UO‘K 81'373.46:81'322.2

DENOTATIV BARQARORLIK INDEKSI (DBI) — SUN’IY INTELLEKT ATAMALARINING TERMINOLOGIK BARQARORLIGINI O‘LCHASHNING YANGI MIQDORIY VOSITASI
ИНДЕКС ДЕНОТАТИВНОЙ УСТОЙЧИВОСТИ (ИДУ) — НОВЫЙ КОЛИЧЕСТВЕННЫЙ ИНСТРУМЕНТ ИЗМЕРЕНИЯ ТЕРМИНОЛОГИЧЕСКОЙ УСТОЙЧИВОСТИ ТЕРМИНОВ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА
THE DENOTATIVE STABILITY INDEX (DSI): A NEW QUANTITATIVE TOOL FOR MEASURING THE TERMINOLOGICAL STABILITY OF ARTIFICIAL INTELLIGENCE TERMS

Solijonova Mahbubaxon Bo‘riboy qizi, mustaqil izlanuvchi, Andijon davlat universiteti filologiya fakulteti xorijiy tillar filologiyasi kafedrasi, Andijon
Солижонова Махбубахон Бўрибой қизи, самостоятельный исследователь, кафедра филологии иностранных языков филологического факультета Андижанского государственного университета, Андижан
Solijonova Makhbubakhon Boriboy qizi, Independent Researcher, Department of Foreign Languages Philology, Faculty of Philology, Andijan State University, Andijan
ORCID iD: 0009-0000-0618-9172; e-mail: makhbuba.solijanova@gmail.com

Annotatsiya. Maqolada muallif tomonidan ishlab chiqilgan Denotativ barqarorlik indeksi (DBI) taqdim etiladi. Ko‘rsatkich ikki kompozit komponentni birlashtiradi: dominant variantning umumiy chastotaga nisbati (terminologik konvergensiya) va variantlilik intensivligining qarama-qarshi ko‘rsatkichi (divergensiya). Indeks 2 275 000 so‘zlik ikki tilli korpus materialida 150 ta atamaga tatbiq etildi. Natijalar shuni ko‘rsatdiki, o‘zbek SI terminologiyasining o‘rtacha DBI ko‘rsatkichi 0,58 ni tashkil etadi, ya’ni tizim me’yorlashuv bosqichini to‘liq bosib o‘tmagan. Indeks korpus-ichki qiyosiy o‘lchov sifatida talqin etiladi va terminograf uchun me’yorlashtirishni talab etuvchi atamalarni aniqlash vositasi bo‘lib xizmat qiladi.

Kalit so‘zlar: denotativ barqarorlik indeksi, terminologik variantlilik, konvergensiya, divergensiya, miqdoriy terminologiya, korpus tahlili, me’yorlashtirish.

Аннотация. В статье представлен разработанный автором Индекс денотативной устойчивости (ИДУ). Показатель объединяет два композитных компонента: отношение доминантного варианта к общей частоте (терминологическая конвергенция) и обратный показатель интенсивности вариативности (дивергенция). Индекс применён к 150 терминам на материале

двуязычного корпуса объёмом 2 275 000 слов. Результаты показали, что средний показатель ИДУ узбекской терминологии ИИ составляет 0,58, то есть система не прошла полностью этап нормализации.

Ключевые слова: индекс денотативной устойчивости, терминологическая вариативность, конвергенция, дивергенция, количественная терминология, корпусный анализ, нормализация.

Abstract. The article presents the Denotative Stability Index (DSI) developed by the author. The measure combines two composite components: the ratio of the dominant variant to total frequency (terminological convergence) and the inverse indicator of variation intensity (divergence). The index was applied to 150 terms using a bilingual corpus of 2,275,000 words. The results show that the mean DSI for Uzbek AI terminology is 0.58, indicating that the system has not fully completed the normalisation stage. The index is interpreted as a corpus-internal comparative measure.

Key words: denotative stability index, terminological variation, convergence, divergence, quantitative terminology, corpus analysis, standardisation.

Terminologik tizimning holatini baholash uchun sifatli tavsif yetarli emas. Muayyan atamaning me’yorlashgan yoki me’yorlashmagani, terminologik tizimning qaysi qismi aralashuvni talab etishi kabi masalalar obyektiv, takrorlanadigan o‘lchov vositasini talab qiladi.

O‘zbek terminologiyashunosligida shu paytgacha terminologik barqarorlikni miqdoriy o‘lchashning yagona vositasi ishlab chiqilmagan. Mavjud tadqiqotlarda atamalarning variantliligi sanoq yo‘li bilan qayd etiladi, biroq bu yondashuv variantlarning chastota nisbatini hisobga olmaydi. Ikki variantga ega, lekin ulardan biri 95 foiz hollarda qo‘llaniladigan atama bilan ikki varianti teng chastotada uchraydigan atama bir xil baholanadi — bu esa metodologik jihatdan asosli emas.

