



GRAFEMA-FONEMA (G2P) MODEL: NAZARIY ASOSLARI VA O‘ZBEK TILI UCHUN AHAMIYATI

Maxmudjonova Gulshaxnoz Ulug‘bek qizi

*Alisher Navoiy nomidagi Toshkent davlat o‘zbek tili va adabiyoti universiteti
Kompyuter lingvistikasi va raqamli texnologiyalar kafedrasi tayanch doktoranti*

Annotatsiya Grafema-fonema (G2P) konversiyasi yozma matnni uning fonetik ifodalariga aylantirish jarayonidir. Ushbu jarayon tabiiy tilni qayta ishlash (NLP), matndan ovozga (TTS) va avtomatik nutqni tanish (ASR) tizimlarida muhim rol o‘ynaydi. Maqolada G2P konversiyasining nazariy asoslari, turli tillarda ishlab chiqilgan modellari va ularning afzalliklari ko‘rib chiqiladi. Shuningdek, o‘zbek tili uchun grafema-fonema modelining zarurati, uning yaratish yondashuvlari va texnik jihatlari muhokama qilinadi.

Kalit so‘zlar: nutq sintezatori, grafema-fonema konversiyasi, alignment, tabiiy tilni qayta ishlash, o‘zbek tili.

GRAPHEME-TO-PHONEME (G2P) MODEL: THEORETICAL FOUNDATIONS AND PRACTICAL APPLICATION FOR THE UZBEK LANGUAGE

Abstract: Grapheme-to-Phoneme (G2P) conversion is the process of transforming written text into its phonetic representations. This process plays a crucial role in natural language processing (NLP), text-to-speech (TTS), and automatic speech recognition (ASR) systems. The article examines the theoretical foundations of G2P conversion, various models developed for different languages, and their advantages. Additionally, it discusses the necessity of a G2P model for the Uzbek language, the approaches to its development, and its technical aspects.

Keywords: speech synthesizer, grapheme-to-phoneme conversion, alignment, natural language processing, Uzbek language.

МОДЕЛЬ ГРАФЕМЫ-ФОНЕМЫ (G2P): ТЕОРЕТИЧЕСКИЕ ОСНОВЫ И ЗНАЧЕНИЕ ДЛЯ УЗБЕКИСТАНСКОГО ЯЗЫКА

Абстрактный. Преобразование графемы в фонему (G2P) — это процесс преобразования письменного текста в его фонетическое представление. Этот процесс играет важную роль в системах обработки естественного языка (NLP), преобразования текста в речь (TTS) и автоматического распознавания речи (ASR). В статье рассматриваются теоретические основы преобразования G2P, модели, разработанные на

разных языках, и их преимущества. Также обсуждаются необходимость графемно-фонемной модели для узбекского языка, подходы к ее созданию и технические аспекты.

Ключевые слова: синтезатор речи, графемно-фонемное преобразование, выравнивание, обработка естественного языка, узбекский язык.

Kirish. G2P konversiyasi yozma matndagi **grafemalarni** (ya’ni, orfografik belgilarni) **fonemalarga** (tilning tovush birliklariga) o‘tkazish jarayonidir[1]. Bu jarayon, ayniqsa, matndan ovozga (TTS) va avtomatik nutqni tanish (ASR) tizimlari uchun muhim hisoblanadi. Ingliz va fransuz tillari kabi murakkab fonetik qoidalarga ega tillar uchun konversiya yetarlicha murakkab jarayon bo‘lib, statistik va neyron tarmoq yondashuvlari talab etiladi. G2P modeli harflar va tovushlar o‘rtasidagi bog‘liqlikni lingvistik qoidalar asosida o‘rganadi. Bu bog‘liqlik tilga qarab o‘zgarib turadi, ayrim tillarda grafemalar va fonemalar orasidagi moslik deyarli bir-biriga to‘g‘ri keladi, boshqalarida esa murakkabroq qonuniyatlar mavjud.

