Phonological Classification and Acoustic Analysis of Vowel Phonemes in the Uzbek Language

Bakhtiyor Mengliyev
Department of Uzbek Language and Literature
Tashkent University of Applied Sciences
Tashkent, Uzbekistan
0000-0002-6757-0476

Shahlo Khamroeva
Computational linguistic and digital technology
Tashkent State University of Uzbek language and Literature
Tashkent, Uzbekistan
0000-0002-5429-4708

Gulshakhnoz Makhmudjonova
Computational linguistic and digital technology
Tashkent State University of Uzbek language and Literature
Tashkent, Uzbekistan
009-0002-8536-0680

Saidislombek Abdusamatov
Internet and information communication
Korea International University in Ferghana
Fergana, Uzbekistan
0009-0009-9437-1315

2025 10th International Conference on Computer Science and Engineering (UBMK) | 979-8-3315-9975-1/25/$31.00 ©2025 IEEE | DOI: 10.1109/UBMK67458.2025.11206747

***Abstract*—This paper investigates the articulatory and acoustic properties of Uzbek vowel and consonant phonemes. Using data from native speaker recordings and analysis via Praat software, we measure formant frequencies (F1, F2, F3) and propose a revised classification of Uzbek phonemes. Particular attention is given to allophonic variation, acoustic correlates of articulation, and phoneme distribution in literary Uzbek. The study supports phonetic modeling for Uzbek TTS and G2P systems.**

Keywords—Uzbek phonetics, vowel formants, consonant acoustics, G2P, speech synthesis, Praat

I. Introduction

Phonetics is a branch of linguistics that studies the articulatory and acoustic aspects of speech sounds. It focuses on how sounds are produced by the speech organs and how their physical properties can be measured and analyzed. Within Uzbek linguistics, phonetic and phonological research primarily focuses on the speech sounds of the Uzbek language – in particular, on their articulatory (physiological), acoustic, and perceptual characteristics – as well as on their functional roles in communication. Alongside the segmental sounds, the prosodic features of Uzbek speech (such as syllable structure, stress patterns, and intonation) constitute an important part of the phonological system and are given due attention. The vowel inventory of modern Uzbek literary language consists of six phonemic vowels (/a/, /e/, /i/, /o/, /u/, /oʻ/), and these vowel phonemes are traditionally classified by the position and shape of the speech organs: namely, by tongue height (high, mid, low), tongue advancement (front vs. back), and lip rounding (rounded vs. unrounded) [1].

However, a purely phonological classification does not fully capture the physical realization of vowel sounds. To gain a more comprehensive understanding of Uzbek vowels, it is necessary to examine their acoustic profiles using instrumental measurements. An acoustic analysis of vowel phonemes – considering parameters such as formant frequencies, intensity, and duration – provides objective data on how these sounds are realized in actual speech. By combining phonological classification with acoustic analysis, we can correlate the abstract distinctive features of each vowel with concrete acoustic signals, thereby deepening our understanding of how Uzbek vowel phonemes function both in theory and in practice.

II. Literature Review

In the field of Uzbek phonological studies, there is ongoing debate regarding the number of vowel phonemes in the language. While the Uzbek alphabet officially includes six vowel letters, various phonetic sources report between six and fourteen vowel phonemes. This variation stems from differences in articulatory and acoustic classification. Some sources maintain the presence of only six core vowel phonemes, whereas others extend the inventory by including additional allophonic variants[2], [3]. Miraziz Mirtojiyev proposes eleven articulatory vowel positions in modern Uzbek and attributes the presence of additional phonemes to Russian loanwords. Based on this, he argues for the existence of fourteen vowel phonemes in the standard literary Uzbek language [4]

Historically, vowel harmony has played a central role in the phonological structure of Turkic languages. Notably, in Old Turkic, Middle Turkic, and Early Uzbek written sources, the vowel harmony system was active, with researchers documenting eight vowels in Old Turkic, nine in Middle Turkic, and ten in Early Uzbek [5].

