IRSTI 16.21.31; 16.41.25

Z.M. Bazarbayeva¹*, T.K. Chukayeva²
¹The Institute of Linguistics named after A. Baitursynov, Almaty, Kazakhstan
²Kazakh Ablai Khan University of International Relations and World Languages,
Almaty, Kazakhstan
* Corresponding author: zeynepmb@mail.ru

The new Kazakh latinized script

Abstract. The article presents results of the scientific research carried out to construct the most effective Kazakh alphabet based on Latin graphics. The idea of switching the Kazakh alphabet to Latin graphics related to enhancing the state language and helping it to become a part of informational community. The methods of phonetic and contrastive analysis were used to reveal the features of using Latin graphics to represent nationally specific sounds. The research showed that the usage of diacritics was a feature of successfully switching to Latin graphics. The usage of diacritical marks has been pointed out as an effective way to adopt Latin letters to the phonetic system of Kazakh.

Keywords: grapheme; phoneme; phonetic system; diacritical mark; phonematic principle; orthography.

Introduction

As Kasym-Jomart Kemelevich pointed out at the meeting of National committee of public trust, the switch of Kazakh alphabet from Cyrillic script to Latin script requires the modernization of the linguistic system, which implies thorough scientific approach. At the same time, it was said that the switch of the state language to Latin script needs to be carried out stage by stage, considering the public opinion when it comes to the imperfections of the alphabet. Linguists were asked to erase such imperfections of the script.

Since the emergence of the idea of switching the state language to Latin graphics, various linguistic research has been handled in this field. As a result, an alphabet was created. It was tested both by researchers and by citizens. Polls and questionaries helped to view the general picture of how Kazakhstani citizens were reacting to the emergence of this alphabet. In most cases, the participants of such polls noted the imperfection and the impracticality of the mentioned graphics. The way it was constructed, and the diacritical marks it included were inconsistent in terms of practicality.

The primary alphabet contained 32 letters, with the national letters being specified via the use of acute. Acute is a diacritical mark that occurs above the letter it modifies. It was utilized for modifying both vowels and consonants. There were also two digraphs which stood for the sounds [ʃ] and [ʧ] (ш and ч). Those digraphs and the use of acute were the reason why this alphabet was criticized by a considerable number of specialists and the citizens of the Republic of Kazakhstan. Considering the experience of European countries which have been using Latin script for a long time and the experience of Turkic nations which recently switched to this script, the linguists from the Institute of Linguistics named after A. Baitursynov enhanced the mentioned alphabet. This article presents the justification of the changes which were made and the results of the work of the Institute.

24   № 3(136)/2021

Л.Н. Гумилев атындағы Еуразия ұлттық университетінің ХАБАРШЫСЫ.
Филология сериясы
ISSN: 2616-678X, eISSN:2663-1288

Methodology of the research

The main methods utilized in the preparation of new Kazakh alphabet based on Latin script were the methods of phonetic analysis. They allowed revealing the correlation between phonemes and correspondent letters. Besides, the contrastive analysis of various alphabets and phonetic systems was utilized. The languages of Europe and Turkic languages were compared in terms of their experience in using Latin alphabet for conveying their nationally specific sounds. The results of such a contrast were used to construct a brand new and successful alphabet for the Kazakh language.

Discussion

As it is well-known from the history of script, the European languages used the same Latin-Roman alphabet despite the differences in pronunciation. Borrowing the alphabet of classic Latin resulted in the discrepancy between phonemic (sounds) and graphic (letters) representations of the languages of Western Europe. In contemporary European languages, there are more phonemes than graphemes: 35 phonemes in French and 44 in English, for example. Nonetheless, such a great variety of phonemes can be only represented by 26 Latin letters (6 vowels and 20 consonants) [1].

The gap between phonemic system and script is explicated as follows: phonetic systems are in the state of constant development. External (cultural factors, neologisms, borrowings) and internal factors (phonologization, divergence and convergence of phonemes) influence the expansion of phonetic systems of various languages [2]. Additionally, the scripts of these languages did not undergo considerable orthographic reforms for a significant period. Many letters are written but not pronounced (for instance, the silent "e" in the end of French and English words); other letters are pronounced distinctively according to their position in the word; the same sounds are represented by various letters and combinations of them [3].

