IRSTI 16.21.37
UDC 81.33 **DOI 10.70239/arsu.2025.t80.n2.28**

THE SIGNIFICANCE OF G2P MODELS FOR THE LOW-RESOURCE UZBEK LANGUAGE

HAMROYEVA SH.M., MAKHMUDJONOVA G.U.*

Hamroyeva Shahlo Mirdzhonovna - Doctor of philology, docent, Tashkent State University of Uzbek language and literature named after A. Navoi, Republic Uzbekistan
E-mail: hamroyeva81@mail.ru, https://orcid.org/0000-0002-5429-4708
***Makhmudjonova Gulshakhnoz Ulugbekqyzy** – PhD student, Tashkent State University of Uzbek Language and Literature named after A. Navoi, Tashkent Uzbekistan.
E-mail: gulshaxnozmahmudjonova@gmail.com, https://orcid.org/0009-0002-8536-0680

Abstract. This paper explores the challenges and advancements in Grapheme-to-Phoneme (G2P) conversion for the Uzbek language, a low-resource language in speech technology. G2P conversion is a fundamental component in text-to-speech (TTS), automatic speech recognition (ASR), and other natural language processing (NLP) applications. The accuracy of G2P directly affects speech synthesis quality and intelligibility. The study reviews traditional rule-based, statistical, and modern deep learning approaches to G2P modeling, highlighting their advantages and limitations. Particular attention is paid to the phonological features of Uzbek, such as vowel harmony, loanwords from Russian, Arabic, and Persian, and homographs, which pose specific difficulties in phoneme prediction. Existing tools like Phonetisaurus, Sequitur-G2P, and Muxlisa AI are evaluated, noting their potential and current shortcomings. Muxlisa AI represents a significant step in developing TTS systems for Uzbek, yet its reliance on rule-based methods limits pronunciation accuracy. The paper emphasizes the importance of integrating morphological analysis, prosodic modeling, and the International Phonetic Alphabet (IPA) for improved phoneme alignment and synthesis. Furthermore, it suggests adopting hybrid models combining rule-based systems with neural networks to overcome dataset scarcity and phonological complexity. The findings underscore the need for continued research in Uzbek-specific G2P modeling to improve speech technology solutions for underrepresented languages. By enhancing phonetic precision and naturalness, robust G2P models will significantly contribute to digital inclusivity and the broader application of AI-driven linguistic tools in Uzbek.
Key words: uzbek language, grapheme-to-phoneme (G2P), speech synthesis, phonetics, NLP, text-to-speech (TTS), automatic speech recognition (ASR)

Introduction. G2P conversion is the process of mapping graphemes (orthographic symbols) to phonemes (sound units) in a given language [1]. It plays a vital role in converting written text into phonetic transcriptions, which are essential for TTS, ASR, and other linguistic applications. While many rich-resource languages have robust G2P models, low-resource languages like Uzbek face significant challenges due to limited datasets and phonological structures. The accuracy of the G2P model directly affects the naturalness of speech synthesis and speech intelligibility. One of the essential components of TTS systems is G2P conversion, which ensures the naturalness, clarity, and phonetic harmony of synthesized speech. TTS systems are more common in a number of applications. Probably the main use of TTS today is in call-centre automation, where a user calls to pay an electricity bill or book some travel and conducts the entire transaction through an automatic dialogue system. Beyond this, TTS systems have been used for reading news stories, weather reports, travel directions and a wide variety of other applications. TTS systems generally consist of two primary components: Text processing based on natural language processing (NLP). Speech generation. The overall performance of TTS systems heavily depends on this NLP component. In linguistically complex languages, text normalization, G2P conversion, and prosodic modeling accuracy are crucial for ensuring high-quality speech synthesis. This paper analyzes the significance of G2P conversion in TTS systems for the Uzbek language, existing methods, and their effectiveness. Additionally, key challenges in the G2P process and the potential of statistical and deep-learning-based approaches to address these issues are explored.