In addition to historical and structural analyses, recent research has increasingly focused on the acoustic and phonological properties of Uzbek vowel phonemes. MacMillan's descriptive phonetic study [6] presents a comprehensive analysis of vowel articulation in Uzbek, with particular emphasis on articulatory parameters such as tongue height distribution. The study specifically highlights that high vowels occur more frequently than low vowels within the Uzbek vowel system. The study identifies seven core vowel phonemes alongside ten phonetic variants. Although the audio data was recorded from a student originally from Samarkand,

the study bases its phonological generalizations on the Ferghana and Tashkent dialects, which are typically regarded as representative of the standard literary Uzbek language. More recently, Ido [7] has investigated ongoing variation in the realization of Uzbek vowel phonemes. These observations highlight the evolving nature of Uzbek vowel phonology and underscore the need for any comprehensive classification to integrate both diachronic developments and contemporary phonetic variability.

III. Results And Analysis

A. Phonological classification of Uzbek vowel phonemes

A phoneme is the smallest distinctive sound unit in a language, serving to construct morphemes and words. Nurmonov [8] emphasizes two essential aspects of the phoneme: its articulatory-acoustic dimension and its functional-semantic role. These dual dimensions provide the basis for understanding the phoneme as an independent unit within the overall linguistic system.

In all major Uzbek phonetics textbooks [9][10] four vowels, namely /i/, /u/, /a/, /o'/, are described as having allophonic variations when occurring adjacent to velar or uvular consonants. These variations are commonly analyzed as back vowel phonemes due to their coarticulatory influence. Moreover, a distinction is made between "hard" and "soft" vowels, which also contributes to phonemic differentiation. According to Chomsky and Halle's generative phonology framework, the phonological distinction between so-called "hard" and "soft" vowels is formally represented by the features [+/- back] and [+/- high]. Front vowels are thus considered "soft," while back vowels are treated as "hard." The distinction between "hard" and "soft" vowels plays a critical role, as it allows these sounds to function as independent phonemes, contributing to semantic contrasts [11]. This type of phonemic distinction is also illustrated in Shavkat Rahmatullayev's *Dictionary of Homonyms*. Representative examples of such contrasts are provided in Table I.

TABLE I. Uzbek Homonymous Word Pairs Whose Semantic Contrast Is Determined By Vowel Phoneme Quality

Homonyms	Meaning	Vowel characteristics	Part of speech
is	smell of sweat	soft, long	noun
is	smoke	hard, short	noun
tuz	to regulate	soft	verb
tuz	salt	hard	noun
tur	type, form	soft	noun
tur	to stand upright	hard	verb
uch	number	soft	numeral
uch	tip, edge	hard, long	noun
o'z	oneself	soft, long	pronoun
o'z	to overflow	hard, short	verb
to'r	the upper end of the room	soft	noun
to'r	trap	hard	noun
o'ng	to wake up	soft	verb
o'ng	comfortable	hard	adjective
o'r	to braid	soft	verb
o'r	to reap with a sickle	hard	verb
o't	to pass	soft, short	verb
o't	fire	hard, long	noun

B. Phonological descriptions of Uzbek vowel phonemes

Uzbek vowels are analyzed in terms of their phonological characteristics and IPA (International Phonetic Alphabet) transcription.

The vowel phoneme /a/ [a, IPA] is predominantly realized as near-open central unrounded vowel. In the Tashkent and Fergana dialects, it functions as an open, combinatory front vowel, as in a*k*a *(brother)*, *m*a*k*t*a*b *(school)*, *m*a*n*a *(that)*. It can appear in any syllable position and frequently co-occurs with the consonants /k/ and /g/, as in *agar (if)*, *dangal (clear-cut)*, *gavhar (pearl)*, *shakar (sugar)*, *kapalak (butterfly)*. When occurring in initial position (anlaut), its duration ranges between 90-110 milliseconds; in non-stressed medial syllables (inlaut), it shortens to 50 milliseconds; and in final position (auslaut), it may be extended up to 210 milliseconds [4]. The back allophone of this phoneme represents a more posterior articulatory variant. It becomes clearly perceptible in specific phonetic environments, particularly when occurring adjacent to post-velar consonants. In the Andijan, Fergana, and Kokand dialects, as well as in borrowings from Russian, when adjacent to deep back consonants /q/, /g'/, and /x/, the back open unrounded vowel [ɑ, IPA] is used, as in *qalam (pen)*, *qanaqa (which)*, *xabar (news)*, *mag'iz (kernel)*, *g'ayrat (diligence)* [10].