Ilmiy muammo quyidagicha ifodalanadi: terminologik barqarorlikni ham konvergensiya (dominant variantning ustunligi), ham divergensiya (variantlar soni) jihatidan bir vaqtda hisobga oluvchi kompozit ko‘rsatkichni qanday qurish mumkin?

Maqolaning maqsadi — Denotativ barqarorlik indeksining matematik tuzilishini asoslash, uning lingvistik talqinini belgilash va 150 ta SI atamasi materialida sinovdan o‘tkazishdan iborat.

J.Freixa terminologik variantlilikning sabablarini tizimli tahlil qilib, uni tasodifiy hodisa emas, balki qonuniyatli jarayon sifatida tavsifladi [1, 55-b.]. Uning ishida variantlilikning dialektal, funksional, kognitiv va stilistik turlari farqlanadi.

M.T.Cabré kommunikativ terminologiya nazariyasi doirasida terminning uch o‘lchamli tabiatini asosladi [2, 81-b.]. Uning yondashuvida atama barqarorligi shaklda emas, real qo‘llanishda shakllanadi — bu tamoyil mazkur tadqiqotning nazariy poydevorini tashkil etadi.

J.Sinclair korpus lingvistikasida chastota va kollokatsion barqarorlik o‘rtasidagi bog‘liqlikni ko‘rsatdi [3, 110-b.]. Uning kuzatuvlari chastota asosidagi o‘lchovlarni qurish uchun metodologik asos beradi.

J.Pearson terminologik birliklarni korpus materialida aniqlash mezonlarini ishlab chiqdi [4, 42-b.]. Uning ishida qo‘llanilgan chastota mezonlari mazkur tadqiqotda tanlanmani shakllantirishda moslashtirilgan holda ishlatildi.

P.Drouin va M.-C.L’Homme terminologik ma’lumotlarni yarim avtomatik ajratishda statistik va lingvistik mezonlarni birlashtirgan gibrid yondashuvni taklif etdilar [5, 71-b.].

K.Krippendorff kontent-tahlil metodologiyasida ishonchlilik ko‘rsatkichlarini hisoblashning standart usullarini bayon etdi [6, 277-b.]. Koen κ koeffitsienti mazkur tadqiqotda tasniflash izchilligini tekshirishda qo‘llanildi.

V.M.Leychik terminologik tizimning ichki tuzilishini tavsiflashda termin — tushuncha — tizim uchligini taklif etdi [7, 112-b.]. Bu model atamalarni makromaydonlarga guruhlashda metodologik asos bo‘lib xizmat qiladi.

P.Faber va P.León-Araúz ixtisoslashgan bilimlar tasviri va kontekstni parametrlashtirish masalalarini tadqiq etdilar [8, 8-b.]. Ularning natijalari kontekstga bog‘liq o‘zgaruvchanlikni o‘lchash imkoniyatini ko‘rsatadi.

Shunga qaramay, ko‘rib chiqilgan ishlarning hech birida terminologik barqarorlikni kompozit tarzda — konvergensiya va divergensiyani birgalikda — o‘lchovchi indeks taklif etilmagan. Mazkur maqola ana shu bo‘shliqni to‘ldirishga qaratilgan.

Tadqiqot ADU-SI Corpus materialida o‘tkazildi (2 275 000 so‘z, sakkiz segment, 2018–2024-yillar). Tanlanmaga chastota mezoni asosida 150 ta yadro atama kiritildi: atamaning korpusda kamida besh marta uchrashi va IEEE Std 2894-2022 hamda ISO/IEC 22989:2022 standartlarida qayd etilgan bo‘lishi shart qilib belgilandi.

Denotativ barqarorlik indeksi quyidagi formula bo‘yicha hisoblanadi:

$$DBI = (f_i / F) \times (1 - Vnorm / Vmax),$$

bunda f_i — atamaning dominant, ya’ni eng ko‘p qo‘llaniladigan variantining korpusdagi chastotasi; F — atamaning barcha variantlari umumiy chastotasi; Vnorm — ushbu atama uchun korpusda haqiqatan qayd etilgan farqli variantlar soni; Vmax — ushbu atama uchun nazariy mumkin bo‘lgan maksimal variantlar soni.

Formulaning matematik tuzilishi ikkita kompozit komponentni birlashtiradi. Birinchi ko‘paytuvchi (f_i / F) dominant variantning umumiy chastotaga nisbatini, ya’ni terminologik **konvergensiya** darajasini ko‘rsatadi. Ikkinchi ko‘paytuvchi ($1 - Vnorm / Vmax$) variantlilik intensivligining qarama-qarshi ko‘rsatkichini, ya’ni **divergensiya**ning teskari qiymatini ifodalaydi. Bu ikki omilning ko‘paytmasi atamaning ham konvergensiya, ham divergensiya jihatidan hisobga olingan kompozit barqarorligini beradi.

