



Marjona AMINOVA,
Fan va texnologiyalar universiteti o'qituvchisi
E-mail: aminovamarjona@mail.uz

Fan va texnologiyalar universiteti dotsenti, PhD H.Ortiqova taqrizi asosida

LEXICAL-STATISTICAL METHODS OF TEXT ANALYSIS

Annotation

This article provides information on lexical-statistical methods of text analysis. The methods of psychological examination of the text are taught. This analysis provides valuable information about the content, tone, and meaning of the text, as well as the language and style of writing. Statistical analysis is a universal method, as it is used equally in all layers of the language. Lexical-statistical analysis helps to study the lexical and grammatical forms of the text. In this method, the appearance indicators of each word in the text, that is, the information, meanings, and concepts related to them, are analyzed.

Key words: Text, subtext, psychological analysis, process, linguistic expertise.

ЛЕКСИКО-СТАТИСТИЧЕСКИЕ МЕТОДЫ АНАЛИЗА ТЕКСТА

Аннотация

В данной статье представлена информация о лексико-статистических методах анализа текста. Преподают методы психологического исследования текста. Этот анализ дает ценную информацию о содержании, тоне и значении текста, а также о языке и стиле письма. Статистический анализ является универсальным методом, поскольку одинаково используется во всех слоях языка. Лексико-статистический анализ помогает изучить лексические и грамматические формы текста. В этом методе анализируются показатели появления каждого слова в тексте, то есть связанная с ними информация, значения и понятия.

Ключевые слова: Текст, подтекст, психологический анализ, процесс, лингвистическая экспертиза.

MATN TAHLIL QILISHNING LEKSIK-STATISTIK USULLARI

Аннотация

Ushbu maqolada matnни tahlil qilishning leksik-statistik usullari haqida ma'lumot keltirilgan. Matnни psixologik ekspertiza qilish usullari o'rgatilgan. Bu tahlil matnning mazmuni, ohangi va ma'nosi, shuningdek, tilning o'ziga xosligi va yozilish uslubi haqida qimmatli ma'lumotlar beradi. Statistik tahlil tilning barcha qatlamlarida birdek qo'llanishi sababli universal uslub hisoblanadi. Leksik-statistik tahlil, matnning leksik va grammatik ko'rinishlarini o'rganishga yordam beradi. Ushbu usulda, matndagi har bir so'zning ko'rinish ko'rsatkichlari, ya'ni, ularga oid ma'lumotlar, ma'nolari, tushunchalari, tahlil qilinadi.

Kalit so'zlar: Matn, submatn, psixologik tahlil, jarayon, lingvistik ekspertiza.

Kirish. Leksik-statistik tahlil – bu matnlarni yoki nutqni analiz qilishda qo'llaniladigan usul bo'lib, asosan so'zlarning tarqalishi, ulardan qanday ishlatilish xususiyatlari, va o'zaro bog'liqliklarni aniqlashga qaratilgan. Ushbu tahlil, matndagi so'zlar va iboralar o'rtasidagi aloqalarni, ularning chastotasini, tez-tez uchraydigan so'z birikmalarini va stilistik xususiyatlarini o'rganadi. Leksik-statistik tahlil yordamida matnning struktura va mazmunini yaxshiroq tushunish, maqsadli auditoriya bilan qanday aloqalar o'rnatilishini aniqlash va tilning kommunikativ xususiyatlarini o'rganish mumkin.

Bunday tahlilning asosiy vazifalari orasida matndagi so'zlar va frazeologizmlarning vaqti-vaqti bilan qanday ishlatilishini va qanday kontekstlarda ularning ma'nosi o'zgarishini aniqlash bor. Shu bilan birga, leksik-statistik tahlil matndagi so'z turkumlarini, ular orasidagi aloqalarni va tilning pragmatik jihatlarini o'rganish orqali kommunikativ jarayonni chuqurroq tushunishga yordam beradi. Ushbu yondashuv, masalan, adabiyotshunoslikda, lingvistika va nutq tahlilida, shuningdek, tilni ishlatishning zamonaviy usullarini o'rganishda samarali qo'llaniladi.