The orthography of German, Finnish, Czech, Polish, Lithuanian, Hungarian, Swedish, and Norwegian languages is closer to the contemporary literary norm. This is the result of the reforms related to religious-reformative and revolutionary movements that these nations underwent.

These nations constructed their orthography taking into consideration the literary norms of their languages. As the Latin alphabet was quite restricted, additional letters were used to represent the specific national sounds; these letters were created with the help of diacritical marks [4]. As a result, the alphabets of these languages contain from 29 to 40-42 graphemes.

There are 29 letters in Swedish, Finnish, and Norwegian; 30 letters in German; 31 letters in Romanian and Moldovan, 32 letters in Polish, Lithuanian and Estonian: 33 letters in Latvian; 40 and 42 letters in Hungarian and Czech, respectively. For instance, the contemporary German alphabet consists of 26 letters of classic Latin alphabet with the addition of three umlaut marks (Ää, Öö, Üü) and one ligature ß. The umlaut allophones of the main phonemes, having been liberated from their positional dependence, turned into independent phonemes, and were represented with separate graphemes.

Latvian can be taken as an example of Eastern European languages. This language lacks four letters of the classic Latin alphabet (Q, W, X, Y), but there are 11 additional letters which represent the nationally specific sounds: Āā, Čč, Ēē, Ģ ģ, Īī, Ķķ, Ļļ, Ņņ, Šš, Ūū, Žž. These letters are created with the help of superscript and subscript diacritics. Prolonged vowels are represented via macron, a superscript diacritic, while the marks of breve and cedilla. Subsequently, there are 33 graphemes in the Latvian alphabet.

The Turkish alphabet consists of 29 letters. Three graphemes of classic Latin are not in use (Q, W, X), while there are six additional letters: Çç, Ğğ, Iı, Öö, Şş, Üü. Native Turkish sounds are represented with the help of superscript and subscript diacritics. Umlaut is used to convey some front vowels, and the additional consonants are made with the help of breve and cedilla.

ВЕСТНИК Евразийского национального университета имени Л.Н. Гумилёва.
Серия Филология
BULLETIN of L.N. Gumilyov Eurasian National University. Philology Series

№ 3(136)/2021   25

Among the Turkic nations of CIS, Azerbaijanian has the most successful experience of switching to Latin graphics. The alphabet of this language consists of 32 letters. Seven nationally specific sounds are represented by the graphemes Çç, Şş, Ğğ, Iı, Öö, Əə, Üü, and the only extracted classic Latin letter is W. Diphthongoids are usually conveyed with the help of umlaut diacritical marks, while consonants are represented via superscript and subscript diacritical marks of breve and cedilla.

The improved and enhanced alphabet, considering the experience of other languages (Turkish, Azerbaijanian, Turkmen, German, Finnish, Swedish, Latvian etc.) utilized the diacritical marks of umlaut, breve, cedilla, and macron to convey the nationally specific sounds. Every diacritical mark has its own phonetic meaning and accurately transmits the content of the sound. The enhanced alphabet based on Latin script was discussed at the meetings of orthographic group, at the seminars, round tables, and conferences. It was approved not only by the scholars specializing in narrow fields, but also by the representatives of public.

The next stage of the linguistic reform was to construct the rules of spelling for Latin graphics. While developing the orthographic rules, the theoretical aspects of Kazakh linguistics, dialectological development of the language, and the rules of phonology, orthography and grammatology were considered. The language was viewed from the perspective of dialectological development. In fact, no language develops in an isolated manner. The influence of the neighbor tongues is always present, which results in mutual enrichment of lexis and pronunciation. For instance, English and French are the neighboring languages which have been interacting for a considerable period. Hence English has many borrowings from French, which impacted the enrichment of this language and its high status in contemporary society.

The emergence of borrowed phonemes is the direct result of borrowing the words, notions, and terms. Those phonemes are indicated by specific letters when written. According to the rules of language, divergence divides a single phoneme into two. This is a common process for all the developed languages of the world: the external and internal factors influenced the emergence of new phonemes which were represented via diacritical marks utilized for clarifying or changing the phonetic meaning.

