The Rise and Fall and Rise again of the Seventh Letter: A Technological Story of the Russian Alphabet

Alexander Doroshev(✉) ⓘD and Anna Polyakova ⓘD
Peter the Great St. Petersburg Polytechnic University, St. Petersburg, Polytechnicheskaya, 29, 195251, Russia
doroshevam@icloud.com

Abstract

It is generally understood that the letter ё is not obligatory in the rules of Russian language – even though there are many names and words that include it. This lack of obligation led to the absence of that letter in the typewriters, keyboards and code systems. The difficulty and impossibility of using that letter then led to its slow disappearance in printed texts, even though there are still many names and words that include it. Therefore, persons who have the letter ё in their name will try to obtain this letter with two dots in programs and technical devices. Currently, the use of inbuilt dictionaries without having a possibility of variational typing and machine learning led to the tendency of using the letter ё due to autocorrection and systems of hints in the smartphones. When using keyboard applications in electronic devices, some words which contain the letter ё, are typed with the letter е. In particular, some words are changed automatically to е, others are not. Only time will tell what will happen with the letter the seventh letter of Russian alphabet in the future, whether it will lose its relevance or whether there will be a resurgence of its use.

Keywords: Technologies of writing; Letter ё; Autocorrect; T9; Keyboard; Typewriters

Взлет, падение и снова взлет седьмой буквы: Технологическая история русского алфавита

Александр Дорошев (✉) 🆔 и Анна Полякова 🆔
Санкт-Петербургский политехнический университет Петра Великого, Политехническая, 29, 195251, Санкт-Петербург, Россия
doroshevam@icloud.com

Аннотация
Буква ё не является обязательной согласно правилам русского языка, хотя существует множество имен и слов, содержащих её. Необязательность привела к отсутствию этой буквы в пишущих машинках, клавиатурах и кодовых системах. Трудность или невозможность использования буквы затем привели к ее медленному исчезновению в печатных текстах. Наибольшие сложности возникают у людей, имеющих в своем имени букву ё, так как существует требование её использования в официальных документах, несмотря на проблематичность получения буквы с двумя точками в программах и технических устройствах. В настоящее время использование встроенных словарей без возможности вариационной типизации и машинного обучения привело к тенденции использования буквы ё за счёт автокоррекции и систем подсказок в смартфонах. При использовании клавиатурных приложений в электронных устройствах некоторые слова, содержащие букву ё, набираются с буквой е. В одних словах происходит автоматическая замена, в других – нет. Только время покажет, что будет с седьмой буквой русского алфавита в будущем, потеряет ли она свою.

Ключевые слова: Технологии письма; Буква ё; Автозамена; Т9; Клавиатура; Пишущие машинки

INTRODUCTION

It would seem to hold for many, perhaps all languages that alphabets have become simplified and standardized as writing became increasingly mediated by technology. Handwriting and manual typesetting still allowed for many variants, including diacritics. But these were first to go with the advent of typewriters and their limited keyboards, aggravated by the speed of electric typewriting which militated against backtracking to add a diacritic. For Europeans this meant first of all a neglect of diacritics which were replaced by conventionalized variants: č and ć become c, š becomes s, ž becomes z, and đ becomes d, ä, ö, ü becomes ae, oe, ue, ß becomes ss. However the technological trend did not continue to work against these special signs. There are now many interesting solutions where auto-correction or proposals for word and sentence completion provide opportunities for technological restoration of extra signs in Romanian (Grozea, 2012; Ruseti et al., 2020), Hungarian (Novák & Siklósi, 2015), Vietnamese (Hung, 2018), in Turkish (Adali & Eryiğit, 2015), Slavic (Ljubešić et al., 2016), Arabic texts (Alkhatlan et al., 2020; Orife et al., 2020; Zitouni & Sarikaya, 2009), Māori (Cocks & Keegan, 2011), Igbo (Ezeani et al., 2016).

The seventh letter of the Russian alphabet ё [jo] differs from the sixth letter e only by two dots on top. The special position of the letter in the alphabet is connected with its status of being unnecessary and replaceable in most cases by the letter e according to the rules of the modern Russian language. It is understood that the readers must independently know that in certain words they should pronounce the "e" with the sound [jo] of the "ё".

"Ё" is considered dispensable because it is, in fact, the youngest letter in the Russian alphabet. It did not exist in the Old Church Slavonic alphabet which was the predecessor of the Russian language. In the course of language-reforms by Peter the Great in the years 1708-1711, superscript signs were abolished. Much later, however, princess Ekaterina Romanovna Dashkova suggested in the 1783 meeting 'The academy of Literature' that certain letter-combinations which produced the sound [jo] should be replaced by the e with two dots would be located above the letter. This proposal did not catch on initially. One could see the letter ё in print for the first time only in the 1795 book "My trifles" by Ivan Dmitrievich. The letter became famous in 1797 when the poet Nikolay Karamzin decided to replace [jo] in the word слиозы ("tears") for слёзы.

