кўрилган, умумлаштирилган. Шунинг учун биз исм+исм қолипли сўз бирикмаларнинг моделини шу қолиплар асосида тузишимиз мумкин. Бунинг учун, аввало, қолип таркибидаги исмларнинг турини

англатувчи қисмларни маълум белги билан, тобеланишни кўрсатиб турувчи морфологик воситаларнинг махсус белгиларини танлаб оламиз. Бунда от = N, сифат = Adj, сон = Num, турдош от = N^{sub}, олмош = Pr, равиш = Prv, ҳаракат номи = Ger теглари билан; қаратқич келишиги = Case (ёки Cs), эгалик кўшимчаси = Possessive (ёки Pos) теглари билан белгиланади. Шундан келиб чиқиб, тил корпуси учун сўз бирикмаларни синтактик теглашнинг исм+исм қолипи учун қуйидаги моделларни таклиф қилиш мумкин:

- 1) [N^{Cs} → N^{Pos}]: *китобнинг варағи*;
- 2) [N^{Cs} → Adj^{Pos}]: *дарахтнинг мўрти*;
- 3) [Adj^{Cs} → N^{Pos}]: *гулнинг/қизилининг ҳиди*;
- 4) [Adj^{Cs} → Adj^{Pos}]: *олманинг/каттасининг чучуги*;
- 5) [N турдош^{Cs} → Adj^{Pos}]: *гулнинг биттаси*;
- 6) [Num^{Cs} → Num^{Pos}]: *ўннинг ярми*;
- 7) [N^{Cs} → Ger^{Pos}]: *Отабекнинг қайтиши*;
- 8) [Ger^{Cs} → N^{Pos}]: *уялишининг ўрни*;
- 9) [Ger^{Cs} → Ger^{Pos}]: *олмоқнинг бермоғи*;
- 10) [N^{Cs} → Adjдош^{Pos}]: *юракнинг тўхтагани*;
- 11) [Adjдош^{Cs} → N^{Pos}]: *кўрққаннинг кўзи*;
- 12) [N^{Cs} → Prv^{Pos}]: *меҳнатнинг кеча-кундузи*;
- 13) [Pr^{Cs} → N^{Pos}]: *менинг ватаним*;
- 14) [Prv^{Cs} → N^{Pos}]: *ҳозирнинг ҳузури*.

Шунингдек, замонавий синтаксисда сўз бирикмасининг [исм^{келишик} кўшимчаси + феъл] қолипи ҳам ажратилади⁸⁴. Сўз бирикмасининг бу ЛСҚи биринчи узви морфологик шаклланган ва шунинг учун ШМЖ бирикув омилига эга ва иккинчи қисмда грамматик кўрсаткич боғланиш учун нисбатан аҳамиятсиз бўлганлиги боис, бирикув омилларининг МШЖ комбинацияси амал қилади. Қолипнинг

⁸⁴Замонавий ўзбек тили: Синтаксис./ Муаллифлар жамоаси. Масъул муҳаррирлар Ҳ.Ғ.Неъматов, Р.Сайфуллаева. ЎзР Олий ва ўрта махсус таълим вазирлиги, Мирзо Улуғбек номидаги Ўзбекистон Миллий университети. – Тошкент.: Мумтоз сўз, 2011. – 312 б. – 67-71.

биринчи узвини исм туркумига оид сўзлар тўлдириб, улар ҳоким узвга тушум, жўналиш, ўрин-пайт, чиқиш келишиклари шакллари билан боғланади. Биринчи узви асосида ЛСҚ қуйидаги кўринишларга эга бўлади⁸⁵:

- 1) [исм^{тушум келишиги} → феъл];
- 2) [исм^{жўналиш келишиги} → феъл];
- 3) [исм^{ўрин-пайт келишиги} → феъл];
- 4) [исм^{чиқиш келишиги} → феъл].

ЛСҚнинг ҳар бир кўриниши исм узвининг бўлиниши асосида яна тармоқланади. Масалан, [исм тушум келишиги → феъл] қолипи қуйидаги бўлинишларни беради⁸⁶:

- 1) [от^{тушум келишиги} → феъл];
- 2) [сифат^{тушум келишиги} → феъл];
- 3) [сон^{тушум келишиги} → феъл];
- 4) [олмош^{тушум келишиги} → феъл];
- 5) [ҳаракат номи^{тушум келишиги} → феъл].

Ушбу қолипларга *китобни/яхшини/ўнни айир, уни гапириш/ўқимоқ/сўрамоқ/ бошламоқ/ни бас қил* кабиларни мисол келтириш мумкин.

Юқоридаги ЛСҚлар асосида синтактик теглашнинг исм+феъл қолипи учун [Sub^{Cs3} → V]; [Sub^{Cs4} → V]; [Sub^{Cs5} → V]; [Sub^{Cs6} → V] моделларини таклиф этсак, исмнинг кўринишларига кўра турлари бўйича [N^{Cs3} → V]; [Adj^{Cs3} → V]; [Num^{Cs3} → V]; [Pr^{Cs3} → V]; [Ger^{Cs3} → V] каби моделлар сўз бирикмаларни фарқлашга ёрдам беради. Бунда ^{Cs3} тушум, ^{Cs4} жўналиш, ^{Cs5}

⁸⁵Замонавий ўзбек тили: Синтаксис./ Муаллифлар жамоаси. Масъул муҳаррирлар Ҳ.Ғ.Неъматов, Р.Сайфуллаева. ЎзР Олий ва ўрта махсус таълим вазирлиги, Мирзо Улуғбек номидаги Ўзбекистон Миллий университети.-Тошкент.: Мумтоз сўз, 2011.-312 б.

⁸⁶Замонавий ўзбек тили: Синтаксис./ Муаллифлар жамоаси. Масъул муҳаррирлар Ҳ.Ғ.Неъматов, Р.Сайфуллаева. ЎзР Олий ва ўрта махсус таълим вазирлиги, Мирзо Улуғбек номидаги Ўзбекистон Миллий университети. – Тошкент.: Мумтоз сўз, 2011. – 312 б. – 67-71.

ўрин-пайт, ^{Cs6}чиқиш келишигини; Sub исмни, V феълни ифодаловчи тег ҳисобланади.

Хулоса сифатида айтиш жоизки, сўз бирикмалари теглар тизимини ишлаб чиқишда синтактик алоқа ва синтактик муносабатни фарқлаш, икки бирликдан иборат гап ва сўз бирикмаси фарқланиши лозим. Шунингдек, морфологик кўрсаткичсиз ва морфологик кўрсаткичли сўз бирикмасини аниқлашда лисоний-синтактик қолиплар асосида тузилган моделлардан фойдаланиш синтактик теглар тизимининг мукаммал бўлишига олиб келади.

2.2. Ўзбек тилида содда йиғиқ гапларни синтактик теглаш:

лисоний-синтактик қолип ва лингвистик модел

2.2.1. *Синтактик разметкалашда бирликларни ажратиш тамойиллари.* ХАНКО тузувчилари мавжуд ёндашувларни таҳлил қилар экан, синтактик разметканинг икки альтернатив синтактик разметка схемаси: тобе-ҳокимлик грамматикаси ҳамда анъанавий гап бўлаклари синтаксиси энг кулай разметка усули эканлиги ҳақида хулосага келдилар. Бу икки ёндашувдан аралаш ҳолда фойдаланиш куйидаги муаммолар ечимига олиб келади⁸⁷:

- 1) синтактик курилманинг аъзоси, улар орасидаги алоқани батафсил изоҳлаш;
- 2) профессионал тилшунос ва таълим берувчи талабини қондириш;
- 3) фойдаланувчи истагига кўра альтернатив разметка натижаларини мустақил/ биргаликда кўрсатиш.

Бундай типдаги аннотациялашни ташкил этиш мантиқий жиҳатдан икки босқичга бўлинади⁸⁸.

⁸⁷ Копотев В., Гурин Г.Б., Принципы синтаксической разметки хельсинкского аннотированного корпуса русских текстов ХАНКО // https://www.mv.helsinki.fi/home/kopotev/dialog_2006.pdf

⁸⁸ Копотев В., Гурин Г.Б., Принципы синтаксической разметки хельсинкского аннотированного корпуса русских текстов ХАНКО // https://www.mv.helsinki.fi/home/kopotev/dialog_2006.pdf

Ишнинг олдинги бўлимларида айтилганидек, ҳамма корпусда ҳам синтактик разметка мавжуд эмас: бу ҳолат синтактик разметка воситасининг йўқлиги билан изоҳланади. Рус тилининг ХАНКО корпусида синтактик разметка мавжуд бўлиб, қуйида ундаги айрим хусусиятларга тўхталамиз.

ХАНКО корпусида анъанавий синтаксисдан фойдаланилган бўлиб, унинг афзалликлари сифатида оддийлик, умумэътироф этилганлик; тадқиқот мақсадида билвосита материал қидириш имконининг мавжудлиги кўрсатилади. Ушбу ёндашувнинг камчилиги сифатида замонавий синтактик структуралар ҳақидаги тасаввурларга мос келмаслик; синтактик аъзонинг тавсифланганлиги, аммо синтактик муносабатнинг четлаб ўтилгани; тавсифда изчилликнинг йўқлиги, ечилмаган қарама-қарши фикрларнинг мавжудлиги (иккинчи даражали бўлакни аниқлаш чегарасининг қатъий тизими мавжуд эмаслиги); автоматик ишлов беришнинг мураккаблиги таъкидланади⁸⁹. ХАНКО тузувчилари, биринчи навбатда, у ёки бу конкрет топшириқни бажаришда маълум бир синтактик ахборот қай даражада зарурлиги, маълумотга автоматик ишлов беришнинг қанчалик мураккаблигига алоҳида эътибор қаратишади.

Қўлда ишлов бериладиган ишлар ҳамда натижалар аҳамиятини прогнозлаш, одатда, бир-бирига зид: бирлик сифатида детерминантнинг танланиши кўл ишини сунъий равишда оширади, чунки детерминант қидирувини автоматлаштириш имкони мавжуд эмас. Бу ишни изчил давом эттириш қийин бўлса-да, кўпинча, синтактик теглар мажмуидан чиқарилади. Бу ҳолат детерминант тушунчасининг турли лингвистик тадқиқотларда турлича талқин қилиниши билан изоҳланади⁹⁰.

Интерфейснинг қулайлиги, албатта, инобатга олиниши лозим. Синтактик ахборот турли бирликларга, хусусан, морфологик ахборотга эга бўлган матн қисмларига ҳам бириктирилади. Шундай пайтда, икки марта

⁸⁹ Копотев В., Гурин Г.Б., Принципы синтаксической разметки хельсинкского аннотированного корпуса русских текстов ХАНКО // https://www.mv.helsinki.fi/home/kopotev/dialog_2006.pdf

⁹⁰ Копотев В., Гурин Г.Б., Принципы синтаксической разметки хельсинкского аннотированного корпуса русских текстов ХАНКО // https://www.mv.helsinki.fi/home/kopotev/dialog_2006.pdf

тегланган бирликлар корпусда амал қилувчи “+” белгиси билан белгиланади. Табиийки, синтактик разметканинг ўзи ҳам икки марталик бўлади. Масалан, матн таркибидаги бирлик (сўзшакл) контекстига қараб ҳам ҳол, ҳам тўлдирувчи бўлиб келиши мумкин: бундай бирликларга икки хил тег бириктирилиши табиий. Корпусда билвосита белгилар асосида аниқланиши мумкин бўлган бирликлар тегланмаслиги ҳам мумкин: масалан корпусда ҳис-ҳаяжонли ва ҳис-ҳаяжонсиз гаплар тегланмайди, чунки бундай гаплар тегишли тиниш белгилари орқали фарқланади⁹¹.

Ушбу кузатишларимиз асосида гап бўлаклари ҳамда гапнинг маъно муносабатига кўра турларини теглаш борасида қуйидаги хулосаларга келиш мумкин:

биринчидан, тиниш белгилари орқали фарқлаш имкони мавжуд бўлган гапга шундай ахборотдан иборат тег бириктирмаслик ҳам мумкин;

иккинчидан, икки хил гап бўлаги бўлиши мумкин бўлган бирликларга иккилик тег бириктириш талаб этилади;

учинчидан, синтактик теглар тизимини белгилашда интерфейснинг қулайлигини инобатга олиш зарурати мавжуд;

тўртинчидан, синтактик теглар тизимини тузишда матнга автоматик ишлов беришга эришиш, қўлда ишлов беришни имкон қадар камайтириш принципига амал қилиш лозим.

2.2.2. *Ўзбек тилшунослигида йиғиқ гап ва ихчам гап ҳақидаги назарий умумлашмалар хусусида.* Йиғиқ гап атамасини ўзбек тилшунослигига дастлаб тилшунос А.Ғуломов киритган, унга муфассал таъриф берган⁹². А.Ғуломов таълимотида йиғиқ гап эга ва кесимдан иборат бўлган содда гапнинг бир тури сифатида талқин этилади. Бундай эга ва кесимдан иборат содда гаплар икки таркибли содда гапнинг энг характерли типи эканлиги

⁹¹ Ғуломов А.Ғ. Содда гап. – Тошкент: ЎзФА, 1955. – Б. 47.

⁹² Ғуломов А.Ғ. Содда гап. – Тошкент: ЎзФА, 1955. – Б. 49.

кўрсатиб ўтилади⁹³. Тилшунослар О.Азизов⁹⁴, М.Мирзаев, С.Усмонов, И.Расулов⁹⁵, Ғ.Абдурахмонов⁹⁶, Н.Маҳмудов, А.Нурмоновнинг асарларида⁹⁷ ҳам *йиғиқ гап* атамаси “эга ва кесимдан иборат гап” ёки “фақат бош бўлакдан тузилган гап” сифатида талқин этилади. Фақат кесимдан иборат бўлган гаплар назардан четда қолади. Фақат XX аср охирларида бундай гаплар *ихчам гап* атамаси остида берилди бошлади. УЎТ ДТС асосидаги дастурга мувофиқ яратилган 5-синф она тили дарслигидан “Кесим. Ихчам гап”⁹⁸ мавзулари ўрин олганлиги фикримиз далилидир. “Ихчам гап” мавзуси остида фақат кесимдан иборат бўлган гаплар талқин этилиб, улар содда йиғиқ гап таркибига киритилмади. Аммо “Содда йиғиқ гап” мавзусида *Олдим. Кўргансиз. Учувчи бўлади. Синфдошимиз. Ёзгансиз.* (Ҳ.Неъматов, А.Фуломов, М.Қодиров, М.Абдураимова. Она тили. 5-синф. 14-бет, 28-машқ.)⁹⁹ каби мисоллар ҳам келтирилади.

Гапнинг бош бўлаги бўлган кесим, унинг мавқеи ҳақида кейинги йилларда ўзбек тилшунослигида қатор тадқиқотлар юзага келди. Айниқса, В.Эгамбердиев¹⁰⁰, Ж.Омонтурдиев¹⁰¹ тадқиқотларидан сўнг ҳамда улар асосида формал-функционал тадқиқ усулларида иш олиб борган Р.Сайфуллаева,

⁹³Фуломов А., Асқарова М. Ҳозирги ўзбек адабий тили. Синтаксис: университетлар ва педагогика институтларининг филология факультетлари учун дарслик. – Тошкент: Ўқитувчи, 1987. – Б. 67.

⁹⁴ Азизов О. Тилшуносликка кириш. – Тошкент: ЎзССР “Ўрта ва олий мактаб” давлат нашриёти, 1963. – Б.111.

⁹⁵ Мирзаев М., Усмонов С., Расулов И. Ўзбек тили. – Тошкент: Ўқитувчи, 1970. – Б. 181.

⁹⁶ Абдурахмонов Ғ. Ўзбек тили грамматикаси. – Тошкент: Ўқитувчи, 1996. –Б. 8.

⁹⁷ Маҳмудов Н, Нурмонов А. Ўзбек тилининг назарий грамматикаси. – Тошкент: Ўқитувчи, 1995. – Б. 65.

⁹⁸ Қаранг: Неъматов Ҳ., Фуломов А. ва бошқ. Она тили: 5-синф учун дарслик. – Тошкент: Ижод дунёси, 2001. – Б.14.

⁹⁹ Неъматов Ҳ., Фуломов А. ва бошқ. Она тили 5-синф учун дарслик. – Тошкент, 2001. – Б. 14.

¹⁰⁰Эгамбердиев В. Кесим таърифи ҳақида баъзи мулоҳазалар. Профессор-ўқитувчилар илмий конференциясининг тезислари (Қарши пед.институту). – Қарши, 1967.

¹⁰¹Омонтурдиев Ж. Кесимнинг структурасига кўра классификацияси. // Совет мактаби, 1965. – №4. – Б. 26-27.

М.Абузалова, Л.Раупова, Р.Бобокалонов¹⁰² каби тадқиқотчиларнинг ишларида кесим ҳақида баҳс юритилди. Мазкур тадқиқотларда кесимнинг ўзига хос морфологик белгиларга эга эканлиги, кесимни ифода этадиган сўз ўзининг грамматик структураси билан бошқа бўлаклардан ажралиб туриши, баъзи бир гапларда эга махсус сўз билан ифодаланмасда, кесим орқали эгани аниқлаш мумкинлиги алоҳида таъкидланди. Шунингдек, фақат кесимдан иборат сўзлар гап бўлиб келиши мумкинлиги, бунинг учун у кесимлик мақомига эга бўлиши лозимлиги уқтирилди¹⁰³. Айнан мана шу ишларда келтирилган хулосалар ўзбек тилидаги содда гапларни синтактик теглашда назарий асос бўлиб хизмат қилади. Чунки ушбу ишларда содда гап, хусусан, кесимнинг ифодаланишига кўра турларининг лисоний-синтактик қолиплари яратилган. Ушбу ЛСҚлар асосида синтактик теглар тизимини шакллантириш мумкин. Ишнинг кейинги қисмларида бунга батафсил тўхталамиз.

Гаплар турли-туман бўлганидек, лисоний синтактик қолиплар ҳам ҳар хил бўлади. Бу тилда шаклланган, сўзловчининг коммуникатив нияти, узатилаётган ахборотнинг ўзига хослиги умумийликларининг барқарорлашуви натижасида шаклланади. Гапнинг минимал қолипларини тиклашда ана шу жиҳатлар асос қилиб олинади¹⁰⁴.

Ўзбек тилининг яхлит синтактик таълимотини ишлаб чиқишда гапнинг лисоний сатҳдаги энг кичик қурилиш қолипи,

¹⁰² Сайфуллаева Р. Ҳозирги ўзбек тилида қўшма гапларнинг формал-функционал талқини: монография. – Тошкент: Фан, 1994. – Б.8.; Абузалова М. Ўзбек тилида содда гапнинг энг кичик қурилиш қолиплари ва унинг нутқда вокеланиши: филол. фан. номз. дисс. – Бухоро, 2004. – 22 б.; Раупова Л. Ўзбек тилида номустақил кесим масаласи ва [WP_m-WP_m] қурилишли гаплар: филол. фан. номз. ... дисс. – Тошкент, 1999. – Б.46.; Бобокалонов Р. Ўзбек тилида гап синтаксиси ва сўз-гапларнинг систем-структур талқини: монография. – Тошкент: Фан, 2006. – Б.35.

¹⁰³ Қурбонова М, Сайфуллаева Р., Боқиева Г, Менглиев Б. Ўзбек тилининг структурал синтаксиси. – Тошкент, 2004. – Б. 53.

¹⁰⁴ Сайфуллаева Р. ва б. Замонавий ўзбек тили. Синтаксис. – Тошкент, 2010. – 312 б.

яъни $[WP_m]$ асос қилиб олинди¹⁰⁵. Бундай субстанциал ёндашув ўзбек тили грамматик қурилиши талқинини тубдан янгилашга олиб келди. Гапнинг энг кичик қурилиш қолипи ўзида унинг умумий грамматик маъносини мужассамлаштириб, қурилиш элементлари орасидаги муносабатни ифодалайди. Гапнинг лисоний қолипи – нутқда намоён бўлаётган гапларнинг умумий шартли белгиси. Шунинг учун гапнинг энг кичик қурилиш қолипи – $[WP_m]$ ҳамда унинг тармоқланишидан вужудга келган қуйи лисоний қурилиш қолиплари нутқда воқеланаётган гаплардан умумийлик ва хусусийлик муносабати даражасида фарқ қилади.

$[WP_m]$ қурилишли йиғиқ гаплар. $[WP_m]$ қолипи ўзбек нутқида воқеланган аксарият гапларнинг барчаси учун умумий ва мажбурий; у аслида йиғиқ гап. $[WP_m]$ қурилишли йиғиқ гапларда W ва P_m компонентлари ёйиқ гаплардан шаклан фарқланмайди. P_m йиғиқ гапларда турлича структур кўринишларга эга. WP_m даги шаклий фарқнинг мавжуд эмаслиги синтактик теглар тизимида ҳам ноаниқлик келтириб чиқариши мумкин. Демак, айрим ҳолларда синтактик теглар тизимини шакллантиришда фақат ЛСҚларнинг ўзи етарли эмас.

$[WP_m]$ қурилишли йиғиқ гапларда кесимлик категорияси маъноси кесимлик кўшимчаларининг фақат ўзи билан ёки боғлама ва кесимлик кўшимчалари воситасида ифодаланиши мумкин. Содда гапларни синтактик теглашда кесимлик кўшимчаси асосида кесимни, шу асосда содда гапни аниқлаш яхши натижа бериши мумкин. Чунки корпуснинг морфологик теглар тизимида кесимликни билдирувчи шакллар морфологик

¹⁰⁵ Нигматов Х. и др. Структура предложения и актуальные вопросы синтаксиса тюркских языков (Тезисы формально-функционального исследования) // Советская тюркология, 1984.– №5. – С.5-10.

жихатдан таърифланган. Берилган таъриф (тег)га қўшимча ахборот қўшилса, синтактик тег шаклланади.

“Ўзбек тилининг структурал синтаксиси” китоби муаллифлари таъкидлашганидек, ўзбек тили йиғиқ гапларда кесимлик қўшимчаларининг қуйидаги турлари мавжуд:

1. [P_m]нинг маънолари яхлит ифодаланади: *Бор-ай. Ишлайлик. Юр-ди.* Белгиланган қўшимчалар(-й, -йлик, -ди)дан ҳар бирида шахс-сон, замон, тасдиқ-инкор, майл маънолари биргаликда умумий ифодаланмишга эга. Масалан, *Бор-ай* гапидаги *-ай* қўшимчаси биринчи I шахс (II шахс *Бор*), бирлик (кўпликда *Борайлик*) келаси замон, тасдиқ (*-май* унинг инкор шакли), буйруқ-истак майли (бунда *-син, -нг*, (масалан, *Борсин. Боринг*) билан ассоциатив муносабатга киришади) маъноларига эга. Бундай ҳолатни *Борса. Ёзса* гаплари P_m воқелантирувчилари таркибидаги шарт майли қўшимчасида ҳам кўриш мумкин. Бундай синкретик шаклли маънолар ноль морфемаларда ҳам мавжуд: 1. *Бор!* 2. *Қиш.*

2. [P_m] таркибида тасдиқ-инкор алоҳида шакл орқали, модаллик, замон, шахс-сон маънолари эса яхлит ифодаланади: *Бор - ма – ди. Кел -ди.*

3. [P_m] таркибида шахс-сон маъноси алоҳида шакл орқали, тасдиқ-инкор, модаллик, замон маънолари яхлит ифодаланади: *Бор - са -нг. Кел - са - нгиз. Бил - ди - ми?*

4. [P_m] таркибида тасдиқ-инкор, шахс-сон алоҳида-алоҳида шакл орқали, модаллик ва замон маънолари яхлит ифодаланади: *Бор-ма-ди-нгиз.*

5. [P_m]нинг ҳар бир маъноси алоҳида шакл орқали ифодаланади: *Бор-ма-са-йди-нгиз.*

6. [P_m] маъноларини ифодалашга боғлама ёки тўлиқсиз феъл хизмат қилади. Масалан: *Ишчи бўламан. Ҳамон ёдимдадир: гул чоғи эрди* (Уйғ., И.Султ.)

7. [P_m] маъносини ифодаловчи воситалар ҳаракат тарзи шакллари таркибида келади: *Бақуриб юборди. Бора кўрма*.¹⁰⁶

Юқоридаги [P_m] таснифи шуни кўрсатяптики, уларнинг воқеланиши турлича, аммо ЛСҚда P_mбилан ифодаланади. Бизнингча, синтактик разметкада, гапнинг кесимини автоматик аниқлаш учун юқорида кўрсатилган 7 хил кўринишдаги тузилиш синтактик теглар мажмуига киритилиши талаб қилинади. Ушбу морфемалар морфологик тегланган, яъни дастур уларни автоматик “танийди”. Синтактик ахборотнинг аниқланиши учун уларнинг юқорида кўрсатилган позициялари ҳақида ахборотнинг бўлиши етарли.

[WP] қурилишли *йиғиқ гаплар*. Ўзбек тилшунослигида гапнинг энг кичик қурилиш қолипи [WP_m] сифатида талқин этилгач, баъзи ўринларда зиддиятли, изоҳталаб нуқталар ҳам вужудга келди. Чунки тилимизда шундай сўзлар борки, улар ўзининг алоҳида белгилари билан мустақил сўзлардан ҳам, ёрдамчи сўзлардан ҳам, содда ёки қўшма гаплардан ҳам фарқ қилади. Бу сўзлар – бўлакларга ажралмайдиган гап ёки сўз-гапдир. Сўз-гап мустақил грамматик ҳодиса бўлиб, ўзбек тилшунослигида улар турли хил атамалар билан талқин этилган.

Тилшунос олим, профессор А.Фуломов сўз-гапларни “ундаш гаплар” сифатида эътироф этса¹⁰⁷, М.Мирзаев, С.Усмонов, И.Расуловлар ўзларининг “Ўзбек тили” дарсликларидида сўз-гапларни “бўлакларга бўлинмас гаплар”

¹⁰⁶Қурбонова М., Сайфуллаева Р., Боқиева Г., Менглиев Б. Ўзбек тилининг структурал синтаксиси. – Тошкент, 2004. – Б. 68-69.

¹⁰⁷ Фуломов А.Ф., Асқарова М. Ҳозирги замон ўзбек адабий тили. Синтаксис. – Тошкент, 1965. – Б.177-178.

тарзида талқин этадилар¹⁰⁸. Ўзбек тилшунослигида сўз-гаплар 50-йилларга қадар дарслик ва қўлланмаларга киритилмаган. 50-йиллардан сўнг чоп этилган тилшунос С.Усмоновнинг “Ундов гаплар” номли қўлланмаси бу соҳадаги илк қадам бўлди. С.Усмонов ўзининг мазкур китобида ундов сўзларнинг синтактик вазифасини, яъни улар мустақил сўз-гап вазифасида келишини биринчилардан бўлиб асослаб беради¹⁰⁹.