The vowel phoneme /i/ manifests in two forms based on its acoustic and articulatory properties and the horizontal movement of the tongue: the close front unrounded vowel [i, IPA] and the close central unrounded vowel [ɨ, IPA]. Depending on sound combinations and syllable structure, the /i/ phoneme appears in various positional variants. In native Uzbek vocabulary, /i/ occurs in syllables in odd positions [4], and in closed syllables between front consonants such as /r, l, n, m, z/, it is pronounced as a weak front vowel and realized as the allophone [ɪ, IPA], e.g., bɪlak (wrist), tɪlak (wish), sɪz (you), tɪl (tongue) [10]. When combined with the consonant /y/, the /i/ phoneme is pronounced as a long vowel /i:/, as in *si:ymo (appearance)*, and *si:yrak (sparse)* [11]. In word-initial position, its duration ranges from 110-140 ms, as an inlaut it lasts 40-70 ms, and in final position about 110 ms [4]. Notably, when adjacent to deep back consonants (q, g', x), it is heard as /ɨ/, as in qɨz (girl), qɨl (hair), xɨrmon (harvest). Additionally, in many polysyllabic words borrowed from Arabic and Persian, the /i/ vowel in final syllables becomes a long unrounded vowel, e.g., *di:n (religion)*, *i:mon (faith)*, *nasi:b (destiny)*, *mudi:ra (meneger)*, *zai:f (weak)* [11].

The vowel phoneme /u/ in Uzbek literary language occurs in two main forms. These forms differ based on the horizontal movement of the tongue. The first variant is a labialized, close, central soft vowel [ʉ, IPA]. During its articulation, the tip of the tongue touches the lower teeth, and the body of the tongue moves slightly forward. The second variant is a labialized, close, back hard vowel [u, IPA]. It is pronounced softly in conjunction with the consonants /k/ and /g/, whereas it acquires a hard quality when adjacent to deep back consonants such as /q/, /g'/, /x/, and pharyngeal /h/, as seen in examples like *quloq (ears)*, *qumloq (sand)*, *xurmo (date)*, and *g'uldur (babbling)* [4]. When pronounced in the context of front or back consonants, the /u/ vowel tends to be realized as soft. In such cases, derivational suffixes added to the root also reflect this softness, as in *tʉzʉk (good)*, *tʉrkʉm (category)*, *kʉndʉz (daytime)*, *kʉlgi (laugh)*, and *gʉrʉch (rice)*. The presence of /u/ in initial syllables is particularly evident in dialects of the Fergana Valley and the Karluk and Chigil branches [4]. In

addition, Ne'matov and Abduazizov argue that the /u/ vowel phoneme may undergo reduction, weakening, or manifest as allophonic variants depending on its intra-syllabic position.