G2P modeling. Grapheme-to-Phoneme (G2P) conversion transforms a target word from its written form (graphemes) into its phonetic representation (phonemes) by applying word segmentation and text normalization techniques [3]. Automatic grapheme-to-phoneme (G2P) conversion was initially developed for text-to-speech (TTS) applications. Following text normalization - where abbreviations, numerals, and special symbols are expanded—the input text must be converted into a sequence of phonemes, which serves as the foundation for controlling a speech synthesizer [4]. As shown in Figure 1, the G2P model converts words into phonemes by applying learned phonological rules.

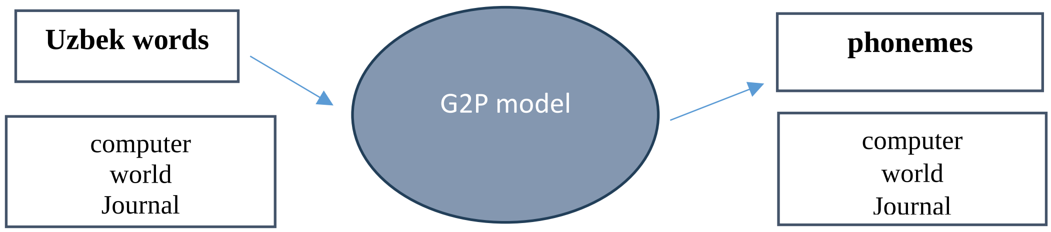

Figure 1. Grapheme-to-Phoneme (G2P) model

Related work. The development of G2P models has been extensively studied in various languages, with methodologies evolving from rule-based approaches to statistical and deep-learning models. The study by [5] explores the theoretical foundation of G2P conversion and its application to transliteration, highlighting the importance of phoneme alignment techniques. Recent advances in neural network-based G2P models have been presented in, where Transformer-based architectures were found to outperform traditional sequence-to-sequence models in phoneme prediction tasks. A comprehensive survey of G2P methods [6] categorizes existing approaches into rule-based, statistical, and neural models. Rule-based systems rely on predefined phonological rules but struggle with loanwords and irregular phoneme mappings. Statistical models such as Hidden Markov Models (HMMs) and Conditional Random Fields (CRFs) improve accuracy by learning from annotated datasets. More recent research in deep-learning models [7] demonstrates that Transformer-based G2P systems can generalize better, particularly for morphologically complex languages. In the context of low-resource languages, research [8] shows that phonetic transcription tools such as Phonetisaurus and Sequitur G2P are effective for training G2P models when annotated phonetic data is limited. Additionally, multilingual phoneme models, including those integrated into the Montreal Forced Aligner (MFA), have been explored in [9] to extend G2P coverage to underrepresented languages.

Several open-source G2P implementations have been developed, including:

- Phonetisaurus [10] - a WFST-based G2P conversion tool.
- Piper-Phonemize [12] - a phoneme conversion library for multilingual TTS.
- CharsiuG2P [12] - a ByT5-based multilingual G2P model, though it lacks correct phoneme mappings for Uzbek.
- Sequitur-G2P [13] - a data-driven tool widely used for grapheme-to-phoneme conversion.

Phonetisaurus and Sequitur-G2P employ statistical and deep learning-based approaches for phoneme conversion, whereas Muxlisa AI primarily relies on a rule-based phonetic system. Integrating an advanced G2P model could significantly enhance its pronunciation accuracy and overall speech synthesis performance.

Muxlisa AI [14] is a text-to-speech (TTS) system specifically developed for the Uzbek language. It leverages artificial intelligence (AI) and deep learning to generate high-quality speech synthesis. As Uzbek is a low-resource language in the field of speech technology, Muxlisa AI represents a significant step toward enhancing digital accessibility and communication in Uzbek.

Further research is needed to refine Uzbek-specific G2P models, incorporating hybrid approaches that combine rule-based phonology with machine-learning techniques. The integration of morphological analysis and prosody-aware phoneme prediction is expected to improve pronunciation accuracy in Uzbek

TTS systems.