Ўзбек тилидаги сўз-гапларни формал-функционал таҳлил қилган Р.Бобокалонов уларнинг семантик-структур моҳиятини очар экан, бундай гапларнинг структурасида кесимлик қўшимчалари мутлақо йўқ бўлиб, атов бирлиги ўзида ҳам луғавий маънони, ҳам кесимлик маъноларини яхлит ҳолда ташишини таъкидлайди. “Кесимлик маъно ва вазифалари [W]да мужассамлашган бўлади, улар шундай лисоний бирликларни ташкил этадики, гап ёки гап маркази вазифасида келиш ҳам, кесимлик маъносини ифодалаш ҳам уларнинг онтологик табиатида бирлашган бўлади. Шу сабабли грамматик шаклланган гапларнинг минимал қолипни [WP_m] сифатида берсак, семантик-функционал шаклланган сўз-гапларнинг энг кичик қурилиш қолипи рамзини [WP] сифатида бериш маъқул. Бу рамз кесимлик маъно ва вазифасини [W]нинг ичида, унинг таркибий қисми сифатида мавжудлигига ишорадир. Шу билан бирга, бу рамз худди шундай гапларда кесимлик қўшимча(маркер)ларининг бўлмаслигига, шунинг учун бундай гапларда [WP_m] қурилишли содда гапларда бўладиган кесимлик маъноларининг ранг-баранглиги ҳам хос эмаслигига

¹⁰⁸ Мирзаев М., Усмонов С., Расулов И. Ўзбек тили. –Тошкент, 1966. – Б. 224-225.

¹⁰⁹ Усмонов С. Ундов гаплар. –Тошкент: ЎзФА нашриёти, 1955.-Б.56

айни бир замонда ишора қилиб туради”¹¹⁰. Семантик-функционал шаклланган сўз-гапларнинг кесимлик қўшимчаларидан холилиги уларнинг кенгаймаслигига асос. Бу эса, ўз ўрнида, унинг маъно имкониятлари чегараланганлиги, грамматик шаклларни талаб этмаслиги, ҳар қандай сўзлар билан бирикиш имкониятидан маҳрумлиги, гапда бажарадиган вазифаси чекланганлиги хусусиятларига эгаллигини келтириб чиқаради¹¹¹.

“Сўз-гаплар шундай сўзларки, гап вазифасида қўлланиш, уларнинг луғавий, ички маъносидан ўрин олган ва улар бошқа сўзлар каби кесимлик шакли қўшимчаларига муҳтож эмас. Лекин қўлланишда сўз-гаплар оҳанг, матн билан боғлиқ бўлади»¹¹². “Махсус оҳанг билан талаффуз қилинадиган, аммо бирор ҳукм ёки мавжудлик ифодаламаган сўз/сўз бирикмалари атов гап бўла олмайди. Улар сўз-гапни ташкил этади... Бундай гаплар тобе ёки ҳоким бўлақларни талаб ҳам этмайди, уларда бирор бўлақнинг “яширингани” ҳам сезилмайди. Улар айтиб ўтилган фикрга модал муносабатни ифодалайди”¹¹³.

1. XX асрнинг 90-йилларидан бошлаб ўзбек тилшунослигида сўз-гаплар атрофлича ўрганила бошланди. Тилшунос олим Ғ.Абдурахмонов сўз-гапларга алоҳида тўхталиб, уларни маъно ва вазифасига кўра (1) тасдиқ билдирувчи сўз-гап; (2) инкор сўз-гап; (3) сўроқ сўз-гап; (4) ҳис-туйғу сўз-гаплар каби гуруҳларга бўлади¹¹⁴. Р.Бобокалонов уларни (1) модаллар; (2) ундовлар; (3) тасдиқ-инкорлар; (4)

¹¹⁰ Бобокалонов Р. Ўзбек тилида гап синтаксиси ва сўз-гапларнинг систем-структур талқини: монография. – Тошкент: Фан, 2006. – Б.38.; Шодмонов Э. Ҳозирги замон ўзбек тилида сўз-гаплар: филол. фан. номз. ... дисс. – Тошкент, 1970. – Б.230.

¹¹¹ Қаранг: Шу асарлар. – Б.39; Б.230.

¹¹² Неъматов Ҳ, Ғуломов А. ва бошқалар. Она тили: 7-синф учун дарслик. – Тошкент, 2000. – Б.106.

¹¹³ Абдурахмонов Ғ. Ўзбек тили грамматикаси. –Тошкент: Ўқитувчи, 1996. –Б.35.

¹¹⁴ Юқоридаги дарслик. –Б.35.

таклиф-хитоблар каби тўрт гуруҳга бўлади¹¹⁵. Юқорида сўз-гапларнинг батафсил изоҳини келтиришимизга сабаб шуки, тил корпусларининг синтактик теглари лингвистик таъминотида ушбу гапларнинг барчаси махсус “сўз-гап” <SG> теги билан белгиланиши лозим.

Демак, айтиш мумкинки, ўзбек тилида гапнинг туркона – кесим марказли қурилиши [WP_m] қолипи асосида шаклланган. Шу қолип асосида шаклланган гапларни содда йиғиқ гап деб аташ имкони мавжуд. Бунда гап кесимлик кўрсаткичлари билан шаклланиши шарт. Бироқ нутқимизда от(исм) кесимли гаплар ҳам қадимдан мавжуд бўлганлиги туфайли содда йиғиқ гапнинг эгаси ифодаланиши шарт бўлган кўриниши ҳам мавжуд. Агар бундай гаплар қурилишидаги эга тушириб қолдирилса ҳамда *Ўқитувчи. Талаба* шаклида берилса, фикр фақат қайд моҳиятини акс эттириб, атов гап сифатида намоён бўлиши мумкин ва гапнинг коммуникатив субстанцияси бутунлай ўзгариб кетади. Шунга кўра, ўзбек тилининг субстанциал синтаксисидида [эга+от-кесим] кўринишида намоён бўлувчи *Сен – шифокор* типидидаги йиғиқ гапларни [Э+WP] қолипи асосида ифодалаш мумкин. Бунда *Сен* сўзи қолипнинг [Э] узви воқеланиши бўлса, *шифокор* бирлиги [WP]нинг воқеланишидир.

Кўринадики, ўзбек тилшунослигининг субстанциал синтаксисидида содда йиғиқ гапларнинг қурилиш қолипи фақат [WP_m] кўрсаткичли эмас, балки [Э-Р] тарзида ҳам берилиши мумкин. Бунда [Р_m] унсурунинг [Р] қисми қолдирилиб, [m] қисми туширилганлигининг сабаби шундаки, кесим вазифасида келаётган сўзда марказлик хоссаси морфологик кўрсаткичсиз

¹¹⁵ Бобокалонов Р. Ўзбек тилида гап синтаксиси ва сўз-гапларнинг систем-структур талқини: монография. –Тошкент: Фан, 2006.

намоён бўлган, у эга ва кесимнинг ўзаро муносабатидан юзага чиқади¹¹⁶.

2.3. Ўзбек тилидаги содда ёйиқ гапларни синтактик теглаш масалалари.

Фикр турлича: содда ёки кўшма гап орқали; дарак, сўроқ, буйруқ, ундов гап; икки таркибли ёхуд бир таркибли гап; тўлиқ ёки тўлиқсиз гап орқали ифодаланиши мумкин. Бу ҳол фикрнинг характериға, маълум мақсад ёки ниятға қараб белгиланади. Аммо гап қандай шаклда бўлмасин, унда кишининг маълум бир мақсади реаллашади (ҳис-ҳаяжон ёки сўроқ ифодаловчи гаплар ҳам бундан мустасно эмас)¹¹⁷.

Гапнинг асосий белгилари, унинг сўз ёки сўз бирикмаларидан фарқи нимада? Гапнинг асосий белгилари – унда нисбий фикр тугаллиги ва предикативликнинг мавжуд бўлиши, грамматик жиҳатдан маълум қонун-қоидалар асосида шаклланиши, ўзига хос интонацияға эга бўлиши шарт. Гапнинг бу хусусиятлари кўпчилик тиллар учун умумий, аммо бу белгиларнинг турли тилларда намоён бўлиши ҳамда уларнинг аҳамияти, ўрни турлича. Гап бўлиши учун, албатта, ўз эгаси билан шахсда, сонда мослашган феъл бўлиши шарт эмас. Маълум мақсад, фикр ифодаланган, предикатив муносабат мавжуд бўлган конструкциялар, ҳатто, айрим сўзлар ҳам гап саналади (масалан, бир таркибли номинатив гаплар)¹¹⁸.

Гаплар структурасига кўра, тўрт хил бўлади: содда, мураккаб кўшма гап ва периодлар¹¹⁹.

¹¹⁶Замонавий ўзбек тили: Синтаксис./ Муаллифлар жамоаси. Масъул муҳаррирлар Ҳ.Ғ.Неъматов, Р.Сайфуллаева. ЎзР Олий ва ўрта махсус таълим вазирлиги, Мирзо Улуғбек номидаги Ўзбекистон Миллий университети. – Тошкент.: Мумтоз сўз, 2011. – 312 б. – 88-93.

¹¹⁷Ўзбек тили грамматикаси. II том. Синтаксис. [Масъул муҳаррирлар: Ғ. А.Абдурахмонов ва бошқ.]. – Т., «Фан», 1976. – 560 б. – Б. 95.

¹¹⁸Ўзбек тили грамматикаси. II том. Синтаксис. [Масъул муҳаррирлар: Ғ. А.Абдурахмонов ва бошқ.]. – Т., «Фан», 1976. – 560 б. – Б. 96.

¹¹⁹Ўзбек тили грамматикаси. II том. Синтаксис. [Масъул муҳаррирлар: Ғ. А.Абдурахмонов ва бошқ.]. – Т., «Фан», 1976. – 560 б.

Содда гап –маълум фикр, мақсадни ифодаловчи, грамматик/интонацион томондан шаклланган синтактик бирлик. Содда гап тузилишига кўра йиғиқ ёки ёйиқ бўлиши мумкин. Бош бўлақлардан таркиб топган гаплар –йиғиқ гап, бош бўлақдан ташқари иккинчи даражали бўлақ ҳам иштирок этса, ёйиқ гап саналади¹²⁰.

Мураккаб гап таркибида асосий гап бўлақларидан ташқари ажратилган бўлақ, ундалма/кириш бўлақ ва бирикмалари бўлган гап конструкциясидир. Бундай конструкция содда гап конструкциясидан мазмуни ва структурасига кўра фаркланади. Мураккаб гапларда мураккаб мазмун англашилади: *Йиғим-теримни олганимиздан кейин, баҳоргача икки юз йигирма гектар янги ер очишимиз керак* (А.Қаҳҳор). Бу гапда, биринчидан, *йиғим-теримнинг олиними*, иккинчидан, *янги ер очилми* зарурлиги ифодаланади. Бундан ташқари, бу гапнинг сифатдошли конструкцияси ўз интонацияси ва структурасига кўра, содда гапнинг оддий гап бўлақларидан фарқланиб туради ва эргаш гапларга яқинлашади. Шунингдек, равишдош, шарт феъли, ҳаракат номи бошқарган конструкциялар иштирок этган гапда ҳам юқоридаги хусусиятлар мавжуд. Ундалмали, кириш бўлақ/бирикмали гаплар мураккаб фикрни ифодалаши ҳамда структуралари билан содда гаплардан фарқланиб туради. Шунинг учун ҳам бундай конструкцияли гапларни мураккаб гаплар сифатида алоҳида ўрганиш мақсадга мувофиқ¹²¹.

Қўшма гаплар –турли грамматик, лексик, интонацион воситалар билан икки гапнинг бирикувидан тузилган гап конструкциялари¹²².

Период – маълум мақсад ёки фикр тўлиқ ифодаланган, грамматик ва интонацион жиҳатдан тўлиқ шаклланган, иккидан ортиқ гапларнинг бирикувидан тузилган синтактик конструкция¹²³.

¹²⁰Ўзбек тили грамматикаси. II том. Синтаксис. [Масъул муҳаррирлар: Ғ. А.Абдурахмонов ва бошқ.]. – Т., «Фан», 1976. – 560 б. – Б. 98-99.

¹²¹Ўзбек тили грамматикаси. II том. Синтаксис. [Масъул муҳаррирлар: Ғ. А.Абдурахмонов ва бошқ.]. – Т., «Фан», 1976. – 560 б. – Б. 98-99.

¹²²Ўзбек тили грамматикаси. II том. Синтаксис. [Масъул муҳаррирлар: Ғ. А.Абдурахмонов ва бошқ.]. – Т., «Фан», 1976. – 560 б. – Б. 98-99.

¹²³Ўзбек тили грамматикаси. II том. Синтаксис. [Масъул муҳаррирлар: Ғ. А.Абдурахмонов ва бошқ.]. – Т., «Фан», 1976. – 560 б. – Б. 98-99.

Демак, содда гапларни синтактик теглашда гапнинг (юқорида кўрсатилган) тузилишига кўра турларига алоҳида эътибор қаратиш лозим.

“Замонавий ўзбек тили. Синтаксис” дарслигида гапларнинг [W]нинг, [Pm]нинг ифодаланиши ва [Pm] асосида кенгайишига кўра қуйидаги турлари фарқланади¹²⁴:

I. Бир таркибли гаплар.

1. От марказли йиғиқ кесимлик қўшимчали йиғиқ содда гап.
2. От марказли ёйиқ кесимлик қўшимчали йиғиқ содда гап.
3. От марказли аналитик кесимли йиғиқ содда гап.
4. Феъл марказли йиғиқ кесимли йиғиқ содда гап.
5. Феъл марказли ёйиқ кесимли йиғиқ содда гап.
6. Феъл марказли аналитик кесимли йиғиқ содда гап.

Юқоридаги таснифдаги гап турларига қуйидагилар мисол бўлади: *Баҳор. Фасли баҳор. Баҳордир. Фасли баҳордир. Баҳор эди. Фасли баҳор эди.*

Борилсин. Тошкентга борилсин. Борилмасин. Бу ерга борилмасин. Ҳаёт экан. Ўзиям бола-чақалик бўлгандан кейин ёшликда – юрак самимий истақлар, эзгуниятлар билан тўлиб-тошган пайтларда берган ваъдалар, ичган қасамлар – ҳаммаси ёддан кўтарилиб, дунёнинг ишлари билан ўралашиб қоларкан. Апрель эди. Олам ям-яшиллашиб, гулу муаттар исларга бурканган. Эҳтиёт бўлиш керак каби гапларда кесимлик қўшимчалари тафсиллий усулда берилгандир.

II. Икки таркибли гаплар¹²⁵.

1. Икки таркибли эгасиз от марказли йиғиқ кесимли йиғиқ содда гап.
2. Икки таркибли эгасиз от марказли ёйиқ кесимли йиғиқ содда гап.
3. Икки таркибли эгасиз от марказли аналитик кесимли йиғиқ содда гап.
4. Икки таркибли эгасиз феъл марказли йиғиқ кесимли йиғиқ содда гап.
5. Икки таркибли эгасиз феъл марказли ёйиқ кесимли йиғиқ содда гап.

¹²⁴ Сайфуллаева Р. Ва б. Замонавий ўзбек тили. Синтаксис. – Тошкент, 2010. – 312 б.

¹²⁵ Сайфуллаева Р. Ва б. Замонавий ўзбек тили. Синтаксис. – Тошкент, 2010. – 312 б.

6. Икки таркибли эгасиз феъл марказли аналитик кесимли йиғиқ содда гап.
7. Икки таркибли эгали от марказли йиғиқ кесимли содда гап.
8. Икки таркибли эгали от марказли ёйиқ кесимли содда гап.
9. Икки таркибли эгали от марказли аналитик кесимли содда гап.
10. Икки таркибли эгали феъл марказли йиғиқ кесимли содда гап.
11. Икки таркибли эгали феъл марказли ёйиқ кесимликли содда гап.
12. Икки таркибли эгали феъл марказли аналитик кесимли содда гап.

Икки таркибли гаплар таснифига *Ўқитувчиман. Янги ўқитувчиман. Ўқитувчидирман. Янги ўқитувчидирман. Ўқитувчи бўлай. Яхши ўқитувчи бўлай. Келдим. Кеча келдим. Келмагандим. Кеча келмагандим. Келмаган эдим. Кеча келмаган эдим. Мен ўқитувчиман. Мен янги ўқитувчиман. Мен ўқитувчи бўлай. Мен яхши ўқитувчи бўлай. Мен келдим. Мен кеча келдим. Мен кеча келмагандим. Мен кеча келмаган эдим* каби жумлаларни мисол қилиб келтириш мумкин. Тўғри, нутқда гаплар бундай кўринишда учрамайди: гаплар иккинчи даражали бўлақлар билан келади. Шундай бўлса-да, автоматик синтактик теглаш дастури – парсинг тиниш белгилари чегараси ва эга-кесим белгилари асосида синтактик бирлик чегарасини аниқлайди. Синтактик теглар ва лингвистик моделлар эса содда гапларнинг қайси турига тегишли эканлигини аниқлашга ёрдам беради.

Тадқиқотчи Ш.Хамроева ўзбек тили муаллифлик корпусини тузишнинг лингвистик асосларини тадқиқ этар экан, корпус материалларини синтактик теглаш масаласига йўл-йўлакай муносабат билдиради ҳамда содда гапларни синтактик теглаш бўйича айрим тавсияларни беради¹²⁶. Унинг фикрича, матн синтактик разметкасининг энг катта ахборот базасини ташкил этувчи изоҳлар гап қурилишига оид маълумотлар йиғиндисидир. Гап синтаксиси гапни қайси жиҳатдан ўрганса, теглаш жараёнида шу белгиларнинг барчасини камраб олиш ўринли. Чунки тузилган корпус кейинчалик синтаксис, бошқа

¹²⁶ Хамроева Ш.М. Ўзбек тили муаллифлик корпусини тузишнинг лингвистик асослари: Филол.фан.бўйича фалсафа доктори (PhD)...дис. – Бухоро, 2018. – 250 б.

бўлимлар билан боғлиқ турли тадқиқотлар олиб боришда маълумотлар базаси сифатида хизмат қилиши керак.

Ш.Ҳамроева гапнинг тузилишига кўра турларини фарқлаш учун теглар тизимини қуйидагича таклиф этади¹²⁷:

1. Гап билан боғлиқ энг биринчи тег гапнинг тузилишига кўра турини ифодаловчи изоҳдан иборат. Бу тег “содда гап” = <СГ>, </СГ>, “уюшган гап”= <УГ>, </УГ> ёки “қўшма гап” = <ҚГ>, </ҚГ> каби белгилардан иборат.

2. Гапнинг ифода мақсадига кўра турини изоҳлаш ҳам теглардан бирини ташкил этади: “дарак гап” = <дг>, “сўроқ гап”= <сг>, “буйруқ гап”= <бг>.

3. Гап лисоний қурилишида эганинг ифодаланиш ёки ифодаланмаслигига кўра “эғали гап” = <Е+>, “эғасиз гап” = <Е->; эғасиз гапларнинг “шахси номаълум гап” = <ш.н.г>, “атов гап” = <а.г>, “семантик-функционал шаклланган гап” = <с.фш.г> каби белгиларни қўшиш мумкин.

4. Бош, иккинчи даражали бўлакнинг иштирокига кўра “йиғиқ гап” = <йг>, “ёйиқ гап” = <ёг> каби изоҳ ҳам тег сирасидан ўрин олади.

5. Гап билан грамматик алоқага киришмайдиган бўлакларнинг мавжудлиги ҳам изоҳ талаб қилади. “Ундалма” = <у>, </у>, “киритма” = <к>, </к> каби белгилар мажмуидан иборат бўлади; жуфт қўлланадиган тег. Чунки бу бирликнинг чегарасини кўрсатиш лозим бўлади.

Шунингдек, тадқиқотчи теглар тизимини шакллантирар экан, синтактик теглар тизимини соддалаштириш ҳақида шундай ёзади: “Одатда, содда гап таҳлилида гап бўлаклари таҳлилига алоҳида эътибор қаратилади. Корпус лингвистикасида сўнгги корпусларнинг синтактик разметкаси гап бўлақларига оид изоҳ анъанаси йўқолди. Бу ҳар қандай корпус разметкасини соддалаштириш кераклиги нуқтаи назаридан келиб чиқди. Биз ҳам содда

¹²⁷Ҳамроева Ш., Менглиев Б. Тил корпусларини синтактик разметкашнинг лингвистик асослари // “Осиё мамлакатлари тамадуни ва ипак йўли”. Халқаро илмий анжуман материаллари. – Самарқанд, 2019 йил 21 август. – 236 б. – Б. 36-39.

гапга оид теглар қаторини юқоридаги изоҳлар сираси билан яқунлаймиз”¹²⁸.

Тадқиқотчининг синтактик теглар тизимини соддалаштириш хусусидаги фикрларини инобатга олган ҳолда, Ш.Хамроева томонидан ишлаб чиқилган теглар тизимини мукамаллаштириш бўйича қуйидаги таклифларни берамиз:

1. Грамматик марказлар сонига кўра гапнинг турларини аниқлаш учун <СГ>, </СГ> ҳамда <ҚГ>, </ҚГ> дан бири олинади. Содда гаплар синтактик теглар тизимини шакллантираётганимиз боис фақат <СГ>, </СГ> теги бизга етарли бўлади.

2. Гапнинг ифода мақсадига кўра турини фарқлаш учун “дарак гап” = <дг>, “сўроқ гап” = <сг>, “буйруқ гап” = <бг> тегларини киритамиз.

3. Эганинг иштирокига кўра турини аниқлаш учун “эғали гап” = <Е+>, “эғасиз гап” = <Е->; эғасиз гапларнинг “шахси номаълум гап” - <ш.н.г>, “атов гап” = <а.г>, “семантик-функционал шаклланган гап” = <с.фш.г> каби белгиларни қўшиш мумкин.

4. Бош, иккинчи даражали бўлакнинг иштирокига кўра “йиғиқ гап” = <йг>, “ёйиқ гап” = <ёг> каби изоҳ ҳам тег сирасидан ўрин олади.

5. Гап билан грамматик алоқага киришмайдиган бўлакларнинг мавжудлигига кўра “Ундалма” = <у>, </у>, “киритма” = <к>, </к> теглари киритилиши мумкин.

Боб бўйича қисқа хулосалар

1. Тил корпусларида синтактик бирликларни фарқлашнинг назарий асослари сифатида қуйидагиларни ажратиш мумкин:

биринчи, корпусда қидирувнинг содда кўринишлари сифатида лемма, сўзшакл, ибора ва коллокация фарқланса-да, аннотацияли корпусларда морфологик белги асосидаги қидирув етакчилик қилади. Синтактик ва семантик белгилар асосидаги қидирув нисбатан мураккаб жараён бўлганлиги

¹²⁸ Хамроева Ш.М. Ўзбек тили муаллифлик корпусини тузишнинг лингвистик асослари: Филол.фан.бўйича фалсафа доктори (PhD)...дис. – Бухоро, 2018. – 250 б.

сабабли ҳамма корпусларни ҳам синтактик/семантик теглаш имкони, воситаси мавжуд эмас;

иккинчи, ўзбек тили корпусларининг парсер дастури, унинг синтактик теглар тизимини ишлаб чиқишда замонавий автоматик таҳлил дастурлари ҳамда ўзбек тилшунослигида синтаксис бўйича мавжуд назарий материалларни ўрганиш, умумлаштириш ҳамда синтактик теглар яратишда ёндашувни тўғри танлаш муҳим;

учинчи, корпусда гап ёки бирикма эканлиги ташқи белгилари асосида фарқланмайдиган бирикувларни аниқлашнинг махсус алгоритми, фильтри ишлаб чиқилиши лозим. Бунда гап/сўз бирикмасини аниқлашнинг бир компонентли ёки икки компонентли эканлигига асосланиш натижа бера олади.

2. Синтактик алоқа сўз бирикмаси ва гапдаги сўзлар, нутқдаги гапларнинг ўзаро грамматик боғланишини билдиради. Синтактик муносабат эса сўз бирикмаси, гапдаги сўзларнинг ҳамда нутқдаги гапларнинг бири-бирига нисбатан қандай грамматик маъно ва функцияда қўлланишини билдиради. Синтактик теглашда синтактик алоқанинг тури инобатга олинади ҳамда теглар тизимидан синтактик алоқани ифодаладиган тег жой олади. Тил корпуси материалини синтактик теглашда синтактик алоқа ва синтактик муносабат фарқланиши лозим. Шаклан фарқлашнинг имкони йўқ: боғловчи воситалар бир хил. Шу сабабли корпус разметкасида синтактик алоқа ва синтактик муносабатни фарқловчи тег киритилиши талаб этилади.

3. Корпус синтактик разметка тизимида коллокатив қидирув жуда муҳим: қидирув натижасининг сифати қидирувнинг мукамаллиги билан тўғри пропорционал. Коллокатив қидирувни ташкил этиш учун корпус бирликлари сўз бирикмаси ва сўз қўшилмасини ажратувчи тегга эга бўлиши талаб қилинади.

4. Морфологик кўрсаткичсиз тобе-ҳокимлик муносабатини аниқлаш учун лисоний-синтактик қолип ҳамда морфологик теглар тизимидан фойдаланиш ўринли, чунки лисоний-синтактик қолипда қайси сўз туркуми

ўзаро бирикиши кўрсатиб берилган. Корпус бирликлари морфологик тегланган бўлса, тег ва лисоний-синтактик қолип асосида морфологик кўрсаткичсиз синтактик муносабатли бирикмаларни ҳам автоматик теглаш имкони пайдо бўлади.

5. Тил корпусида сўз бирикмаларни теглаш муаммоси умумий (инвариант) ЛСҚлар билан ҳал этилмайди: сўз бирикмаларни аниқлашда нисбатан аниқроқ қолиплар талаб этилади. Шундан келиб чиқиб, сўз бирикмаларни синтактик теглашнинг исм+исм қолипи учун 14 та модел ишлаб чиқилди.

6. Синтактик разметканинг икки тури мавжуд. Матн таркибидаги бирлик контекстга қараб икки хил гап бўлаги (ҳол ва тўлдирувчи) бўлиб келиши мумкин; бундай бирликларга икки хил тег бириктирилиши табиий. Корпусда билвосита белгилар асосида аниқланиши мумкин бўлган бирликлар тегланмаслиги ҳам мумкин: чунки бундай бирликлар тегишли тиниш белгилари орқали фарқланади. Шундан келиб чиқиб:

тиниш белгилари орқали фарқлаш имкони мавжуд бўлган гапга шундай ахборотдан иборат тег бириктирмаслик ҳам мумкин;

икки хил гап бўлаги бўлиши мумкин бўлган бирликларга иккилик тег бириктириш талаб этилади.

7. Синтактик теглар тизимини белгилашда интерфейснинг қулайлигини инобатга олиш зарурати мавжуд. Синтактик теглар тизимини тузишда матнга автоматик ишлов беришга эришиш, кўлда ишлов беришни имкон қадар камайтириш принципига амал қилиш лозим.