The vowel phoneme /e/ [e, IPA] in standard Uzbek is classified as an unrounded mid-open front vowel. In words of Turkic origin, the phoneme is consistently pronounced as a front soft vowel. In native Uzbek words, it typically co-occurs with front vowels and back consonants /k/ and /g/, and is most commonly found in monosyllabic words and in the first syllable of polysyllabic words, such as *bek (lord)*, *bel (waist)*, *ket (to go)*, *kel (to come)*, *sezgi (intuition)*, *sen (you)*, *sep (dowry)*, and *teng (equal)*. In closed syllables, especially when occurring between consonants, it tends to be slightly raised and pronounced more narrowly, as in *mehnat (labour)*, *sevgi (love)*, *telba (crazy)*, *belgi (sign)*, and *chehra (face)*. In closed syllables that begin with a vowel and end in a consonant, the /e/ phoneme is articulated more openly, as in *egri (curve)*, *ezma (chatterbox)*, and *ehson (charity)* [12]. Acoustic measurements indicate that in the initial syllable following consonants, /e/ is realized with a duration of approximately 60-85 milliseconds, while at the beginning of a word, its duration may reach 80-85 milliseconds [4].

The vowel phoneme /o'/ displays both soft and hard variants depending on its position: for instance, *qol (hand)* and *kөl (lake)* illustrate a labialized, mid-central soft vowel [ɵ, IPA], and a labialized, mid-back hard vowel [o, IPA]. Compared to the front vowel /ɵ/, the back vowel /o/ is articulated with a more open quality [11]. In two-phoneme words or syllables, it is slightly more open in articulation, as in *өt (fire)*, *өch (revenge)*, and *өtkir (sharp)* [12]. When adjacent to sonorant consonants, it tends to be realized with a slightly lengthened allochron, as in *ɵ:rin (place)*, *sɵ:ng (after)*, and *tɵ:n (robe)* [13]. This soft vowel variant is especially found in the dialects of the Fergana Valley, notably in the Qarluq and Uyghur dialects. In all Uzbek dialects, the labialized, mid-back hard vowel [o] occurs and is typically found adjacent to deep or pharyngeal consonants, as seen in *qozi* (lamb) and *qol (hand)* [4].

The vowel phoneme /o/ [ɒ, IPA] is articulated as a back, open vowel with passive labial involvement. It typically occurs in monosyllabic words such as *nɒn (bread)*, *dɒn (grain)*, *bɒsh (head)*, *ɒsh (the national food)*, *qɒsh (eyebrow)*, and *ɒz (little)*. Regardless of whether the syllable is open or closed, it frequently appears in the second syllable of words like *kitɒb (book)*, *ɒbɒd (prosperous)*, *bɒrmɒq (to go)*, *dɒnɒ (wise)*, and *ammɒ (but)* [5]. In words borrowed from Arabic and Persian, the phoneme appears in any syllable and is pronounced as a long vowel, as in *ɒ:shiq* (beloved) and *ɒ:lim (scientist)*. When appearing in initial position (anlaut), it is pronounced with a duration of approximately 60-90 milliseconds, while in stressed syllables, it ranges from 95 to 140 milliseconds in length [4].

Based on the analyses, it can be concluded that the Uzbek language contains nine vowel phonemes, while the remaining vowel sounds are considered allophonic variants. The following table (Table II) was compiled accordingly and aligned with the International Phonetic Alphabet (IPA).

TABLE II. UZBEK VOWEL PHONEMES (IPA-BASED).

Vowel backness	**Front**		**Central**		**Back**	
Vowel height/ Vowel rounding	Unrounded	Rounded	Unrounded	Rounded	Unrounded	Rounded
High Close	i		ɨ	ʉ		u
Upper-mid Close-mid	*e*			ɵ		
Lower-mid open-mid						o
Low Open	a					ɒ

C. Acoustic analysis of Uzbek vowel phonemes

The phonological classification of vowel phonemes in Uzbek is based not only on articulatory descriptions but also on acoustic measurements. In his analysis of literary Uzbek vowels, Abduazizov measured formant frequencies (F1, F2, F3), spectral amplitude, harmonic structure, and vowel duration, with lengths reaching up to 2.4 seconds. Formants defined as peaks in the acoustic spectrum resulting from the resonances of the vocal tract are crucial in determining the perceptual quality of vowels. They are also used to evaluate the fluency and intelligibility of speech [13]. Among these, the F1-F2 formant relationship has been shown to be a particularly strong indicator of vowel quality and has been applied in pronunciation assessment, phonetic training, and developmental studies [14]. Most vowel sounds in the world's languages are characterized by three formants: F1, F2, and F3. Of these, the first two formants (F1 and F2) are considered sufficient for determining the vowel's articulatory and acoustic properties [15]. The first formant (F1) is inversely related to vowel height: a low F1 corresponds to a high close vowel, while a high F1 indicates a low (open) vowel. The second formant (F2) is directly associated with tongue advancement: higher F2 values reflect front vowels, while lower F2 values indicate back vowels.

