

## TEZAURUS TUSHUNCHASI VA ILMIY YONDASHUV

**Ubaydullayeva Ziroatxon**

Mirzo Ulugʻbek nomidagi Oʻzbekiston

Milliy universiteti magistranti.

E-mail: [ubaydullayevaziroatxon9@gmail.com](mailto:ubaydullayevaziroatxon9@gmail.com)

**Annotatsiya:** Ushbu maqolada “tezaurus” tushunchasi, uning kelib chiqishi va oʻrganishga boʻlgan turlicha qarashlar ilmiy yondashuvlar asosida ochib berilgan. Tezaurusning hayotning har bir jabhasida, har bir sohada qanchalik muhim ahamiyat kasb etishi toʻgʻrisida fikrlar bayon etilgan.

**Kalit soʻzlar:** tezaurus, indekslash, iyerarxik, adekvat, assotsiativ, sinonimik, anaforistik, elliptik, konvolyutsiya.

**Abstract:** In this article, the concept of “thesaurus”, its origin and different views on learning are revealed based on scientific approaches. Opinions are expressed about how important the thesaurus is in every aspect of life, in every field.

**Keywords:** thesaurus, indexing, hierarchical, adequate, associative, synonymous, anaphoristic, elliptical, convolution.

**Аннотация:** В данной статье раскрываются понятие «тезаурус», его происхождение и различные взгляды на обучение, основанные на научных подходах. Высказываются мнения о том, насколько важен тезаурус во всех аспектах жизни, в каждой области.

**Ключевые слова:** тезаурус, индексация, иерархическая, адекватная, ассоциативная, синонимическая, анафористическая, эллиптическая, свертка.

Tezaurus tarixi haqidagi dastlabki maʼlumot Alan Gilchristning 1971-yilda chop etilgan “Tezaurus in Retrieval” asarida uchraydi hamda muallif bu soʻzni yunon tilidan olinganligini taʼkidlaydi. Shorter Oksford Dictionary lugʻatiga koʻra, ingliz tilida birinchi marta 1736-yilda paydo boʻlgan va “lugʻat sifatidagi bilimlar xazinasi yoki ombori” maʼnosini bildirgan. Eng mashhur tezaurus 1852-yilda Peter Mark Roget tomonidan ishlab chiqilib, “Thesaurus of English Words and Phrases” deb nomlangan. Bu alifbo tartibida emas, balki tizimli ravishda, soʻzlar ifodalaydigan tushunchalarga koʻra va fikrga eng mos keladigan soʻz yoki soʻzlarni topishga qaratilgan. Taxminan bir asr oʻtgach, bu atama maʼlumot qidirishda ishlatiladigan lugʻat roʻyxatlariga nisbatan qoʻllanila boshlandi [Abdurakhmonova, N.2020,2021,2022].

Maʼlumot qidirish maʼnosida tezaurus hozirgi xalqaro standart ISO 2788da “boshqariladigan indekslash tilining lugʻati” boʻlib, unda tushunchalar oʻrtasidagi munosabatlar aniq ifodalab, indekslovchi ishlatadigan lugʻatni qidiruvchi tiliga moslashtirish va shu bilan tegishli hujjatlarni qidirishni yaxshilaydi. Dastlab, tezaurus kordinatalarini indekslash va kordinatalarini indekslashdan foydalanib qidirishda qoʻllash, shuningdek, qoʻlda postkordinatalar tizimi yordamida indekslash Batten, Calvin Mooers va boshqalar tomonidan ishlab chiqilgan. Postkordinatalar tizimida qidiruvchi aniq bir soʻzni topish uchun soʻzlarning bir qancha kombinatsiyalaridan foydalanishi mumkin [Abdurakhmonova, N.2020,2021,2022].