Ш БОБ. ЎЗБЕК ТИЛИДА ҚЎШМА ГАПЛАРНИ СИНТАКТИК ТЕГЛАШНИНГ НАЗАРИЙ АСОСЛАРИ

3.1. Ўзбек тилида боғланган ва боғловчисиз қўшма гапларни синтактик теглаш ва моделлаштириш

3.1.1. *Боғланган қўшма гапларни синтактик теглаш моделлари ва теглар тизими.* Маълумки, ҳар бир тилда синтактик бирликлар ўзига хос характерга эга. Шу сабабли ўзбек тилидаги қўшма гаплар учун ҳам синтактик теглар тизимини ишлаб чиқиш талаб этилади. Жаҳон компьютер лингвистикаси тажрибасида синтактик теглаш муаммолари ёритилган махсус тадқиқотлар мавжуд бўлмаса-да, айрим масалалар таҳлиliga бағишланган мақолалар мавжуд¹²⁹. Аммо катта ҳажмли корпуслар таркибида ҳам қўшма гапларнинг тегланмагани – қўшма гап турларига оид сўров/қидирувни амалга ошириш имконининг йўқлиги ҳали бу борада қилиниши керак бўлган, ечимини кутаётган масалаларнинг кўплигини кўрсатади. Ш.Хамроева қўшма гапларнинг умумий теглари ҳақида қуйидагиларни ёзади¹³⁰: “Қўшма гапнинг боғловчи воситасига кўра “боғланган ҚГ” = <БҚГ>, </БҚГ>, “эргашган ҚГ” = <ЭҚГ >, </ЭҚГ>, “боғловчисиз ҚГ” = <Б-сиз ҚГ>, </Б-сиз ҚГ> турларини изоҳлаш бундай бирликларни ажратиш ҳамда изоҳлаш учун алоҳида белги вазифасини ўтайди. **Теглаш кўриниши:**

1) <Б-сиз ҚГ>*Аллақандай бир хотин келиб толнинг хипчини билан савалади, товуқ сўйиб қонлади...</Б-сиз ҚГ>.*

¹²⁹Апресян Ю. Д., Богуславский И. М., Иомдин Б. Л. и др. Синтаксически и семантически аннотированный корпус русского языка: современное состояние и перспективы // Национальный корпус русского языка: 2003-2005. М.:Индрик, 2005, 193-214.;Зобнин А.И., Сахарова А.В. Универсальная система разметки текста ObjectATE // Национальный корпус русского языка: 2006-2008. Новые результаты и перспективы. СПб.: Нестор-История, 2009, 283-297.; Поляков А.Е. Технология подготовки информации в Национальном корпусе русского языка // Национальный корпус русского языка: 2003-2005. Результаты и перспективы. – М., 2005, 175-192.; Сичинава Д. В. Обработка текстов с грамматической разметкой: инструкция разметчика // Национальный корпус русского языка: 2003-2005. Результаты и перспективы. – М., 2005, 136-154.; Lashevskaja O. Corpus-aided Construction Grammar: Semantic Tools in the Russian National Corpus // Cognitive-linguistic approaches: What can we gain by computational treatment of data? A theme session at the Second International Conference of the German Cognitive Linguistics Association, 7 October 2006, Munich. Electronic publication /http://www1.uni-hamburg.de/metaphern/GCLA06_Theme/Lashevskaja_GCLA06.pdf

¹³⁰Хамроева Ш.М. Ўзбек тили муаллифлик корпусини тузишнинг лингвистик асослари: Филол. фан. бўйича фалсафа доктори (PhD) дис. – Бухоро, 2018. – 250 б. – Б. 155.

2) <БҚГ>Сотиболди хўжайинининг олдига арзга борди, аммо бу боршдан муддаоси нима эканини аниқ билмас эди.</БҚГ>.

3) <ЭҚГ >, </ЭҚГ> .

Шунингдек, Ш.Ҳамроева таркибидаги содда гапларнинг ўзаро мазмун-муносабати асосида боғланган ҚГларнинг бир неча турга бўлинишини таъкидлаган ҳолда теглар тизимини қуйидагича шакллантиради¹³¹:

1) бириктирув муносабатли БҚГ = <б.БҚГ>;

2) зидлов муносабатли БҚГ = <з.БҚГ>;

3) айирув муносабатли БҚГ = <а.БҚГ>;

4) инкор муносабатли БҚГ = <и.БҚГ>.

Қўшма гапларнинг мукамал теглар тизимини ишлаб чиқиш учун ўзбек тилшунослигида ҳозирга қадар тўпланган назарий материал ва таснифларга таяниш мақсадга мувофиқ. Қуйида ўзбек тилида қўшма гап ҳақидаги айрим маълумотларни келтирамиз ва теглар тизими ҳақида хулоса қиламиз.

Бирдан ортиқ содда гапнинг мазмун, грамматик ва оҳанг жиҳатидан бирикувидан тузилган гап қўшма гап дейилади: *Эшик секин очилди-ю* (биринчи содда гап), *Қаландаровнинг юзи кўринди* (иккинчи содда гап)¹³². Содда гап таркибида бир грамматик марказ (кесим бирлиги) иштирок этса, қўшма гап таркибида бирдан ортиқ грамматик марказ қатнашади. Қўшма гап таркибидаги содда гап ўзаро боғловчи, юклама, кўмакчи, феъл шакли ҳамда оҳанг орқали боғланади. Шунга кўра, қўшма гапнинг 3 тури: боғланган қўшма гап; эргашган қўшма гап; боғловчисиз қўшма гап фарқланади. Юқорида қўшма гапнинг бу уч тури Ш.Ҳамроева таснифи асосида теглар сирасига қўшилган.

Айрим дарсликда қўшма гап боғловчи воситасига кўра қуйидаги турларга ажратилади¹³³:

¹³¹Ҳамроева Ш.М. Ўзбек тили муаллифлик корпусини тузишнинг лингвистик асослари: Филол. фан. бўйича фалсафа доктори (PhD) дис. – Бухоро, 2018. – 250 б. – Б. 155.

¹³²Ўзбек тили грамматикаси. II том. Синтаксис. [Масъул муҳаррирлар: Ғ. А.Абдурахмонов ва бошқ.]. – Т., «Фан», 1976. – 560 б. – Б. 300.

¹³³Замонавий ўзбек тили: Синтаксис./ Муаллифлар жамоаси. Масъул муҳаррирлар Ҳ.Ғ.Неъматов, Р.Сайфуллаева. ЎЗР Олий ва ўрта махсус таълим вазирлиги, Мирзо

1. Боғловчилар: тенг боғловчилар (*ва, ҳамда, аммо, лекин, бироқ, ё, ё... ё, дам..., дам...*); эргаштирувчи боғловчилар;

2. Боғловчи вазифасидаги воситалар: *бўлса, эса, деб; -у, (-ю), -да, -ки* юкламалари; нисбий сўзлар; *шунинг учун, шу туфайли, шу сабабли* сингари кўмакчили курилмалар; шарт майли қўшимчаси (*-са*);

3. Оҳанг.

Юқорида кўрсатилган воситалар қўшма гап турини автоматик аниқлашнинг бирламчи воситаси саналади, чунки синтактик теглаш дастури (парсер) содда гаплар чегарасини улар орасидаги тиниш белги ҳамда боғловчи восита орқали аниқлайди. Боғловчи восита оҳанг бўлганда, ёзувда тиниш белги билан ифодаланади, натижада, содда гаплар чегараси топилади.

3.1.2. *Боғланган қўшма гапларни синтактик теглаш.* Қисмлари тенг боғловчилар, *бўлса, эса* сўзлари, *-у (-ю), -да* юкламалари ёрдамида боғланган гаплар боғланган қўшма гаплар ҳисобланади: *Кечаси қалин қор ёғди, лекин ҳаво совимади.* Таркибидаги содда гапнинг ўзаро мазмун муносабатига кўра боғланган қўшма гап қуйидаги турга бўлинади¹³⁴:

- 1) бириктирув боғловчиси ёрдамида БҚГ;
- 2) зидлов боғловчилари ёрдамида БҚГ;
- 3) айирув боғловчилари ёрдамида БҚГ;
- 4) бўлса, эса сўзлари ёрдамида БҚГ;
- 5) инкор муносабатли боғланган БҚГ.

Демак, ўзбек тилида боғланган қўшма гапларнинг синтактик теглар тизимида дастлаб <b.BQG>, <z.BQG>, <a.BQG>, <yo.s.BQG>, <i.BQG> каби белгилар киритилади.

Бириктирув боғловчиси ёрдамида боғланган қўшма гапларда бириктирув боғловчиси қўшма гапнинг пайт, сабаб-натижа муносабатларида бўлган

Улуғбек номидаги Ўзбекистон Миллий университети. – Тошкент.: Мумтоз сўз, 2011. – 312 б. – 106-109.

¹³⁴Замонавий ўзбек тили: Синтаксис./ Муаллифлар жамоаси. Масъул муҳаррирлар Ҳ.Ғ.Неъматов, Р.Сайфуллаева. ЎзР Олий ва ўрта махсус таълим вазирлиги, Мирзо Улуғбек номидаги Ўзбекистон Миллий университети. – Тошкент.: Мумтоз сўз, 2011. – 312 б. – Б. 226-235.

қисмларини боғлаб келади. Қисмлари ўртасида сабаб-натижа муносабати ифодаланган қўшма гапларда қисм олдида натижада сўзини қўйиш мумкин бўлади. -у, -ю, -да юкламалари ҳам бириктирув боғловчилари бажарган вазибаларни бажаради: *Машгулот тугади ва ҳамма ўз уйига тарқалди. Баҳор келди ва ўз сепини ёя бошлади. Одам қўли тегди-ю, ташландиқ ерларни обод қилди. Чорпоянинг бериги ёнида қариган Абдурахмонбой, унинг ёнида Абдуҳаким ва Мулласобит номли икки катта ўғли ўтирар ва буларнинг рўпараларида, ...туманнинг бойларидан бир неча киши ўтирмоқда эди (С.Айний). Мажлиснинг тўрида татарча нўстин, татарча телпак, татарча кийгиз этик кийган қорни катта бир татар бойи ерлашган ва унинг ёнида салласини катта қилиб қишлоқ имоми қўнқайиб ўтирмоқда эди (С.Айний). Буларнинг бири қишлоғимизнинг хатиби эди ва унинг ёнида ўрта бўйли қоқ гавдали, ориқ туси қорамтил, оқарган ўртача соқолли иккинчи одам борар эди (С.Айний). Хурясун устига кўрпача тўшаб бобом (эшакка) минди ва мени ҳам унинг орқасига мингаштирдилар, отам ва бошқалар пиёда юриб боғ томон йўналдик (С.Айний). Гапни юклама боғласа, ундан кейин вергул қўйилади¹³⁵. Чироқ ўннга қайрилиб зойиб бўлди-ю, ҳаял ўтмай яқинроқдан кўзни қамаштирадиган ёғду сочиб чиқди ва босинқи мотор товуши эшитилди (А.Қаҳҳор).*

Зидлов боғловчилари боғланган қўшма гап қисмларини боғлаш билан бирга улар ўртасида зидлик муносабати мавжудлигини ҳам ифодалайди. Зидлов боғловчилари қўшма гапнинг иккинчи қисми бошида қўлланади ҳамда қўшма гап қисмлари вергул билан ажратилади. -у(-ю), -да юкламалари боғланган қўшма гап қисмларини боғлаш билан бирга улар ўртасида зидлаш, кетма-кетлик муносабати мавжудлигини билдиради: *Юртимнинг бу куни чиройли, лекин эртаси, индини яна чиройлироқ, бахтлироқ бўлади. Ҳаво очилди-ю, ҳарорат сезилмади. Яна ўйлашиб кўрамиз, ҳар ҳолда сўзингни ерда қолдирмасман, лекин ундай ножўя сўздан қайт (Ойбек). Қоринлар тўлди, товоқлар бўшади,*

¹³⁵ Ўзбек тили грамматикаси. II том. Синтаксис. [Масъул муҳаррирлар: Ғ. А. Абдурахмонов ва бошқ.]. – Т., «Фан», 1976. – 560 б. – Б.412.

*лекин сўфи «савоб» бўлади деб, ҳали ҳам дастурхонга тўкилган гурунчларни териб, оззига солмоқда эди (С.Айний). Фатҳиллабой ҳовлисига етди, ҳовли дарвозаси очиқ, аммо у дарвоза олдида бир қизил аскар соқчилик қилиб турар эди (С.Айний)*¹³⁶.

Баъзан зидлик мазмунини кучайтириш учун зидлов боғловчиси ҳамда-у(-ю) юкламаси бирга ишлатилади: *Кечаси қор ёғди-ю, лекин кундузи дарров тўхтади*. Бундай қўшма гапда зидлов боғловчисидан олдин, юкламадан эса кейин вергул қўйилади.

Кўриниб турибдики, зидлов ва бириктирув муносабатли боғловчилар боғлаган қўшма гапларни боғлашда -у(-ю) юкламалари бирдек қўлланади. Бундай пайтда боғловчи восита орқали қўшма гап турини аниқлаш моделлари грамматик омонимия ҳодисасига дуч келади. Шу сабабли бундай ҳолларда қўшма гапларни аниқлашнинг моделларини қуйидагича шакллантириш лозим.

1. Агар [Wpm] + ва/ ҳам/ -у/ -ю/ -да+[Wpm] бўлса = <a.BQG>. Бунда [Wpm] = содда гап.

2. Агар [Wpm] + аммо/ лекин/ бироқ/ -у/ -ю/ -да + [Wpm] бўлса = <z.BQG>. Бунда [Wpm] = содда гап.

Бу икки моделда ҳам омонимлик юзага келади, натижада автоматик синтактик таҳлил -у, -ю, -да юкламалари билан боғланган қўшма гапларни фарқлай олмайди. Бунинг учун фойдаланувчи автоматик таҳлил натижасини бирма-бир таҳлил қилиши керак. Келажакда ўзбек тилининг семантик таҳлил моделлари яратилса, шаклий ноаниқлик мазмуний моделлар асосида аниқланишига эришилади.

Айирув боғловчилари қўшма гап қисмларини боғлаш билан бирга уларда ифодаланган воқеа-ҳодисаларнинг галма-гал рўй беришини ёки уларнинг фақат биттасининг юзага чиқишини таъкидлайди¹³⁷. *Ё (ёки)* боғловчиси ёлғиз ҳолда

¹³⁶Замонавий ўзбек тили: Синтаксис./ Муаллифлар жамоаси. Масъул муҳаррирлар Ҳ.Ғ.Неъматов, Р.Сайфуллаева. ЎзР Олий ва ўрта махсус таълим вазирлиги, Мирзо Улуғбек номидаги Ўзбекистон Миллий университети. – Тошкент.: Мумтоз сўз, 2011. – 312 б. – 289-291.

¹³⁷Ўзбек тили грамматикаси. II том. Синтаксис. [Масъул муҳаррирлар: Ғ. А.Абдурахмонов ва бошқ.]. – Т., «Фан», 1976. – 560 б. – Б.412.

хам, такрор ҳолда ҳам қўлланилиши мумкин. Қолган айирув боғловчилари такрор қўлланилади ва ёзувда такрор қўлланаётган боғловчидан олдин вергул қўйилади: *Гоҳ осмонни тутиб ашула янграр, гоҳ аллақадан гармон товуши эшитилиб қолар эди.* Айирув боғловчиси, одатда, гапнинг бошида келади. Такрорланиб қўлланган айирув боғловчисининг кейингисидан олдин вергул қўйилади, бундай боғловчи яқка қўлланса, тиниш белгиси ишлатилмайди. Тиниш белгиларининг бундай қўлланиши қўшма гап таркибидаги содда гапни аниқлашда кийинчилик туғдиради. Шу сабабли юқорида кўрсатилган моделлардан фарқли ўлароқ, айирув боғловчилари билан боғланган қўшма гаплар чегараси тиниш белгилари ва боғловчилар билан эмас, фақат боғловчи воситалар ёрдамида фарқланади. Мисолларни таҳлил қиламиз. *Катта ҳовузлардаги сувларда кўланкалар ва шуълалар жим-жим ўйнашади, гоҳ кўланкалар ёниб кетади, гоҳ шуълаларни кўланкалар ютади* (Ойбек). *Ё подшо бирор ерга чиқади, ёки элчи келиши кутилади* (Ойбек). Мисоллардан кўриниб турибдики, айирув боғловчиси такрор келганда қўшма гап таркибидаги содда гаплар тиниш белги билан ажратилади. Модел ва тегларини қуйидагича белгилаш мумкин бўлади:

1. [Гоҳ...Wpm] + [гоҳ...Wpm] бўлса = <a.BQG>.
2. [Ё/ёки...Wpm] + [ёки...Wpm] бўлса = <a.BQG>.

Бўлса, эса сўзлари боғланган қўшма гап қисмларини боғлаш билан бирга, улар ўртасида қиёслаш ва зидлаш муносабати мавжудлигини ифодалайди. *Бўлса, эса* сўзлари қиёсланувчи бўлакдан сўнг келади: *Режадаги зарур ишларни амалга оширишга вақт етказа олмаяпман, сен бўлса фақат дам олайлик дейсан («Гулх.») Ким бўлса бўлсин, мен эса архитектор бўламан. («Фун.»)*¹³⁸. Кўриниб турибдики, бундай гаплардаги боғловчи воситалар [.....са +бўлса/эса] шаклида жойлашади. Синтактик

¹³⁸Замонавий ўзбек тили: Синтаксис./ Муаллифлар жамоаси. Масъул муҳаррирлар Ҳ.Ғ.Неъматов, Р.Сайфуллаева. ЎзР Олий ва ўрта махсус таълим вазирлиги, Мирзо Улуғбек номидаги Ўзбекистон Миллий университети. – Тошкент.: Мумтоз сўз, 2011. – 312 б. – Б. 256-264.

моделли ва теги куйидагича белгилаш мумкин: [Wpm...ca] + [бўлса/эса... Wpm] = <yo.s.BQG>.

Инкор муносабатли боғланган қўшма гап қисми ўзаро такрор қўлланган *на* инкор юкламаси орқали боғланади; гап орасига вергул қўйилади: *На сув бор, на бирон емиш қолибди*. Инкор боғловчиси ҳам, одатда, содда гапнинг бошида келади¹³⁹. Модел ва теги куйидагича бўлади: [Ha...Wpm]+ [на ... Wpm] = <i.BQG>.

Боғланган қўшма гап қисмларининг ўзаро алоқаси орқали куйидаги мазмун-муносабатлари ифодаланади¹⁴⁰:

1. Қиёслаш муносабати.
2. Бириктирув муносабати.
3. Айирув муносабати.
4. Сабаб ва натижа муносабати.
5. Изоҳлаш муносабати.

Афсуски, корпуснинг синтактик белгилар асосидаги қидируви ёки синтактик теглар тизимига бундай муносабатни кўрсатувчи параметрни кирита олмаймиз.

3.1.2. Боғловчисиз қўшма гапларнинг синтактик моделлари ва теглар тизими. Таркибида содда гаплар мазмунан бир-бири билан боғлиқ бўлиб, махсус боғловчи воситаларсиз, фақат оҳанг ёрдамида боғланган қўшма гаплар боғловчисиз қўшма гап саналади¹⁴¹. Шунингдек, бошқа манбаларда икки ва ундан ортиқ содда гапларнинг ўзаро грамматик боғловчилар ёрдамисиз бирикувидан тузилган қўшма гаплар боғловчисиз қўшма гаплар бўлиши таъкидланади: *Яхши отга бир қамчи, ёмон отга минг қамчи (Мақол). Юрагида қандай дард бор –мана бу менга сир (Ойбек)*¹⁴².

¹³⁹Замонавий ўзбек тили: Синтаксис./ Муаллифлар жамоаси. Масъул муҳаррирлар Х.Ғ.Неъматов, Р.Сайфуллаева. ЎзР Олий ва ўрта махсус таълим вазирлиги, Мирзо Улуғбек номидаги Ўзбекистон Миллий университети. – Тошкент.: Мумтоз сўз, 2011. – 312 б. – Б. 256-264.

¹⁴⁰Кўрсатилган манба. Б. 264.

¹⁴¹Кўрсатилган манба. Б. 266.

¹⁴²Ўзбек тили грамматикаси. II том. Синтаксис. [Масъул муҳаррирлар: Ғ. А.Абдурахмонов ва бошқ.]. – Т., «Фан», 1976. – 560 б. – Б. 467.

Боғловчисиз кўшма гапларни ташкил этган содда гаплар ўзаро тобеланиш ҳолати ёки тенг боғловчисиз бирикади. Аммо бундай содда гапларнинг структурасида лексик, морфологик, синтактик параллеллик бўлади. Демак, боғловчисиз кўшма гапни ташкил этган қисмлар мазмунан бирикиб, маълум воқеа-ходисалар ўртасидаги мазмун-муносабатини ифодалаб, умумий бир мақсадни англатибгина қолмай, бу қисмларнинг гап бўлакларида ва грамматик структурасида маълум бир параллелизм бўлади. Бу ҳол қисмларни бир-бирига жипс боғлаб, маълум кўшма гап турининг тузилишига сабаб бўлади. Боғловчисиз кўшма гапларнинг тузилишидаги асосий воситалардан бири интонация. Интонация боғловчисиз кўшма гапнинг тузилишида ҳамда бу хил гапларда турлича маъно ва оттенкаларнинг ифодаланишида асосий фактор бўлиб қолади¹⁴³.

1) агар боғловчисиз кўшма гап қисми бир пайтда ёки кетма-кет юз берган воқеа-ходисани ифодаласа, вергул қўйилади: *Арава гижирлаб борар, аравакаш эса хиргойисини бир зум ҳам тўхтатмас эди. Булар роҳатда, мен ҳасратда яшабман (Ойбек). Яхши иш битирар, ёмон иш йитирар (Мақол). Унинг руҳи дўстга қуёшдай илиқ, душманга даҳшат (Ойбек).*

2) агар боғловчисиз кўшма гап қисми мазмунан бир-биридан анча узоқ бўлса, нуқтали вергул қўйилади: *Ёмғир челақдан қуйгандай шариллаб ёғарди; унинг хаёли узоқ, ёшлик пайтларида кезмоқда;*

3) боғловчисиз кўшма гапнинг иккинчи қисми биринчи қисмнинг сабабини билдирса, изоҳласа, тўлдирса, улар орасига икки нуқта қўйилади: *Гап шу: эртага мен билан йўлга чиқасиз. Ёз охирида Самарқандда кечалари салқин, кундузлари иссиқ бўлади: туш пайтида шабада тоғдан салқин ва роҳатбахш ҳаво келтирмайди (Ойбек). Ҳамма айб ўғлимда: гўл, бўшанг.*

4) агар боғловчисиз кўшма гап қисми ўзаро зидлик ёки ўхшашлик, шарт муносабатига киришган бўлса, тире қўйилади: *Қор ёғди – дон ёғди. Ёшим етмиш иккида – ўзим йигитман. Мақтанчоқ бўлманг – ҳижолат тортмайсиз (Мақол).*

¹⁴³Ўзбек тили грамматикаси. II том. Синтаксис. [Масъул муҳаррирлар: Ғ. А. Абдурахмонов ва бошқ.]. – Т., «Фан», 1976. – 560 б. – Б. 300.

Давлат тинч –халқ тинч (Ойбек). Юрт бетинч –сен бетинч (Мақол). Дангаса бўлманг –бахтсиз бўлмайсиз (Мақол). Ўзи юзсиз –сўзи тузсиз (Мақол). Кўшининг тинч –сен тинч (Мақол). Ишининг кўзини бил –дала ҳам обод, чойхона ҳам... (Ойбек). Кўзим тирик –мен билан муомала қиласиз. (Ойбек). Ўлимга ҳали вақт бор. Тириксан –кўрганнинг ганимат(Ойбек).

Боғловчисиз қўшма гапларни (оҳанг) тиниш белги боғлашини инобатга олсак, уларнинг модел ва тегларини қуйидаги тартибда белгилаш мумкин бўлади:

1. [Wpm] <,> [Wpm] бўлса = <Б-сиз ҚГ>.
2. [Wpm] <:> [Wpm] бўлса = <Б-сиз ҚГ>.
3. [Wpm] <-> [Wpm] бўлса = <Б-сиз ҚГ>.
4. [Wpm] <;> [Wpm] бўлса = <Б-сиз ҚГ>.

Ажратилган кесимли содда гаплар борки, улар [Wpm] <-> [Wpm] қолипига тўғри келадиган шаклда бўлади, яъни биринчи содда гап <->кесим. Бундай ажратилган кесимлар дастур томонидан йиғиқ гап сифатида аниқланади, натижада грамматик/ структур омонимлик келиб чиқади. Масалан: *Каттагина, чирик тўнгатк орқасида мукка тушган гавдага кўзи тушди –югурди (Ойбек)...пайқамай қолди: бакларга тегдим, моторигами ёки бомбалар солинган кассеткаларгами –билмади (Ойбек).* Ҳозирча бундай тузилишли қўшма гапларни теглаш ва моделлаштириш орқали автоматик аниқлашнинг имкони йўқ. Келгусида ўзбек тилида онтологик моделларнинг мукамаллашиши структур/грамматик кўпмаънолилик ва омонимияни аниқлашга ёрдам бериши мумкин.

Боғловчисиз қўшма гап қисмларининг тузилиши. Боғловчисиз қўшма гап конструкциясини ташкил этган қисмлар турли шаклда бўлиши мумкин. Кўпинча, улар дарак гап шаклида бўлади. Қисмларнинг бири ёки ҳар иккиси сўроқ шаклида бўлиши ҳам мумкин. Аммо бундай гаплар шаклан сўроқ, мазмунан риторик сўроқ гап бўлиб, қўшма гапнинг бошқа қисми уни

изохлайди; сўроқ гап асосан биринчи ўринда келади¹⁴⁴: *Чекинсинми –йўқ, гурури йўл бермади (Ойбек). Бўзами, арақми, мусалласми, коньякми – қўймайди эшон (Ойбек). Миянг борми ўзи, Сидиқжон колхозга киргани билан, ер колхозга ўтиб кетмайди, ўртада бўлади! (А.Қаҳҳор). Мактаб боғи қанақа бўлади –ҳозир бутун гап шу ҳақда (Космодемьянская). Бека синглисига жангчини таништирди –қаердан, ким у, қандай қабул қилди уни ва ҳоказо (Ойбек). «Салом жангчи ўртоқ! Мен билмайман, сиз ким, русми, украинми, ўзбекми, қозоқми, қирғизми...» (Ойбек). Бу гапда биринчи содда гап сўроқ мазмунли бўлса ҳам, сўроқ эмас, балки ҚГ мазмуний муносабатига боғлиқ равишда тире қўлланилган¹⁴⁵.*

Қўшма гапнинг иккинчи қисми соф сўроқ гап бўлиб шаклланиши мумкин: *Нима қиласан, яна қишлоққа қайтасанми? (Ойбек)*. Боғловчисиз қўшма гапни ташкил этган риторик сўроқ гапнинг айрим турлари барқарор фразеологик гап тарзида бўлади: *Дунёда нима кўп –кулки кўп. Нима кўп –иш кўп (Ойбек). ...ғариблар қорнини тўйдирган ким –Холназарбой (Б.Кербобоев)*. Бундай ҳолатда ҳам ҚГлар орасидаги тиниш белгилар қоида асосида бўлади, истисно кузатилмайди.

Боғловчисиз қўшма гапнинг муҳим белгиларидан бири шуки, унинг қисмлари содда гапнинг турли тип ва кўринишларидан ташкил топиши мумкин. Бунинг сабаби қўшма гап қисмларининг жипс муносабати, бир қисмнинг мазмуни ва таркибидаги бўлак маъносидан иккинчи қисмнинг таркиби аён бўлиб туришида. Боғловчисиз қўшма гаплар жонли нутқни акс эттиргани учун ҳам структураси анча эркин тузилади¹⁴⁶.

¹⁴⁴Ўзбек тили грамматикаси. II том. Синтаксис. [Масъул муҳаррирлар: Ғ. А. Абдурахмонов ва бошқ.]. – Т., «Фан», 1976. – 560 б. – Б. 330.

¹⁴⁵Ўзбек тили грамматикаси. II том. Синтаксис. [Масъул муҳаррирлар: Ғ. А. Абдурахмонов ва бошқ.]. – Т., «Фан», 1976. – 560 б.