G2P methods. G2P models can be broadly classified into **three main categories**: rule-based methods, statistical models, and deep learning approaches. Each of these methods has unique characteristics, advantages, and limitations, as illustrated in Figure 2. Rule-based G2P systems rely on an extensive set of grapheme-to-phoneme rules. The development of such a system requires substantial linguistic expertise to ensure accurate phoneme mapping and rule application. Statistical models improve upon rule-based approaches by learning probabilistic relationships between graphemes and phonemes from labeled data. These models are more **flexible** but require **large phonetic corpora** for effective training. Deri and Knight [15] propose statistical grapheme-to-phoneme (G2P) models using weighted finite-state transducers (WFST) trained on multilingual data from Wiktionary. Their approach demonstrates how statistical methods can leverage high-resource languages to enhance pronunciation accuracy for related low-resource languages.

Recent advancements in deep learning have led to significant improvements in G2P modeling by leveraging neural networks to **automatically learn phoneme structures** from large datasets. Sharma (2021) explored novel NLP approaches for enhancing Text-to-Speech (TTS) synthesis, particularly focusing on improving the grapheme-to-phoneme (G2P) conversion component. The research employed deep learning architectures, including Long Short-Term Memory (LSTM) networks combined with attention mechanisms. Sharma's work demonstrated that integrating advanced neural architectures significantly improves phoneme prediction accuracy and overall naturalness of synthesized speech. The study provides strong evidence supporting the effectiveness of deep learning-based G2P methods over traditional statistical and rule-based approaches, especially in handling linguistic complexities and context-aware pronunciation variations inherent in natural speech [16].

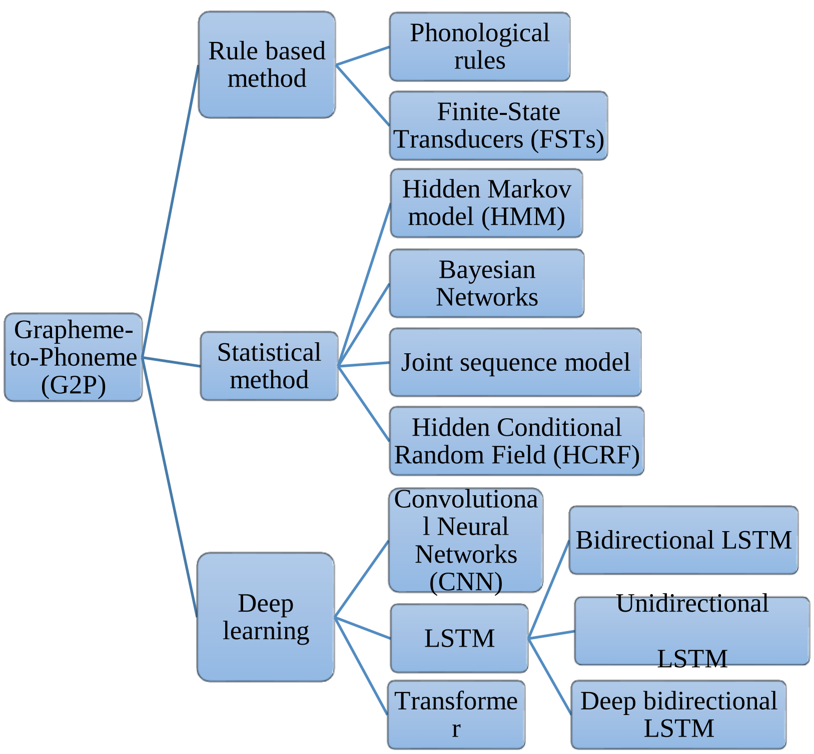

Figure 2. G2P methods

Materials and methods of research

This study focuses on the development and analysis of a Grapheme-to-Phoneme (G2P) conversion model for the Uzbek language, primarily applied in Text-to-Speech (TTS) synthesis, Automatic Speech Recognition (ASR), and Natural Language Processing (NLP) applications. Given the limited resources available for the Uzbek language in speech technology, our approach combines both rule-based and neural network-based models to address the challenges associated with phonetic transcription.