Using **Praat software [16]**, we measured the formant values of vowel phonemes extracted from a series of homonymous Uzbek word pairs, as listed in Table III. These pairs, which contrast only in vowel quality, offer a controlled context for analyzing phonemic distinctions. For example, the contrast between /i/ and /ɨ/ in *is* 'smell' vs. *ɨs* 'smoke', or between /o/ and /ɵ/ in *ot* 'fire' vs. *өt* 'to pass', reflects differences in tongue height and backness, supported by corresponding shifts in F1 and F2 values.

The measured formant values confirm that each Uzbek vowel occupies a distinct position in the acoustic vowel space. This separation supports the phonemic status of variants such as /ɨ/, /ʉ/, /ɵ/, and /ɒ/, which are often considered allophonic in traditional grammars but exhibit consistent and distinctive acoustic patterns. A spectrographic comparison of selected pairs further highlights these contrasts, particularly in F1-F2 configurations.

TABLE III. FORMANT FREQUENCIES (F1, F2, F3) OF VOWEL PHONEMES IN SELECTED UZBEK HOMONYMOUS WORDS

Homonyms (meaning)	F1	F2	F3
is (smell of sweat)	383	2664	3230
ɨs (smoke)	502	1442	4748
tʉz (to regulate)	474	1610	2831
tuz (salt)	418	1035	2811
tʉr (type, form)	538	1716	2831
tur (to stand upright)	422	956	2901
ʉch (number)	435	1034	2406
uch (tip,edge)	456	1869	2841
ɵz (oneself)	473	1437	2926
oz (to overflow)	531	850	3214
tɵr (the upper end of room)	450	1590	2758
tor (trap)	490	1036	2958
ɵng (to wake up)	473	1600	2675
ong (comfortable)	534	1009	2192
ɵr (to braid	473	1500	2642
or (to reap with a sickle)	518	1018	2872
ɵt (to pass)	511	1688	2757
ot (fire)	572	1011	2999

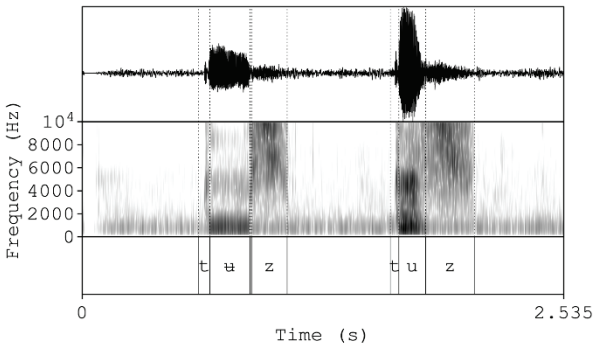

Fig. 1. Spectrogram and waveform of the word pair is (smell of sweat) vs. ɨs (smoke).

Differences in F1 and F2 values reflect a phonemic contrast between front and central vowels.

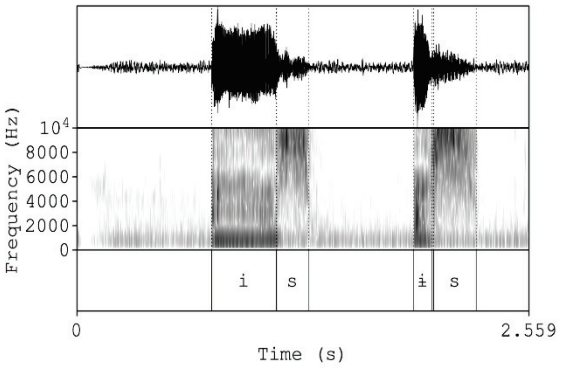

Fig. 2. Spectrogram and waveform of the word pair tʉz (to regulate) vs. tuz (salt).