¹⁴⁶Ўзбек тили грамматикаси. II том. Синтаксис. [Масъул муҳаррирлар: Ғ. А. Абдурахмонов ва бошқ.]. – Т., «Фан», 1976. – 560 б.

1. Боғловчисиз қўшма гап қисмларидан бири ёки ҳар иккиси бир составли гап шаклида бўлиши мумкин. Бу хил гапларнинг ўринлашишида ҳам эркинлик кўринади:

а) шахссиз гаплар: *Эшакни яйдоқ миниб бўлмайди –йиқитади кишини* (А.Қахҳор). *Азиз дўстим, сени жуда кўргим келади, согиндим* (Ш.Рашидов).

б) шахси номаълум гаплар: *Ётган ҳўкизга ем йўқ, дейдилар.*

в) шахси умумлашган: *Мингларча чақирим шимолга юр, мингларча чақирим жанубга юр –одам салтанатидасан...* (Ойбек).

г) атов гаплар: *Камол мудроқ босган кўзларини ишқалаб, боши қуйи, хаёл сурди: яна Салима* (Ойбек). *Мана, ўрмон ҳиди –бундан ёқимли нарса борми! Бу ерда ўлиш ҳам –бахт* (Ойбек). *Ё ўлим, ё ҳаёт бошқа чора йўқ!* (Ойбек). *Ҳовли-жой, рўзгор, машина, портрет... нима қилса-ю, мени қанотимдан уйга михлаб қўйса!* (А.Қахҳор). *Ия дедим, секин бориб қулоқ солдим. Аёл овози... галати гаплар. Турқичдан мўраладим: Йўлчи билан етук бир қиз* (Ойбек).

Бундай мазмуний муносабатдаги гапларни кузатувимиз шуни кўрсатяптики, айрим ҳолларда ноаниқлик келиб чиқади. Масалан, *Ё ўлим, ё ҳаёт бошқа чора йўқ!* гапида уюшиқ бўлақлар орасига кўйиладиган ҳамда боғловчисиз қўшма гаплар орасидаги тиниш белгилар омонимлиги юзага чиқади. Бундай пайтда натижага чиққан мисоллар бирма-бир сараланади.

Боғловчисиз қўшма гапнинг қисмларидан бири ёки ҳар иккиси, кўпинча, тўлиқсиз гап шаклида бўлади¹⁴⁷. Қисмлардан бирининг таркибида эгадан ташкил топган тўлиқсиз гап бўлади (ёлғиз эгадан ташкил топган тўлиқсиз гап боғловчисиз қўшма гап таркибида кам қўлланади): *Мен гоҳ ҳазиллашсам, ёнидан кичкина суратни чиқаради –учувчи кийимида бир йигит* (Ойбек). *Бектемир ўлим даласида боради –кўзларида газаб яшини, оқарган юзида пўлатнинг қаҳри ва қаттиқлиги* (Ойбек). *Эртасига ҳамма одатдагича далага чиқди – Тоштемир отанинг ўзи ҳам* (И.Ўктамов). *Бугдой устида бир*

¹⁴⁷Ўзбек тили грамматикаси. II том. Синтаксис. [Масъул муҳаррирлар: Ғ. А.Абдурахмонов ва бошқ.]. – Т., «Фан», 1976. – 560 б. – Б. 331.

қиз, кўзини уйқу босган (Чехов). Энг муҳими –ҳар қанча қийин бўлса ҳам қутқармоқ керак.

Бундай бирликларни фарқлаш моделларини ишлаб чиқиш махсус тадқиқотни талаб қилади: ажратилган бўлакли, уюшиқ бўлакли ҳамда мураккаблашган содда гапларда тиниш белгилар, уларнинг моделлари ўрганилса, боғловчисиз қўшма гаплардаги мана шундай истисно ҳолатга аниқлик киритилади.

Боғловчисиз қўшма гаплар уч типга бўлинади:

- 1) пайт муносабатини ифодалаган боғловчисиз қўшма гаплар;
- 2) киёслаш муносабатини ифодалаган боғловчисиз қўшма гаплар;
- 3) изоҳлаш муносабатини ифодалаган боғловчисиз қўшма гаплар¹⁴⁸.

Боғланган қўшма гаплардаги бундай мазмуний муносабатни ифодаловчи тегларни ишлаб чиқиш ҳозирча имконсиз. Юқорида айтилганидек, шаклий восита (боғловчи восита ва тиниш белги) ёрдамида боғланган қўшма гаплардан фарқли равишда, мазмуний муносабатларни ифодаловчи бирликларнинг теглари тизимини ишлаб чиқиш учун гапларни автоматик рефератлаш, автоматик аннотация ҳосил қилиш ёки WordNet тизими базаси каби дастурий воситаларни талаб этади.

3.2.Ўзбек тилида эргашган қўшма гапларни синтактик теглаш ва моделлаштириш

3.2.1.Ўзбек тилида эргашган қўшма гапларнинг боғловчи воситаларига кўра турларининг лингвистик теглар тизими. Синтактик теглаш масаласи, аввал айтганимиздек, жаҳон компьютер лингвистикасида жуда кам ўрганилган муаммо. Шунингдек, қўшма гапларнинг синтактик теглари ҳам ишлаб чиқилмаган, чунки бошқа тилларда автоматик парсерлар ишга туширилган. Ўзбек компьютер лингвистикасида йўл-йўлакай қилинган айрим қайдларни учратиш мумкин. Ш.Ҳамроева қўшма гап қисмлари орасидаги

¹⁴⁸ Ўзбек тили грамматикаси. II том. Синтаксис. [Масъул муҳаррирлар: Ғ. А.Абдурахмонов ва бошқ.]. – Т., «Фан», 1976. – 560 б. – Б. 461.

функционал муносабатга кўра турларини ажратар экан, разметкада куйидаги белги(тег)ларни таклиф қилади¹⁴⁹:

- 1) кесим тобе гапли ҚГ = <КҚГ>;
- 2) эга тобе гапли ҚГ = <ЭҚГ>;
- 3) ҳол тобе гапли ҚГ = <ХҚГ>;
- 4) тўлдирувчи тобе гапли ҚГ = <ТҚГ>;
- 5) аниқловчи тобе гапли ҚГ = <АҚГ>.

Таъкидлаш жоизки, эргашган кўшма гапларнинг тег тизимини ишлаб чиқишда ўзбек тилшунослигида эргаш гапларнинг таснифлари ҳақидаги хулосалар асос бўлади.

Қисмлари эргаштирувчи боғловчилар ёки шундай боғловчи вазифасидаги сўзлар ёрдамида боғланган кўшма гап эргашган кўшма гап ҳисобланади: *Илдиз озиқ берса, новда кўкарар*. Маълумки, эргашган кўшма гап, таркибида нечта гап бўлишидан қатъи назар, доимо икки қисмдан ташкил топади: 1) бош гап; 2) эргаш гап. Мазмуни изоҳланаётган гап *бош гап*, бош гап мазмунини изоҳлаётган гап *эргаш гап* ҳисобланади¹⁵⁰. Эргаш гапли кўшма гап таркибидаги гапни ифодаловчи тегни куйидагича белгилаймиз: бош гап=[bG]; эргаш гап =[eG].

Ишнинг олдинги қисмларида айтилганидек, кўшма гап синтактик тег ва моделларини ишлаб чиқиш учун асосий шаклий воситалар – кўшма гап таркибидаги содда гапларни боғловчи воситалар муҳим рол ўйнайди. Шу сабабли эргаш гапни боғловчи воситалар уларнинг турини автоматик аниқлаш учун асосий восита вазифини бажаради. Ўзбек тилида эргаш гап бош гапга куйидаги воситалар ёрдамида боғланади¹⁵¹:

¹⁴⁹Хамроева Ш., Менглиев Б. Тил корпусларини синтактик разметкалашнинг лингвистик асослари // “Осиё мамлакатлари тамадуни ва ипак йўли”. Халқаро илмий анжуман материаллари. – Самарқанд, 2019 йил 21 август. – 236 б. – Б. 36-39.

¹⁵⁰Ўзбек тили грамматикаси. II том. Синтаксис. [Масъул муҳаррирлар: Ғ. А. Абдурахмонов ва бошқ.]. – Т., «Фан», 1976. – 560 б. – Б. 462.

¹⁵¹ Ўзбек тили грамматикаси. II том. Синтаксис. [Масъул муҳаррирлар: Ғ. А. Абдурахмонов ва бошқ.]. – Т., «Фан», 1976. – 560 б. – Б. 466.

- 1) эргаштирувчи боғловчи: *Биз китобни севадим, чунки у билим манбаи.Юртимиз тинч, шунинг учун кўнглимиз хотиржам.*
- 2) кўмакчи қурилма: *шунинг учун, шу сабабли, шу туфайли;*
- 3) юкламалар (-ки):*Мақсадим шуки, ватанимга содиқ фарзанд бўлиш.*
- 4) нисбий сўзлар (*ким... ўша, нима... шу, қаер... шу ер*);
- 5) *деб* сўзи ва бошқалар.

Бош гапда ифодаланган мазмуннинг сабаби, мақсади, шарти ёки қиёсини билдирган эргаш гаплар бош гапга сабаб (*чунки, негаки*), мақсад (*токи*), шарт (*агар, агарда, башарти, мабодо*), қиёслаш-чоғиштирув (*гўё, гўёки,худди*) эргаштирувчи боғловчилари ёрдамида боғланади. Сабаб, мақсад, қиёслаш-чоғиштирув боғловчилари эргаш гап бошида қўлланади, эргаш гап бош гапдан сўнг келади. Шундай вақтда ёзувда эргаш гап бош гапдан вергул билан ажратилади. Шарт боғловчилари гап бошида келади. Эргаш гап кесими шарт майлидаги феъллар билан ифодаланади. Бундай эргаш гаплар бош гапдан олдин келади. Кўпинча, шарт боғловчилари қўлланмаса ҳам, шарт майли шакли қўшма гап қисмларини боғлайверади¹⁵². Ушбу ўрнатилган қоидалар эргаш гапни автоматик теглаш, модел ва тегларини ишлаб чиқишда мустаҳкам назарий асос бўлади. Шу сабабли эргаш гапни боғловчи воситаларига кўра қуйидаги турларини кўриб чиқамиз:

- 1) кўмакчили қурилма ёрдамида ЭҚГ;
- 2) *деб* сўзи ёрдамида ЭҚГ;
- 3) шарт майли воситасида ЭҚГ;
- 4) кўрсатиш олмоши ёрдамида ЭҚГ;
- 5) нисбий сўз ёрдамида ЭҚГ.

1.Кўмакчили қурилмалар ёрдамида эргашган қўшма гаплар. Бош гапда ифодаланган мазмуннинг сабабини билдирувчи эргаш гаплар бош гапга *шунинг*

¹⁵²Замонавий ўзбек тили: Синтаксис./Муаллифлар жамоаси. Масъул муҳаррирлар Ҳ.Ғ.Неъматов, Р.Сайфуллаева. ЎзР Олий ва ўрта махсус таълим вазирлиги, Мирзо Улуғбек номидаги Ўзбекистон Миллий университети. – Тошкент.: Мумтоз сўз, 2011. – 312 б. – 288-293.

учун, шу сабабли, шу туфайли сингари кўмакчили қурилмалар воситасида боғланади. Улар бош гап таркибида келади. *Халқ Юсуфбек ҳожидан зарра ёмонлик кўрмаган, шунинг учун у қаяққа юрса, шу ёққа боришига ҳозир турар эди. (Абдулла Қодирий)* 2. Ёзи билан тинимсиз меҳнат қилдик, **шу сабабли** мўл ҳосил олдик. (Отимжон Холдор) 3. Тузумлар ўзгарса ҳам, Инсон ўзгармас экан, **Шул боисдан** Пайғамбар Уммат замин еганлар; Аллоҳдан бошқасига сизинмагин еганлар. (А.Орипов) 4. Камтарин бўл, ҳатто бир қадам Ўтма зурур остонасидан, Пиёлани **шунинг-чун** инсон ўпар доим пешонасидан. (Э.Воҳидов) 5. Ўзимиз жуда чўллаб келган эдик, **шунинг учун** кечирдик. (Абдулла Қодирий). Мисоллардан кўриниб турибдики, бадий услубда кўмакчили қурилмаларнинг шунинг учун, шу сабабли, шул боисдан, шунинг-чун, шунинг учун турли шакллари қўлланади. Бундай боғловчиларнинг тўлиқ рўйхати шаклантирилади ва улар умумий <KQ> теги билан белгиланади. Шу асосда бундай қўшма гапларнинг синтактик моделларини қуйидагича тузиш мумкин. Агар: [eG] + <KQ>[bG] бўлса, [eG] + [bG] = <EQG> / <сабаб-EQG>. Ушбу тегда эргашган қўшма гап ҳамда унинг маъно жиҳатдан сабаб маъносини ифодалаб эргашган қўшма гап эканлиги ҳақидаги ахборот мавжуд.

2. *Деб* сўзи ёрдамида эргашган қўшма гаплар. Бош гапда ифодаланган мазмуннинг мақсади, сабабини билдирувчи эргаш гаплар бош гапга кўпинча **деб** сўзи ёрдамида боғланади ва эргаш гапларнинг кесими III шахс буйруқ майли шаклидаги феъллар билан ифодаланади. Бундай гаплар таркибидаги *деб* сўзи учун кўмакчиси билан маънодош саналади, шунинг учун бир-бири билан эркин алмаша олади¹⁵³: *Гул барглари учишади, тушимайин деб қўлингга (Ҳамид Олимжон). Шу мўъжасз уйимда ёнсун деб чироқ, Не азиз зотларга ёндошдим гоҳи (А.Орипов). Камалакдек ранго-ранг Бўлсин деб, санъат, тиллар, Асрларча қилдим жанг (Э.Воҳидов). Сен оташ ичида урасан жавлон, Душман қолмасин деб севган элимда (Ҳамид Олимжон). Ёвлар яксон бўлсин деб,*

¹⁵³Замонавий ўзбек тили: Синтаксис./ Муаллифлар жамоаси. Масъул муҳаррирлар Ҳ.Ғ.Неъматов, Р.Сайфуллаева. ЎзР Олий ва ўрта махсус таълим вазирлиги, Мирзо Улуғбек номидаги Ўзбекистон Миллий университети. – Тошкент.: Мумтоз сўз, 2011. – 312 б. – 288-293.

Йигитлар жанг бошлади, Ёвлар қоча бошлади (Ҳамид Фулом). Бундай восита ёрдамида боғланишни <debQ> теги билан белгилаймиз. Шу асосда бундай кўшма гапларнинг синтактик моделларини қуйидагича тузиш мумкин. Агар: [bG] + <...sin deb Q>[eG] бўлса, [eG] + [bG] = <EQG> / <мақсад-EQG>. Ушбу тегда эргашган кўшма гап ҳамда унинг маъно жиҳатдан мақсад маъносини ифодалаб эргашган кўшма гап эканлиги ҳақидаги ахборот мавжуд.

3. Шарт майли воситасида эргашган кўшма гаплар. Эргаш гапнинг кесими шарт майли шаклидаги феъллар орқали ифодаланганда, шарт майли кўшимчаси эргаш гапни бош гапга боғловчи восита ҳам саналади. Бундай эргаш гаплар бош гапда ифодаланган мазмуннинг юзага чиқиш ёки чиқмаслик шартини, пайтини билдиради: *Агар дарахт шохлари ёз фаслида сарғайса, куз эрта келади. Олма, олча каби мевали дарахтлар иккинчи бор гулласа, бу кузнинг илиқ келишидан даракдир. Қалдирғоч настлаб учса, ёгингарчилик бўлади. Чиннигул ўзидан ёқимли ҳид таратса, тез орада шамолни кутаверинг. Балиқ тез-тез сув бетига сапчиб чиқса, ёмғир ёғади. Мевали дарахтлар кўп ҳосил қилса, қиш қаттиқ келади*¹⁵⁴. Бундай восита ёрдамида боғланишни <shartQ> теги билан белгилаймиз. Шу асосда ушбу кўшма гапларнинг синтактик моделини қуйидагича тузиш мумкин. Агар: [eG...ca] + [bG] бўлса, [eG] + [bG] = <EQG> / <шарт-EQG>. Ушбу тегда шарт маъносини ифодалаб келган эргашган кўшма гап эканлиги ҳақидаги ахборот мавжуд.

4. Кўрсатиш олмошили эргашган кўшма гаплар. Бош гап таркибидаги кўрсатиш олмошининг маъносини изоҳлаш учун қўлланган эргаш гаплар бош гапга *-ки* юкламаси ёрдамида боғланади. Бу юклама бош гап кесими таркибида бўлади ва эргаш гап бош гапдан вергул билан ажратилади. Баъзан бу бўлак қўлланмаслиги ҳам мумкин, аммо унинг ўрни билиниб туради. *Бахтим шуки, сени учратдим (З.Саидносирова).* Шундай ўзгаришлар

¹⁵⁴Замонавий ўзбек тили: Синтаксис./ Муаллифлар жамоаси. Масъул муҳаррирлар Ҳ.Ғ.Неъматов, Р.Сайфуллаева. ЎзР Олий ва ўрта махсус таълим вазирлиги, Мирзо Улуғбек номидаги Ўзбекистон Миллий университети. – Тошкент.: Мумтоз сўз, 2011. – 312 б. – 294.

бўлмоқдаки, ақлинг бовар қилмайди. Шуниси қувонарлики, синфдаги ўқувчиларнинг ҳаммаси бир-бири билан жуда аҳил. Шунни билингки, ҳеч бир янгилик бехосият бўлмайди (Асқад Мухтор). Биз шундай кутубхона бино қилайликки, бутун эл қошида манзур ва мўътабар бўлсин (Ойбек). Бундай восита ёрдамида боғланишни <olmoshQ> теги билан белгилаймиз. Ушбу тартибдаги боғланиш тури синтактик тегларини қуйидагича тузиш мумкин. Агар: [bG\shu\shuni\shunga\shunday...ki] + [eG] бўлса, [bG] + [eG] = <EQG> / <шарт-EQG>. Ушбу тегда бош гапдаги кўрсатиш олмоши маъносини ифодалаган эргашган кўшма гап эканлиги ҳақидаги ахборот мавжуд.

5. Нисбий сўзли эргашган кўшма гаплар. Эргаш гап таркибида кўлланувчи *ким, нима, қанча, қанчалик, қандай, қаер* каби сўроқ олмошлари ва бош гап таркибида унга жавоб бўлиб келувчи *шу, ўша, шунча, шунчалик, шундай* каби олмошлар бир-бирига нисбатан кўлланганлиги, бири иккинчисини тақозо этганлиги учун нисбий сўзлар ҳисобланади. Эргаш гапнинг кесими шарт майли шаклидаги феъллар билан ифодаланади¹⁵⁵. *Сиз Ҳиндистонда неники кўрган бўлсангиз, шуларнинг ҳаммасини жам этиб, бир китоб ёзмагингиз даркор (А.Қаюмов). Қаландаров югурдагини «Эшон ҳой!» деб чақирганда, Саида қанчалик хурсанд бўлган бўлса, хужрани кўриб шунчалик таъби тирриқ бўлди (Абдулла Қаҳҳор). Кимнинг кўнгли тўғри бўлса, унинг ўзи ҳам тўғри бўлади (Мақол). Кимки ўзининг яхши ишидан қувонса, ёмон ишидан хафа бўлса, демак, у ҳақиқий мўминдир (Ҳадис). Бу Ватанни жон билан сақлашга, Чўлпон, ҳозир ўл, Кимки қасд этса анга кийсин пушаймондин кафан!(Чўлпон).*¹⁵⁶ Бундай восита ёрдамида боғланишни <nisbiyso`zQ> теги билан белгилаймиз. Ушбу тартибдаги боғланиш тури синтактик тегларини қуйидагича тузиш мумкин. Агар: [eG\ ким\ нима\ қанча\

¹⁵⁵Замонавий ўзбек тили: Синтаксис./ Муаллифлар жамоаси. Масъул муҳаррирлар Ҳ.Ғ.Неъматов, Р.Сайфуллаева. ЎзР Олий ва ўрта махсус таълим вазирлиги, Мирзо Улуғбек номидаги Ўзбекистон Миллий университети. – Тошкент.: Мумтоз сўз, 2011. – 312 б. – 295-296.

¹⁵⁶Замонавий ўзбек тили: Синтаксис./ Муаллифлар жамоаси. Масъул муҳаррирлар Ҳ.Ғ.Неъматов, Р.Сайфуллаева. ЎзР Олий ва ўрта махсус таълим вазирлиги, Мирзо Улуғбек номидаги Ўзбекистон Миллий университети. – Тошкент.: Мумтоз сўз, 2011. – 312 б. – 297.

қанчалик\ қандай\ қаер...sa] + [eG\ шу\ ўша\ шунча\ шунчалик\ шундай] бўлса, [eG] + [bG] = <EQG> / <nisbiy so`z-EQG>. Ушбу тегда бош гапдаги нисбий сўз (олмош)ни изоҳлаш маъносини ифодалаган эргашган кўшма гап эканлиги ҳақидаги ахборот мавжуд.

Юқоридаги таҳлиллардан келиб чиқиб, эргашган кўшма гапларнинг қидирув ойнасида қуйидаги параметрни жойлаштириш мумкин:

- Кўмакчили қурилмалар ёрдамида эргашган кўшма гаплар.
- Деб* сўзи ёрдамида эргашган кўшма гаплар.
- Шарт майли воситасида эргашган кўшма гаплар.
- Кўрсатиш олмошили эргашган кўшма гаплар.
- Нисбий сўзли эргашган кўшма гаплар.

Шунингдек, эргашган кўшма гапларнинг қандай боғловчи воситалар асосида шаклланишига кўра автоматик таҳлилни амалга ошириш учун модел ва теглар тизимини қуйидаги шаклда ифодалаймиз:

- 1) Агар: [eG] + <KQ>[bG] бўлса, [eG] + [bG] = <EQG> / <сабаб-EQG>.
- 2) Агар: [bG] + <...sin deb Q>[eG] бўлса, [eG] + [bG] = <EQG> / <мақсад-EQG>.
- 3) Агар: [eG...sa] + [bG] бўлса, [eG] + [bG] = <EQG> / <шарт-EQG>.
- 4) Агар: [bG\shu\shuni\shunga\shunday...ki] + [eG] бўлса, [bG] + [eG] = <EQG> / <шарт-EQG>.
- 5) Агар: [eG\ ким\ нима\ қанча\ қанчалик\ қандай\ қаер...sa] + [eG\ шу\ ўша\ шунча\ шунчалик\ шундай] бўлса, [eG] + [bG] = <EQG> / <nisbiy so`z-EQG>.

Хулоса қилиш мумкинки, ўзбек тилидаги эргашган кўшма гапларни автоматик қидирув тизими учун тег ва моделларни эргаш гапни боғловчи воситалар асосида ишлаб чиқиш мумкин.

3.2.2. *Ўзбек тилида эргашган кўшма гапларни моделиштиришда маъно тури асосидаги ёндашув.*

1. *Эга эргаш гап*–бош гапдаги олмош билан ифодаланган эгани изоҳлаган эргаш гап. Эга эргаш гап *-ки* боғловчиси, феълнинг шарт майли кўшимчаси *-са* ҳамда сўрок/кўрсатиш олмошлари (*кимки... у, кимки... ўша,*

кимки... ўзи, нима... ўша, нимаики... ҳаммаси) ёрдамида боғланади¹⁵⁷: *Кимки ёмонлар суҳбатидан қочса, у яхшилар суҳбатига эришади. Шу нарса қизиқки, уйқуда ҳам мия ўз фаолиятини тўхтатмайди. Ким изланса, у, албатта, мақсадига эришади. Кимки ростгўй, самимий бўлса, у эл ўртасида қадр топади. Кўзига нима кўринса, шу уни қизиқтира бошлади. Ким эрта баҳордан далага чиқиб пешона тери тўкса, у кўзда меҳнатининг роҳатини кўради.* Баъзан бош гап таркибидаги кўрсатиш олмоши билан ифодаланган эга туширилиши мумкин¹⁵⁸: *Сизга аёнки, сиз ўз оилангиздасиз.* Бу гапда шу (эга) туширилган. Бундай гапларнинг боғловчилар тизими қуйидагича:

- | | |
|---------------------------|-------------------------|
| 1. Ким... -са, у ... | 5. Ким... -са, ўзи... |
| 2. Кимки... -са, у ... | 6. Кимда-ки...-са, у... |
| 3. Нимаики ...-са, ўша... | 7. Нима... -са, ўша... |
| 4. Шуниси... -ки, ... | |

Юқорида кўрсатилган боғловчилар тизими асосида эга эргаш гапларни аниқловчи моделларни қуйидагича шакллантириш мумкин бўлади:

1. Агар: [eG] + <Ким...-са, у...>[bG] бўлса, [eG] + [bG] = <ega-EQG>.
2. Агар: [eG] + <Кимки...-са, у...>[bG] бўлса, [eG] + [bG] = <ega-EQG>.
3. Агар: [eG] + <Нимаики...-са, ўша...>[bG] бўлса, [eG] + [bG] = <ega-EQG>.
4. Агар: [eG] + <Шуниси...-ки, ...>[bG] бўлса, [eG] + [bG] = <ega-EQG>.
5. Агар: [eG] + <Ким...-са, ўзи...>[bG] бўлса, [eG] + [bG] = <ega-EQG>.
6. Агар: [eG] + <Кимда-ки...-са, у...>[bG] бўлса, [eG] + [bG] = <ega-EQG>.
7. Агар: [eG] + <Нима...-са, ўша...>[bG] бўлса, [eG] + [bG] = <ega-EQG>.

2. *Кесим эргаш гап* бош гапдаги олмош билан ифодаланган кесимни изоҳлаган эргаш гап бўлиб, у бош гап билан биргаликда *кесим эргаш гапли*

¹⁵⁷Замонавий ўзбек тили: Синтаксис./ Муаллифлар жамоаси. Масъул муҳаррирлар Ҳ.Ғ.Неъматов, Р.Сайфуллаева. ЎзР Олий ва ўрта махсус таълим вазирлиги, Мирзо Улуғбек номидаги Ўзбекистон Миллий университети. – Тошкент.: Мумтоз сўз, 2011. – 312 б. – 299-300.

¹⁵⁸Замонавий ўзбек тили: Синтаксис./ Муаллифлар жамоаси. Масъул муҳаррирлар Ҳ.Ғ.Неъматов, Р.Сайфуллаева. ЎзР Олий ва ўрта махсус таълим вазирлиги, Мирзо Улуғбек номидаги Ўзбекистон Миллий университети. – Тошкент.: Мумтоз сўз, 2011. – 312 б. – 265-266.