Language needs to be viewed as a living organism which develops and improves alongside its speakers. For a long period of time, the society of Kazakhstan has been surrounded by various cultures and civilizations, which resulted in mutual penetration of traditions, knowledge, and skills that accumulated in the language. "Clear" languages do not exist since no nation lives in complete isolation. Thus, with the influence of neighboring languages, new terms and notions enrich the vocabulary, which has an impact not only on the lexis, but also on the phonetic system of the languages.

The establishment of new phonemes is connected to the development of language and thinking. This is a natural process of enriching the phonetic system via phonemes which come from other languages and, as time passes, get adapted to the articulatory apparatus of the speakers. There is an opinion that new Kazakh alphabet based on Latin script should not contain the letter "f", as the sound which this letter represents is not native to Kazakh language. However, it is necessary to consider that the phoneme [f] is not native to Russian either: it came from European languages, that is why there are few nationally specific Russian words containing this letter. These words are mainly onomatopoeic notions which repeat after the sound of releasing air through one's teeth. Arabic, however, contains native words with this phoneme. As a result, this letter is utilized both on the east and west of Kazakhstan, and it is also present in the alphabet of Turkic nations, which is why we consider the preservation of this letter in the enhanced alphabet more than relevant.

The development of the phonetic system is carried out not only with the help of external factors such as borrowed sounds, but also based on the internal enhancement of the language. It is necessary to point out that the developed languages are characterized with the enriched lexical and phonetic fund; that is, a significant number of phonemes relates to a developed, perfected language. The goal of

26   № 3(136)/2021

Л.Н. Гумилев атындағы Еуразия ұлттық университетінің ХАБАРШЫСЫ.
Филология сериясы
ISSN: 2616-678X, eISSN:2663-1288

entering the list of thirty developed countries requires the improvement of the language. As Elbasy noted, the switch to Latin script is aimed at "the further development of Kazakh language and high technologies". Switching to Latin graphics is a historically important event which is necessary for modernization and enhancement of the state language. This is a civilized choice for the integration into the advanced world of informational technologies. This is one of the most significant elements of the program of modernizing the public conscience.

The main principles of orthography are defined based on the national alphabet and spelling rules. Contemporary Kazakh script is constructed on the phonematic principle, though the morphological, traditional-historical, and phonetic principles can be also traced. The script must be understandable and convenient for the users, that is why orthographic rules were compiled with the consideration of the main tendencies of spoken and written text. Also, the universal rules of ergonomics, orthographic memory, consistency, and practicality of the script were taken into account.

The ergonomic or practical writing is based on representing various phonemes with the help of restricted number of letters. Such a script, "tote jazu", based on Arabian graphics, was introduced by A. Baitursynov who is the founder of Kazakh linguistics. Hence the most important principle is to prefer phonematic type of script to transcriptional type. In transcriptional script, all the pronounced sounds are written down. At the same time, phonematic principle allows to document only those sounds which can distinguish the meaning and influencing the content of the utterance – the phonemes.

Speech and text as the forms of communication are two representations of the same language. Speech has a sound and serves for the immediate communication between people and implies the presence of an interlocutor. Unlike rapid and immediately fading speech, text or script can be stored for a long period of time, thus connecting the representatives of different nations, and conveying the fruits of civilization from generation to generation. Script by itself is conservative. In fact, it aims at preserving the continuity, unity, and stability, which is represented in the so-called orthographic memory of spelling.

Results

To conserve the continuity of script, the orthographic rules developed by outstanding Kazakh linguists such as A. Baitursynov, K. Zhubanov, S. Amanzholov, N. Sauranbayev, M. Balakayev, I. Kenesbayev, R. Syzdykova, A. Kaidarov etc. were taken into consideration [5]. The objective of the linguists at the Institute of linguistics named after A. Baitursynov was not only conveying the Kazakh sounds in script, but also preserving the orthographic norms which help accurately perceive the speech content and improve the culture of latinized graphics.

These norms include the rules of vowel harmony and the basic rules which have been compiled since the ancient times. Light will be shed on the scientific interpretation of changes and additions to the rules of orthography. The results of scientific research will be published in linguodidactic manuals, orthographic dictionaries, toolkits, and various types of instructions.