The pair illustrates acoustic differences between the central vowel [ʉ] and the back vowel [u].

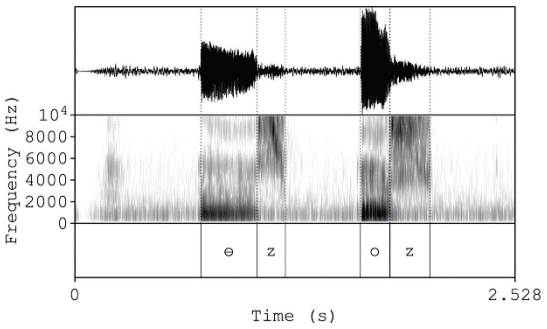

Fig. 3. Spectrogram and waveform of the word pair ɵz (oneself) vs. oz (to overflow).

A clear shift in F2 values reflects vowel frontness variation.

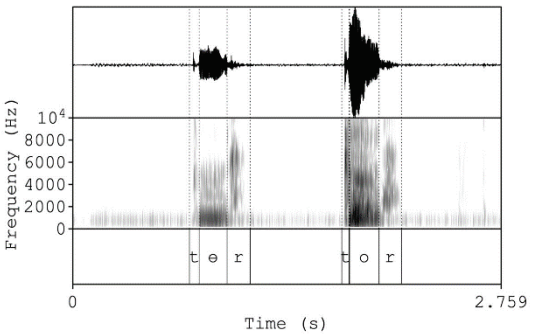

Fig. 4. Spectrogram and waveform of the word pair tɵr (upper end of room) vs. tor (trap).

Contrast between [ɵ] and [o] highlights the rounding and backness distinction in mid vowels.

Fig. 5. Acoustic vowel space of selected Uzbek words plotted based on mean formant frequencies.

The graph displays vowels according to their F1 (y-axis, vowel height) and F2 (x-axis, vowel frontness) values, measured in Hertz (Hz) using Praat software. Each point

represents a vowel extracted from a native Uzbek word, with color coding indicating vowel categories (e.g., /i/, /e/, /u/, /o/, /a/, etc.) following IPA transcription. The plot illustrates the articulatory-acoustic distribution of Uzbek vowel phonemes, where lower F1 values correspond to higher tongue positions and higher F2 values correspond to more fronted articulations.

Based on this observation, we propose an extended vowel phoneme table for Uzbek to be used in TTS and G2P systems, taking into account not only phonemes but also their context-dependent allophonic realizations. The following table presents the suggested phonemes, their articulatory descriptions, and lexical examples:

TABLE IV. PROPOSED UZBEK VOWEL PHONEME INVENTORY FOR TTS AND G2P SYSTEMS

IPA	Description	Example
i	Close front unrounded	Ip, it, ich, miya
ɨ	Close central unrounded	xɨrmon, qɨzil, g'ɨsht, xɨl,
ɪ	Open-mid central unrounded	bɪlak, tɪlak, tɪl, bɪl, sɪz
ʉ	Close central rounded	tʉzuk, kʉnduz, gʉruch, sʉt, pʉch
u	Close back rounded	quloq, qumloq, g'uldur, xunuk, xurmo
e	Close-mid front unrounded	bek, kel, egri, ezma, sezgi
ɵ	Close-mid central rounded	chɵpon, bɵlak, kɵl, sɵl gɵl
o	Close-mid back rounded	qozi, qol, qoshni xorak, g'orlik
ɒ	Open back rounded	ɒna, tɒr, yɒsh, bɒla, bɒzor
a	Open front unrounded	aka, kasb, danak, gavhar, baraka
ɑ	Open back unrounded	ɑqcha, ɑql, vɑqt, g'ɑflat, xɑ

The acoustic measurements of vowel phonemes extracted from Uzbek homonymous words confirm clear distinctions in formant frequency distributions, supporting an expanded phonemic inventory for the Uzbek language.