Asosiy qism. G2P konversiyasi nutq sintezatorining asosiy tarkibiy qismi bo‘lib, u matndagi harflarni fonetik transkripsiyaga o‘tkazish orqali so‘zlarning to‘g‘ri talaffuz qilinishini ta’minlaydi[1-rasm]. Bu jarayon nutq sintezatori tizimlariga lug‘atlarga tayanmasdan, yangi yoki kam uchraydigan so‘zlarni ham tabiiy ovoz bilan ifodalash imkonini beradi. Shu sababli, G2P konversiyasi TTS tizimlarining aniq, moslashuvchan va yuqori sifatli nutq hosil qilishiga yordam beradi[2].

Grafema-fonema modeli bir tilli (monolingual) va ko‘p tilli (multilingual) bo‘lishi mumkin. Bir tilli modellarga qozoq tili uchun qilingan kz_g2p grafema-fonema modelini misol keltirish mumkin. Qozoq tilidagi so‘zlarni fonetik transkripsiyaga o‘tkazish uchun ishlab chiqilgan ushbu model arab va kirill yozuvlaridagi matnlarni fonemalarga o‘zgartiradi[3]. Ko‘p tilli modellarga Phonetisaurus[4], G2P tizimi bor bo‘lib, har qanday til uchun moslashtirilishi mumkin. Ingliz tilida so‘zlarning yozilishi va talaffuzi o‘rtasidagi farqlar tufayli, so‘zning talaffuzini aniqlash qiyin. Kyubyong Park tomonidan ishlab chiqilgan model esa ingliz tilidagi grafemalarni fonemalarga avtomatik ravishda aylantirish uchun mo‘ljallangan[5]. MFA (Montreal Forced Aligner) turli tillar uchun oldindan tayyorlangan grafema-fonema modellarini taqdim etadi.[6] Ularning asosiy farqlari quyidagicha:

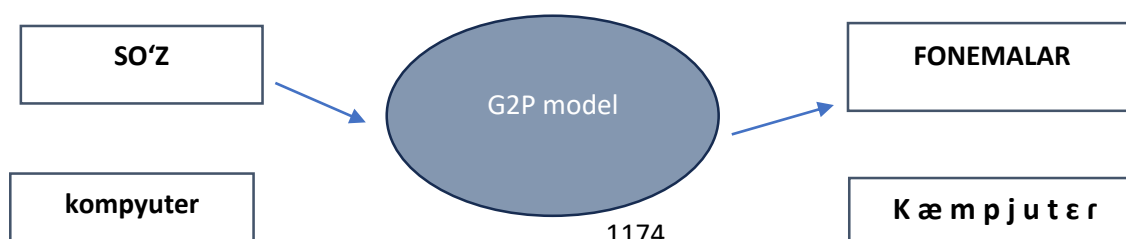
Til qamrovi	Faqat bitta til uchun	Bir nechta tilni qo‘llab quvvatlaydi
Moslashuvchanlik	Har bir tilga moslashtiriladi	Ko‘p tillar uchun umumiy model
Aniqlik	Yuqori aniqlik bilan ishlaydi	Aniqlik foizi past bo‘lishi mumkin
Resurs	Keng hajmli dataset	Kam resurs bilan ham ishlaydi

Amaliy tatbiqi

Matndan ovozga (TTS), Ko‘p tilli TTS va
avtomatik nutqni tanish transliteratsiya
(ASR)

Grafema-fonema rivojlanishi qoidaga asoslangan yondashuvlardan boshlab mashinali o‘rganish modellariga o‘tish bilan davom etdi. Chomsky va Halle [7] tomonidan ilgari surilgan fonologik nazariyalar G2P konversiyasining asosiy tamoyillarini shakllantirgan. Elovitz va boshqalar tomonidan ishlab chiqilgan qoidaga asoslangan ilk avtomatik tizimlar 329 fonologik qoidaga tayanib, harflarni fonemalarga o‘tkazishda qat’iy qoidalarga amal qilgan. Keyingi yillarda, qoidalarga asoslangan tizimlardan mashina o‘rganish yondashuvlariga o‘tildi. Bu usullar so‘z-fonema juftliklaridan o‘rganish orqali ancha moslashuvchan G2P tizimlarini yaratish imkonini berdi. Generativ yondashuvlar esa grafemalarning turli kontekstlarda qanday fonetik mosliklarga ega bo‘lishini statistik usullar bilan model qildi[8].