In all the ABC's until 1917 the letter ё then found a place at the end of the alphabet near the forgotten letters phita and izhitsa. In 1904 a commission set out to solve the problem of writing the letter ё. It decided to declare it an advisable but not obligatory letter. By 1917 and thanks to the October Revolution, however, the letter ё had not the status of a letter, but was considered to be an old-fashioned option for orthography – since a lot of poets decided to distort the orthography purposely in order to find new ways of writing. The government appreciated these new ways of writing as well as the attendant literary points of view. The greatest poet, Marina Tsvetaeva, in her own poems wrote "чорт" instead of "чёрт" (devil in Russian) on purpose. Furthermore, the poet, Andrey Bely, used the word "жолтый" instead of "жёлтый" (yellow in Russian) in his writings. In 1922, the famous poet, Sergey Esenin, published his poem "Пугачов" without using

the letter ё (Bozhko, 2011; Chumakov, 2010).

On December 23rd, 1917, a decree was authored and published by the Soviet commissar of education Anatolii Lunacharsky. It agreed with the 1904 commission and said: "The letter ё is advisable but it is not obligatory".[1] Thus, the letter ё officially entered the seventh place of the Russian alphabet in the period of the USSR.[2]

In 1942, another decree was issued by the People's Commissar of Education for the Russian-speaking parts of the USSR. Now the usage of the letter ё was obligatory to be taught in school programs. The following ten years were the only time in history when the letter ё was widely used.

However, after some years publishers decided to revert to the old practice. They decided to use letter ё only where it was helpful to have. It could have happened if the language is ambiguous (the ability to read incorrect) or that kind of method was given to person who had not have enough skills in the language (children, who are studying the language) (Korbut, 2017).

1952 saw the publication of the second edition of *The Dictionary of Orthography and Punctuation for Print Workers* by Konstantin Iakiyefovich Bylinsky and Nikolay Nikolaevich Nikolsky[3]. It declared that in printed texts the letter ё is usually replaced by the letter e and recommended to use ё in these cases:

1) When we have to prevent the wrong reading of the word. For example:

- узнаём instead of узнаем:
  Мы узнаём на уроках в школе много нового. (*We learned a lot during our lessons at school.*)
  Мы все узнаем завтра. (*We will know everything tomorrow*).
- всё (*everything*) instead of все (*everybody*);
- передохнёте (*will take a break*) instead of передохнете (*will die*);
- поёшь (*sing*) instead of поешь (*eat*)

2) When we have to point out to the pronunciation of an unknown word, for example, the river Олёкма.

3) In dictionaries and orthographic dictionaries, in textbooks for foreigners, in the books for children of the primary schools and other special varieties of literature.

The latest decree concerning the use of ё was published in 2007 ("On the Decisions of the Interdepartmental Commission of the Russian language"). Here, the Ministry of education and science of the Russian Federation states that the letter ё should be used if the word cannot otherwise be read correctly – for example, in the case of personal names, because ignoring that case would break the Federal Law (FL) "On the State Language of the Russian Federation." Beyond that, the rules state that officials are free to use or not use the letter ё as they see fit.[4]

---

[1] Decree of national commissariat from the 23rd of December, 1917 "About the Introduction of a New Type of Spelling" // https://base.garant.ru/187916/

[2] One should also mention the "New Alphabet" by Lev Tolstoy from 1874-1875 where the letter "ё" was in the 34th place, between "э" and izhitsa as an optional letter, see http://tolstoy-lit.ru/tolstoy/pedagogika/novaya-azbuka.htm

[3] Bylinsky, Nikolsky https://obuchalka.org/2015051684760/spravochnik-po-orfografii-i-punktuacii-dlya-rabotnikov-pechati-bilinskii-k-i-nikolskii-n-n-1970.html

[4] Law about the national language http://www.consultant.ru/document/cons_doc_LAW_53749/