Shuningdek, tezaurus 1950-yillar davomida turli shakllarda, shu jumladan, Taube tomonidan ishlab chiqilgan Uniterm tizimida rivojlandi. Bunda hujjat matnidan erkin saylash yo‘li bilan, bir marta bo‘lsa ham ishlatilgan so‘zlar olingan. Albatta, sinonimlar, omonimlar, turkumlarni qidirish va boshqa muammolarni hal qilish uchun faqat bitta so‘zli atamalar mavjud bo‘lgan vaziyatda qiyinchiliklar yuzaga keldi. Unitermlar ko‘pincha kontekstga, qo‘llanilishiga qarab turli xil ma‘nolarda ishlatiladi, biroq filtr yo‘qligi sababli ular bu o‘ziga xos farqlarni ifoda eta olmaydi. Misol uchun, izolyatsiya termini orqali ifodalangan “issiqlik izolyatsiyasi” tushunchasi “elektr izolyatsiyasi” dan juda katta farq qiladi va bir so‘zdan iborat “izolyatsiya” atamasi ularni ajratish uchun yetarli bo‘lmaydi. Unifikatsiyalashgan terminlar yetarli emasligi sababli, sof shakldagi tizim uzoq davom etmaydi. Oxir-oqibat, uning o‘rnini ko‘p sonli “assotsiatsiyalangan” yoki qo‘shma terminlarni o‘z ichiga olgan lug‘atlar egalladi. Bu ilk tezauruslar sinonimlar va omograflarni nazorat qila boshladi, atamalar o‘rtasidagi iyerarxik va assotsiativ munosabatlarni ko‘rsata boshladi.[J.Aitchison, SD.Clarke. 2004:]

Shu bilan birga, Adam Kilgarriff va Colin Yallop ham o‘zining maqolasida [Adam Kilgarriff, Colin Yallop. 2000:5] tezaurusning to‘rt turini tafsiflab, ular uchun umumiy bo‘lgan xususiyat ularning iyerarxik yoki tarmoq tashkil etilishini asoslab berishgan.

O‘zbekiston Milliy Ensiklopediyasida [O‘zbekiston Milliy Ensiklopediyasi, 2004: 329] tezaurus terminiga quyidagicha izoh berilgan: tezaurus (yun. thesaurus – xazina, boylik) – 1) muayyan tildagi barcha so‘zlarni qamrab oladigan, ularning matnda qo‘llanish holatlarini to‘la-to‘kis aks ettiradigan lug‘atdir.

Avvalambor, rezausus termini tilning lug‘at fondini maksimal darajada to‘liqlik bilan ifodalaydigan xazina sifatida talqin etiladi. Matnlarning adekvat talqini va ma‘lumotlari uning sistem ma‘nolarini o‘zida mujassam etgan va soha vakillari tomonidan e‘tirof etilgan tezaurusda aniqlanadi. [Valitova, 2012: 19]. Tezausus – bu umumiy ma‘noda maxsus bilim sohasi yoki faoliyat sohasining tushunchalari, ta‘riflari, atama yoki terminlarini to‘liq qamrab oladigan maxsus terminologiya, yanada qat‘iy va obyektiv ravishda lug‘at, ma‘lumotlar to‘plami, korpus yoki kod demakdir.

Jumladan, tezaurus atamasini L.V. Shcherba “Опыт теории общей лексикографии” maqolasida [Л.В. Щерба,1940:10] tezaurus deyishganda, hozirgi paytda ko‘pincha “Thesaurus linguae latinae”ni nazarda tutishini, aslida u beshta nemis akademiyasining hamkorlik faoliyati bo‘lib, 1900-yilda boshlanganini aytib, muhim xarakterli xususiyati xususida ma‘lum bir tilda ro‘y bergan barcha muhim so‘zlarni o‘z ichiga olib, har bir so‘z ostida, ushbu tilda mavjud bo‘lgan matnlardan misollar keltiriladi. Olimlar A.N. Baranov va D.O. Dobrovolskiy [Добровольский, Баранова.1965: 4] tezaurusga quyidagi ta‘rifni bergan: “Maxsus tur lingvistik materialni tartibga solish usuli bilan boshqalardan (xususan, tushuntirish, ikki tilli va hokazo) farq qiladigan lug‘atdir. Tezaususda til birliklari oddiy lug‘atda bo‘lgani kabi alifbo tartibida berilmagan, lekin ularning ma‘nosiga qarab guruhlangan”. Shuningdek, axborot-qidiruv tezauruslarini qo‘llash sohasida ikkita asosiy yo‘nalish ajratiladi: matnning semantik ma‘nosidan foydalanadigan ma‘lum qoidalardan foydalangan holda hujjatlarni indekslash va foydalanuvchilarning qidiruv so‘rovlarini qayta ishlashda ierarxik, assotsiativ va sinonimik) foydalanishdir. Bundan tashqari, tavsiflovchilar orasidagi semantik aloqalar :