G2P konversiyasining aniq ishlashi uchun so‘zning nafaqat fonemalarga ajratilishi, balki bo‘g‘inlarga ajratilishi (syllabification) va urg‘u taqsimoti (word stress assignment) ham muhim. Ushbu uch jarayon bir-biri bilan bog‘liq bo‘lib, biri ikkinchisining aniqligini oshirishga yordam beradi[9]. Bo‘g‘in chegaralarini bilish G2P modelining aniq ishlashiga yordam beradi, chunki ayrim fonetik o‘zgarishlar bo‘g‘in chegaralariga bog‘liq bo‘ladi. Marchand va Damper [10] tadqiqotlariga ko‘ra, ingliz tilida bo‘g‘inlash ma’lumotlari qo‘shilganda so‘z xatosi (Word error rate) taxminan 5% kamaygan. Fonemalar asosida bo‘g‘inlash ishlatilganda, WER 50% ga kamaygan. Bu fonetik ma’lumotlarning bo‘g‘inlash jarayonini ancha yaxshilashini ko‘rsatadi. Urg‘u odatda bo‘g‘in darajasida taqsimlanadi, shuning uchun bo‘g‘inlash to‘g‘ri bo‘lmasa, urg‘u noto‘g‘ri aniqlanishi mumkin. Muller [11] tadqiqotlariga ko‘ra, urg‘u taqsimoti va bo‘g‘inlash ma’lumotlari modelning samaradorligini oshiradi (ayniqsa nemis va golland tillari uchun). G2P konversiyasining asosiy maqsadi harflarni to‘g‘ri fonetik transkripsiyaga aylantirishdir. Biroq, turli tillarda bir xil yozilgan harflarning talaffuzi turlicha bo‘lishi mumkin. Shu sababli, G2P modelining aniq ishlashi uchun standart fonetik transkripsiya tizimidan foydalanish talab etiladi. Xalqaro Fonetik Alifbo (IPA) butun dunyodagi tillarning fonetik tovushlarini standartlashtirilgan shaklda ifodalash imkonini beradi. IPA [<https://www.internationalphoneticassociation.org>] G2P tizimlari uchun turli tillardagi fonemalarni yagona formatda ifodalash, talaffuzdagi tafovutlarni aniqroq aks ettirish, mashinali o‘rganish modellarini sifatli o‘qitish uchun ishonchli fonetik annotatsiyalar yaratish kabi muhim xususiyatlarni ta’minlaydi. Shu sababli, o‘zbek tili uchun ishlab chiqilayotgan G2P modelida ham IPA dan foydalanish zarur, chunki u tovushlarni aniq va universal formatda ifodalash imkoniyatini beradi va modelning moslashuvchanligini oshirishga yordam beradi.