Materials. This study focused on the development and evaluation of the Grapheme-to-Phoneme (G2P) model for the Uzbek language, using a variety of materials. The primary data source was a phonetic corpus comprising annotated Uzbek texts, which served for both training and testing the model. To support model evaluation and comparison, several pre-existing G2P tools, such as Phonetisaurus, Sequitur-G2P, and Piper-Phonemize, were used. Additionally, both statistical models (e.g., Hidden Markov Models (HMMs)) and neural network models (including Transformer-based and LSTM-based architectures) were incorporated to improve G2P conversion accuracy. Model performance was assessed using standard evaluation metrics, including phoneme error rate (PER) and word error rate (WER).

Methods. The study applied a comprehensive methodology for the development and evaluation of the Grapheme-to-Phoneme (G2P) conversion model for the Uzbek language. Data preparation involved normalizing the raw text to address issues such as abbreviations, numerals, and special symbols. Abbreviations were expanded, numerals converted to their written forms, and unnecessary symbols were removed. This normalization was carried out using Python scripts specifically designed for Uzbek text, ensuring consistency across the dataset before G2P conversion. For the G2P conversion, the process was performed in two stages. Initially, rule-based models were used, applying predefined phonological rules specific to Uzbek. Subsequently, neural network models (including LSTM and Transformer architectures) were trained on the phonetic corpora to predict phoneme sequences. The models were optimized iteratively to enhance phoneme prediction accuracy. The evaluation and analysis of model performance was conducted using standard metrics, phoneme error rate (PER) and word error rate (WER), by comparing the predicted phoneme sequences to the ground truth transcriptions. In addition, qualitative analysis was performed to evaluate the naturalness and intelligibility of the synthesized speech in TTS applications. Finally, statistical analysis was carried out using descriptive statistics to compare the performance of the models based on PER and WER.

Phonetic Analysis. The phonetic analysis of Uzbek vowel phonemes highlights the significance of articulatory characteristics, particularly tongue positioning and lip rounding, in differentiating vowels. Lip rounding notably impacts the production of specific vowel qualities. This analysis [17] demonstrates that rounded vowels are subject to constraints arising from physiological and phonological factors, thus limiting their phonetic distribution. Additionally, it underscores the interplay between tongue position and the expansion of the oral cavity, both critical in determining vowel acoustics. Moreover, the study suggests that lip rounding does not establish a distinct articulatory space but rather modifies the acoustic properties of existing vowels, particularly open and mid-open vowels. A detailed understanding of Uzbek vowel phonetics is therefore essential for accurate phonetic modeling, especially in Grapheme-to-Phoneme (G2P) conversion and speech synthesis applications. Insights from this analysis can significantly enhance the precision and reliability of Uzbek language processing tools by accounting for articulatory and phonological variations. To further improve G2P accuracy, aligning the Uzbek alphabet with the International Phonetic Alphabet (IPA) is recommended, as the IPA [18] provides a standardized phonemic representation that facilitates precise speech synthesis and linguistic processing within TTS and ASR systems.

Degree of Mouth Opening	Front Vowels	Back Vowels
Narrow	**i (ʏ)**	**ь (y)**
Mid-Open	**e (ɵ)**	**(o)**
Open	ə	**a ɒ**

Table 1. Uzbek vowel phonemes

Based on our extensive discussions, analyses, and research on Grapheme-to-Phoneme (G2P) conversion, it is clear that developing a robust G2P model for the Uzbek language is crucial for several reasons:

Factors	***Description***
Phonetic Complexity and Variability	Uzbek phonetics involve vowel harmony, dialectal variations, and loanwords.
Text-to-Speech (TTS) Systems	High-quality TTS relies on accurate phoneme predictions. Precise G2P conversion ensures natural prosody, rhythm, and pronunciation, essential for intelligible speech synthesis.
Automatic Speech Recognition (ASR)	ASR requires accurate phoneme modeling. A robust G2P model enhances transcription accuracy by reducing phoneme recognition errors, particularly in spontaneous speech contexts.
Machine Translation and NLP Applications	Accurate phonetic transcriptions improve machine translation, text alignment, keyword search, and sentiment analysis by ensuring consistency between written and spoken forms.

