

## HUMBOLDTNING LINGVOFALSAFIY MEROSI VA SUN'IY INTELLEKT TILSHUNOSLIGIDA UNING METODOLOGIK AHAMIYATI

**Raxmonqulov Mavlon**

*Samarqand davlat chet tillar instituti magistranti*

**Annotatsiya.** Ushbu maqolada Wilhelm von Humboldtning tilga oid lingvofalsafiy qarashlari — xususan, "energeia" va "ergon" dialektikasi — zamonaviy sun'iy intellekt (AI) tilshunosligiga metodologik ta'siri nuqtai nazaridan tahlil qilinadi. Muallif katta til modellari (LLM — Large Language Models) ning til birliklarini "o'rganish" jarayoni bilan Humboldtning energeia tushunchasini solishtirib o'rganadi va ularning fundamental o'xshashliklari hamda farqlarini aniqlaydi. Maqolada, shuningdek, Humboldtning "tilning ichki shakli" (innere Sprachform) konsepsiyasi zamonaviy semantik vektorlar (word embeddings) nazariyasi bilan qanday munosabatda ekanligi yoritiladi.

**Kalit so'zlar:** Humboldt lingvofalsafasi, energeia, ergon, katta til modellari, sun'iy intellekt va til, semantik vektorlar, ichki til shakli, kognitiv tilshunoslik.

**Kirish.** Wilhelm von Humboldt (1767–1835) tilni ijtimoiy va kognitiv faoliyat sifatida talqin etgan birinchi mutafakkirlardan biridir. Uning "til — harakatdagi jarayon" degan asosiy g'oyasi zamonaviy lingvistikada doimo muhokamalarga sabab bo'lib kelmoqda. Ammo so'nggi o'n yillikda bu munozaraning xarakteri tubdan o'zgardi: endi gap faqat inson nutqi haqida emas, balki mashina nutqi — ya'ni sun'iy intellekt orqali hosil bo'lgan til haqida ham ketmoqda.

Humboldtning "Ichki til shakli" va semantik vektorlar

Humboldt tildagi "ichki shakl" (innere Sprachform) tushunchasi orqali har bir tilning o'ziga xos tafakkur tarzini ifodalovchi chuqur tuzilmani nazarda tutgan. Bu tuzilma grammatik qoidalardan farqli o'laroq, tilning ma'no qatlamida yashiringan bo'ladi va uni to'g'ridan-to'g'ri ko'rish mumkin emas.

Zamonaviy kompyuter tilshunosligida "so'z vektorlari" (word embeddings) deb ataluvchi texnologiya aynan shu fikrga yaqin bir kashfiyotni ifodalaydi. Word2Vec, GloVe, FastText kabi modellar har bir so'zni ko'p o'lchovli matematik fazoda nuqta sifatida tasvirlaydi, bunda ma'nodosh so'zlar bir-biriga yaqin joylashadi. Masalan, "shoh" va "qirol" vektorlari fazoda yaqin tursa, "ona" va "ota" ham o'zaro yaqin bo'ladi.

Bu parallellizmtasodifiy emas. Humboldtning ichki shakl g'oyasi ham til birliklarining bir-biriga nisbatan ma'noviy munosabatlarini — ya'ni ma'no fazosidagi "joylashuvini" — muhim deb hisoblaydi. Biroq muhim farq ham bor: Humboldt uchun ichki shakl jonli tafakkur

mahsuli, semantik vektorlar esa statistik korrelyatsiyaning matematik ifodasi. Bu ikki tushuncha o'rtasidagi chegarani aniqlashtirish zamonaviy AI falsafasining asosiy vazifalaridan biriga aylanmoqda.

LLM va “energeia” mashina nutqida jarayon

Humboldtning energeia tushunchasi tilni har safar qayta yaratiladigan, doimiy ijod holati sifatida tavsiflaydi. Inson har gapirganida grammatik qoidalardan foydalanadi, lekin ayni paytda yangi ma'nolar, yangi kontekstlar, yangi uslublar vujudga keltiradi. Bu jarayon takrorlanmas va ijodiy.

LLM modellari ham so'z birikmalarini hech qachon so'zma-so'z yoddan chiqarmaydi — aksincha, ularning har bir javobi kontekst asosida, ehtimollik taqsimotlari orqali yangicha tarzda shakllanadi. Texnik nuqtai nazardan bu jarayon Humboldtning energeia sxemasiga o'xshaydi: har safar yangi, takrorlanmas chiqish hosil bo'ladi.

Biroq bu o'xshashlik ham chekli. Humboldt energeiasining asosi — niyat, motivatsiya va ijtimoiy muloqot istagi. LLM modellari esa bu ma'noda "niyatsiz" tizimlar: ular ifodalaydilar, lekin ifodalashni "xohlashmaydi". Shu bois, Humboldtning energeia konsepsiyasini AI tilshunosligiga to'liq tatbiq etish uchun bu tushunchani ikki qismga ajratish taklif etiladi:

- Protsessual energeia — til birliklarining har safar qayta generatsiya qilinishi (LLMda mavjud);
- Intentional energeia — tilning ongli maqsad va kommunikativ niyat bilan bog'liqligi (LLMda hali to'liq amalga oshirilmagan).

Bu farqlash sun'iy intellektning til imkoniyatlarini baholashda muhim metodologik asos bo'la oladi.