Mavzuga oid adabiyotlar tahlili. Leksik-statistik tahlil bo'yicha adabiyotlar juda keng va turli sohalarida, xususan, lingvistika, psixolingvistika, kommunikatsiya, adabiyotshunoslik kabi yo'nalishlarda mavjud. Ushbu adabiyotlar, asosan, matnlarni tahlil qilishda, so'zlarning o'zaro aloqalari, ularning ta'siri va kontekstdagi o'rinlarini aniqlashga yordam beradi. Quyida leksik-statistik tahlil bo'yicha bir nechta asosiy adabiyotlar va ularning mazmunini tahlil qilib beraman:

John L. V. McCreeryning "Statistical Methods for Linguistics" asari lingvistik tahlil uchun statistik metodlarni o'rgatadi va ular qanday qilib tilshunoslikda qo'llanishi mumkinligini tushuntiradi. Kitobda statistik usullarni, jumladan,

korrelyatsiya, regressiya, va klasterlash kabi texnikalarni tilshunoslik matnlarida qo'llash misollari bilan keltiradi. Tahlil jarayonida, so'zlar o'rtasidagi o'zaro aloqalar va til strukturasi qanday statistik ko'rsatkichlar bilan o'lchanishi mumkinligi yoritiladi. Kitobda tahlil qilingan metodlar zamonaviy tilshunoslikda tilni o'rganishning yangi usullari sifatida ahamiyatga ega. McCreery lingvistik tahlilda statistik usullarni integratsiyalashning qiyinchiliklari va imkoniyatlarini ko'rsatadi, bu esa leksik-statistik tahlilni kengroq tushunishga yordam beradi.

Rusan Mitkovning "The Oxford Handbook of Computational Linguistics" kitobi, tilshunoslik va kompyuter fanlarining kesishgan sohalariga oid eng muhim adabiyotlardan biri hisoblanadi. Kitobda kompyuter lingvistikasi, statistik til tahlili, va leksik analitika kabi mavzular bo'yicha ko'plab maqolalar mavjud. Ayniqsa, matnни tahlil qilishda statistik metodlardan foydalanish, masalan, so'zlarning chastotasini hisoblash, semantik aloqalarni aniqlash, va til modellari yaratish kabi masalalar batafsil bayon etilgan. Ushbu manba kompyuter lingvistikasi va statistik tahlilning amaliy jihatlarini muhokama qilishda juda foydalidir. Leksik-statistik tahlilni zamonaviy texnologiyalar bilan uyg'unlashtirish, xususan, mashina o'rganish va sun'iy intellekt texnologiyalarini qo'llash bo'yicha yaxshi qo'llanma bo'lishi mumkin.

McEneary va Hardie ning asari leksik-statistik tahlilni o'rganishni xohlaganlarning ehtiyojlarini qondiradi, chunki korpus metodologiyasi so'zlarni kontekstda tahlil qilish uchun eng samarali yondashuvlardan biridir. Bu metod tilning tabiiy rivojlanishidagi o'zgarishlarni tahlil qilishda, shuningdek, ma'lum bir til va madaniyatga xos xususiyatlarni aniqlashda muhim rol o'ynaydi. Matnlarni statistik tahlil qilishning eng yaxshi metodlari taqdim etilgan. Asar tilshunoslikda statistika metodlarini qo'llashda ko'plab misollarni o'z ichiga oladi, masalan, matndagi

so'zlarning taqsimoti, semantik tahlil, va leksik aloqalar. Leye matnlar orqali tilni yaxshiroq tushunish va tahlil qilish uchun zarur bo'lgan statistik vositalarni taqdim etadi.

Leksik-statistik tahlil o'tkazish, matni tasniflash, semantik to'plamlarni yaratish, tilni o'rganishda yordamlashish, yoki eventual ravishda mashinasozlik va sun'iy intellektga asoslangan tizimlarda (masalan, sentiment analizi, avtomatik tarjima) qo'llanishi mumkin. Shuningdek, bu tahlil kommunikatsiya jarayonlarini, ijtimoiy tarmoqlarda yuzaga keladigan yangi so'zlar va iboralarning tezkor tarqalishini o'rganishda ham keng foydalaniladi. Matnning leksik statistik tahlili deganda matni alohida birliklar yoki leksemalarga, masalan, so'z yoki iboralarga bo'lish, so'ngra bu birliklarning paydo bo'lish chastotasini tekshirish yo'li bilan tahlil qilish jarayonidir.