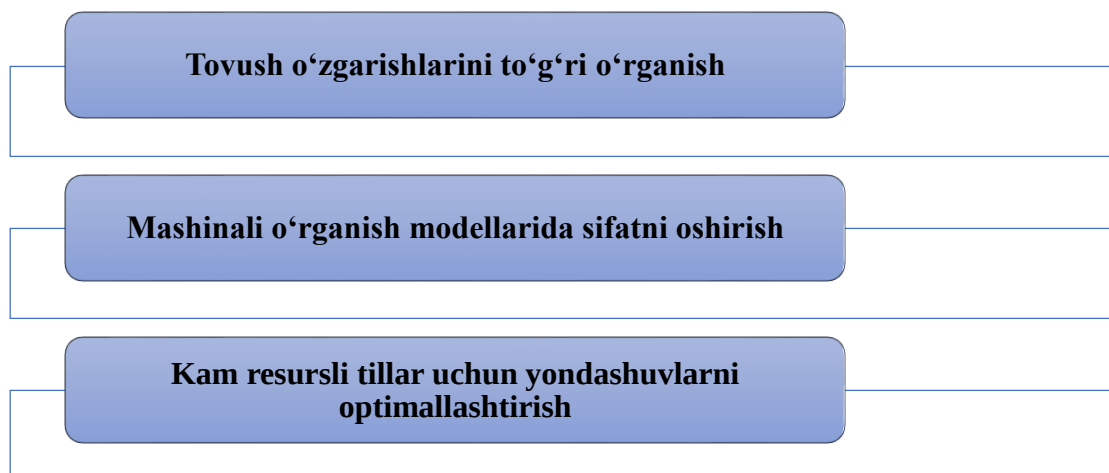
1-rasm. Grafema-fonema modeli

G2P modellarining asosiy vazifalari yozma matndan to‘g‘ri fonetik ifodani hosil qilish, talaffuzi murakkab so‘zlarni avtomatik ravishda fonemalarga ajratish, nutqni avtomatik tanish tizimlarida aniq fonetik model yaratish, transliteratsiya tizimlarida yozuv tizimlari orasidagi moslikni yaxshilashlardan iborat. Quyidagi sohalarda foydalaniladi:

QO‘LLANILISH SOHASI

N	Yozilgan so‘zlarning to‘g‘ri talaffuzini aniqlash orqali tabiiy tovushli sintez yaratish.
U	Nutqni tanish tizimlarida fonetik transkripsiyalar orqali tanib olish sifatini yaxshilash.
AVTOMATIK NUTQNI TANISH (ASR) TIZIMLARI	
IMLO TEKSHIRISH VA TALAFFUZ YORDAMI	Imlo tuzatish va til o‘rganishda to‘g‘ri talaffuzni o‘rgatish. Fonologik munosabatlarni o‘rganish va til rivojlanishini tahlil qilish.

G2P modelida aligmentning ahamiyati. Grapema va fonema o‘rtasidagi mosliklarni (alignment) aniqlash G2P modellarida muhim rol o‘ynaydi. Aligment jarayoni quyidagi sabablar tufayli ahamiyatlidir:

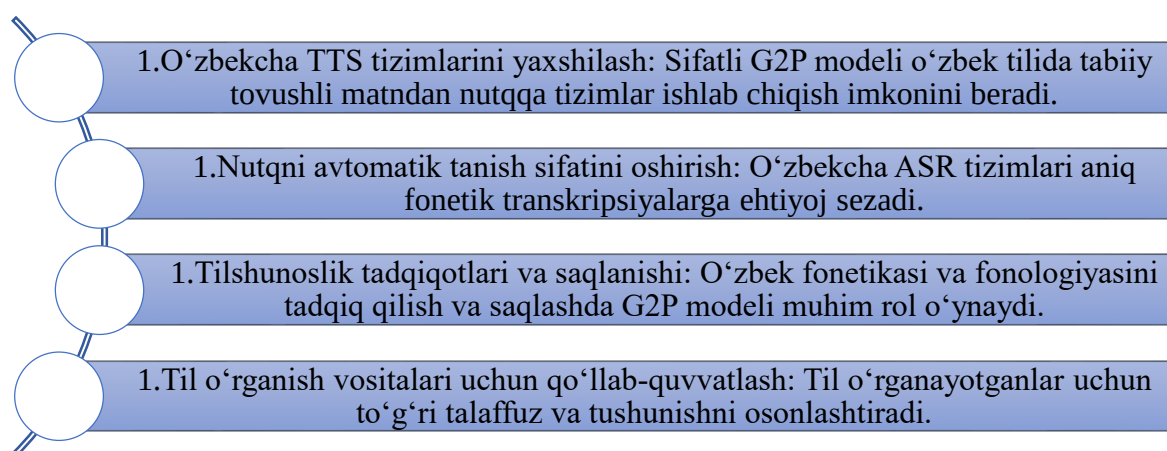


Har bir grafemaga mos keladigan fonemalarni aniq belgilash uchun aligner mexanizmi zarur. Aligment to‘g‘ri belgilangan taqdirda, model aniqroq bashorat qila oladi. Qoidaga asoslangan ma’lumotlardan avtomatik ravishda fonetik qoidalarni chiqarish imkonini beradi. Ilk tadqiqotlarda alignerlar qoidalarga tayanib qo‘lda aniqlangan, lekin keyinchalik Expectation Maximization (EM) algoritmi [12] yordamida avtomatlashtirilgan. G2P tizimlarida ikkita asosiy komponent mavjud:

1. *Grapema va fonema o‘rtasidagi yashirin bog‘liqlikni aniqlash (alignments)*. Bu jarayon G2P tizimining har bir grafema uchun mos fonemani aniqlashiga imkon beradi. Ushbu mosliklarni aniqlash uchun alohida model yoki generativ modellar ichiga kiritilgan mexanizmlar ishlatiladi [13].