қўшма гапни ҳосил қилади¹⁵⁹: *Иккинчиси шуки, халқ бошига кўп офат келтирган Андижон zilzilаси Қорабулоқни сувдан маҳрум қилиб қўйди. Энди сизларга муборак топшириқ шуки, ҳар бирингиз алоҳида-алоҳида кетасиз. Толеим шулки, жаҳонда бир гулистон танладим. Муҳаммад Юсуф ижодининг ўзига хос хусусияти шундаки, у жуда осон ўқилади, содда, халқчил. Бунинг натижаси шу бўлдики, Тўлаганнинг қанишарига мушт тушди. Кераги шуки, ана шу миққийни ҳам боқиш керак. Қизиги шундаки, шу кўринишига овози мулойим эди. Кесим эргаш гаплар бош гапга -ки боғловчиси, ўрин-пайт ёки бош келишигидаги шу олмоши ёрдамида боғланади. Баъзан бош гапдаги олмош билан ифодаланган кесим туширилади: *Яхшиси, сен дарсингни тайёрлагин*¹⁶⁰. Бундай гапларнинг боғловчилар тизими қуйидагича:*

- | | |
|----------------------|----------------------|
| 1. ... шуки,..... | 3. ... шундаки, ... |
| 2. ... шуердаки, ... | 4. ... шундайки, ... |

Юқорида кўрсатилган боғловчилар тизими асосида кесим эргаш гапларни аниқловчи моделларни қуйидагича шакллантириш мумкин бўлади:

- 1) агар: [bG] + <шуки>[eG] бўлса, [eG] + [bG] = <kesim-EQG>;
- 2) агар: [bG] + <шуердаки>[eG] бўлса, [eG] + [bG] = <kesim-EQG>;
- 3) агар: [bG] + <шундаки>[eG] бўлса, [eG] + [bG] = <kesim-EQG>;
- 4) агар: [bG] + <шундайки>[eG] бўлса, [eG] + [bG] = <kesim-EQG>.

3. *Аниқловчи эргаш гапли қўшма гаплар.* Бош гап таркибида аниқловчи вазифасида қўлланган кўрсатиш олмошининг маъносини изоҳлаб келувчи эргаш гапга *аниқловчи эргаш гап* дейилади. Аниқловчи эргаш гаплар бош гапга нисбий сўзлар ёки бош гап кесими таркибида қўлланувчи *-ки* юкламаси ёрдамида боғланади. Биринчи ҳолатда эргаш гап бош гапдан олдин, иккинчи

¹⁵⁹ Замоनावий ўзбек тили: Синтаксис./ Муаллифлар жамоаси. Масъул муҳаррирлар Ҳ.Ғ.Неъматов, Р.Сайфуллаева. ЎзР Олий ва ўрта махсус таълим вазирлиги, Мирзо Улуғбек номидаги Ўзбекистон Миллий университети. – Тошкент.: Мумтоз сўз, 2011. – 312 б.

¹⁶⁰ Замоनावий ўзбек тили: Синтаксис./ Муаллифлар жамоаси. Масъул муҳаррирлар Ҳ.Ғ.Неъматов, Р.Сайфуллаева. ЎзР Олий ва ўрта махсус таълим вазирлиги, Мирзо Улуғбек номидаги Ўзбекистон Миллий университети. – Тошкент.: Мумтоз сўз, 2011. – 312 б. – 268-269.

ҳолатда кейин келади¹⁶¹. *Аниқловчи эргаш гап бош* гап билан *-ки* боғловчиси билан (бунда бош гап таркибида кўпинча аниқловчи бўлиб келган *шундай, айрим, баъзи* сўзлари иштирок этади) ёки шарт майли кўшимчаси *-са* билан (бунда эргаш гап таркибида *қандай, қайси, кимнинг* каби олмошлар, бош гап таркибида эса *шу, шундай, ўша, унинг, баъзи, бир хил, айрим, бир, бир қанча* каби олмошлар бўлади)боғланади¹⁶²: *Шундай инсон ҳақида хабар келтирдимки, унинг ҳар бир сўзи гавҳардир. Кимнинг кўнгли тўғри бўлса, унинг йўли ҳам тўғри.* Баъзан бош гапдаги аниқловчи туширилиши мумкин: *Одам борки, ҳайвон ундан яхшироқ.* Бундай ҳолатни автоматик аниқлаш қийин: бош гап охиридаги *-ки* юкламаси ёрдам бериши мумкин, шунда ҳам дастур 100 фоиз аниқ натижа кўрсата олмайди. Баъзан бош гапнинг аниқловчиси вазифасида келган *шундай, шунақа* каби сўзлар қўлланмаслиги мумкин, аммо унинг ўрни билиниб туради. Бундай гапларнинг боғловчилар тизими қуйидагича:

1. Шундай ... -ки, ...
- 2 ...Шундай... -ки, ...
3. ... шундай...-ни...-ки, ундай ...
4. ... шундай (аниқл.)...-ки, ...
5. ... шундай(ҳол)...-ки, ...
6. ... қайси... -са, ўша...

Юқорида кўрсатилган боғловчилар тизими асосида аниқловчи эргаш гапларни аниқловчи моделларни қуйидагича шакллантириш мумкин бўлади:

- 1) агар: <Shunday>[bG]<-ки> + [eG] бўлса, [eG] + [bG] = <aniqlovchi-EQG>.
- 2) агар: [bG]<Shunday><-ки> + [eG] бўлса, [eG] + [bG] = <aniqlovchi-EQG>.

¹⁶¹ Замоनावий ўзбек тили: Синтаксис./ Муаллифлар жамоаси. Масъул муҳаррирлар Ҳ.Ғ.Неъматов, Р.Сайфуллаева. ЎзР Олий ва ўрта махсус таълим вазирлиги, Мирзо Улуғбек номидаги Ўзбекистон Миллий университети. – Тошкент.: Мумтоз сўз, 2011.-312 б

¹⁶² Замоनावий ўзбек тили: Синтаксис./ Муаллифлар жамоаси. Масъул муҳаррирлар Ҳ.Ғ.Неъматов, Р.Сайфуллаева. ЎзР Олий ва ўрта махсус таълим вазирлиги, Мирзо Улуғбек номидаги Ўзбекистон Миллий университети. – Тошкент.: Мумтоз сўз, 2011. – 312 б.

3) агар: [bG]<Shunday><ни><-ки> + [eG] бўлса, [eG] + [bG] = <aniqlovchi-EQG>.

4) агар: <Shunday>[bG]<-ки> + [eG] бўлса, [eG] + [bG] = <aniqlovchi-EQG>.

5) агар: <Shunday>[bG]<-ки> + [eG] бўлса, [eG] + [bG] = <aniqlovchi-EQG>.

6) агар: <qaysi>[eG]<-са> + [bG] бўлса, [eG] + [bG] = <aniqlovchi-EQG>.

4. *Тўлдирувчи эргаш гап*–бош гапдаги кўрсатиш олмоши билан ифодаланган тўлдирувчини изоҳлаган эргаш гап. Тўлдирувчи эргаш гап бош гап билан биргаликда *тўлдирувчи эргаш гапли қўшма гап*ни ташкил этади. Тўлдирувчи эргаш гап, одатда, бош гапга *–ки* юклагаси, шарт майли қўшимчаси *-са* ёрдамида боғланади. *Шуни билингки, ҳар бир янгилик бехосият бўлмайди. Кимда гумонинг бўлса, уни кўздан қочирма.(О.)* Бош гап таркибида *шуни, шунга, шундан* каби тўлдирувчилар қатнашади. Баъзан бу тўлдирувчилар туширилиши мумкин. *Илтимос қиламанки, олдин раис сўзласинлар*¹⁶³. Бу гапда *шуни* сўзи туширилган. Бундай гапларнинг боғловчилар тизими қуйидагича:

- | | |
|--------------------------|--------------------------------|
| 1. Шуни...-ки, ... | 5. Шундан...-ки, ... |
| 2. Шунга... -ки, ... | 6. Нима(ни)... -са, шуни ... |
| 3. Кимда... -са, уни ... | 7. Нима билан... -са, ўшани... |
| 4. ... -ки, ... | |

Кўрсатилган боғловчилар тизими асосида тўлдирувчи эргаш гапни аниқловчи моделларни қуйидагича шакллантириш мумкин:

1) агар: <shuni>[bG]<-ki> + [eG] бўлса, [eG] + [bG] = <to`ldirovchi-EQG>.

2) агар: <shundan>[bG]<-ki> + [eG] бўлса, [eG] + [bG] = <to`ldirovchi -EQG>.

3) агар: <shunga>[bG]<-ki> + [eG] бўлса, [eG] + [bG] = <to`ldirovchi -EQG>.

4) агар: <nimani>[eG]<-sa> + <shuni>[bG] бўлса, [eG] + [bG] = <to`ldirovchi -EQG>.

¹⁶³Замонавий ўзбек тили: Синтаксис./ Муаллифлар жамоаси. Масъул муҳаррирлар Ҳ.Ғ.Неъматов, Р.Сайфуллаева. ЎзР Олий ва ўрта махсус таълим вазирлиги, Мирзо Улуғбек номидаги Ўзбекистон Миллий университети. – Тошкент.: Мумтоз сўз, 2011. – 312 б. – 271.

5) агар: <kimda>[eG]<-sa> + <uni>[bG] бўлса, [eG] + [bG] = <to`ldirovchi -EQG>.

6) агар: <nima bilan>[eG]<-sa> + <o`shani> [bG] бўлса, [eG] + [bG] = <to`ldirovchi -EQG>.

7) агар: [bG]<-ki> + [eG] бўлса, [eG] + [bG] = <to`ldirovchi -EQG>.

5. *Равиш эргаш гап* бош гапдан англашилган иш-ҳаракатнинг кайтарзда бажарилганлигини билдирган эргаш гапдир. Равиш эргаш гап бош гап билан бирга равиш эргаш гапли қўшма гапни ташкил этади¹⁶⁴: Боғловчи воситалар: *-б(-иб), -май, -масдан, деб: Юлдузлар битта-битта сўниб, уфқ оқара бошлади. Наимийнинг бирор иши ўнгидан келмасдан, умри бекорга ўтиб кетди. (Ас.М.) Чамбил обод бўлди деб, келаётир Авазхондай зўравон. (Ф.Й.)* Бундай гапларнинг боғловчилар тизими қуйидагича: 1. <...-иб, ...>; 2. <... -май, ...>; 3. <... -масдан, ...>; 4. <... деб, ...>.

Юқорида кўрсатилган боғловчилар тизими асосида равиш эргаш гапни аниқловчи моделларни қуйидагича шакллантириш мумкин бўлади:

1) агар: [eG]<-иб> + [bG] бўлса, [eG] + [bG] = <ravish-EQG>.

2) агар: [eG]<-май> + [bG] бўлса, [eG] + [bG] = <ravish-EQG>.

3) агар: [eG]<-масдан> + [bG] бўлса, [eG] + [bG] = <ravish-EQG>.

4) агар: [eG]<-деб> + [bG] бўлса, [eG] + [bG] = <ravish-EQG>.

6. *Пайт эргаш гап* бош гапдаги иш-ҳаракатнинг пайтини билдирган эргаш гапдир. Пайт эргаш гап бош гап билан *пайт эргаш гапли қўшма гапни* ташкил этади: *У келгунча, ҳамма тарқалган еди. Отлиқлар жарлик оша тепаликлардан кўтарилгандан кейин, чўл бошланди. Собиржон приёмникни энди қўйган эди, эшик тақиллаб қолди. Қушлар тўп-тўп бўлиб*

¹⁶⁴Замонавий ўзбек тили: Синтаксис./ Муаллифлар жамоаси. Масъул муҳаррирлар Ҳ.Ғ.Неъматов, Р.Сайфуллаева. ЎзР Олий ва ўрта махсус таълим вазирлиги, Мирзо Улуғбек номидаги Ўзбекистон Миллий университети. – Тошкент.: Мумтоз сўз, 2011. – 312 б. – 282-299.

учса, ҳаво очиқ бўлади¹⁶⁵. Бундай гапларнинг боғловчилар тизими куйидагича бўлади:

- | | |
|-----------------------------|----------------------------|
| 1. ... -ганда, ... | 10. ... – ар экан,... |
| 2. ... –ган захоти, | 11. ... - са,..... |
| 3. ... –ган вақтда, | 12. ... –масдан аввал,... |
| 4. –ган пайтда,.... | 13. ... -иб,..... |
| 5. ... –гани замон | 14. ... -гач,... |
| 6. ... –гандан кейин,... | 15. ... -гунча,... |
| 7. ... –гунга қадар, ... | 16. ... -май,... |
| 8. ... –ган эди, | 17. ... –иши билан(оқ),... |
| 9. ... –ганича йўқ эди,.... | |

Юқорида кўрсатилган боғловчилар тизими асосида пайт эргаш гапни аниқловчи моделларни куйидагича шакллантириш мумкин бўлади:

- 1) агар: [eG]<-ганда> + [bG] бўлса, [eG] + [bG] = <payt-EQG>.
- 2) агар: [eG]<-ган захоти> + [bG] бўлса, [eG] + [bG] = < payt -EQG>.
- 3) агар: [eG]<-ган вақтда\ ган пайтда> + [bG] бўлса, [eG] + [bG] = < payt -EQG>.
- 4) агар: [eG]<-гани замон> + [bG] бўлса, [eG] + [bG] = < payt -EQG>.
- 5) агар: [eG]<-гандан кейин> + [bG] бўлса, [eG] + [bG] = < payt -EQG>.
- 6) агар: [eG]<-гунга қадар> + [bG] бўлса, [eG] + [bG] = < payt -EQG>.
- 7) агар: [eG]<-ган эди \ ганича йўқ эди> + [bG] бўлса, [eG] + [bG] = < payt -EQG>.
- 8) агар: [eG]<-ар экан> + [bG] бўлса, [eG] + [bG] = < payt -EQG>.
- 9) агар: [eG]<-са> + [bG] бўлса, [eG] + [bG] = < payt -EQG>.
- 10) агар: [eG]<-масдан аввал> + [bG] бўлса, [eG] + [bG] = < payt -EQG>.
- 11) агар: [eG]<-иб> + [bG] бўлса, [eG] + [bG] = < payt -EQG>.
- 12) агар: [eG]<-гунча\ гач\ -май-> + [bG] бўлса, [eG] + [bG] = < payt -EQG>.

¹⁶⁵Замонавий ўзбек тили: Синтаксис./ Муаллифлар жамоаси. Масъул муҳаррирлар Ҳ.Ғ.Неъматов, Р.Сайфуллаева. ЎзР Олий ва ўрта махсус таълим вазирлиги, Мирзо Улуғбек номидаги Ўзбекистон Миллий университети. – Тошкент.: Мумтоз сўз, 2011. – 312 б. – 282-299.

13) агар: [eG]<-иши билан(ок)> + [bG] бўлса, [eG] + [bG] = <райт -EQG>.

Булар орасида -иб воситаси билан боғланиш модели равиш эргаш гапли қўшма гап билан омонимликни келтириб чиқаради. Қидирилган натижа фойдаланувчи томонидан сараланиши лозим бўлади.

7. Ўрин эргаш гап бош гапдан англашилган иш-ҳаракатнинг бажарилиш ўрнини билдирган эргаш гапдир. У бош гап билан биргаликда *ўрин эргаш гапли қўшма гап* дейилади: *Қаерда интизом кучли бўлса, шу ерда тартиб бўлади.* Боғловчи воситалар: *-са (қаердан, қайга, қаерга, қаерда, шу ерда, ўша ерда, шунда, шу ерга, шу ердан, у ерда)-манг, -масин, экан: Қаерда сув бўлса, ўша ерда ҳаёт бўлади.* Бундай гапларнинг боғловчилар тизими қуйидагича бўлади:

- 1) қаерга ...-са, ўша ерга...
- 2) қайга ...-са, ўша ерда,
- 3) қанчаманг, у ерда ...
- 4) қаерда ...-са, ўша ерда ...
- 5) -масин, ўша ерда...
- б) қаерда....., шу ерда.....

Юқорида кўрсатилган боғловчилар тизими асосида ўрин эргаш гапларни аниқловчи моделлари нисбий сўзлар билан боғланган қўшма гап моделлари билан бир хил бўлади, чунки бундай мазмунли қўшма гаплар фақат нисбий сўзлар билан боғланади.

8. Сабаб эргаш гап бош гапдан англашилган иш-ҳаракатнинг сабабини билдирган эргаш гапдир. *Қайрилишлар тобора кўпайиб, машинанинг тезлиги сусая бошлади. Бу уйлар илгари бедахона бўлганидан, дарчаси ҳам йўқ эди.* Бундай гапларнинг боғловчилар тизими қуйида кўрсатилган 18 шаклда бўлади¹⁶⁶:

- 1) ... шунинг учун ...
- 2) ... -ганидан ...

¹⁶⁶Замонавий ўзбек тили: Синтаксис./ Муаллифлар жамоаси. Масъул муҳаррирлар Ҳ.Ғ.Неъматов, Р.Сайфуллаева. ЎзР Олий ва ўрта махсус таълим вазирлиги, Мирзо Улуғбек номидаги Ўзбекистон Миллий университети. – Тошкент.: Мумтоз сўз, 2011. – 312 б. – 282-299.

- | | |
|---|-----------------------------|
| 3)... -гани учун, ... | 4)... -ганидан кейин ... |
| 5) ... -ганидан сўнг, ... | 6)... шекилли, ... |
| 7) ... -ганидан сўнг, ... | 8) ... -ган бўлса керак ... |
| 9) ... -ми, ... | 10) ..., чунки ... |
| 11) ..., негаки ... | 12) ... -масдан аввал, ... |
| 13) ..., сабабки... .. | 14) ... -гач, ... |
| 15) ... шунинг учун...-ки, 16)-май, ... | |
| 17) ... -ки, ... | 18)... –иб, ... |

Юқорида кўрсатилган боғловчилар тизими асосида сабаб эргаш гапларни аниқловчи моделларни қуйидагича шакллантириш мумкин бўлади:

- 1) агар: [eG]<-шунинг учун\ шунинг учун...-ки \-ки> + [bG] бўлса, [eG] + [bG] = <sabab -EQG>;
- 2) агар: [eG]<- гани учун\-ганидан сўнг\ -ганидан сўнг> + [bG] бўлса, [eG] + [bG] = <sabab -EQG>;
- 3) агар: [eG]<-ми\> + [bG] бўлса, [eG] + [bG] = <sabab -EQG>;
- 4) агар: [eG]<- негаки\ сабабки> + [bG] бўлса, [eG] + [bG] = <sabab -EQG>;
- 5) агар: [eG]<-ганидан \-ганидан кейин\ шекилли\ -ган бўлса керак> + [bG] бўлса, [eG] + [bG] = <sabab -EQG>;
- 6) агар: [eG]<-чунки\ -масдан аввал \-гач \-май \-иб\ ми\> + [bG] бўлса, [eG] + [bG] = <sabab -EQG>.

Булар орасида -иб, -май, -гач воситалари билан боғланиш модели равиш, пайт ва б. эргаш гапли қўшма гап билан омонимликни келтириб чиқаради, бундай пайтда ҳам қидирилган натижа фойдаланувчи томонидан сараланиши лозим бўлади.

9. Мақсад эргаш – гап бош гапдан англашилган иш-ҳаракатнинг нима мақсадда юзага келишини билдирган эргаш гап. У бош гап билан бирга *мақсад эргаш гапли қўшма гап* деб номланади. Боғловчи воситалар: *деб, -син деб, -ди деб, -ар деб, -са деб, -ми экан (деб, дея, учун), токи: Одамлар яхши яшасин деб, тинчликка имзо чекдик. Ғўза мириқиб ичсин учун, сув*

текис оқизилди. Бундай гапларнинг боғловчилар тизими қуйида кўрсатилган 10 шаклда бўлади¹⁶⁷:

- 1)...деб, ...2)...-ми экан деб...
- 3) ...-син деб,...4) -дея ...
- 5)... учун... б)..... токи ...
- 7) ...-са деб,...

Юқорида кўрсатилган боғловчилар тизими асосида мақсад эргаш гапни аниқловчи моделларни қуйидагича шакллантириш мумкин бўлади:

- 1) агар: [eG]<-дея \- деб\ -син деб\ -ми экан деб\ -са деб > + [bG] бўлса, [eG] + [bG] = <maqsad-EQG>.
- 2) агар: [eG]<-учун> + [bG] бўлса, [eG] + [bG] = < maqsad-EQG>.
- 3) агар: [eG]<-токи> + [bG] бўлса, [eG] + [bG] = < maqsad-EQG>.

10. Шарт эргаш гап – бош гапдан англашилган иш-ҳаракатнинг қандай шарт билан бажарилишини билдирган эргаш гап. У бош гап билан бирга *шарт эргаш гапли қўшма гапни* ҳосил қилади. Боғловчи воситалар: -са, -са эди, агар, мабодо, борди-ю, -ганда(эди), -р(-ар) экан, -ган экан, -ми, бўлмаса, йўқса, -май, бўлай десанг, -синки: *Агар сув бўлса, чўл-у саҳролар бўстонга айланади.* Бундай гапларнинг боғловчилар тизими қуйида кўрсатилган 14 шаклда бўлади¹⁶⁸:

- 1) (Агар)... -са, ...
- 2) (Агар) ...-са эди,...
- 3) (Агар).... -ганда, ...
- 4) (Агар) ... -ганда эди ...
5. (Борди-ю) ... -р(-ар) экан, ...
- 6) (Мабодо)... -ган экан, ...
- 8) ... -ми,...
- 9) ... бўлмаса, ...
- 10) ..., йўқса ...
- 11) ... -май,...
- 12) ... -масдан аввал, ...
- 13) ... бўлай десанг, ...

¹⁶⁷Замонавий ўзбек тили: Синтаксис./ Муаллифлар жамоаси. Масъул муҳаррирлар Ҳ.Ғ.Неъматов, Р.Сайфуллаева. ЎзР Олий ва ўрта махсус таълим вазирлиги, Мирзо Улуғбек номидаги Ўзбекистон Миллий университети. – Тошкент.: Мумтоз сўз, 2011. – 312 б. – 282-299.

¹⁶⁸Кўрсатил Замонавий ўзбек тили: Синтаксис./ Муаллифлар жамоаси. Масъул муҳаррирлар Ҳ.Ғ.Неъматов, Р.Сайфуллаева. ЎзР Олий ва ўрта махсус таълим вазирлиги, Мирзо Улуғбек номидаги Ўзбекистон Миллий университети. – Тошкент.: Мумтоз сўз, 2011. – 312 б. – 282-299.

7) (Агар) ... -моқчи экан, ... 14) ... -синки, ...

Юқорида кўрсатилган боғловчилар тизими асосида шарт эргаш гапни аниқловчи моделларни қуйидагича шакллантириш мумкин бўлади:

1) агар: <Агар>[eG]<-са\ -са эди\ -ганда\ -ганда эди\ -моқчи экан> + [bG]

бўлса, [eG] + [bG] = <shart-EQG>.

2) агар: < Борди-ю >[eG]<-p/-ap) экан> + [bG] бўлса, [eG] + [bG] = <shart-EQG>.

3) агар: < Мабодо>[eG]<-ган экан> + [bG] бўлса, [eG] + [bG] = <shart-EQG>.

4) агар: [eG]<-ми\ бўлмаса\ йўқса\ -май\ -масдан аввал\ бўлай десанг\ ган экан> + [bG] бўлса, [eG] + [bG] = <shart-EQG>.

11. *Тўсиқсиз эргаш гап* бош гапнинг мазмунига зид бўлса ҳам, унда ифодаланганвоқеа-ҳодисанинг юзага келишига тўсиқ бўла олмайдиган фикрни англатган эргаш гапдир. Боғловчи воситалар: *-са ҳам(ки), қанча, қанчалик, нақадар, гарчи, гарчанд, -ганда ҳам, -са-да, -ганда-да, -гани билан, -масин, қарамай, қарамасадан, -и(иб), -гани ҳолда, -ди ҳамки, хоҳ...хоҳ (-са/-маса), -са: Қуёш яширинган бўлса ҳам, кундузнинг ёруғлиги ҳали сўнмаган эди. Баъзан гарчи сўзи қўлланиши мумкин*¹⁶⁹. Бундай гапларнинг боғловчилар тизими қуйида кўрсатилган 14 шаклда бўлади:

1. Қанча... -са ҳам, ...

8. ... қарамасдан, ...

2. Қанча-са ҳамки,...

9. ... -гани ҳолда, ...

3. ... -ганда ҳам...

10. ... ҳамки, ...

4. ... -са-да,...

11. Хоҳ..., хоҳ...

5. ... -гани билан,...

12. ... -са/-маса, ...

6.... -масин,...

13. ... -са... -дики, ...

7. ... -қарамай,...

14. ... -са, ...

Юқорида кўрсатилган боғловчилар тизими асосида тўсиқсиз эргаш гапларни аниқловчи моделларни қуйидагича шакллантириш мумкин бўлади:

¹⁶⁹Замонавий ўзбек тили: Синтаксис./ Муаллифлар жамоаси. Масъул муҳаррирлар Ҳ.Ғ.Неъматов, Р.Сайфуллаева. ЎзР Олий ва ўрта махсус таълим вазирлиги, Мирзо Улуғбек номидаги Ўзбекистон Миллий университети. – Тошкент.: Мумтоз сўз, 2011. – 312 б. – 282-299.

1) агар: <Қанча>[eG]<-са ҳам\ -са ҳамки> + [bG] бўлса, [eG] + [bG] = <to`siqsiz-EQG>.

2) агар: [eG]<қарамасдан\ -гани ҳолда\ -ганда ҳам\ҳамки\ -са-да\ -гани билан\ -са/-маса\ -масин\ -са... -дики\ -қарамай\ -са> + [bG] бўлса, [eG] + [bG] = <to`siqsiz-EQG>.

3) агар: <Хоҳ>[eG]<хоҳ> + [bG] бўлса, [eG] + [bG] = <to`siqsiz-EQG>.

12. *Миқдор-даража эргаш гап* – бош гапдан англашилган иш-ҳаракатнинг бажарилиш миқдор-даражасини билдирган эргаш гап: *У қанча қизишса, мен шунча совуққон тус олардим.* Бундай гапларнинг боғловчилар тизими куйида кўрсатилган шаклларда бўлади:

1. ... қанча.....-са, ... шунча ...
2. ... қанчалик...-са, ... шунчалик ...
3. ... қанча...-са, ...шунчалик...
4. Қанчалик...-ган сайин, ...шунчалик...

Юқорида кўрсатилган боғловчилар тизими асосида миқдор-даража эргаш гапларни аниқловчи моделларни куйидагича шакллантириш мумкин бўлади:

1).... агар: <Қанчалик>[eG]< -ган сайин, шунчалик> + [bG] бўлса, [eG] + [bG] = <miqdor-daraja-EQG>;

2) ... агар: [eG]<қанча\ қанчалик ><- са, шунча \ шунчалик> + [bG] бўлса, [eG] + [bG] = <miqdor-daraja-EQG>.

13. *Ўхшатиш эргаш гап* – бош гапдан англашилган иш-ҳаракатнинг қай тарзда бажарилишини ўхшатиш йўли билан билдирган гап: *Чирилдоқларнинг майин мусиқаси ҳамма ёқни тўлдирган, гўё кечанинг ўзи куйлайди.* Ўхшатиш эргаш гапли қўшма гапларнинг боғловчилар тизими куйида кўрсатилган шаклларда бўлади¹⁷⁰:

1. ... каби, ...

7. Гўё... -гандай ...

¹⁷⁰Замонавий ўзбек тили: Синтаксис./ Муаллифлар жамоаси. Масъул муҳаррирлар Ҳ.Ғ.Неъматов, Р.Сайфуллаева. ЎзР Олий ва ўрта махсус таълим вазирлиги, Мирзо Улуғбек номидаги Ўзбекистон Миллий университети. – Тошкент.: Мумтоз сўз, 2011. – 312 б. – Б. 301.