Tadqiqot metodologiyasi. So'z chastotasini tahlil qilish berilgan matnda qaysi so'zlar ko'proq ishlatilishini aniqlash jarayonidir. Ushbu tahlil matnning asosiy mavzulari va mavzularini ochishga yordam beradi. Masalan, Yevropa tarixi haqidagi kitobda so'z chastotasi tahlili o'tkazilsa, eng ko'p ishlatiladigan so'zlar "Yevropa", "tarix", "madaniyat" va "siyosat" ekanligini aniqlashi mumkin. Ushbu ma'lumot o'quvchiga kitobning asosiy yo'nalishini tushunishga yordam beradi. Bundan tashqari, so'z chastotasi tahlili tildan foydalanishning qiziqarli naqshlarini aniqlashi mumkin. Misol uchun, agar kimdir romandagi turli qahramonlar tomonidan qo'llaniladigan so'zlarning chastotasini tahlil qiladigan bo'lsa, unda bir qahramon uzoqroq va murakkabroq so'zlarni ishlatishga moyil bo'lsa, boshqa bir qahramon oddiyroq va tushunarliroq tildan foydalanadi. Masalan, yangilik maqolalarining leksik-statistik tahlili shuni ko'rsatadiki, eng ko'p qo'llaniladigan so'zlar "prezident", "hukumat", "siyosat" va "iqtisod".

Siyosiy matnlarda chastota tahlilini o'tkazish matnda eng ko'p uchraydigan so'z yoki iboralarni aniqlash va hisoblashni o'z ichiga oladi. Ushbu uslub muayyan siyosiy nutqda muhokama qilinadigan asosiy mavzular, muammolar va g'oyalarni aniqlashga yordam beradi.

Chastotali tahlilni amalga oshirish uchun bir necha qadamlar:

Matni to'plash: nutqlar, maqolalar yoki siyosiy partiyalar manifestlari kabi siyosiy matn yoki matnlar to'plamini tanlash.

Matni tozalash va qayta ishlash: To'xtash so'zlari (masalan, "va"), tinish belgilari va maxsus belgilar kabi ahamiyatsiz ma'lumotlarni olib tashlash.

Tahlil va natijalar. Chastotani tahlil qilish o'quvchiga nutqda aytilgan asosiy mavzular va g'oyalarni tushunishga yordam beradi, bu ma'ruzachining muayyan masalalar bo'yicha pozitsiyasini izohlashda foydali bo'lishi mumkin.

Kalit so'zlarni tahlil qilish: Bu usul matndagi eng muhim yoki tez-tez ishlatiladigan kalit so'zlarni aniqlashni o'z ichiga oladi. Masalan, veb-saytning leksik statistik tahlili "sog'liq", "fitness" va "oziqlanish" kalit so'zlari eng ko'p ishlatilishini ko'rsatishi mumkin. Matndagi izchillik tahlili matndagi bog'lanish va birlik darajasini bildiradi. Bu matnning mo'ljallangan ma'nosini yetkazish va g'oyalar va tushunchalar o'rtasidagi mantiqiy munosabatlarni saqlab turish qobiliyatidir. Agar matnda keltirilgan fikrlar mantiqiy tartibga solingan bo'lsa va o'tish moslamalari mohirlik bilan qo'llanilsa, matn izchil deb hisoblanadi. Misol quyidagicha bo'lishi mumkin: "Yertalab men

qahva ichaman, nonushta qilaman va elektron pochta xabarlarimni tekshiraman. Shundan so'ng men ishga ketaman". Bu parcha xronologik tartibga solinganligi sababli izchil va voqealar ketma-ketligi mantiqiy ma'noga ega. "Yertalab" va "keyin" kabi o'tish moslamalaridan foydalanish o'quvchilarga yozuvchining mo'ljallangan ma'nosini kuzatishga yordam beradi.

Kollokatsiyani tahlil qilish: Bu usul matnda so'zlarning birgalikda yuzaga kelish yo'llarini aniqlash va tahlil qilishni o'z ichiga oladi. Masalan, romanning leksik statistik tahlili shuni ko'rsatadiki, "sevgi" va "yurak" so'zlari tez-tez bir-biriga mos keladi. Matn birikmasini tahlil qilish - matn yoki korpusda tez-tez uchraydigan so'z birikmalarini aniqlashning lingvistik usulidir. Bu ma'lum bir kontekstda so'zlarning birgalikda ishlatilishini tushunishga yordam beradi. Kollokatsiyalar tasodifan kutilganidan ko'ra tez-tez birga paydo bo'ladigan so'zlardir. Misol uchun "o'tkir qahva" - bu birikma, chunki "o'tkir" va "qahva" so'zlari tasodifiy tasodifdan ko'ra ko'proq birga keladi. 1-misol: tibbiy matnda "yurak xuruji" so'zining birikmasi "yurak urishi" dan ko'ra ko'proq uchraydi. Bu shuni ko'rsatadiki, yurak xuruji yurak urishidan ko'ra tez-tez uchraydigan tibbiy holatdir. 2-misol: badiiy matnda qo'rqinchli muhit yaratish uchun "qorong'u o'rmon" yoki "arvoh qal'a" so'zlarining birikmasi tez-tez ishlatilishi mumkin. Bu muallif o'quvchida hissiy munosabatni yaratish uchun qanday qilib aniq so'z birikmalaridan foydalanilganligini ko'rsatadi.