Results and its discussion

The performance of the Grapheme-to-Phoneme (G2P) models was evaluated using rule-based phonological rules specific to the Uzbek language. The rule-based G2P model performed effectively with regular and standard words but showed limitations with more complex linguistic constructs. Its reliance on predefined phonological rules made it suitable for straightforward phoneme mapping but less adaptable to loanwords and words with irregular phonetic patterns. Although the rule-based approach provided a solid foundation, it encountered challenges in processing loanwords from languages like Russian, which do not always adhere to the phonetic conventions of Uzbek. Consequently, the model struggled with accurately mapping graphemes to phonemes in these cases.

Problems. Although Uzbek employs a phonemic orthography, several linguistic challenges complicate Grapheme-to-Phoneme (G2P) conversion. The presence of vowel harmony significantly impacts phoneme mapping accuracy, requiring context-sensitive phonetic analysis. Additionally, the incorporation of numerous loanwords from languages like Russian, Arabic, and Persian often diverges from standard Uzbek phonetic conventions, complicating consistent phoneme representation. The existence of homographs, where words share the same spelling but differ in pronunciation, further complicates the process, necessitating advanced context-aware processing. Lastly, there is a scarcity of comprehensive phonetic datasets for Uzbek, making it difficult to develop and train robust G2P models.

Conclusion. This study highlights the importance of Grapheme-to-Phoneme (G2P) modeling for the Uzbek language, emphasizing its critical role in text-to-speech (TTS) synthesis, automatic speech recognition (ASR), and broader NLP applications. Given Uzbek's phonological complexity and low-resource status in speech technology, developing an accurate G2P system is essential to enhancing pronunciation accuracy and the naturalness of synthesized speech.

Developing a high-quality G2P model for Uzbek requires expanding phonetic and textual corpora, implementing hybrid models that combine rule-based and neural network approaches, and integrating morphological analysis to improve context-based pronunciation predictions. As Uzbek language technologies continue to advance, an optimized G2P model will significantly bridge the gap between written and spoken language, thereby enhancing the effectiveness and accuracy of AI-driven linguistic applications.

References

1. https://docs.nvidia.com/nemo-framework/user-guide/24.09/nemotoolkit/tts/g2p.html
2. Taylor P. Text-to-Speech Synthesis. Cambridge University Press, 2009.

3. Yi L., Li J., Hao J., & Xiong Z. Improved Grapheme-to-Phoneme Conversion for Mandarin TTS. Tsinghua Science & Technology, (2009), 14, 606-611. doi: 10.1016/S1007-0214(09)70124-5.

4. Bisani M., & Ney H. Joint-Sequence Models for Grapheme-to-Phoneme Conversion. Speech Communication, (2008).

5. Jiampojamarn S. Grapheme-to-Phoneme Conversion and Its Application to Transliteration. University of Alberta, 2009. Retrieved from https://era.library.ualberta.ca

6. Jolchieva S., Nemet G., & Giresh B. Preobrazovanie grafem v fonemy na osnove transformera. Trudy Interspeech, (2019).

7. Cheng S., Zhu P., Liu J., & Wang Z. A Survey of Grapheme-to-Phoneme Conversion Methods. Applied Sciences, (2024), 14(24), 11790. doi:10.3390/app142411790.

8. Novak J.R., Minematsu N., & Hirose K. WFST-Based Grapheme-to-Phoneme Conversion: Open Source Tools for Alignment, Model-Building and Decoding. Proceedings of the 10th International Workshop on Finite State Methods and Natural Language Processing, Donostia–San Sebastián, Spain, (2012).

9. https://github.com/MontrealCorpusTools/mfa-models

10. https://github.com/AdolfVonKleist/Phonetisaurus

11. https://github.com/rhasspy/piper-phonemize/tree/master

12. https://github.com/lingjzhu/CharsiuG2P

13. https://github.com/sequitur-g2p/sequitur-g2p

14. https://muxlisa.uz

15. Deri A., & Knight K. Grapheme-to-Phoneme Models for (Almost) Any Language. Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics, Berlin, Germany, (2016), 399–408.