IV. CONCLUSION

This study provides phonological and acoustic evidence supporting a revision of the Uzbek vowel system. Although the Uzbek alphabet officially recognizes only six vowels, the analysis identifies nine core vowel phonemes based on distinct formant patterns. The remaining variants are classified as allophones. For practical applications in speech technologies such as text-to-speech (TTS) and grapheme-to-phoneme (G2P) modeling a total of eleven vowel units have been selected to accurately reflect the phonetic variation present in spoken Uzbek. Furthermore, the study highlights that vowels, especially /i/, undergo significant coarticulatory shifts depending on adjacent consonants, necessitating a dynamic approach to vowel classification. These insights contribute to building more precise and linguistically informed ASR, TTS, and G2P systems for the Uzbek language.

REFERENCES

[1] S. Otamirzayeva and M. Yusupova, "O'zbek tili (Fonetika, grafika, orfoepiya va orfografiya)". Tashkent: O'zbekiston Milliy Ensiklopediyasi, 2004. – Б. 17.

[2] U. Tursunov, A. Muxtorov, and Sh. Rahmatullaev, "Hozirgi o'zbek adabiy tili". Tashkent: O'zbekiston, 1992.

[3] M. T. Irisqulov, "Tilshunoslikka kirish", Tashkent: Yangi asr avlodi, 2009.

[4] M. M. Mirtojiyev, "O'zbek tili fonetikasi". Tashkent: Fan, 2013.

[5] Q. Mahmudov, "O'zbek tilining tarixiy fonetikasi". Tashkent: Ixtiyor, 2006.

[6] D. MacMillan, Descriptive Phonetics of Uzbek. https://www.yumpu.com/en/document/view/33134773/descriptive-phonetics-of-uzbek

[7] S. Ido, "Uzbek," Journal of the International Phonetic Association, pp. 1-17, 2025. doi:10.1017/S0025100324000148

[8] A. Nurmonov, O'zbek tilining fonologiyasi va morfonologiyasi, Toshkent: O'qituvchi, 1990.

[9] F.Kamal, Hozirgi zamon o'zbek tili: Leksikologiya, fonetika, grafika, orfografiya, morfologiya. Tashkent: Fanlar Akademiyasi nashriyoti, 1957.

[10] Hozirgi o'zbek adabiy tili. Jild I: Fonetika, leksikologiya, morfologiya. Tashkent: O'zbekiston SSR "Fan" nashriyoti, 1966.

[11] H. Ne'matov, O'zbek tili tarixiy fonetikasi: Pedagogika institutlari va universitetlarning filologiya fakultetlari uchun qo'llanma. Tashkent: O'qituvchi, 1992.

[12] A. A. Abduazizov, "O'zbek tilining fonologiyasi va morfonologiyasi". Toshkent: O'qituvchi, 1992.

[13] A. Chen, A. R. Bradlow, N. Minematsu, et al., "Vowel space metrics and intelligibility in non-native speech," in Proc. Interspeech, 2010.

[14] H. K. Vorperian and R. D. Kent, "Vowel acoustic space development in children: A synthesis of acoustic and anatomic data," J. Speech Lang. Hear. Res., vol. 50, no. 6, pp. 1510–1545, Dec. 2007.

[15] Y. Korkmaz and A. Boyacı, "Classification of Turkish Vowels Based on Formant Frequencies," in Proc. Int. Artif. Intell. Data Process. Symp. (IDAP), 2018, pp. 1–5, doi: 10.1109/IDAP.2018.8620877.

[16] P. Boersma and D. Weenink, Praat: Doing phonetics by computer [Computer program]. Version 6.4.10, 2025. https://www.fon.hum.uva.nl/praat/