- | | |
|----------------------|----------------------------------|
| 2. ... -дай... | 8. ... дегандек,... |
| 3. ... -сингари ... | 9. ... кандай....-са, шундай ... |
| 4. ... янглиғ, ... | 10. ... -гунча ... |
| 5. ... -ки, гүё... | 11. ... -гандан кўра, ... |
| 6. ... -ки, худди... | 12. ...-ишдан кўра ... |

Юқорида кўрсатилган боғловчилар тизими асосида ўхшатиш эргаш гапни аниқловчи моделларни қуйидагича шакллантириш мумкин бўлади:

- 1) агар: [eG] <каби \ -дай \ сингари \ янглиғ\ -ки, гүё \-ки, худди> + [bG] бўлса, [eG] + [bG] = <o`xshatish-daraja-EQG>;
- 2) агар: < Гүё>[eG] <-гандай > + [bG] бўлса, [eG] + [bG] = <o`xshatish-daraja-EQG>;
- 3) агар: [eG] < дегандек \ қандай....-са, шундай > + [bG] бўлса, [eG] + [bG] = <o`xshatish-daraja-EQG>;
- 4) агар: [eG] < гунча \ гандан кўра \ -ишдан кўра > + [bG] бўлса, [eG] + [bG] = <o`xshatish-daraja-EQG>.

14. *Натижа эргаш гап*– бош гапдан англашилган иш-ҳаракатнинг натижасини билдирган эргаш гап. Натижа эргаш гап бош гапга *-ки* боғловчиси ёрдамида боғланади. Бош гап таркибида *шундай, шу қадар, шунчалик, шу даражада, шунча, чунон* каби сўзлар, эргаш гаплар таркибида эса *натижада, оқибатда, ҳатто* каби сўзлар қўлланиши мумкин: *Дўл бир зумда шундай ёғдики, ер оппоқ бўлиб кетди. Азимбой шундай зулм ўтказдики, оқибатда халқнинг сабр-косаси тўлди.* Бундай кўшма гапларнинг боғловчилар тизими қуйида кўрсатилган шаклларда бўлади:

1. ...шундай ... -ки, натижада ...
2. ...шуқадар ... -ки, оқибатда...
3. ...шунчалик... -ки, ...
5. ...шунча ... -ки, ...
4. ...шу даражада...-ки,...
6. ...чунон ...-ки, ...

Юқорида кўрсатилган боғловчилар тизими асосида натижа эргаш гапни аниқловчи моделларни қуйидагича шакллантириш мумкин бўлади:

агар: <шундай \ шу қадар \ шунчалик \шунча \шу даражада\ чунон> [eG]< -ки, -ки, натижада \ -ки, оқибатда> + [bG] бўлса, [eG] + [bG] = <natija-daraja-EQG>.

Хулоса ўрнида айтиш жоизки, боғланган, боғловчисиз ҳамда эргашган қўшма гапларнинг жами 87 та модел, жумладан, эргаш гапнинг бош гапдаги қайси бўлакни изохлашига кўра эргашган қўшма гап турларининг қидирувини йўлга қўйиш учун уларнинг боғловчи воситаларига асосланган 67 та модел ишлаб чиқилди.

Юқорида ишлаб чиқилган моделлар асосида корпус интерфейси учун қидирув тизими ишлаб чиқиш мумкин. Бунда ўзбек тили матнларини синтактик таҳлил қилиш ҳамда синтактик параметрлар асосидаги қидирувни қуйидагича шакллантириш мумкин:

- Боғланган қўшма гаплар
- ✓ бириктирув муносабатли қўшма гаплар
- ✓ зидлов муносабатли қўшма гаплар
- ✓ айирув муносабатли қўшма гаплар
- ✓ инкор муносабатли қўшма гаплар
- Боғловчисиз қўшма гаплар
- ✓ Изохлаш муносабатини ифодаловчи қўшма гаплар
- ✓ Кетма-кет ёки бир вақтда содир бўлган ҳодисани ифодаловчи қўшма гаплар
- ✓ Мазмунан бир-бирин узоқ ҳодисани ифодаловчи қўшма гаплар
- ✓ шарт, зидлик муносабатини ифодаловчи қўшма гаплар
- Эргашган қўшма гаплар (Боғловчи воситасига кўра тури)
- ✓ Кўмакчили қурилмалар ёрдамида эргашган қўшма гаплар.
- ✓ *Деб* сўзи ёрдамида эргашган қўшма гаплар.
- ✓ Шарт майли воситасида эргашган қўшма гаплар.
- ✓ Кўрсатиш олмошили эргашган қўшма гаплар.

- ✓ Нисбий сўзли эргашган қўшма гаплар.
 - Эргашган қўшма гаплар (Бош гаптаги мазмунни ифодалашига кўра тури)
- ✓ Эга эргаш гапли қўшма гап
- ✓ Кесим эргаш гапли қўшма гап
- ✓ Аниқловчи эргаш гапли қўшма гап
- ✓ Тўлдирувчи эргаш гапли қўшма гап
- ✓ Равиш эргаш гапли қўшма гап
- ✓ Микдор-даража эргаш гапли қўшма гап
- ✓ Мақсад эргаш гапли қўшма гап
- ✓ Сабаб эргаш гапли қўшма гап
- ✓ Ўрин эргаш гапли қўшма гап
- ✓ Пайт эргаш гапли қўшма гап
- ✓ Натижа эргаш гапли қўшма гап
- ✓ Шарт эргаш гапли қўшма гап

3.3. Ўзбек тилидаги мураккаб қўшма гаплар ва кўчирма гапли қўшма гап конструкциясининг автоматик анализ ва синтези

3.3.1. Ўзбек тилида мураккаб қўшма гапларни автоматик синтактик таҳлил қилишга доир. Ҳозирги ўзбек адабий тилида эргаш гаплар бош гапта бирдан ортиқ ҳолда тобеланиб келиб, қўшма гапнинг мураккаб турини ҳосил қилади. Бундан ташқари, эргаш гапгина эмас, эргашган қўшма гаплар ҳам мазмунан бир-бирига боғланган ҳолда бирдан ортиқ келиб, мураккаб компонентли қўшма гапни ташкил қилиши мумкин¹⁷¹. Мактаб дарслигида уч ва ундан ортиқ содда гапнинг бирикишидан таркиб топган гап мураккаб қўшма гап дейилади. Мураккаб қўшма гаплар тўрт гуруҳга бўлинади¹⁷²:

- 1) бир неча эргаш гапли мураккаб қўшма гаплар;

¹⁷¹Ўзбек тили грамматикаси. II том. Синтаксис. [Масъул муҳаррирлар: Ғ.А.Абдурахмонов ва бошқ.]. – Т., «Фан», 1976. – 560 б. – Б. 462.

¹⁷²Замонавий ўзбек тили: Синтаксис./ Муаллифлар жамоаси. Масъул муҳаррирлар Ҳ.Ғ.Неъматов, Р.Сайфуллаева. ЎзР Олий ва ўрта махсус таълим вазирлиги, Мирзо Улуғбек номидаги Ўзбекистон Миллий университети. – Тошкент.: Мумтоз сўз, 2011. – 312 б.– Б. 463.

- 2) бир неча бош гапли мураккаб қўшма гаплар;
- 3) аралаш мураккаб қўшма гаплар;
- 4) қисмлари уюшган мураккаб қўшма гаплар.

Тил корпуси қидирув тизимида мураккаб қўшма гапларга ишлов бериш параметрларини аниқлашда Ғ.А.Абдурахмонов масъул муҳаррирлигида нашр этилган грамматикада келтирилган таснифга асосланиш маъқул топдик, чунки бу манбада қуйидаги аниқ таснифлар мавжуд.

1. Мураккаб компонентли қўшма гаплар икки, уч ёки ундан ортиқ эргашган қўшма гапларнинг боғланган қўшма гап қисмларига ўхшаш тенгланиш орқали бирикишидан ҳосил бўлади. Шунинг учун бундай компонентлар орасида боғланган қўшма гапларда учрайдиган кўпгина ҳолатларни кузатиш мумкин:

а) мураккаб компонентлар орасида тенгланиш интонацияси мавжуд бўлади: *Камтар бўлсанг, обрў ортар; Гердайсанг, энса қотар (Мақол). Истасанг, қишлоқдан аямни олдириб келай; истасанг, синглим келиб турсин (Саид Аҳмад)*¹⁷³.

б) бундай компонентлар орасида боғланган қўшма гап қисмларини бириктирувчи тенг боғловчилар ёки шу вазифада қўлланувчи юкламалар ишлатилиши мумкин: *Агар дарс вақти тугаган ёки ўтиб кетган бўлса, шу билан бугунги дарс тамом қилинар ва агар ҳали вақт бўлса, яна бирор жумла ўқиларди*¹⁷⁴.

Кўриняптики, қидирув тизимида бундай гапларни аниқлашда боғловчисиз қўшма гап ёки боғланган қўшма гапларни таҳлил қилиш учун тузилган моделлардан фойдаланиш мумкин: мураккаб гапнинг чегараси битта бутунлик сифатида тиниш белги асосида белгиланади. Фақат боғловчисиз қўшма гапда

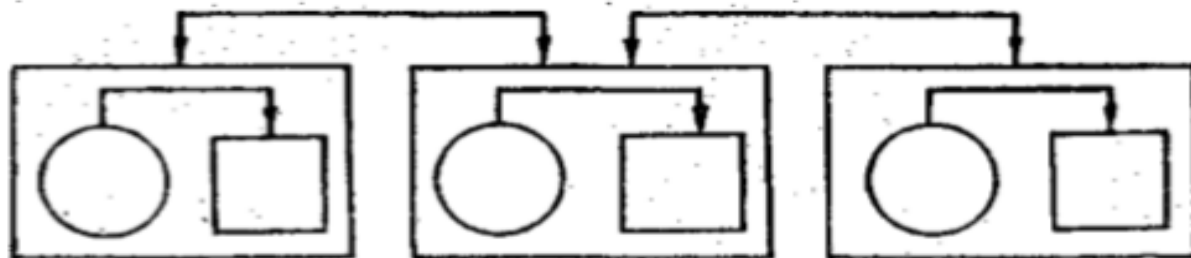
¹⁷³Замонавий ўзбек тили: Синтаксис./ Муаллифлар жамоаси. Масъул муҳаррирлар Ҳ.Ғ.Неъматов, Р.Сайфуллаева. ЎзР Олий ва ўрта махсус таълим вазирлиги, Мирзо Улуғбек номидаги Ўзбекистон Миллий университети. – Тошкент.: Мумтоз сўз, 2011. – 312 б.– Б. 463.

¹⁷⁴Замонавий ўзбек тили: Синтаксис./ Муаллифлар жамоаси. Масъул муҳаррирлар Ҳ.Ғ.Неъматов, Р.Сайфуллаева. ЎзР Олий ва ўрта махсус таълим вазирлиги, Мирзо Улуғбек номидаги Ўзбекистон Миллий университети. – Тошкент.: Мумтоз сўз, 2011. – 312 б.

“нуқтадан нуқтагача” принципи иккита содда гапдан иборат қўшма гапни аниқласа, мураккаб қўшма гапда уч ва ундан ортиқ содда гапдан ташкил топган бирликни ажратиш олади. Бундай усул тенг боғловчилар билан боғланган мураккаб қўшма гапларга нисбатан ҳам қўлланади: мураккаб ҚГ учун махсус модел тузишга зарурат қолмайди. Тенг боғловчилар билан боғланган бир неча қўшма гап “нуқтадан нуқтагача” оралиқда аниқланади. Аммо кидирув ойнасида мураккаб қўшма гапларнинг турини ажратиш мумкин. Бунда дастур икки босқичли таҳлилни қуйидаги алгоритм асосида амалга оширади:

- 1) “нуқтадан нуқтагача” оралиқдаги бирликни ажратади;
- 2) боғловчи воситалар миқдорига кўра “қўшма гап” (шунингдек, унинг тури) ёки “мураккаб қўшма гап” эканлиги ҳақида хулоса чиқаради.

Мураккаб компонентли қўшма гаплар уч ёки ундан ортиқ эргаш гапнинг қўшилувидан тузилади. Эргашган қўшма гаплардан тузилган мураккаб компонентли қўшма гапнинг боғланиш схемасини қуйидагича кўрсатиш мумкин¹⁷⁵:



1. Бир неча эргаш гапли қўшма гап. Бирдан ортиқ эргаш гап бир бош гапга тобеланиб келса, бир неча эргаш гапли қўшма гап дейилади¹⁷⁶.

- 1) тўғридан-тўғри тобеланиш (биргалик эргашиш);
- 2) кетма-кет эргашиш.

2. Бир неча бош гапли мураккаб қўшма гап. Таркибида икки ва ундан ортиқ бош гап мавжуд бўлган қўшма гаплар бир неча бош гапли мураккаб қўшма гап ҳисобланади.

¹⁷⁵Ўзбек тили грамматикаси. II том. Синтаксис. [Масъул муҳаррирлар: Ғ.А.Абдурахмонов ва бошқ.]. – Т., «Фан», 1976. – 560 б.
– Б. 497.

¹⁷⁶УҚ

3. Аралаш мураккаб қўшма гаплар. Таркибидаги гап эргашиш йўли билан ҳам, тенгланиш йўли билан ҳам боғланган бўлса, аралаш турдаги қўшма гап дейилади. Бундай қўшма гап учдан ортиқ содда гапдан ҳам ташкил топади.

4. Қисмлари уюшган мураккаб қўшма гаплар. Бир турдаги уч ва ундан ортиқ гаплардан ташкил топган гаплар қисмлари уюшган қўшма гаплар дейилади.

Юқоридаги таснифдан келиб чиқиб, синтактик таҳлил дастури мураккаб қўшма гаплар қидирув интерфейсига қуйидаги параметрларни киритиш мумкин:

- Бир неча эргаш гапли қўшма гап
- ✓ тўғридан-тўғри тобеланиш (биргалик эргашиш)
- ✓ кетма-кет эргашиш
- Бир неча бош гапли мураккаб қўшма гап
- Аралаш мураккаб қўшма гап
- Қисмлари уюшган мураккаб қўшма гап.

Булар орасида аралаш мураккаб қўшма гап ҳам тиниш белги, ҳам боғловчи восита асосида таҳлил қилинадиган бирликлар қаторига киради. Бундай қўшма гапларнинг автоматик қидируви ҳам ажратишнинг “нуқтадан нуқтагача” принципи, шунингдек, қўшма гапларнинг боғловчи воситасига асосланган моделларига таянади.

3.3.2. *Ўзбек тилида кўчирма гапли қўшма гапларнинг синтактик теглар тизими.* Маълумки, муаллиф ўз нутқида бошқа киши ифодалаган фикрдан ҳам фойдаланиши мумкин. Бундай фикр кўчирма гап, ўзлаштира гап, диалог шаклида берилади. Сўзловчи баъзан ўз нутқида бировнинг гапини айнан келтиради. Ўзганинг айнан келтирилган гапи *кўчирма гап*, сўзловчининг гапи эса *муаллиф гапи* дейилади. Ушбу гаплар биргаликда кўчирма гапли қўшма гапни ҳосил қилади¹⁷⁷. Тил корпусларида, бошқа

¹⁷⁷Ўзбек тили грамматикаси. II том. Синтаксис. [Масъул муҳаррирлар: Ғ.А.Абдурахмонов ва бошқ.]. – Т., «Фан», 1976. – 560 б. – Б. 506-508.

синтактик бирликларда бўлганидек, кўчирма гапли қўшма гапнинг автоматик кидирувини ҳам йўлга қўйиш мумкин. Бунинг автоматик таҳлили моделларини тузишда кўчирма гапли конструкция таркибидаги содда гапнинг кесимлик кўрсаткичи ҳамда тиниш белгилар позициясидан ўринли фойдаланиш мумкин.

Кўчирилган гап муаллиф гапидан аввал, кейин, унинг ичида ёки икки томонида келиши мумкин¹⁷⁸:

1. “Меҳнат иштаҳа очар”, – дейди бобом.
2. Чўпондан сўрадик: “Бу қўйлар кимнинг қўйи?”
3. “Мен, – деди у, – эртага келаман”.
4. Бош инженер босиқлик билан: “Бу гапга Худойқуловнинг алоқаси йўқ”, – деди.

Кўчирилган гап қўштирноқ ичига олинади, бадий асар матнида эса тире билан ажратиб берилади. Уларнинг умумий кўриниши қуйидагича:

- 1) М: «К». М: «К?» М: «К!»
- 2) М: «К», – М. М: «К?» – М. М: «К!» – М.
- 3) «К», – М. «К?» – М. «К!» – М.
- 4) «К, – М, – К».

Кўчирма гапли қўшма гапнинг ушбу турғун (аниқ) конструкцияси автоматик таҳлил жараёнини бир мунча осонлаштирилади.

Кўчирма гапда, кўчирилган гапнинг ўрнига қараб, тиниш белгисининг ишлатилиши турлича:

1) кўчирилган гап дарак гап бўлиб, муаллиф гапидан олдин келса, ундан кейин вергул ва тире қўйилади: “*Юринг, мен ўша томонга бораман*”, – деди *Комила*. Сўроқ ва ундов белгиси қўштирноқ ичида қолади. Ушбу позициянинг модели ва тегини қуйидагича белгилаш мумкин: “WPm”, –WPm = <К>, <М>.

а) кўчирилган гап муаллиф гапидан кейин келса, муаллиф гапидан кейин икки нуқта қўйилади: *Маърузачи бундай деди: “Ўсиб бораётган авлодга китоб худди*

¹⁷⁸Ўзбек тили грамматикаси. II том. Синтаксис. [Масъул муҳаррирлар: Ғ.А. Абдурахмонов ва бошқ.]. – Т., «Фан», 1976. – 560 б.

мактаб каби керак”. Ушбу позициянинг модели ва тегини қуйидагича белгилаш мумкин: WP_m : “ WP_m ” = $\langle M \rangle$: $\langle K \rangle$.

б) муаллиф гапи кўчирилган гап ичида келса, тиниш белгиси қуйидагича қўйилади: “Бизнинг қишлоғимизда, – деди Фазлиддин, – киши зерикмайди”. “Ҳаво булут, пахта очикда қолмасин, – деди бригадир. – Ҳар қайси звенодан биттадан одам чақиринг, пахтани саройга ташиб киритсинлар”. Бунда муаллиф гапи ва кўчирма гап тугалланмай қолгани автоматик таҳлилда чалкашлик келтириб чиқариши мумкин. Шу сабабли кўштирноқ ичидаги қисм муаллиф гапи, яъни “ WP_m ”, $-WP_m$, $-“WP_m”$ = $\langle M \rangle$, $\langle K \rangle$, $\langle M \rangle$ деб белгиланади. Аммо ушбу тег (эҳтимолий белги)нинг қанчалик тўғри натижа бериши парсинг дастури ишга тушгач билинади.

Хулоса ўрнида айтиш жоизки, мураккаб қўшма гаплар ҳамда қўшма гапли конструкциялар, асосан, тиниш белгилар воситасида таҳлил қилинади: гап чегараси тиниш белгидан тиниш белгигача тамойили билан белгиланади.

Боб бўйича қисқа хулосалар

1. Қўшма гапни боғловчи воситалар қўшма гап турини автоматик аниқлашнинг бирламчи воситаси бўла олади, чунки синтактик теглаш дастури (парсер) содда гаплар чегарасини улар орасидаги тиниш белги ҳамда боғловчи восита орқали аниқлайди.

2. Автомтик таҳлил дастури боғловчи воситалари омонимияни келтириб чиқарувчи қўшма гапларни фарқлай олмайди. Бунинг учун фойдаланувчи автоматик таҳлил натижасини бирма-бир таҳлил қилиши керак. Ўзбек тилининг семантик таҳлил моделлари яратилса, шаклий ноаниқлик мазмуний моделлар асосида аниқланиши мумкин.

3. Ажратилган кесимли содда гаплар дастур томонидан йиғиқ гап сифатида аниқланади, натижада грамматик/ структур омонимлик келиб чиқади. Ҳозирча бундай тузилишли қўшма гапларни теглаш ва моделлаштириш орқали автоматик аниқлашнинг имкони йўқ. Келгусида ўзбек тилида онтологик моделларнинг мукамаллашиши

структур/грамматик кўп маънолилик ва омонимияни аниқлашга ёрдам бериши мумкин.

4. Уюшиқ бўлақлар орасига кўйиладиган ҳамда боғловчисиз кўшма гаплар орасидаги тиниш белгилар автоматик таҳлилда омонимияга дуч келади. Бундай пайтда натижа фойдаланувчи томонидан бирма-бир сараланади. Бундай бирликларни фарқлаш моделларини ишлаб чиқиш махсус тадқиқотни талаб қилади: ажратилган бўлақли, уюшиқ бўлақли ҳамда мураккаблашган содда гапларда тиниш белгилар, уларнинг моделлари ўрганилса, боғловчисиз кўшма гаплардаги мана шундай истисно ҳолатга аниқлик киритилади.

5. Боғланган кўшма гапларнинг мазмуний муносабат ифодаловчи турларини ифодаловчи тегларни ишлаб чиқиш ҳозирча имконсиз. Шаклий восита (боғловчи восита ва тиниш белги) ёрдамида боғланган кўшма гаплардан фарқли равишда, мазмуний муносабатларни ифодаловчи бирликларнинг теглари тизимини ишлаб чиқиш учун гапларни автоматик рефератлаш, автоматик аннотация ҳосил қилиш ёки WordNet тизими базаси каби дастурий воситаларни талаб этади.

6. Мураккаб гапнинг чегараси битта бутунлик сифатида тиниш белги асосида белгиланади. Фақат боғловчисиз кўшма гапда “нуқтадан нуқтагача” принципи иккита содда гапдан иборат кўшма гапни аниқласа, мураккаб кўшма гапда уч ва ундан ортиқ содда гапдан ташкил топган бирликни ажратиб олади. Бундай усул тенг боғловчилар билан боғланган мураккаб кўшма гапларга нисбатан ҳам қўлланади: мураккаб ҚГ учун махсус модел тузишга зарурат қолмайди. Тенг боғловчилар билан боғланган бир неча кўшма гап “нуқтадан нуқтагача” оралиқда аниқланади.

УМУМИЙ ХУЛОСА

1. Синтактик теглаш матннинг автоматик синтактик таҳлилини амалга оширишга ёрдам берувчи асосий жараён бўлиб, синтактик теглар мажмуидан иборат бўлади, синтактик қурилмалар орасидаги синтактик алоқани кўрсатади. Илк корпусларни яратишда синтактик разметкалар қўлда амалга оширилган бўлса, кейинги авлод корпуслари синтактик разметкаси парсинг дастури асосида, автоматик/ярим автоматик тарзда амалга оширилган.

2. Синтактик разметкада қўлланиладиган асосий алгоритм – синтактик декомпозиция бўлиб, у бўшлиқ (пробел), тиниш белгилари асосида жумлани фарқлайди. Аммо бу усул мукамал эмас, чунки ажратувчи сифатида ишлатиладиган белги нафақат жумланинг охирида, балки ўртасида ҳам ишлатилиши мумкин. Сарлавҳа, бўлимни ажратиш кўрсатувчи бирлик, расм, жадвал номи, колонтитул каби бирликлар гап бўлмаса-да, гап кўринишида ифодаланади.

3. Синтактик разметкаланган матнда морфологик разметка билан бирга ҳар бир гапга унинг синтактик структураси тўғрисидаги изоҳ ёзилади. Синтактик разметкада теглар яқка ҳамда контейнер тегга ажратилади. Яқка тег матн бирлиги (сўз) ҳақида ахборот беради, контейнер тег эса разметка тизимида сақланадиган матн структураси тўғрисидаги ахборотни ташиydi. Матнни гапга ажратиш жуфт контейнер теглар воситасида амалга оширилади. Идентификатор синтактик тобе сўзга ишора қилса, алоқа типи ҳоким ва тобе сўз ўртасидаги синтактик муносабат типини акс эттиради.

4. Ҳар бир тилдаги синтактик разметка тизимини ишлаб чиқиш учун ўша тилнинг синтактик қурилиши моделлаштирилиши талаб этилади. Моделлаштиришдан кейинги босқич синтактик теглар тизимини тузиш, сўнгги қадам эса матн тил бирликларига синтактик тегларни бириктиришдир. Шунингдек, синтактик теглар тизимини белгилашда интерфейснинг қулайлигини инобатга олиш зарурати мавжуд. Синтактик теглар тизимини

тузишда матнга автоматик ишлов беришга эришиш, қўлда ишлов беришни имкон қадар камайтириш принципига амал қилиш лозим.

5. Аннотацияланган корпусларда морфологик белги асосидаги қидирув етакчилик қилади. Синтактик ва семантик белгилар асосидаги қидирув нисбатан мураккаб жараён бўлганлиги сабабли ҳамма корпусларни ҳам синтактик/семантик теглаш имкони, воситаси мавжуд эмас.

6. Ўзбек тили корпусларининг парсер дастури, унинг синтактик теглар тизимини ишлаб чиқишда замонавий автоматик таҳлил дастурлари ҳамда ўзбек тилшунослигида синтаксис бўйича мавжуд назарий материалларни ўрганиш, умумлаштириш ҳамда синтактик теглар яратишда ёндашувни тўғри танлаш муҳим. Корпусда гап ёки бирикма эканлиги ташқи белгилари асосида фарқланмайдиган бирикувларни аниқлашнинг махсус алгоритми, фильтри ишлаб чиқилиши лозим. Бунда гап/сўз бирикмасини аниқлашнинг бир компонентли ёки икки компонентли эканлигига асосланиш натижа бера олади.

7. Синтактик теглашда синтактик алоқанинг тури инобатга олинади ҳамда теглар тизимидан синтактик алоқани ифодаладиган тег жой олади. Тил корпуси материалини синтактик теглашда синтактик алоқа ва синтактик муносабат фарқланиши лозим. Шаклан фарқлашнинг имкони йўқ: боғловчи воситалар бир хил. Шу сабабли корпус разметкасида синтактик алоқа ва синтактик муносабатни фарқловчи тег киритилиши талаб этилади. Корпус синтактик разметка тизимида коллокатив қидирув жуда муҳим. Коллокатив қидирувни ташкил этиш учун корпус бирликлари сўз бирикмаси ва сўз қўшилмасини ажратувчи тегга эга бўлиши талаб қилинади.

8. Морфологик кўрсаткичсиз тобе-ҳокимлик муносабатини аниқлаш учун лисоний-синтактик қолип ҳамда морфологик теглар тизимидан фойдаланиш ўринли, чунки лисоний-синтактик қолипда қайси сўз туркуми ўзаро бирикиши кўрсатиб берилган. Корпус бирликлари морфологик тегланган бўлса, тег ва лисоний-синтактик қолип асосида морфологик кўрсаткичсиз синтактик муносабатли бирикмаларни ҳам автоматик теглаш

имкони пайдо бўлади. Албатта, корпусда сўз бирикмасини теглаш муаммоси умумий (инвариант) ЛСҚлар билан ҳал этилмайди: сўз бирикмаларни аниқлашда аниқроқ қолиплар талаб этилади. Шундан келиб чиқиб, сўз бирикмаларни синтактик теглашнинг исм+исм қолипи учун 14 та модел ишлаб чиқилди.

9. Корпусда билвосита белгилар асосида аниқланиши мумкин бўлган бирликлар тегланмаслиги ҳам мумкин, чунки бундай бирликлар тегишли тиниш белгилари орқали фарқланади. Тиниш белгилари орқали фарқлаш имкони мавжуд бўлган гапга шундай ахборотдан иборат тег бириктирмаслик ҳам мумкин; икки хил гап бўлаги бўлиши мумкин бўлган бирликларга иккилик тег бириктириш талаб этилади.