- hujjatlarni tasniflash va turkumlashtirish;
- so‘rov bilan bog‘liq so‘zlar ro‘yxatini tuzish:

-ma'lumotlarni qidirish vazifalari uchun ishlatilishi mumkin. Uning atamalar to'plami matnda qo'llaniladigan barcha terminologik iboralarni o'z ichiga olishi kerak. Ushbu usulni amaliy amalga oshirish uchun tezaurusda ko'pchilik atamalarning borligi yetarli bo'lib, har qanday mavzu sohasining tabiiy til tavsifi matn konvolyutsiyasi uchun ishlatiladigan ba'zi konstruksiyalar bilan tavsiflanadi, masalan, bir xil atamalar, anaforistik va elliptik konvolyutsiyalar. Omonimiya qo'shimcha muammo yaratadi hamda bir xil so'zlar anaforistik va elliptik konvolyutsiyalar bilan bog'liqdir. Ba'zi muammolarni asl tezaurusga sinonimlarning keng bazasini kiritish orqali juda samarali hal qilish mumkin, omonimiya masalalarini hal qilish bilan bog'liq muammolarni faqat statistik tahlil yordamida samarali hal qilish mumkin. [Aktaev, 2015:4]

Xulosa qilib shuni aytish mumkinki, tezaurus tushunchasi haqidagi ilmiy yondashuvlar tahlilidan tezauruslarni yaratish usullari va vositalari sohaning shu kungacha yaxshi o'rganganligini, tadqiqotchilar amaliyotda yaxshi natija beradigan yangi modellarni, yangicha yondashuv va qarashlarni ilgari surganligi ko'rinadi.

### Foydalanilgan adabiyotlar ro'yxati:

1. Kilgarriff, C. Yallop. What's in a Thesaurus? International Conference on Language Resources and Evaluation, 2000.
2. O'zbekiston Milliy Ensiklopediyasi.- Toshkent: 2004.
3. Agostini, A., Usmanov, T., Khamdamov, U., Abdurakhmonova, N., & Mamasaidov, M. (2021, January). Uzwordnet: A lexical-semantic database for the uzbek language. In Proceedings of the 11th Global Wordnet conference (pp. 8-19).
4. Abdurakhmonova, N., Alisher, I., & Toirova, G. (2022, September). Applying Web Crawler Technologies for Compiling Parallel Corpora as one Stage of Natural Language Processing. In 2022 7th International Conference on Computer Science and Engineering (UBMK) (pp. 73-75). IEEE.
5. Abdurakhmonova, N., & Ismailov, A. S. (2022). APPLYING WEB CRAWLER TECHNOLOGIES FOR COMPILING PARALLEL CORPORA AS ONE STAGE OF NATURAL LANGUAGE PROCESSING. In СОВРЕМЕННАЯ ФИЛОЛОГИЯ. СОЦИАЛЬНАЯ И НАЦИОНАЛЬНАЯ ВАРИАТИВНОСТЬ ЯЗЫКА И ЛИТЕРАТУРЫ (pp. 22-27).
6. Abduraxmonova, N., & Abduvaxobov, G. I. (2021). O'quv lug'atini tuzishning nazariy metodologik asoslari. СЎЗ САНЪАТИ ХАЛҚАРО ЖУРНАЛИ, 103.
7. Sulevmanov, D., Gatiatullin, A., Prokopyev, N., & Abdurakhmonova, N. (2020, November). Turkic morpheme web portal as a platform for turkology research. In 2020 International Conference on Information Science and Communications Technologies (ICISCT) (pp. 1-5). IEEE.
8. E. Aktaev. O'quv va axborot jarayoniga avtomatlashtirilgan tezaurusning zamonaviy usullarini joriy etish. Yosh olim- 2015.- B. 3-6.
9. Л.В. Щерба. Опыт теории общей лексикографии. 1940.
10. В. Добровольский. От редакции. С. Баранова. От составителя. - М., 1965. С.4-6
11. N. Valitova. Формирование профессионального тезауруса у студентов физкультурного вуза. Автореферат на соискание учёной степени канд.пед.наук.-Омск, 2012. -В.19.