Ko‘rsatkichning lingvistik talqini quyidagicha belgilandi: 0,80 dan yuqori qiymat terminologik standartlashtirilganlikni, 0,40–0,60 oralig‘i variantlilikning mavjudligini, 0,20 dan past qiymat esa standartlashtirish zaruratini bildiradi. Mazkur chegara qiymatlar korpusdagi 150 ta atamaning DBI taqsimoti asosida empirik tarzda aniqlandi: taqsimotning yuqori choragi 0,80 atrofida, medianasi 0,58 va quyi choragi 0,40 darajasida joylashgan.

Ta’kidlash lozimki, Vmax qiymati korpus tarkibiga bog‘liq bo‘lgani bois DBI mutlaq universal o‘lchov emas, balki **korpus-ichki qiyosiy ko‘rsatkich** sifatida talqin etilishi kerak. Boshqa korpusga tatbiq etilganda indeks qiymatlari qayta kalibrlanadi, biroq atamalarning nisbiy barqarorlik tartibi saqlanib qoladi.

Tasniflash izchilligi test-qayta-test usulida tekshirildi: butun tanlanma muallif tomonidan ikki marta, to‘rt hafta oralig‘ida mustaqil ravishda qayta kodlandi (intra-rater $\kappa = 0{,}82$).

DBI ko‘rsatkichini 150 ta atamaga tatbiq etish natijasida quyidagi taqsimot olindi.

1-jadval
DBI qiymatlarining taqsimoti (n = 150)

DBI oralig‘i	Talqin	Atamalar	Ulush
0,80–1,00	standartlashgan	31	20,7 %
0,60–0,79	yuqori barqaror	38	25,3 %
0,40–0,59	variantlilik mavjud	47	31,3 %
0,20–0,39	beqaror	26	17,3 %
0,00–0,19	standartlashtirish zarur	8	5,4 %

O‘rtacha DBI ko‘rsatkichi 0,58 ni tashkil etadi. Bu qiymat taqsimot medianasiga to‘g‘ri keladi va o‘zbek SI terminologiyasi me’yorlashuv bosqichini to‘liq bosib o‘tmaganidan dalolat beradi.

Eng yuqori DBI ko‘rsatkichiga ega atamalar — *algorithm* (“algoritm”, DBI = 0,94), *robot* (“robot”, 0,91), *data* (“ma’lumot”, 0,87). Bizning fikrimizcha, bu atamalarning barqarorligi ikki omil bilan izohlanadi: ular o‘zbek tiliga SI sohasi paydo bo‘lishidan ancha oldin kirgan va uzoq muddat davomida qo‘llanish orqali mustahkamlangan.

Eng past ko‘rsatkichga ega atamalar — *alignment* (0,21), *explainability* (0,24), *grounding* (0,28). Bularning barchasi uchinchi tarixiy qatlamga, xususan SI axloqi va ishonchlilik sohasiga mansub. Ularning o‘zbek tilida barqaror ekvivalenti hali shakllanmagan.

Alohida diqqatga sazovor natija — *machine learning* atamasining holati. Uning DBI ko‘rsatkichi 0,46 ni tashkil etadi, chunki korpusda uch raqobatdosh variant qayd etildi: “mashinali o‘qitish” (dominant, 52 foiz), “mashinali o‘rganish” (34 foiz) va

“mashina o‘rganishi” (14 foiz). Bu holat sohaning eng markaziy atamasi ham me’yorlashmaganini ko‘rsatadi.

DBI ko‘rsatkichi bilan atamaning tarixiy qatlami o‘rtasidagi bog‘liqlik statistik ahamiyatli bo‘lib chiqdi. Birinchi qatlam atamalari uchun o‘rtacha DBI = 0,71, ikkinchi qatlam uchun 0,61, uchinchi qatlam uchun esa 0,49. Bu tendensiya kutilgan natijadir: atama qanchalik yangi bo‘lsa, uning me’yorlashish darajasi shunchalik past.

Tadqiqotimiz davomida aniqlandiki, DBI va atamaning chastotasi o‘rtasida chiziqli bog‘liqlik yo‘q. Yuqori chastotali atamalar orasida ham beqarorlari uchraydi (*machine learning*), past chastotali atamalar orasida esa barqarorlari bor (*heuristic search*, DBI = 0,83). Bu natija muhim metodologik xulosa beradi: chastota terminologik barqarorlikning ishonchli indikatori emas, shu sababli kompozit ko‘rsatkichga ehtiyoj bor.