“Ergon” paradigmasining AI tilshunosligidagi yangi talqini

Humboldt uchun ergon — tilda mavjud, tayyor tuzilmalar: lug'at, grammatik normalar, orfoepik qoidalar. Bu komponent tilning "don'" yoki "arxiv" tomonini ifodalaydi. Zamona viy AI tilshunosligida bunga mos keladigan tuzilma — model parametrlari (weights) va o'quv ma'lumotlar bazasi (training corpus).

O'quv jarayoni tugagach, LLM modeli statik holga keladi — uning parametrlari qotib qoladi. Bu holat Humboldtning ergon tushunchasiga to'g'ridan-to'g'ri mos keladi: tayyor, o'zgarmas tizim. Ammo model har safar ishga tushirilganda, ya'ni foydalanuvchi bilan muloqotda bo'lganda, yana dinamik jarayon boshlanadi — energeia sahna oladi.

Shunday qilib, zamonaviy LLM arxitekturasida ergon va energeia ikki fazaviy holatni ifodalaydi: birinchisi o'qitish jarayonidan keyin hosil bo'ladigan statik holat, ikkinchisi esa har bir muloqot seansi davomida sodir bo'ladigan dinamik generatsiya jarayoni. Bu dialektika Humboldtning asosiy konsepsiyasining texnologik modeli sifatida talqin etilishi mumkin.

O'zbek tili va AI: milliy tilning energeiasi xavf ostidami?

Humboldtning milliy tillar falsafasida alohida o'rin tutgan fikrlaridan biri shundan iborat: har bir til o'ziga xos tafakkur tarzini shakllantiradi va shu bois tillar soni kamayishi insoniyatning kognitiv xilma-xilligini kambag'allashtiradi. Bu g'oya "lingvistik relyativizm" nomi bilan ma'lum bo'lib, zamonaviy AI kontekstida yangi dolzarblik kasb etmoqda.

Bugungi kunda LLM modellari asosan ingliz tilida o'qitilmoqda. O'zbek tili, qirg'iz tili, tojik tili kabi mintaqaviy tillar bu modellarda juda kam ifodalangan. Bu holat ikki xil xavfni keltirib chiqaradi:

- Birinchidan, kichik tillar AI tizimlarida kam aks etishi ularning raqamli muhitdagi mavqeini kamaytiradi — Humboldt ta'birida, bu tillarning "energetik maydon"i torayadi;
- Ikkinchidan, foydalanuvchilar o'z ona tilida sifatli AI xizmatiga ega bo'lmagach, boshqa tilga o'tishga majbur bo'ladi — bu esa tilning jonli jarayoni sifatidagi faolli giga putur etkazadi.

O'zbek tili uchun bu muammo ayniqsa dolzarb: so'nggi yillarda internet leksikasida ingliz va rus tili ta'siri kuchaymoqda, o'zbek tilidagi LLM modellari esa hali etarlicha rivojlanmagan. Humboldt ta'limotidan kelib chiqib, o'zbek tilining raqamli energieiasini saqlab qolish uchun milliy korpus tilshunoslik infratuzilmasini rivojlantirish strategik zarurat ekanini ta'kidlash lozim.

Xulosa. Humboldtning "energeia" va "ergon" dialektikasi, "ichki til shakli" va lingvistik relyativizm g'oyalari zamonaviy sun'iy intellekt tilshunosligida ham o'z dolzarbligini saqlab qolmoqda. Biroq bu tushunchalar bevosita emas, balki qayta talqin qilingan holda qo'llanilishi zarur. LLM modellarining faoliyati Humboldt energieiasining "protsexual" jihatini aks ettiradi, biroq "intentional" jihat — niyat, ijtimoiy motivatsiya, mavjudlikdagi til — hali sun'iy intellektda to'liq ifodalanmagan. Ichki til shakli esa semantik vektorlar orqali matematiklashtirilib, qisman qayta kashf etilgan. Milliy tillar, xususan o'zbek tili kontekstida Humboldtning lingvistik relyativizm g'oyasi eng kuchli metodologik asos bo'lib qolmoqda: tilning raqamli muhitdagi energieiasini saqlab qolish uchun tilshunoslik va texnologiya birgalikda harakat qilishi shart.

### Foydalanilgan adabiyotlar

1. Humboldt, W. von. (1836). *Über die Verschiedenheit des menschlichen Sprachbaues und ihren Einfluss auf die geistige Entwicklung des Menschengeschlechts*. Berlin: Königliche Akademie der Wissenschaften.
2. Chomsky, N. (2006). *Language and Mind* (3rd ed.). Cambridge University Press.
3. Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). *Efficient Estimation of Word Representations in Vector Space*. arXiv:1301.3781.

4. Vaswani, A. et al. (2017). Attention Is All You Need. *Advances in Neural Information Processing Systems*, 30.
5. Joseph, J. E. (2015). The Reception of Wilhelm von Humboldt's Linguistic Writings in the Anglosphere. *Forum for Modern Language Studies*, 53(1), 7–28.
6. Whorf, B. L. (1956). *Language, Thought, and Reality*. MIT Press.
7. Bender, E. M. et al. (2021). On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? *FACCT '21 Conference Proceedings*, 610–623.
8. Pennington, J., Socher, R., & Manning, C. D. (2014). GloVe: Global Vectors for Word Representation. *EMNLP Proceedings*.
9. Pan, W., & Tan, H. (2020). Contributions of Wilhelm von Humboldt and Benjamin Lee Whorf to the Founding of Contrastive Linguistics. *ECNU Journal of Philosophy and Social Sciences*.
10. Slagle, U. V. (1989). *The Kantian Influence on Humboldt's Linguistic Thought*. John Benjamins Publishing.