10. Қўшма гапни боғловчи воситалар қўшма гап турини автоматик аниқлашнинг бирламчи воситаси бўла олади. Синтактик теглаш дастури (парсер) содда гаплар чегарасини улар орасидаги тиниш белги ҳамда боғловчи восита орқали аниқлайди. Автоматик таҳлил дастури боғловчи воситалари омонимияни келтириб чиқарувчи қўшма гапларни фарқлай олмайди. Бунинг учун фойдаланувчи автоматик таҳлил натижасини бирма-бир таҳлил қилиши керак. Ўзбек тилининг семантик таҳлил моделлари яратилса, шаклий ноаниқлик мазмуний моделлар асосида аниқланиши мумкин.

11. Уюшиқ бўлақлар орасига қўйиладиган ҳамда боғловчисиз қўшма гаплар орасидаги тиниш белгилар автоматик таҳлилда омонимияга дуч келади. Бундай бирликларни фарқлаш моделларини ишлаб чиқиш махсус тадқиқотни талаб қилади: ажратилган бўлақли, уюшиқ бўлақли ҳамда мураккаблашган содда гапларда тиниш белгилар, уларнинг моделлари ўрганилса, боғловчисиз қўшма гаплардаги мана шундай истисно ҳолатга аниқлик киритилади.

12. Боғланган қўшма гапларнинг мазмуний муносабат ифодаловчи турларини ифодаловчи тегларни ишлаб чиқиш ҳозирча имконсиз. Шаклий восита (боғловчи восита ва тиниш белги) ёрдамида боғланган қўшма гаплардан фарқли равишда, мазмуний муносабатларни ифодаловчи

бирликларнинг теглари тизимини ишлаб чиқиш учун гапларни автоматик рефератлаш, автоматик аннотация ҳосил қилиш ёки WordNet тизими базаси каби дастурий воситаларни талаб этади.

Фойдаланилган адабиётлар рўйхати

I. Норматив-ҳуқуқий ҳужжатлар ва ижтимоий-сиёсий адабиётлар

1. Ўзбекистон Республикаси Президенти Шавкат Мирзиёевнинг 2020 йил 20 октябрдаги “Мамлакатимизда ўзбек тилини янада ривожлантириш ва тил сиёсатини такомиллаштириш чора-тадбирлари тўғрисида”ги ПФ-6084-сон фармони // <https://lex.uz/docs/5058351>
2. Ўзбекистон Республикаси Президенти Шавкат Мирзиёевнинг 2019 йил 21 октябрдаги «Ўзбек тилининг давлат тили сифатидаги нуфузи ва мавқеини тубдан ошириш чора-тадбирлари тўғрисида»ги ПФ-5850-сон фармони // www.xabar.uz
3. “Ўзбекистон Республикаси давлат тили ҳақида”ги Қонуни. – Тошкент: Ўзбекистон, 1989.
4. Каримов И.А. Юксак маънавият – енгилмас куч. – Тошкент: Маънавият, 2008. – 176 б.
5. Мирзиёев Ш.М. Буюк келажакимизни мард ва олижаноб халқимиз билан бирга қураимиз. – Тошкент: Ўзбекистон, 2017. – 488 б.
6. Мирзиёев Ш.М. Қонун устуворлиги ва инсон манфаатларини таъминлаш – юрт тараққиёти ва халқ фаровонлигининг гарови. – Тошкент: Ўзбекистон, 2017. – 48 б.

II. Илмий рисола, монография, дарслик, ўқув ва методик қўлланмалар

7. Абдурахмонов Ғ. Ўзбек тили грамматикаси. – Тошкент: Ўқитувчи, 1996. – 11 б.
8. Азизов О. Тилшуносликка кириш. – Тошкент: ЎзССР “Ўрта ва олий мактаб” давлат нашриёти, 1963.- 56 б.
9. Апресян Ю.Д., Богуславский И.М., Иомдин Л.Л. и др. Лингвистическое обеспечение системы ЭТАП-2. – М.: Наука, 1989. – 276 с.

10. Апресян Ю.Д., Богуславский И.М., Иомдин Л.Л., Лазурский А.В., Митюшин Л.Г., Санников В.З., Цинман Л.Л. Лингвистический процессор для сложных информационных систем. М.: Наука, 1992. – 456 с.
11. Апресян Ю.Д., Богуславский И.М., Иомдин Л.Л., Лазурский А.В., Перцов Н.В., Санников В.З., Цинман Л.Л. Лингвистическое обеспечение системы ЭТАП-2. М.: Наука, 1989. – 506 с.
12. Баранов А.Н. Введение в прикладную лингвистику [Текст]: учебное пособие / А.Н. Баранов. – М.: Изд-во Эдиториал УРСС, 2001. – 347 с.
13. Блумфильд Л. Язык. – Москва: Прогресс, 1968. – 608 с.
14. Бобокалонов Р. Ўзбек тилида гап синтаксиси ва сўз-гапларнинг систем-структур талқини: Монография. – Тошкент: Фан, 2006. – 182 б.
15. Валгина Н.С. Современный русский язык. Синтаксис. – М.: Высшая школа, 2003. – 402 с.
16. Ефремова Т.Ф. Толковый словарь служебных частей речи русского языка / Т. Ф. Ефремова. – М.: Изд-во Астрель-АСТ, 2004. – 814 с.
17. Замонавий ўзбек тили: Синтаксис. II том. (Муаллифлар жамоаси). – Тошкент: Мумтоз сўз, 2011. – 316 б.
18. Захаров В.П. Корпусная лингвистика. Учебно-методическое пособие. – Санкт-Петербург, 2005. – 69 с.
19. Кустова Г.И., Мишина К.И., Федосеев В.А. Синтаксис современного русского языка. М., 2005. – 520 с.
20. Махмутова М.К. Проблемы аннотирования (тагирования) текстов в корпусной лингвистике // Выпускная квалификационная работа ЮУрГУ. – Челябинск, 2018. – 94 с.
21. Махмудов Н., Нурмонов А. Ўзбек тилининг назарий грамматикаси. – Тошкент: Ўқитувчи, 1995. – 228 б.
22. Мирзаев М., Усмонов С., Расулов И. Ўзбек тили. – Тошкент: Ўқитувчи, 1970. – 312 б.

- 23.Муҳаммедова С. Ҳаракат феъллари асосида компьютер дастурлари учун лингвистик таъмин яратиш. Методик қўлланма.–Тошкент, 2006. – 45 б.
- 24.Недошивина Е.В. Программы для работы с корпусами текстов: обзор основных корпусных менеджеров. Учебно-методическое пособие. – Санкт-Петербург, 2006. – 26 с.
- 25.Неъматов Ҳ, Ғуломов А. ва бошқалар. Она тили: 7-синф учун дарслик. –Тошкент, 2000. – 106 б.
- 26.Неъматов Ҳ., Ғуломов А. ва бошқ. Она тили: 5-синф учун дарслик. – Тошкент: Ижод дунёси, 2001. – 156 б.
- 27.Норов А. Компьютер лингвистикаси асослари. – Қарши, 2017. – 136 б.
- 28.Нурмонов А., Маҳмудов Н., Аҳмедов А., Солихўжаев С. Ўзбек тилининг мазмуний синтаксиси. -Тошкент: Фан, 1992. - 291 б.
- 29.Пўлатов А. Компьютер лингвистикаси. – Тошкент: Akademnashr, 2011. – 175 б.
- 30.Рогожникова Р.П. Толковый словарь сочетаний, эквивалентных слову / Р.П.Рогожникова. – М.: Изд-во Астрель-АСТ, 2003. – 416 с.
- 31.Сайфуллаева Р. Ҳозирги ўзбек тилида қўшма гапларнинг формал-функционал талқини. Монография. –Тошкент: Фан, 1994. – 112 б.
- 32.Усмонов С. Ундов гаплар. – Тошкент: ЎзФА нашриёти, 1955. – 148 б.
- 33.Френсис Н., Кучера Г. Вычислительный анализ современного американского варианта английского языка. – Москва, 1967.– 342 с.
- 34.Чернякова Т. А. Методика формирования навыков студентов на основе лингвистического корпуса [Текст] / Т. А. Чернякова. – Тамбов, 2012. – 149 с.
- 35.Шахматов А.А. Синтаксис русского языка. Ленинград., 1941, С. 274.
- 36.Ўзбек тили грамматикаси. Синтаксис. II том. – Тошкент, “Фан”, 1976. (Масъул муҳаррирлар: Ғ.А.Абдурахмонов ва б.). – 560 б.

37. Ўринбоева Д.Б. Ўзбек фольклори матнларининг лингвостатистик тадқиқи. – Тошкент: Фан, 2010. – 121 б.
38. Қурбонова М., Сайфуллаева Р., Боқиева Г., Менглиев Б. Ўзбек тилининг структурал синтаксиси. – Тошкент, 2004. – 416 б.
39. Ғуломов А. Синтаксис ва пунктуация бўйича машқлар тўплами. – Тошкент: Ўздавнашр, 1938; 1939; 1947. – 258 б.
40. Ғуломов А. Содда гап. Ҳозирги замон ўзбек тили курси бўйича материаллар. – Тошкент: ЎзФАнашр., 1955. – 386 б.
41. Ғуломов А. Ўзбек тилида аниқловчилар. – Тошкент: Ўздавнашр, 1941.
42. Ғуломов А., Асқарова М. Ҳозирги ўзбек адабий тили. Синтаксис: университетлар ва педагогика институтларининг филология факультетлари учун дарслик. – Тошкент: Ўқитувчи, 1987. – 256 б.
43. Ғуломов А.Ғ. Содда гап. – Тошкент: ЎзФА, 1955. – 192 б.
44. Ғуломов А.Ғ., Асқарова М. Ҳозирги замон ўзбек адабий тили. Синтаксис. – Тошкент, 1965. – 418 б.
45. Biber D., Conrad S., Reppen R. Corpus linguistics. Investigating language structure and use. – Cambridge University Press, 1998. – 486 p.
46. Bongers H. The history and principles of Vocabulary control. – Woerden: WOCOPI, 1947. – 442 p.
47. Finegan E. Language: its structure and use. – N.Y: Harcourt Brace College Publishers, 2004. – 346 p.
48. Fries Ch.C. The structure of English. An introduction to the construction of English sentences. – New York: Harcourt, 1952. – 304 p.

49. III. Матбуот материаллари ва илмий мақолалар

50. Akhmedova D.B., Mengliev B.R. Semantic Tag Categories in Corpus Linguistics: Experience and Examination International Journal of Recent

Technology and Engineering (IJRTE) ISSN: 2277-3878, Volume-8, Issue-3S, October 2019. – P. 208-212.

51. Betsch, M. Automatic Annotation of Russian texts: Evaluation of Different Tagging Methods / R.Meyer, P.Kosta, J.Blaszczak, J.Frasek, L.Geist, M.Zygis // Investigations into Formal Slavic Linguistics. Contributions of the Fourth European Conference on Formal Description of Slavic Languages (FDSLIV) held at Potsdam University, 2003. – 234.
52. Brants, Th., W. Skut, and H. Uszkoreit H. Syntactic annotation of a German newspaper corpus. // Proceedings of the ATALA Treebank Workshop, pp. 69-76, Paris, France.
53. Hajicova E., Panevova J., Sgall P. Language Resources Need Annotations To Make Them Really Reusable: The Prague Dependency Treebank. // Proceedings of the First International Conference on Language Resources & Evaluation, pp. 713-718.
54. Karimov R., Mengliev B. Theoretical fundamentals of uzbek-english parallel corpus / Journal of critical reviews. ISSN- 2394-5125. – VOL 7, ISSUE 17, 2020. – P. 73-76.;
55. Karimov R.A., Mengliev B.R. The Role of the Parallel Corpus in Linguistics, the Importance and the Possibilities of Interpretation International Journal of engineering and Advanced Technology (IJEAT). ISSN: 2249 – 8958, Volume-8, Issue-5S3 July 2019. – P. 388-391.
56. Leech G. Corpus annotation schemes / G.Leech. Literary and Linguistic Computing, 1993. – 8/4. – P. 275-281.
57. Leech G. Corpus annotation schemes // Literary and Linguistic Computing, 1993, № 8/4. Pp. 275-81.
58. Van der Beek, L., Bouma G., Malouf R., Noord van G. The Alpino Dependency Treebank. // Proceedings of LINC 2001 // https://clinjournal.org/CLIN_proceedings/XII/vanderbeek.pdf

59. Xianyao Hu, Richard Xiao. How do English translations differ from non-translated English writings? A multi-feature statistical model for linguistic variation analysis
[//https://www.researchgate.net/publication/294138492_How_do_English_translations_differ_from_non-translated_English_writings_A_multi-feature_statistical_model_for_linguistic_variation_analysis/references](https://www.researchgate.net/publication/294138492_How_do_English_translations_differ_from_non-translated_English_writings_A_multi-feature_statistical_model_for_linguistic_variation_analysis/references)
60. Абжалова М. Матнларга автолингвистик ишлов бериш тизимлари // Шестая Международная конференция по компьютерной обработке тюркских языков «TurkLang-2018». (Труды конференции) – Ташкент, 2018. – 320 с.
61. Мустайоки А. Принципы создания Хельсинкского аннотированного корпуса русских текстов ХАНКО в сети Интернет / Научно-техническая информация. – 2003. – Сер. 2, № 6. – С. 33-37.
62. Нигматов Х. и др. Структура предложения и актуальные вопросы синтаксиса тюркских языков (Тезисы формально-функционального исследования) // Советская тюркология, 1984.– №5. – С.5-10.
63. Омонтурдиев Ж. Кесимнинг структурасига кўра классификацияси. // Совет мактаби, 1965. – №4. – Б. 26-27.
64. Орхун М. Computational analysis of uzbek nouns / Шестая Международная конференция по компьютерной обработке тюркских языков «TurkLang-2018». (Труды конференции) – Ташкент, 2018. – 320 с.
65. Гурсунов А. Вопросы словообразования в формальных моделях тюркских языков (на примере узбекского языка) // Шестая Международная конференция по компьютерной обработке тюркских языков «TurkLang-2018». Труды конференции) – Ташкент, 2018. – 320 с.
66. Ҳамроева Ш , Менглиев Б. Тил корпусларини синтактик разметка лашнинг лингвистик асослари “Осиё мамлакатлари тамаддуни ва ипак йўли” Халқаро илмий анжуман материаллари. –

Самарқанд, 2019 йил 21 август. – 236 б. – Б. 36-39.

67.Эгамбердиев В. Кесим таърифи хақида баъзи мулоҳазалар. Профессор-ўқитувчилар илмий конференциясининг тезислари (Қарши пед.институти). – Қарши, 1967.

IV.Диссертациялар ва диссертация авторефератлари

68.Абдурахмонова Н.З. Инглизча матнларни ўзбек тилига таржима қилиш дастурининг лингвистик таъминоти (Содда гаплар мисолида). Филол.фан.бўйича фалсафа доктори (PhD)...дис. афтореф. – Тошкент, 2018. – 52 б.

69.Абжалова М.А. Ўзбек тилидаги матнларни тахрир ва таҳлил қилувчи дастурнинг лингвистик модуллари (Расмий ва илмий услубдаги матнлар тахрири дастури учун): Филол.фан.бўйича фалсафа доктори (PhD)...диссер. – Фарғона, 2019. – 164 б.

70.Абузалова М. Ўзбек тилида содда гапнинг энг кичик қурилиш қолиплари ва унинг нутқда воқеланиши: филол. фан. номз. дисс. автореф. – Бухоро, 2004. – 22 б.

71.Айымбетов Н.К. Опыт лингвостатистического анализа лексики и морфологии каракалпакского публицистического текста: автореф. дисс... канд. фил. наук.– Ташкент, 1987. – 24 б.

72.Бабанаров А. Разработка принципов построения словарного обеспечения турецко-русского машинного перевода: автореф. дисс... канд. фил. наук. – Ленинград, 1981. – 26 б.

73.Жуманазарова Г.У. Фозил Йўлдош ўғли дostonлари тилининг лингвопоэтикаси: Фил. фан. док. Дис. автореф. – Тошкент, 2017. – 72 б.

74.Муҳаммедов С.М. Статистический анализ лексико-морфологической структуры узбекских газетных текстов: автореф. дисс... канд. фил. наук.–Ташкент, 1980. – 24 с.

75. Ножов И.М. Морфологическая и синтаксическая обработка текста (модели и программы) сегментации русского предложения. – М.: АКД, 2003. – 24 с.
76. Раупова Л. Ўзбек тилида номустақил кесим масаласи ва [WP_m-WP_m] қурилишли гаплар: филол. фан. номз. ... дисс. – Тошкент, 1999. – 145 б.
77. Ризаев С. Ўзбек тилининг лингвостатистик тадқиқи: фил. фан. док. Дис. автореф. – Тошкент, 2008. – 25 б.
78. Хамроева Ш.М. Ўзбек тили муаллифлик корпусини тузишнинг лингвистик асослари: Филол.фан.бўйича фалсафа доктори (PhD) дис. – Бухоро, 2018. – 250 б.
79. Хамроева Ш. Ўзбек тили морфологик анализаторининг лингвистик таъминоти. Филология фанлари доктори диссертацияси автореферати. – Фарғона, 2021. - 78 б.
80. Шодмонов Э. Ҳозирги замон ўзбек тилида сўз-гаплар: филол. фан. номз. дисс. – Тошкент, 1970. – 230 б.
81. Эшмўминов А.А. Ўзбек тили миллий корпусининг синоним сўзлар базаси: Филол.фан.бўйича фалсафа доктори (PhD)...диссер. – Қарши, 2019. – 140 б.

VI. Луғатлар

82. Каримов С., Қаршиев А., Исроилова Г. Абдулла Қаҳҳор асарлари тилининг луғати. Алфавитли луғат. Частотали луғат. Терс луғат. – Тошкент, 2007. – 420 б.

VII. Интернет материаллари, электрон ресурслар

83. Bemova, A., Hajic J., Hladka B., Panevova J. Morphological and Syntactic Tagging of the Prague Dependency Treebank. <http://citeseer.nj.nec.com/296119.html>

84. Бабина О.И., Дюмин Н.Ю. Автоматизация лингвистической разметки корпуса текстов // <http://helling100.narod.ru/pubs/AutomationBabinaDyumin.pdf>
85. Богуславский И.М., Григорьев Н.В., Григорьева С.А., Иомдин Л.Л., Крейдлин Л.Г., Фрид Н.Е. Разработка синтаксически размеченного корпуса русского языка // http://cl.iitp.ru/bibiterns/corpus_SPB.pdf
86. Копотев М.В., Гурин Г.Б. Принципы синтаксической разметки хельсинского аннотированного корпуса русских текстов ХАНКО. (электрон ресурс): www.dialog-21.ru.
87. Кутузов А.Б. Корпусная лингвистика. – (Электрон ресурс): Лицензия Creative commons Attribution Share-Alike 3.0 Unported (Электрон ресурс) - [//lab314.brsu.by/kmp-lite/kmp-video/CL/CorporeLingva.pdf](http://lab314.brsu.by/kmp-lite/kmp-video/CL/CorporeLingva.pdf);
88. Недолужко А., Гаич Я. и др. *Синтаксически аннотированный корпус чешского языка. The Prague Dependency Treebank.* // <http://www.dialog-21.ru/dialog2008/materials/html/62.htm>
89. Плунгян В. Зачем мы делаем Национальный корпус русского языка? “Отечественные записки” 2005, №2. http://magazines.russ.ru/oz/2005/2/2005_2_20-pr.html
90. Рыков В.В. Курс лекций по корпусной лингвистике. URL: <http://rykov-cl.narod.ru/c.html>
91. Синклер Д. Предисловие к книге “Как использовать корпуса в преподавании иностранного языка”/ Д. Синклер [Электронный ресурс]. – Режим доступа: <http://www.ruscorpora.ru/corpora-infro.html>, свободный.

VIII. Интернет манбалари

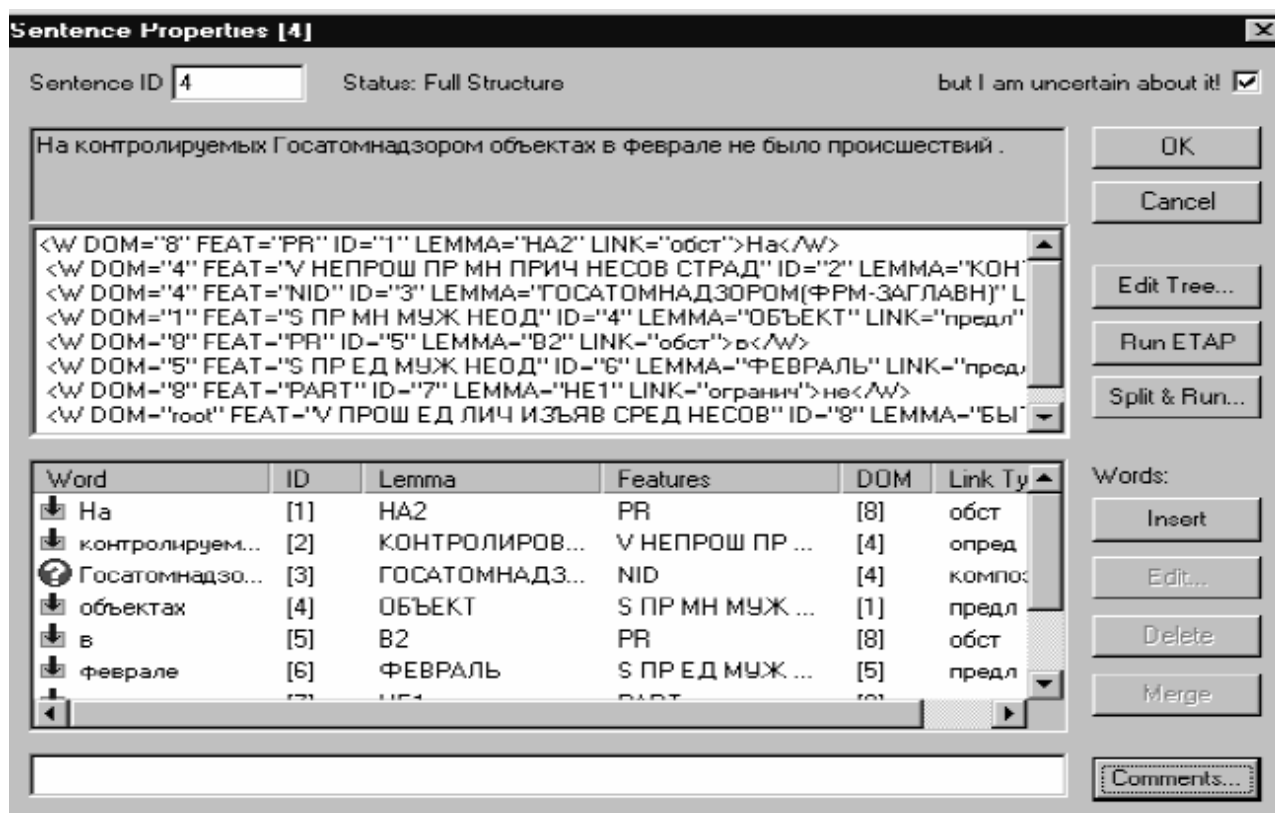
92. <http://aot.ru/demo/synt.html>
93. <http://corp.hum.sdu.dk/>
94. <http://en.wikipedia.org/wiki/Treebank>
95. <http://nlp.stanford.edu:8080/parser/index.jsp>.
96. <http://otipl.philol.msu.ru/~soiza/testsynt/files/info-parse.htm>

97. <http://ufal.mff.cuni.cz/pdt2.0/>
98. <http://www.dictum.ru/ru/syntax/blog>
99. <http://www.evalita.it/>
100. <http://www ldc.upenn.edu/ldc/online/treebank/>
101. <http://www.link.cs.cmu.edu/link/index.html>
102. <http://www.scs.leeds.ac.uk/amalgam/amalgam/amalghome.htm>
103. <https://uz.denemetr.com/docs/769/index-313952-1.html?page=30>
104. www.ruscorpor.ru
105. [http://nactem7.mib.man.ac.uk/geniatagger/.](http://nactem7.mib.man.ac.uk/geniatagger/)
106. <http://www.link.cs.cmu.edu/link/dict/introduction.html>
[http://www.link.cs.cmu.edu/link/submit-sentence-4.html.](http://www.link.cs.cmu.edu/link/submit-sentence-4.html)
107. <http://www.connexor.com/demo/>
108. <http://faculty.washington.edu/dillon/GramResources/GramResources.html>

ИЛОВАЛАР

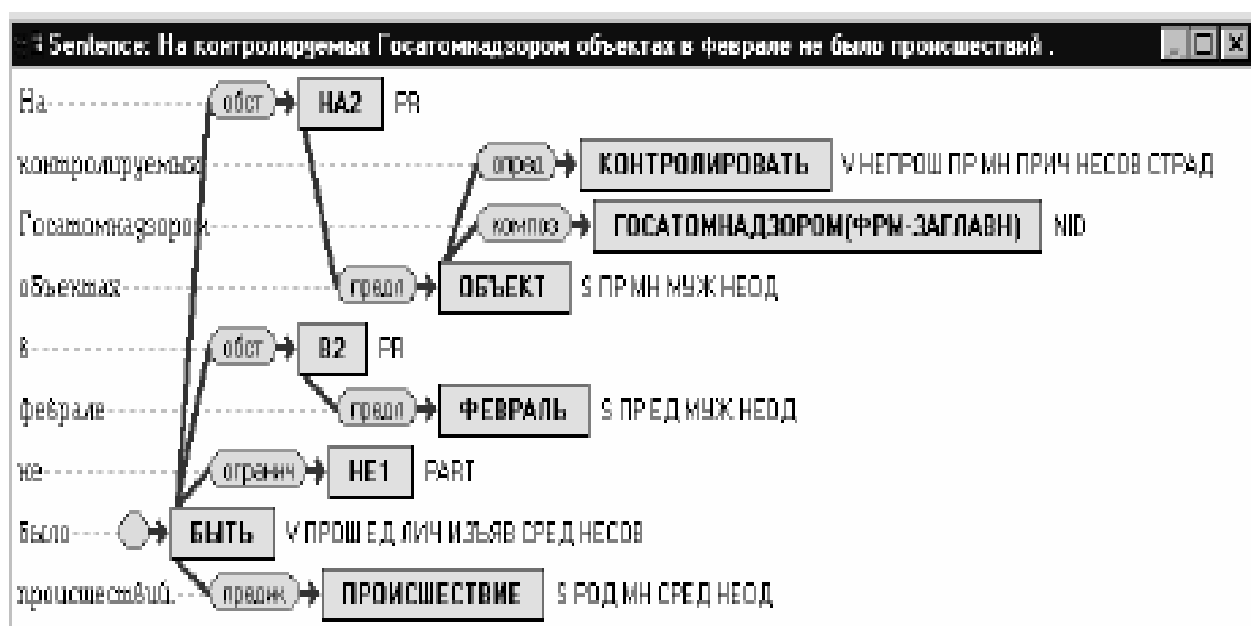
1-расм

ЭТАП-3 машина таржимаси тизими

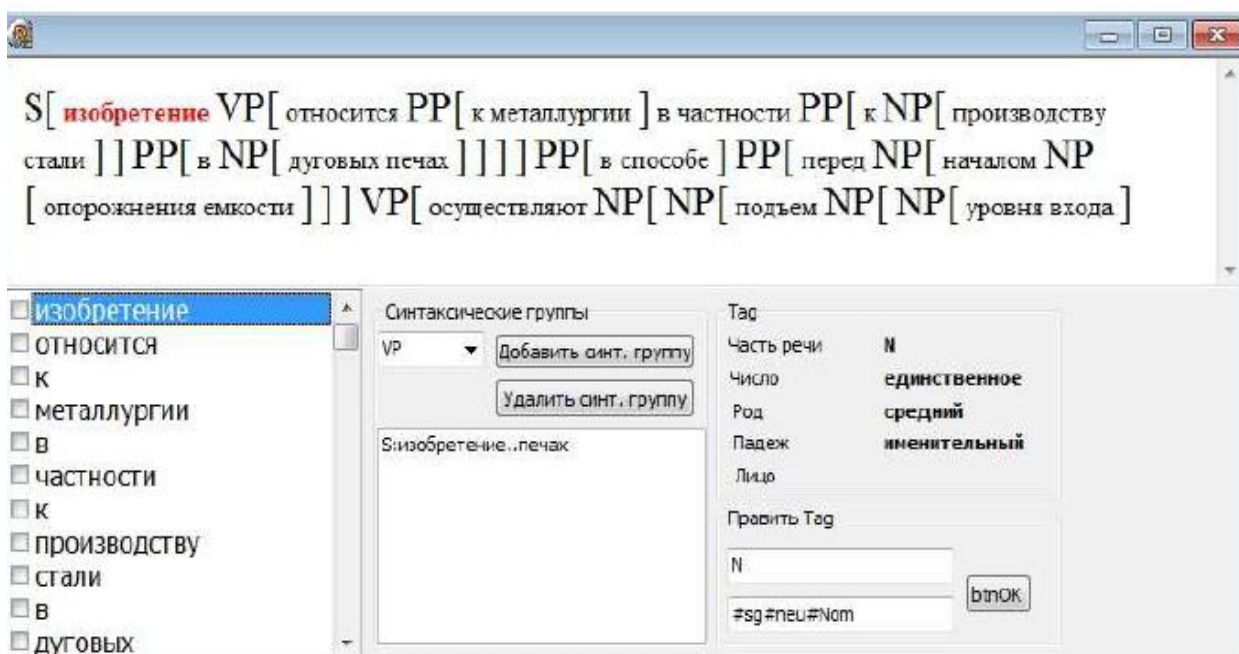


2-расм

Структуриани график тахирлаш



Автоматик синтактик разметка модули (SynTagger)