16. Mirtozhiev M.M. Fonetika uzbekskogo yazyka. Akademiya nauk Uzbekistana, izdatel'stvo «Fan», Tashkent, (2013).

17. Sharma M. Novel NLP Methods for Improved Text-To-Speech Synthesis (Doctoral dissertation), (2021).

18. https://www.ipachart.com

Әдебиеттер тізімі

1. https://docs.nvidia.com/nemo-framework/user-guide/24.09/nemotoolkit/tts/g2p.html

2. Taylor P. Text-to-Speech Synthesis. Cambridge University Press, 2009.

3. Yi L., Li J., Hao J., & Xiong Z. Improved Grapheme-to-Phoneme Conversion for Mandarin TTS. Tsinghua Science & Technology, (2009), 14, 606-611. doi: 10.1016/S1007-0214(09)70124-5.

4. Bisani M., & Ney H. Joint-Sequence Models for Grapheme-to-Phoneme Conversion. Speech Communication, (2008).

5. Jiampojamarn S. Grapheme-to-Phoneme Conversion and Its Application to Transliteration. University of Alberta, 2009. Retrieved from https://era.library.ualberta.ca

6. Йолчиева С., Немет Г., и Гиреш Б. Преобразование графем в фонемы на основе трансформера. Труды Interspeech, (2019).

7. Cheng S., Zhu P., Liu J., & Wang Z. A Survey of Grapheme-to-Phoneme Conversion Methods. Applied Sciences, (2024), 14(24), 11790. doi:10.3390/app142411790.

8. Novak J.R., Minematsu N., & Hirose K. WFST-Based Grapheme-to-Phoneme Conversion: Open Source Tools for Alignment, Model-Building and Decoding. Proceedings of the 10th International Workshop on Finite State Methods and Natural Language Processing, Donostia–San Sebastián, Spain, (2012).

9. https://github.com/MontrealCorpusTools/mfa-models

10. https://github.com/AdolfVonKleist/Phonetisaurus

11. https://github.com/rhasspy/piper-phonemize/tree/master
12. https://github.com/lingjzhu/CharsiuG2P
13. https://github.com/sequitur-g2p/sequitur-g2p
14. https://muxlisa.uz
15. Deri A., & Knight K. Grapheme-to-Phoneme Models for (Almost) Any Language. Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics, Berlin, Germany, (2016), 399–408.
16. Миртожиев М.М. Фонетика узбекского языка. Академия наук Узбекистана, издательство «Фан», Ташкент, (2013).
17. Sharma M. Novel NLP Methods for Improved Text-To-Speech Synthesis (Doctoral dissertation), (2021).
18. https://www.ipachart.com

ӨЗБЕК ТІЛІНІҢ РЕСУРСТАРЫ ШЕКТЕУЛІ ЖАҒДАЙЫНДА ГРАФЕМАДАН ФОНЕМАҒА ТҮРЛЕНДІРУ (G2P) МОДЕЛЬДЕРІНІҢ МАҢЫЗДЫЛЫҒЫ

ХАМРОЕВА Ш.М., МАХМУДЖОНОВА Г.Ұ*

Хамроева Шахло Мирджоновна – Филология ғылымдарының докторы, доцент, А. Навои атындағы Ташкент мемлекеттік өзбек тілі және әдебиеті университеті, Ташкент қ., Өзбекстан Республикасы
E-mail: hamroyeva81@mail.ru, https://orcid.org/0000-0002-5429-4708
***Махмуджонова Гүлшахноз Ұлықбекқызы** – PhD докторанты, А.Навои атындағы Ташкент мемлекеттік өзбек тілі және әдебиеті университеті, Ташкент қ., Өзбекстан Республикасы.
E-mail: gulshaxnozmahmudjonova@gmail.com, https://orcid.org/0009-0002-8536-0680