2. *Fonema hosil qilish modeli*. Model so‘zga mos fonemalarni yaratish uchun o‘qitiladi. Bu ikki asosiy paradigmani o‘z ichiga oladi: *Tasniflashga asoslangan model*: Har bir grafema alohida fonema yoki fonemalar yaratadi. *Ketma-ketlikka asoslangan model*: Avvalgi qarorlar keyingi bosqichlarni belgilash uchun hisobga olinadi.

Ko‘plab yirik tillar uchun G2P modellar mavjud bo‘lsa-da, o‘zbek tili uchun hali yetarlicha rivojlangan model yo‘q. Quyidagi sabablarga ko‘ra o‘zbek tili uchun G2P modeli ishlab chiqilishi zarur:



G2P tizimlarini ishlab chiqishda turli yondashuvlar qo‘llaniladi, jumladan, qoidaga asoslangan (rule-based), ma’lumotlarga asoslangan (data-driven), statistik hamda chuqur o‘rganish (deep learning) modellari. G2P konversiyasida qo‘llaniladigan asosiy yondashuvlar leksik va qoidaga asoslangan yondashuvlar bo‘lib, tayyor fonetik lug‘atlardan foydalanishga asoslanadi, bu esa yuqori aniqlikni ta’minlaydi, biroq yangi so‘zlarning talaffuzi aniqlanmaganda tizim cheklangan bo‘lib qoladi. Qoidaga asoslangan tizimlar fonologik va morfologik qoidalarga asoslanib, cheksiz miqdordagi so‘zlarni talaffuz qilish imkonini beradi, lekin murakkab istisnalar bilan ishlash jarayoni ko‘p resurs talab qiladi. Ma’lumotlarga asoslangan yondashuvlarda analogiyaga asoslangan talaffuz (Pronunciation by Analogy – PbA) texnikasi o‘xshash so‘zlar yoki ularning qismlarining fonetik transkripsiyasidan foydalangan holda yangi so‘zlarning talaffuzini prognoz qiladi. Chuqur o‘rganish va neyron tarmoqlarga asoslangan G2P modellar ma’lumotlarni o‘rganish orqali talaffuzni aniq prognozlash imkonini beradi [15, Bisani, M., & Ney, H. (2008). "Joint-sequence models for grapheme-to-phoneme conversion." *Speech Communication*, 50(5), 434-451.].

	Lingvistik qoidalarga asoslanadi va aniq natija beradi.	Moslashuvchanligi past, murakkab qoidalarga duch kelganda muammolar yuzaga keladi.
	Katta hajmdagi ma’lumotlar asosida natijani yaxshilaydi.	Yetarli ma’lumot to‘plash qiyin bo‘lishi mumkin.
LI O‘RGANISH	Neyron tarmoqlar yordamida yuqori aniqlik bilan talaffuzni bashorat qiladi.	Katta resurs talab qiladi va uzoq vaqt mashg‘ulot talab etadi.
MA’LUMOTLARGA ASOSLANGAN	Yangi so‘zlar uchun moslashuvchan, qoidalarga ehtiyoj yo‘q, katta hajmdagi so‘zlarni qayta ishlash imkoniyati	Katta hajmdagi o‘quv ma’lumotlariga ehtiyoj, turli dialekt va urg‘ularni qamrab olish qiyin