New rules of Kazakh orthography based on Latin graphics consist of nine chapters. In these rules, the spelling of words with diacritical marks are indicated thoroughly.

To represent Kazakh soft vowels, the diacritical mark of umlaut (two dots above the letter) was used. For example: ädıs (method), dän (seed), mädeniet (culture), käsıp (craft); örıs (field), ömırşeŋ (livable), kösem (leader); ürdıs (process), ükımet (government), jüie (system).

To represent the consonant ғ the diacritical mark of breve (an upside-down crescent) was used. For example: ğaryş (space), samğau (to fleet), tabiğat (nature), bağdar (direction), qağaz (paper), sağat (clock). To indicate the consonant ш the diacritical sign of cedilla (resembling a hook) was used. For instance: şaŋyraq (a house or the top of native Kazakh yurt), bürşık (bud), qylyş (sword), qoişy (shepherd), oşaq (fireplace), şara (bowl). To represent the hard vowel ұ a diacritical sign of macron (a

ВЕСТНИК Евразийского национального университета имени Л.Н. Гумилева.
Серия Филология
BULLETIN of L.N. Gumilyov Eurasian National University. Philology Series

№ 3(136)/2021   27

short line above the letter) was used. For example: ūtymdy (profitable), ūstaz (teacher), qūjat (document), maglūmat (data), ūlys (ulus – a block in Kazakh historical government), qūl (slave).

Conclusion

In the present article, there was widely addressed experience of other languages in using Latin alphabet. As the research shows, various languages used different techniques to cope with the discrepancy between their phonetic systems and the graphic representation of the sounds. For instance, English "spelling" refers to saying the word letter by letter, as English is one of the languages where pronunciation does not match the graphic representation of the notion. Such a great difference occurred for the historical reason: when Latin was forcefully introduced as the main language of writing, the native European languages had to adopt their phonetic systems to the restricted number of Latin letters. In case of English, no enhancements were carried out, and the Latin graphemes are used without any diacritical marks. The combinations of letters help identify the nationally specific sounds which did not exist in Latin. When it comes to other languages such as German, French, or Latvian, the presence of diacritical marks has been identified. As the research showed, these languages utilized breve, macron, cedilla, and umlaut to represent their nationally specific sounds (soft vowels, for instance).

The case of Turkic languages is quite different from the European languages. The Turkic languages have recently started to use Latin alphabet, and the experience of Turkish and Azerbaijanian is the most successful. In these languages, diacritical marks are used to convey the sounds which are quite close to the nationally specific phonemes of Kazakh.

Subsequently, the enhanced version of the alphabet based on Latin graphics utilized the diacritical marks of umlaut, breve, cedilla, and macron. This decision was the result of careful discussion and consideration of the experience of European and Turkic languages. The usage of diacritical marks accurately conveys the internal content of the phonemes of the Kazakh language. The present article was prepared within the scope of the project: "Development of a series of updated normative dictionaries and academic publications that provide the transition of the state language into the national Latin graphic alphabet", financed for the period of 2021-2022.

References

1. Burchfield, R. The Cambridge history of English language / R. Burchfield. – Cambridge: Cambridge University press, 1994. – 656 p.

2. Базарбаева З.М. New Kazakh letter based on the Latin chart // Вестник Национальной академии наук РК. - 2019. - № 382. – pp. 355-359.

3. Janson, Tore. A natural history of Latin / Tore Janson. – Oxford: Oxford University press, 2007. – 305 p.

4. Sacks, David. Letter perfect: a marvelous history of our alphabet from A to Z. – Crown, 2003. – 416 p.

5. Байтұрсынов А. Тіл тағылымы / А. Байтұрсынов. – Almaty, 2008. – 325 p.