Аңдатпа. Бұл мақалада жазбаша мәтінді дыбыстық формаға түрлендіруде маңызды рөл атқаратын графема-фонема (G2P) конверсиясының өзбек тіліне қатысты ерекшеліктері қарастырылады. G2P моделінің дәлдігі сөйлеу синтезінің табиғилығы мен түсініктілігіне тікелей әсер етеді. Өзбек тілі – ресурстары шектеулі тілдердің бірі ретінде G2P модельдерін дамытуда бірқатар қиындықтарға тап болады. Мақалада қазіргі қолданыстағы әдістер: ережеге негізделген, статистикалық және терең нейрондық желілерге сүйенетін G2P тәсілдері сипатталып, олардың артықшылықтары мен шектеулері сарапталады. Авторлар G2P конверсиясының мәтінді дыбыстау жүйелеріндегі (TTS), автоматты сөйлеуді тану (ASR), машиналық аударма және лингвистикалық талдау сияқты қосымшалардағы маңызын көрсетеді. Мақалада өзбек тілінің фонетикалық күрделілігі, әсіресе дауысты дыбыстардың артикуляциялық ерекшеліктері мен фонологиялық үйлесімділігі, сондай-ақ өзге тілдерден енген сөздер мен омографтар сияқты факторлар G2P моделін қиындататыны көрсетілген. Сонымен қатар, авторлар Muxlisa AI секілді өзбек тіліне арналған TTS жүйелерінің жетістіктері мен шектеулерін талдап, заманауи G2P модельдерін енгізудің маңыздылығына тоқталады. Болашақта өзбек тіліндегі нақты әрі табиғи дыбыстауды қамтамасыз ету үшін морфологиялық талдау мен IPA жүйесіне сәйкестендіру ұсынылады. Бұл зерттеу ресурсы шектеулі тілдер үшін тиімді G2P модельдерін әзірлеуге жол ашады.
Түйін сөздер: өзбек тілі, графемадан фонемаға түрлендіру (G2P), сөйлеу синтезі, фонетика, табиғи тілді өңдеу NLP, мәтіннен сөйлеуге синтездеу (TTS), автоматты сөйлеуді тану (ASR).

ЗНАЧИМОСТЬ G2P-МОДЕЛЕЙ ДЛЯ МАЛОРЕСУРСНОГО УЗБЕКСКОГО ЯЗЫКА

ХАМРОЕВА Ш.М., МАХМУДЖОНОВА Г.Ұ*

Хамроева Шахло Мирджоновна - Доктор наук, доцент, Ташкентский государственный университет узбекского языка и литературы имени А. Навои, г. Ташкент, Республика Узбекистан.
E-mail: hamroyeva81@mail.ru, https://orcid.org/0000-0002-5429-4708
***Махмуджонова Гулшахноз Улугбек кызы** – Докторант (PhD), Ташкентский государственный университет узбекского языка и литературы им. А. Наваи, г. Ташкент, Республика Узбекистан.
E-mail: gulshaxnozmahmudjonova@gmail.com, https://orcid.org/0009-0002-8536-0680

Аннотация. В статье рассматривается проблема конверсии графем в фонемы (G2P) применительно к узбекскому языку, который относится к числу малоресурсных языков. Процесс G2P играет ключевую роль в системах синтеза речи (TTS), автоматического распознавания речи (ASR) и других лингвистических приложениях. Авторы подробно анализируют существующие подходы к G2P-моделированию: правила, статистические методы, а также современные нейросетевые архитектуры, в частности, трансформеры и LSTM. Особое внимание уделяется специфике фонетики узбекского языка: гармонии гласных, артикуляционным особенностям, наличию заимствованных слов и омографов, что создает определённые трудности при построении точных G2P моделей. Рассматриваются возможности использования открытых инструментов, таких как Phonetisaurus, Sequitur-G2P и CharsiuG2P, а также система Muxlisa AI, предназначенная для синтеза речи на узбекском языке. Указывается на необходимость гибридных моделей, сочетающих правила с методами машинного обучения, а также интеграции морфологического анализа и учёта просодических особенностей речи. Особое значение придается согласованию алфавита узбекского языка с международным фонетическим алфавитом (IPA) для повышения точности транскрипций. Статья подчеркивает актуальность дальнейших исследований в области G2P-моделирования для узбекского языка с целью создания высококачественных речевых технологий и расширения их применения в цифровой среде.

Ключевые слова: узбекский язык, графемно-фонемное преобразование (G2P), синтез речи, фонетика, обработка естественного языка (NLP), преобразование текста в речь (TTS), автоматическое распознавание речи (ASR)