O‘zbek tili uchun G2P modelini ishlab chiqishda gibrid yondashuv yuqori aniqlik va moslashuvchanlikka erishish imkonini beradi. gibrid yondashuvni qo‘llashning asosiy sabablarida fonetik murakkablikni nazorat qilish imkoniyati borligi. Qoidaga asoslangan yondashuv yordamida arabiy-forsiy so‘zlar, morfologik o‘zgarishlar, va urg‘u tizimi kabi murakkab jihatlar modelga kiritiladi. Yana bir eng muhim sabablardan biri katta datasetlarning yo‘qligi. Statistik va neyron tarmoqlar yondashuvi ko‘p hajmli ma’lumotlar bazasi asosida ishlaydi. Kam resursli tillarda neyron tarmoqlar qoidalar bilan birgalikda ishlaganda yanada yaxshi natija beradi. Moslashuvchanlik va yuqori aniqlik olish uchun ham gibrid yondashuvdan foydalanish mumkin. O‘zbek tilida G2P modelini ishlab chiqishda gibrid yondashuv eng samarali usul hisoblanadi, chunki bu qoidalarga asoslangan tahlillarni neyron tarmoqlar bilan to‘ldirib, optimal natijaga erishish imkonini beradi. Bu yondashuv TTS va ASR tizimlarida yuqori aniqlik va moslashuvchanlik bilan ishlaydigan model yaratishga yordam beradi.

Xulosa. G2P modellar allaqachon ko‘plab tillar uchun ishlab chiqilgan bo‘lsa-da, o‘zbek tili bu borada yetarlicha rivojlanmagan. O‘zbek tilida G2P model yaratish uning raqamli rivojlanishi, kirish imkoniyatlari va lingvistik tadqiqotlari uchun katta ahamiyatga ega bo‘lib, kompyuter lingvistikasi sohasida muhim qadam bo‘ladi.

G2P modelning kelajakdagi rivoji esa uning qamrovini kengaytirish, turli dialektlarni qo‘llab-quvvatlash va real vaqt rejimida ishlash imkoniyatlarini kengaytirishga qaratilgan bo‘lishi lozim. Shu bois, o‘zbek tilida G2P model yaratish bo‘yicha tadqiqotlarni yanada kengaytirish va joriy etish zarur.

FOYDALANILGAN ADABIYOTLAR

1. <https://docs.nvidia.com/nemo-framework/user-guide/24.09/nemotoolkit/tts/g2p.html>
2. Nath, C., & Sarma, B. (2023). *A Grapheme to Phoneme Based Text to Speech Conversion Technique in Unicode Language*. **Data and Metadata**, 2(191). <https://doi.org/10.xxxx/xxxxx>

3. [GitHub - Ardaq/kz_g2p: The grapheme to phoneme model converts Kazakh\(Arab|Cyrillic\) characters to phonemes.](#)
4. [GitHub - AdolfVonKleist/Phonetisaurus: Phonetisaurus G2P](#)
5. [GitHub - Kyubyong/g2p: g2p: English Grapheme To Phoneme Conversion](#)
6. [mfa-models/g2p at main · MontrealCorpusTools/mfa-models · GitHub](#)
7. N. Chomsky and M Halle. The sound pattern of English. Harper and Row, New York, 1968.
8. Jiampojarn, Sittichai. *Grapheme-to-Phoneme Conversion and Its Application to Transliteration*. University of Alberta, 2009. <https://era.library.ualberta.ca>.
9. Vera Demberg, Helmut Schmid, and Gregor M'ohler. Phonological constraints and morphological preprocessing for grapheme-to-phoneme conversion. In Proceedings of the 45th Annual Meeting of the Association of Computational Linguistics, pages 96–103, 2007.
10. Y. Marchand and R. I. Damper. 2005. Can syllabification improve pronunciation by analogy of English? Natural Language Engineering.
11. K. Muller. 2001. Automatic detection of syllable boundaries “ combining the advantages of treebank and bracketed corpora training. In Proceedings of ACL, pages 402–409.
12. <https://www.internationalphoneticassociation.org/>
13. Arthur Dempster, Nan Laird, and Donald Rubin. Maximum likelihood from incomplete data via the EM algorithm. In Journal of the Royal Statistical Society, pages B:1–38, 1977
14. Paul Taylor. Hidden Markov Models for grapheme to phoneme conversion. In Proceedings of the 9th European Conference on Speech Communication and Technology, 2005
15. Bisani, M., & Ney, H. (2008). "Joint-sequence models for grapheme-to-phoneme conversion." *Speech Communication*, 50(5), 434-451.