З.М. Базарбаева[1], Т.К. Чукаева[2]
[1]А. Байтұрсынов атындағы тіл білімі институты ҚР БҒМ, Алматы, Қазақстан
[2]Абылай хан атындағы ҚазХКжӘТУ, Алматы, Қазақстан

ЛАТЫН ӘЛІПБИІНЕ НЕГІЗДЕЛГЕН ЖАҢА ҚАЗАҚ ЖАЗУЫ ТУРАЛЫ

28 № 3(136)/2021

Л.Н. Гумилев атындағы Еуразия ұлттық университетінің ХАБАРШЫСЫ.
Филология сериясы
ISSN: 2616-678X, eISSN:2663-1288

Аңдатпа. Мақалада латын әліпбиіне негізделген қазақ жазуының неғұрлым ұтымды нұсқасын құрастыру мақсатында жүргізілген ғылыми зерттеудің нәтижелері көрсетілген. Латын әліпбиіне ауысу ойы тілді дамытумен және әлемдік информациялық кеңістікке ену қажеттілігімен байланысты болды. Фонетикалық және салғастырмалы талдау әдістері латын әліпбиін ұлттық дыбыстарды жеткізу үшін қолданудың ерекшеліктерін айқындауға көмектесті. Зерттеу көрсеткендей, латын әліпбиіне ұтымды түрде ауысудың бірден-бір ерекшелігі – диакритикалық белгілерді қолдану. Сәйкесінше, диакритикалық белгілер латын әріптерін қазақ фонетикалық жүйесіне бейімдеудің тиімді жолы ретінде айқындалды.

Түйін сөздер: графема; фонема; фонетикалық жүйе; диакритикалық белгі; фонематикалық принцип; орфография

З.М. Базарбаева¹, Т.К. Чукаева²
¹*Институт языкознания имени А. Байтурсынова МОН РК, Алматы, Казахстан*
²*КазУМОиМЯ имени Абылай хана, Алматы, Казахстан*

О новом казахском латинизированном письме

Аннотация. В статье представлены результаты научного исследования, проведенного с целью сконструировать наиболее эффективный вариант казахского латинизированного алфавита. Идея перехода на латинскую графику была связана с развитием языка и необходимостью стать частью информационного сообщества. Методы фонетического и сопоставительного анализа помогли выявить особенности использования латинской графики для передачи национально-специфических звуков. Исследование показало, что основной особенностью удачного перехода на латинскую графику было использование диакритических знаков. Таким образом, диакритические знаки были отмечены как эффективный путь адаптирования латинских графем к казахской фонетической системе.

Ключевые слова: графема; фонема; фонетическая система; диакритический знак; фонематический принцип; орфография.

References

1. Burchfield, R. The Cambridge history of English language / R. Burchfield. – Cambridge: Cambridge University press, 1994. – 656 p.

2. Bazarbayeva Z.M. New Kazakh letter based on the Latin chart // Bulletin of National Academy of sciences of the Republic of Kazakhstan. - 2019. - № 382. – pp. 355-359.

3. Janson, Tore. A natural history of Latin / Tore Janson. – Oxford: Oxford University press, 2007. – 305 p.

4. Sacks, David. Letter perfect: a marvelous history of our alphabet from A to Z. – Crown, 2003. – 416 p.

5. Baitursynov A. Til taǵylymy [Language training] / A. Baitursynov. – Almaty, 2008. – 325 p.

Information about authors
Базарбаева Зейнеп Муслимқызы – филология ғылымдарының докторы, профессор, ҚР ҰҒА академигі, А. Байтұрсынов атындағы Тіл білімі институты ҚР БҒМ, Алматы, Қазақстан.

Чукаева Талшын Қанатқызы – шетел филологиясы мамандығы бойынша магистрант, Абылай хан атындағы ҚазХҚжӘТУ, Алматы, Қазақстан.

ВЕСТНИК Евразийского национального университета имени Л.Н. Гумилева.
Серия Филология
BULLETIN of L.N. Gumilyov Eurasian National University. Philology Series

№ 3(136)/2021   29

Bazarbayeva Zeinep Muslimovna – Doctor of Philological Sciences, Professor, Corresponding Member of National Academy of Science Republic of Kazakhstan, Institute of linguistics named after A. Baitursynov of MES RK, Almaty, Kazakhstan.

Chukayeva Talshyn Kanatkyzy – Master's student in Foreign Philology, Kazakh Ablai Khan University of International Relations and World Languages, Almaty, Kazakhstan.

30   № 3(136)/2021

Л.Н. Гумилев атындағы Еуразия ұлттық университетінің ХАБАРШЫСЫ.
Филология сериясы
ISSN: 2616-678X, eISSN:2663-1288