Исм+исм қолипидаги сўз бирикмаларини автоматик теглаш (парсинг) учун моделлар рўйхати

1. $[N^{Cs} \rightarrow N^{Pos}]$
2. $[N^{Cs} \rightarrow Adj^{Pos}]$
3. $[Adj^{Cs} \rightarrow N^{Pos}]$
4. $[Adj^{Cs} \rightarrow Adj^{Pos}]$
5. $[N \text{ турдош}^{Cs} \rightarrow Adj^{Pos}]$
6. $[Num^{Cs} \rightarrow Num^{Pos=}]$
7. $[N^{Cs} \rightarrow Ger^{Pos=}]$
8. $[Ger^{Cs} \rightarrow N^{Pos=}]$
9. $[Ger^{Cs} \rightarrow Ger^{Pos=}]$
10. $[N^{Cs} \rightarrow Adj\text{дош}^{Pos=}]$
11. $[Adj\text{дош}^{Cs} \rightarrow N^{Pos=}]$
12. $[N^{Cs} \rightarrow Prv^{Pos=}]$

13. $[\text{Pr}^{\text{Cs}} \rightarrow \text{N}^{\text{Pos}} =]$

14. $[\text{Prv}^{\text{Cs}} \rightarrow \text{N}^{\text{Pos}} =]$

Исм+феъл қолипли сўз бирикмаларини автоматик теглаш (парсинг) учун моделлар рўйхати

1) $[\text{Sub}^{\text{Cs3}} \rightarrow \text{V}]$;

2) $[\text{Sub}^{\text{Cs4}} \rightarrow \text{V}]$;

3) $[\text{Sub}^{\text{Cs5}} \rightarrow \text{V}]$;

4) $[\text{Sub}^{\text{Cs6}} \rightarrow \text{V}]$;

5) $[\text{N}^{\text{Cs3}} \rightarrow \text{V}]$;

6) $[\text{Adj}^{\text{Cs3}} \rightarrow \text{V}]$;

7) $[\text{Num}^{\text{Cs3}} \rightarrow \text{V}]$;

8) $[\text{Pr}^{\text{Cs3}} \rightarrow \text{V}]$;

9) $[\text{Ger}^{\text{Cs3}} \rightarrow \text{V}]$;

10) Cs3 = тушум ифодаловчи тег;

11) Cs4 = жўналиш ифодаловчи тег;

12) Cs5 = ўрин-пайт ифодаловчи тег;

13) Cs6 = чиқиш келишигини ифодаловчи тег;

14) Sub = исмни;

15) V = феълни ифодаловчи тег.

Содда гап ва унинг турларини ифодаловчи модел ва теглар рўйхати

“содда гап” = $\langle \text{СГ} \rangle, \langle / \text{СГ} \rangle$

“уюшган гап” = $\langle \text{УГ} \rangle, \langle / \text{УГ} \rangle$

“дарак гап” = $\langle \text{дг} \rangle$

“сўроқ гап” = $\langle \text{сг} \rangle$

“буйруқ гап” = $\langle \text{бг} \rangle$.

“эгали гап” = $\langle \text{Е+} \rangle$

“эгасиз гап” = $\langle \text{Е-} \rangle$

“шахси номаълум гап” = $\langle \text{ш.н.г} \rangle$

“атов гап” = $\langle \text{а.г} \rangle$

“семантик-функционал шаклланган гап” = <с.фш.г>

“йиғиқ гап” = <йг>

“ёйиқ гап” = <ёг>

Гап билан грамматик алоқага киришмайдиган бўлақларни ифодаловчи теглар

“ундалма” = <у>, </у>

“киритма” = <к>, </к>

Қўшма гаплар ва унинг турларини кўрсатувчи тег ва моделлар рўйхати

- 1) “боғланган ҚГ” = <БҚГ>, </БҚГ>;
- 2) “эргашган ҚГ” = <ЭҚГ >, </ЭҚГ>;
- 3) “боғловчисиз ҚГ” = <Б-сиз ҚГ>, </Б-сиз ҚГ>.

Боғланган қўшма гап турларини ифодаловчи теглар

- 5) бириктирув муносабатли БҚГ = <б.БҚГ>;
- 6) зидлов муносабатли БҚГ = <з.БҚГ>;
- 7) айирув муносабатли БҚГ = <а.БҚГ>;
- 8) инкор муносабатли БҚГ = <и.БҚГ>.

Боғловчисиз қўшма гап турларини ифодаловчи моделлар

- 1) [Wpm] <, > [Wpm] бўлса = <Б-сиз ҚГ>.
- 2) [Wpm] <:, > [Wpm] бўлса = <Б-сиз ҚГ>.
- 3) [Wpm] <-> [Wpm] бўлса = <Б-сиз ҚГ>.
- 4) [Wpm] <,> [Wpm] бўлса = <Б-сиз ҚГ>.

Эргашган қўшма гап турларини ифодаловчи теглар

- 1) кесим тобе гапли ҚГ = <КҚГ>;
- 2) эга тобе гапли ҚГ = <ЭҚГ>;
- 3) ҳол тобе гапли ҚГ = <ҲҚГ>;
- 4) тўлдирувчи тобе гапли ҚГ = <ТҚГ>;
- 5) аниқловчи тобе гапли ҚГ = <АҚГ>.

Эргашган қўшма гап турларини ифодаловчи моделлар

1. Агар: [eG] + <KQ>[bG] бўлса, [eG] + [bG] = <EQG> / <сабаб-EQG>.

2. Агар: [bG] + <...sin deb Q>[eG] бўлса, [eG] + [bG] = <EQG> / <мақсад-EQG>.
3. Агар: [eG...ca] + [bG] бўлса, [eG] + [bG] = <EQG> / <шарт-EQG>.
4. Агар: [bG\shu\shuni\shunga\shunday...ki] + [eG] бўлса, [bG] + [eG] = <EQG> / <шарт-EQG>.
5. Агар: [eG\ ким\ нима\ қанча\ қанчалик\ қандай\ қаер...sa] + [eG\ шу\ ўша\ шунча\ шунчалик\ шундай] бўлса, [eG] + [bG] = <EQG> / <nisbiy so`z-EQG>.
6. Агар: [eG] + <KQ>[bG] бўлса, [eG] + [bG] = <EQG> / <сабаб-EQG>.
7. Агар: [bG] + <...sin deb Q>[eG] бўлса, [eG] + [bG] = <EQG> / <мақсад-EQG>.
8. Агар: [eG...sa] + [bG] бўлса, [eG] + [bG] = <EQG> / <шарт-EQG>.
9. Агар: [bG\shu\shuni\shunga\shunday...ki] + [eG] бўлса, [bG] + [eG] = <EQG> / <шарт-EQG>.
10. Агар: [eG\ ким\ нима\ қанча\ қанчалик\ қандай\ қаер...sa] + [eG\ шу\ ўша\ шунча\ шунчалик\ шундай] бўлса, [eG] + [bG] = <EQG> / <nisbiy so`z-EQG>.
11. Агар: [eG] + <Ким...-са, у...>[bG] бўлса, [eG] + [bG] = <ega-EQG>.
12. Агар: [eG] + <Кимки...-са, у...>[bG] бўлса, [eG] + [bG] = <ega-EQG>.
13. Агар: [eG] + <Нимаики...-са, ўша...>[bG] бўлса, [eG] + [bG] = <ega-EQG>.
14. Агар: [eG] + <Шуниси...-ки, ...>[bG] бўлса, [eG] + [bG] = <ega-EQG>.
15. Агар: [eG] + <Ким...-са, ўзи...>[bG] бўлса, [eG] + [bG] = <ega-EQG>.
16. Агар: [eG] + <Кимда-ки...-са, у...>[bG] бўлса, [eG] + [bG] = <ega-EQG>.
17. Агар: [eG] + <Нима...-са, ўша...>[bG] бўлса, [eG] + [bG] = <ega-EQG>.
18. агар: [bG] + <шуки>[eG] бўлса, [eG] + [bG] = <kesim-EQG>;
19. агар: [bG] + < шуердаки>[eG] бўлса, [eG] + [bG] = <kesim-EQG>;
20. агар: [bG] + < шундаки>[eG] бўлса, [eG] + [bG] = <kesim-EQG>;
21. агар: [bG] + < шундайки>[eG] бўлса, [eG] + [bG] = <kesim-EQG>.
22. агар: <Shunday>[bG]<-ки> + [eG] бўлса, [eG] + [bG] = <aniqlovchi-EQG>.
23. агар: [bG]<Shunday><-ки> + [eG] бўлса, [eG] + [bG] = <aniqlovchi-EQG>.
24. агар: [bG]<Shunday><ни><-ки> + [eG] бўлса, [eG] + [bG] = <aniqlovchi-EQG>.

25. агар: <Shunday>[bG]<-ки> + [eG] бўлса, [eG] + [bG] = <aniqlovchi-EQG>.
26. агар: <Shunday>[bG]<-ки> + [eG] бўлса, [eG] + [bG] = <aniqlovchi-EQG>.
27. агар: <qaysi>[eG]<-са> + [bG] бўлса, [eG] + [bG] = <aniqlovchi-EQG>.
28. агар: [eG]<-иб> + [bG] бўлса, [eG] + [bG] = <ravish-EQG>.
29. агар: [eG]<-май> + [bG] бўлса, [eG] + [bG] = <ravish-EQG>.
30. агар: [eG]<-масдан> + [bG] бўлса, [eG] + [bG] = <ravish-EQG>.
31. агар: [eG]<-деб> + [bG] бўлса, [eG] + [bG] = <ravish-EQG>.
32. агар: <shuni>[bG]<-ки> + [eG] бўлса, [eG] + [bG] = <to`ldirovchi-EQG>.
33. агар: <shundan>[bG]<-ки> + [eG] бўлса, [eG] + [bG] = <to`ldirovchi -EQG>.
34. агар: <shunga>[bG]<-ки> + [eG] бўлса, [eG] + [bG] = <to`ldirovchi -EQG>.
35. агар: <nimani>[eG]<-са> + <shuni>[bG] бўлса, [eG] + [bG] = <to`ldirovchi -EQG>.
36. агар: <kimda>[eG]<-са> + <uni>[bG] бўлса, [eG] + [bG] = <to`ldirovchi -EQG>.
37. агар: <nima bilan>[eG]<-са> + <o`shani> [bG] бўлса, [eG] + [bG] = <to`ldirovchi -EQG>.
38. агар: [bG]<-ки> + [eG] бўлса, [eG] + [bG] = <to`ldirovchi -EQG>.
39. агар: [eG]<-ганда> + [bG] бўлса, [eG] + [bG] = <payt-EQG>.
40. агар: [eG]<-ган заҳоти> + [bG] бўлса, [eG] + [bG] = <payt -EQG>.
41. агар: [eG]<-ган вақтда \ ганпайтда> + [bG] бўлса, [eG] + [bG] = <payt -EQG>.
42. агар: [eG]<-гани замон> + [bG] бўлса, [eG] + [bG] = <payt -EQG>.
43. агар: [eG]<-гандан кейин> + [bG] бўлса, [eG] + [bG] = <payt -EQG>.
44. агар: [eG]<-гунга қадар> + [bG] бўлса, [eG] + [bG] = <payt -EQG>.
45. агар: [eG]<-ган эди \ ганича йўқ эди> + [bG] бўлса, [eG] + [bG] = <payt -EQG>.
46. агар: [eG]<-ар екан> + [bG] бўлса, [eG] + [bG] = <payt -EQG>.
47. агар: [eG]<-са> + [bG] бўлса, [eG] + [bG] = <payt -EQG>.
48. агар: [eG]<-масдан аввал> + [bG] бўлса, [eG] + [bG] = <payt -EQG>.
49. агар: [eG]<-иб> + [bG] бўлса, [eG] + [bG] = <payt -EQG>.

50. агар: [eG]<-гунча\ гач\ -май-> + [bG] бўлса, [eG] + [bG] = <payt -EQG>.
51. агар: [eG]<-иши билан(ок)> + [bG] бўлса, [eG] + [bG] = <payt -EQG>.
52. агар: [eG]<-шунинг учун\ шунинг учун...-ки \-ки> + [bG] бўлса, [eG] + [bG] = <sabab -EQG>;
53. агар: [eG]<- гани учун\-ганидан сўнг\ -ганидан сўнг> + [bG] бўлса, [eG] + [bG] = <sabab -EQG>;
54. агар: [eG]<-ми\> + [bG] бўлса, [eG] + [bG] = <sabab -EQG>;
55. агар: [eG]<- негаки\ сабабки> + [bG] бўлса, [eG] + [bG] = <sabab -EQG>;
56. агар: [eG]<-ганидан \-ганидан кейин\ шекилли\ -ган бўлса керак> + [bG] бўлса, [eG] + [bG] = <sabab -EQG>;
57. агар: [eG]<-чунки\ -масдан аввал \-гач \-май \-иб\ ми\> + [bG] бўлса, [eG] + [bG] = <sabab -EQG>.
58. агар: [eG]<-дея \- деб\ -син деб\ -ми экан деб\ -са деб > + [bG] бўлса, [eG] + [bG] = <maqsad-EQG>.
59. агар: [eG]<-учун> + [bG] бўлса, [eG] + [bG] = <maqsad-EQG>.
60. агар: [eG]<-токи> + [bG] бўлса, [eG] + [bG] = <maqsad-EQG>.
61. агар: <Агар>[eG]<-са\ -са эди\ -ганда\ -ганда эди\ -моқчи экан> + [bG] бўлса, [eG] + [bG] = <shart-EQG>.
62. агар: <Борди-ю >[eG]<-р/-ар> экан> + [bG] бўлса, [eG] + [bG] = <shart-EQG>.
63. агар: <Мабодо>[eG]<-ган экан> + [bG] бўлса, [eG] + [bG] = <shart-EQG>.
64. агар: [eG]<-ми\ бўлмаса\ йўқса\ -май\ -масдан аввал\ бўлай десанг\ ган экан> + [bG] бўлса, [eG] + [bG] = <shart-EQG>.
65. агар: <Қанча>[eG]<-са ҳам\ -са ҳамки> + [bG] бўлса, [eG] + [bG] = <to`siqsiz-EQG>.
66. агар: [eG]<қарамасдан\ -гани ҳолда\ -ганда ҳам\ ҳамки\ -са-да\ -гани билан\ -са/-маса\ -масин\ -са... -дики\ -қарамай\ -са> + [bG] бўлса, [eG] + [bG] = <to`siqsiz-EQG>.
67. агар: <Хох>[eG]<хох> + [bG] бўлса, [eG] + [bG] = <to`siqsiz-EQG>.

68. агар: <Қанчалик>[eG]< -ган сайин, шунчалик> + [bG] бўлса, [eG] + [bG] = <miqdor-daraja-EQG>;

69. агар: [eG]<қанча\ қанчалик ><- са, шунча \ шунчалик> + [bG] бўлса, [eG] + [bG] = <miqdor-daraja-EQG>.

70. агар: [eG]<каби \ -дай \ сингари \ янглиғ\ -ки, гўё \-ки, худди> + [bG] бўлса, [eG] + [bG] = <o`xshatish-daraja-EQG>;

71. агар: < Гўё>[eG] <-гандай > + [bG] бўлса, [eG] + [bG] = <o`xshatish-daraja-EQG>;

72. агар: [eG] < дегандек \ қандай....-са, шундай > + [bG] бўлса, [eG] + [bG] = <o`xshatish-daraja-EQG>;

73. агар: [eG] < гунча \ гандан кўра \ -ишдан кўра > + [bG] бўлса, [eG] + [bG] = <o`xshatish-daraja-EQG>.

74. агар: <шундай \ шу қадар \ шунчалик \шунча \шу даражада\ чунон> [eG]< -ки, -ки, натижада \ -ки, оқибатда> + [bG] бўлса, [eG] + [bG] = <natiya-daraja-EQG>.

ДИССЕРТАЦИЯДАГИ АТАМАЛАР ВА ШАРТЛИ ҚИСҚАРТМАЛАРНИНГ ҚИСҚАЧА ИЗОҲИ

Тег – лингвистик восита. “Тег” (ингл. tags) – мантдаги ишчи изоҳ, у матн ҳақидаги маълумотни қамраб олади.

Синтактик разметка – синтактик таҳлил, парсинг (ингл. parsing) натижаси, у компонентларининг грамматик структурасига асосланади.

Синтактик тег – синтактик разметка учун изоҳлар мажмуи.

Парсинг – автоматик тарзда синтактик таҳлил қиливчи дастур.

Разметка – изоҳ, тег.

Разметкаланган корпус – (индексланган, аннотацияланган; русча таггированные) – қўшимча изоҳнинг мавжудлигига кўра корпуснинг бир тури. Инглизча tag (изоҳ) сўзидан олинган.

Разметкаланмаган корпус - индексланмаган, аннотацияланмаган; русча нетаггированные) - қўшимча изоҳнинг мавжудлигига кўра корпуснинг бир тури.

Тил корпуси – маълум тилнинг белгиланган даврдаги, хилма-хил жанр, ранг-баранг услуб, ҳудудий ҳамда ижтимоий вариантдаги матнларининг электрон шаклдаги махсус дастурий таъминот асосидаги йиғиндиси.

Тегнинг лингвистик модели – тег қисқартмаси. Корпус таркибидаги сўзга бириктириладиган изоҳ шартли қисқартма шаклида бўлади, корпус тузувчи бу шартли қисқартмани ўзи танлайди. Масалан: аслий сифат = [ас.с.], нисбий сифат = [нис.с.]. разметка кўриниши: разметка кўриниши: **йирок** <йирок> = [сиф.], = [ас.с], [од.д], [с.с], [т.с], [хус. ЛМГ].

Электрон архив – электрон ташувчи қурилмадаги матнлар жамланмаси.

Эркин корпус – эркин фойдаланиш мумкин бўлган корпус; on-line режимда хоҳлаган пайтда исталган матннинг тўлиқ шакли билан ишлаш имконини беради.

SGML – матнни турли томонидан (alternative views) таҳлил қила оладиган разметка типи. В.П.Захаров SGML типи разметкасининг ўзига хос

имкониятидан бири матннинг турли кўринишини намоиш эта олиш эканлигини таъкидлайди.

Экстралингвистик разметка – матн ҳақидаги маълумотлардан иборат изох.

Шеърӣй матнлар корпуси – шеърӣй матнлардан таркиб топган корпус.

Частотали рӯйхат – интерфейсдаги корпусдаги сўзларнинг частотасига ишора берувчи ойна.

Анафорик разметка – “ишора сўзлар” га ёзиладиган изох.

Аралаш корпус – тилнинг барча услубига тегишли бўлган матнлар акс этган миллий корпуслар.

Браун корпуси (АҚШ) – 1 миллион сўз захирасига эга инглиз тили корпуси.

Динамик корпус – таркибининг тўлдирилиб борилишига кўра корпуснинг бир тури.

Ёзма корпус – оғзаки нутқ акс эттирилмаган (Браун корпуси, LOB), фақат ёзма матнлардан иборат корпус.

Ёпиқ корпус – ўзига хос тор (хусусий) мақсадда тузилган, оммавий фойдаланишга мўлжалланмаган корпус.

Идеографик разметка – идеографик луғат ёки тезаурус яратишга асос бўлувчи разметка кўриниши.

Иккинчи авлод корпуслари – электрон шаклдаги разметкаланган корпус.

Иллюстратив корпус – тадқиқот натижаси сифатида пайдо бўладиган корпус; унинг вазифаси фақат янги факт топиш эмас, балки у ёки бу тилга оид фактни аввалроқ бошқа усулда топилган мисолларни далиллашда лингвистик манба сифатида хизмат қилади.

Инкор қилувчи филтер (N-filter) – сўров билан мос келувчи сатр конкорданс рӯйхатидан ўчирилади.

Информацион сўров – маълум ахборотга бўлган эҳтиёжни сўз орқали ифодалаш усули.

Конкорданс – қидирилатган бирликнинг контекстдаги лексик боғланиш ҳолатида манбага илова қилинган контекстлар рўйхати.

Конкорданс қидируви – корпус интерфейсидаги сўз қидируви ойнаси; корпусда мавжуд сўзлар частотасини кўрсатади, уларнинг қўлланиш контекстига ишора қилади.

Корпус аннотацияси – турли лингвистик топшириқ амални бажариш учун матнга лингвистик ва экстралингвистик қўшимча маълумот билан ишлов бериш. Матн аннотацияси – матн ва унинг компонентларига берилган махсус изоҳ (масалан, сўз туркуми ҳақида маълумот). Бу изоҳ матн разметкаси (белги, ишора) деб ҳам аталади.

Корпус интерфейси – фойдаланувчини корпуснинг бутун мазмуни билан таништирувчи, мундарижа вазифасини бажарувчи муҳим таркибий қисм.

Корпус лингвистикаси – тилшуносликнинг корпус, унинг турлари, тарихи, тараққиёти, бугунги ҳолати, корпус тузиш тамойилларини ўрганувчи бўлими.

Корпус материали – корпус таркибидаги матнлар йиғиндиси.

Корпус менежери – лингвистик корпуснинг ажралмас, энг асосий қисми, унинг қидирув тизими, у матн ва лисоний бирликларни бошқарувчи система.

Корпус разметкаси – корпус матнларига бериладиган лингвистик ҳамда экстралингвистик изоҳ.

Корпус фойдаланувчиси – корпусдан фойдаланадиган мутахассислар: тадқиқотчи, филолог, журналист, ўқитувчи, талаба, мактаб ўқувчиси, программист.

Корпус фрагменти – яратилиш босқичида бўлган, тугалланмаган, аммо маълум қисмидаги фойдаланиш мумкин бўлган корпус қисми.

Корпус ҳажми – корпус матнларининг XML тилидаги ҳажми, Мбда ҳисобланади.

Кўп тилли корпус – икки ёки уч тили корпус.

Лемма – корпус матнлардаги лексаманинг корпус лингвистикасида номланиши, лексема.

Лемматизация – сўз шаклининг лемма шаклини ажратиб кўрсатиш.

Лингвистик разметка – матн таркибидаги сўз ёки гапга ёзилган лингвистик изоҳ. Унинг ҳам ўз навбатида морфологик, синтактик, семантик, просодик разметка каби бир канча кўринишлари мавжуд.

Матнлар корпуси – стандартланган, бир хил шаклга келтирилган; матнлари лингвистик воқеликнинг маълум қисмини акс эттиришга мўлжалланган.

Матннинг автоматик разметкаси – парсинг ва таггинг дастурлари томонидан матнни автоматик разметкалаш.

Миллий корпус – жанридан қатъи назар тилнинг миллий бойликлари жамланган электрон қидирув тизими.

Морфологик разметка – синтактик ва семантик разметкага асос бўлувчи, инглиз тилида part-of-speech tagging деб аталувчи, сўзларни туркумга ажратувчи аҳамиятли разметка.

Морфологик тег – морфологик разметка учун изоҳлар мажмуи.

Муаллифлик корпуси – корпус материали алоҳида муаллиф асарларидан иборат корпус.

Оғзаки корпус – фақат оғзаки матндан иборат корпус.

Параллел корпус – бир мавзуга оид икки тилдаги матнли корпус.

Публицистик корпус – публицистик матнлар йиғиндиси.

Таггинг – автоматик морфологик разметкалаш дастури.

Таълимий корпус – хорижий ва она тили таълими воситаси сифатида қараладиган, тил ўрганиш ва ўргатиш усул, имкониятларига эга корпус.

Хельсинки аннотацияланган корпуси – (Ханко) – 2000 йилларда Хельсинки университетида “Рус тилининг функционал синтаксиси” (раҳбар проф. А. Мустайоки) лойиҳасининг бир қисми сифатида иш бошлаган, доимий тўлдириб борилаётган корпус.

Шартли қисқартмалар

N – от

Adj – сифат

Num – сон

N^{sub} – турдош от

Pr – олмош

Prv – равиш

Ger – ҳаракат номи

Case (ёки Cs) – қаратқич келишиги

Possessive (ёки Pos) – эгалик қўшимчаси

ЛСҚ – лисоний синтактик қолип

WP_m – содда гаплар қолипи

<СГ> - содда гап

<ҚГ> - қўшма гап

<УГ> - уюшган гап

<ДГ> - дарак гап

<СГ> - сўроқ гап

<БГ> - буйруқ гап

<Е+> - эгали гап

<Е-> - эгасиз гап

<ш.н.г.> - шахси номаълум гап

<а.г> - атов гап

<йг> - йиғиқ гап

<ёг> ёйиқ гап

<у> - ундалма гап

<к> - киритма сўз

<БҚГ> - боғланган қўшма гап

<ЭҚГ> - эргашган қўшма гап

<Б-сиз ҚГ> - боғлочисиз қўшма гап

<б.БҚГ> - бириктирув муносабатли боғланган қўшма гап

<з.БҚГ> - зидлов муносабатли боғланган қўшма гап

<а.БҚГ> - айирув муносабатли боғланган қўшма гап

<и.БҚГ> - инкор муносабатли боғланган қўшма гап

<еG> - эргаш гап

<bG> - бош гап

<EQG> - эргашган қўшма гап