Indeksning amaliy qiymati shundan iboratki, u me’yorlashtirish ishini ustuvorlashtirish imkonini beradi. DBI qiymati 0,40 dan past bo‘lgan 34 ta atama (22,7 foiz) birinchi navbatda terminologik komissiya muhokamasiga kiritilishi tavsiya etiladi.

Tadqiqotda terminologik barqarorlikni o‘lchash uchun Denotativ barqarorlik indeksi ishlab chiqildi va 150 ta SI atamasi materialida sinovdan o‘tkazildi. Indeksning asosiy ustunligi — konvergensiya va divergensiya omillarini bir vaqtda hisobga olishi.

Empirik natijalar quyidagilardan iborat. O‘zbek SI terminologiyasining o‘rtacha DBI ko‘rsatkichi 0,58 ni tashkil etadi. Atamalarning 20,7 foizi standartlashgan, 31,3 foizida sezilarli variantlilik mavjud, 22,7 foizi esa me’yorlashtirishni zudlik bilan talab etadi. DBI ko‘rsatkichi atamaning tarixiy qatlamiga bog‘liq: uchinchi qatlamda eng past (0,49).

Muhim metodologik xulosa: chastota terminologik barqarorlikning ishonchli indikatori emas — yuqori chastotali atamalar orasida ham beqarorlari uchraydi.

Amaliy tavsiyalar. Birinchidan, DBI ko‘rsatkichi me’yorlashtirish ishini ustuvorlashtirish vositasi sifatida qo‘llanilishi tavsiya etiladi: 0,40 dan past qiymatga ega atamalar birinchi navbatda ko‘rib chiqilishi lozim. Ikkinchidan, indeks milliy terminologik tizimning holatini davriy kuzatish uchun monitoring vositasi bo‘lib xizmat qilishi mumkin. Uchinchidan, Vmax qiymatini ekspert komissiyasi tomonidan belgilanadigan universal chegaraviy qiymat bilan almashtirish indeksni korpuslararo qiyosiy o‘lchovga aylantiradi.

Kelgusi tadqiqotlar uchun istiqbolli yo‘nalish — DBI ko‘rsatkichini boshqa raqamli sohalar terminologiyasida sinovdan o‘tkazish va uning mustaqil validatsiyasini ta’minlash.

FOYDALANILGAN ADABIYOTLAR RO‘YXATI:

1. Freixa J. Causes of Terminological Variation // Terminology. – 2006. – Vol. 12, № 1. – P. 51–77.

2. Cabré M.T. Terminology: Theory, Methods and Applications / Transl. by J.C. Sager. – Amsterdam; Philadelphia: John Benjamins, 1999. – 248 p.

3. Sinclair J. Corpus, Concordance, Collocation. – Oxford: Oxford University Press, 1991. – 179 p.

4. Pearson J. Terms in Context. – Amsterdam; Philadelphia: John Benjamins, 1998. – 246 p.

5. Drouin P., L'Homme M.-C. Semi-automatic Methods for Extracting Terminological Data: Corpus-Based Approaches // Terminology. – 2022. – Vol. 28, № 1. – P. 68–95.

6. Krippendorff K. Content Analysis: An Introduction to Its Methodology. 4th ed. – Thousand Oaks: SAGE Publications, 2018. – 472 p.

7. Leychik V.M. Terminovedeniye: predmet, metody, struktura. – Moscow: LKI, 2009. – 256 s.

8. Faber P., León-Araúz P. Specialized Knowledge Representation and the Parameterization of Context // Frontiers in Psychology. – 2016. – Vol. 7, Art. 196. – P. 1–20.

9. Landis J.R., Koch G.G. The Measurement of Observer Agreement for Categorical Data // Biometrics. – 1977. – Vol. 33, № 1. – P. 159–174.

10. Sager J.C. A Practical Course in Terminology Processing. – Amsterdam: John Benjamins, 1990. – 254 p.

11. IEEE Std 2894-2022. IEEE Guide for an Architectural Framework for Explainable Artificial Intelligence. – New York: IEEE, 2022. – 34 p.

12. ISO/IEC 22989:2022. Information technology – Artificial intelligence – Artificial intelligence concepts and terminology. – Geneva: ISO, 2022. – 62 p.

13. Temmerman R. Towards New Ways of Terminology Description: The Sociocognitive Approach. – Amsterdam: John Benjamins, 2000. – 258 p.

14. Daille B., Romary L. Terminology-Based Knowledge Representation: Corpus-Driven Approaches // Terminology. – 2021. – Vol. 27, № 2. – P. 142–165.

15. Begmatov E.A. O'zbek terminologiyasi asoslari. – Toshkent: Fan, 2001. – 246 b.