Muvofiqlik tahlili: Bu usul matnda so'z yoki ibora paydo bo'lgan kontekstni tahlil qilishni o'z ichiga oladi. Masalan, leksik statistik tahlil shuni ko'rsatadiki, "Apple" so'zi ko'pincha "iPhone" yoki "Macbook" bilan birga keladi.

Birgalikda paydo bo'lish tahlili: Bu usul matnda ikki yoki undan ortiq so'zlarning birga kelishi chastotasini tahlil qilishni o'z ichiga oladi. Masalan, ijtimoiy tarmoqdagi postning leksik statistik tahlili shuni ko'rsatadiki, "yoqdi" va "sharh" so'zlari tez-tez birga keladi. Matni leksik-statistik tahlil qilishning psixolingvistik yondashuvi yozma yoki og'zaki tilni kognitiv va lingvistik jarayonlar nuqtai nazaridan tahlil qilish usulidir. Ushbu yondashuv odamlarning tilni qanday tushunishi va ishlab chiqarishi va turli til omillari bu jarayonga qanday ta'sir qilishini tushunishga qaratilgan. Leksik-statistik tahlil kontekstida bu yondashuv tilni tushunish va ishlab chiqarishga yordam beradigan kognitiv jarayonlar va lingvistik xususiyatlarni tushunishga qaratilgan katta hajmdagi matn ma'lumotlarini tahlil qilish uchun statistik usullardan foydalanadi. Masalan, psixolingvistik tadqiqotchilar ma'lum bir matn yoki tilda ma'lum so'zlar, iboralar yoki grammatik tuzilmalarning chastotasi va tarqalishini aniqlash uchun korpus tahlilidan foydalanishlari mumkin.

Xulosa va takliflar. Xulosa qilib aytganda, matni leksik-statistik tahlil qilish, bir matndagi so'zlar, ularning ko'rinish ko'rsatkichlari va ularga oid boshqa ma'lumotlarni sanash, tahlil qilish va natijalarni ma'lum qilishni o'z ichiga olgan ilmiy usuldir. Leksik-statistik tahlil, matnning leksik va grammatik ko'rinishlarini o'rganishga yordam beradi. Ushbu usulda, matndagi har bir so'zning ko'rinish ko'rsatkichlari, ya'ni, ularga oid ma'lumotlar, ma'nolari, tushunchalari, tahlil qilinadi. Buning uchun, matn statistikasi, so'zlar, ularning tanlangan ko'rsatkichlari, ularning tahlili va boshqa ko'rsatkichlar yordamida ma'lumotlar to'planadi. Leksik-statistik tahlil, matndagi so'zlar va ularning ko'rinish ko'rsatkichlarini hisobga oladi va ularga oid tahlil va natijalar chiqaradi.

ADABIYOTLAR

1. Белянин В. П. Психоллингвистика. –М., 2004. –С. 8.
2. Shaykhislamov, N. (2020). Cognitive Linguistics, the Symbolic and
3. Interactive Functions of Language. Образование и наука в XXI веке, 1(6), 390-393.
4. Nursulton Zamon o'g'li Shayxislamov. Til korpusi yaratishning umumiy tamoyillari xususida –Toshkent.
5. Applying corpus linguistics methodology to psycholinguistics research. Luciane Corrêa FERREIRA.
6. Белянин В.П. Психоллингвистика. Учебное пособие. –М.: Флинта, 2008.
7. Усманова Ш. "Психоллингвистика" фанидан маъруза курслари. –Тошкент: Университет, 2014.
8. The PF STAR Children's Speech Corpus. Anton Batliner1, Mats Blomberg2, Shona D'Arcy3, Daniel Elenius2, Diego Giuliani4, Matteo Gerosa4, Christian Hacker1, Martin Russell3, Stefan Steidl1, Michael Wong3.
9. Ахмедова Ф.Б., М. Х. Психоллингвистиканинг назарий асослари. – Тошкент, 2016
10. Leontiyev A. A. Psixolingvistika asoslari. M.:1999 yil.
11. Gluxov V. P., Kovshikov V.A. Psixolingvistika. Nutq faoliyati nazariyasi. MAST; Astrel, 2006-yil.
12. Nutq psixologiyasi tarixining qisqacha tavsifi. IA: <http://www.scipsychology.ru/ctr/1115.html>