THE EFFECT OF TECHNOLOGIES ON THE Ё

Typewriters

In 1870 the Russian inventor Mikhail Ivanovich Alisov created a writing machine that was known as "Скоропечатник" ("Fast typer" in Russian) (Figure 1). That invention reminds us of classic printing machines. There was no keyboard as we know it today, and letter ё was not accounted for.
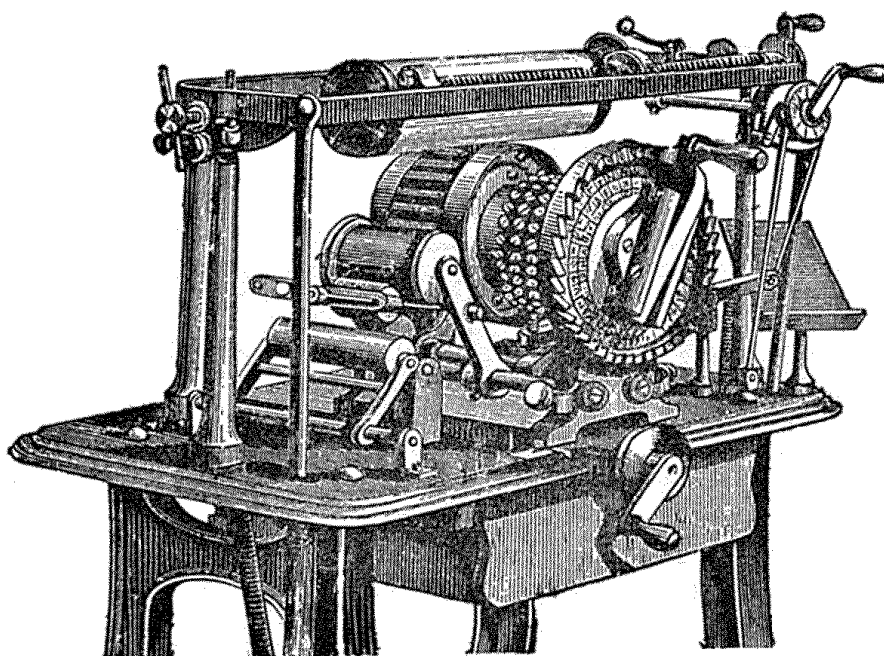
Figure 1. Alisov's "Скоропечатник" ('Fast Typer" in Russian)

The first product-line of typewriters was "Remington №1."[5] It was patented in 1866 by Christopher Scholes and Samuel Soul. In 1877 American engineer George Yost created the typewriter "Yost" which was named after him. In Russia it could be bought in two options: Russian type and Latin type, with the price for Russian type being much higher. In 1889 M.G. Merritt created a smaller version of the typewriter – the „Merritt." It was also sold in Russia, even though there was no letter ё in the machine. In 1892 the German company Guhl & Harbeck created a typewriter which was called "Cosmopolitan" (Figure 2). One of its features was that it had a single flat linear carrier (Michael, 1997; Rehr, 1997; Russo, 2002).

---

[5] On the typewriter Remington № 1 the first book typed was Mark Twain's *The Adventures of Tom Sawyer*.
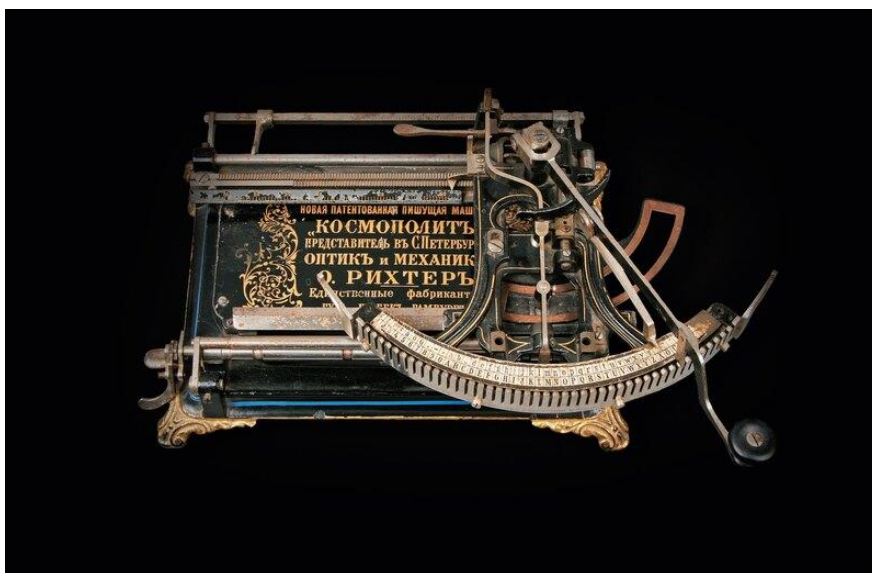
Figure 2. Typewriter "Cosmopolitan"

There was no production of typewriters in Russia until the 1930s. Instead, Western models were used. The first USSR typewriter "Яналиф" was released in the 1929. Initially it had only Latin keys, a version with Russian keys came later. But still there was no letter ё in that kind of machine. In the 1930s, the first national factory for producing typewriters ("ПИШМАШ") was opened in the USSR. The first machine was called "Leningrad," but still without the letter ё.

The situation changed rapidly by the end of the 1930s. The letter ё appeared in the typewriter "Progress" which was produced in a Kazan factory (Figure 3). It was situated on the left side of the keyboard, right to the letter ю.

Figure 3. Typewriter "Progress" (has ё)

However, there were still a lot of typewriters where the letter ё was left out. A special key appeared only in the mid 1950s after the "ЙЦУКЕН" reform.[6]

With classical typewriters, when two adjacent keys were pressed in quick succession, the keys might interlock and thus slow down the speed of typing. Therefore, in the 1950s in the USA a new keyboard was created for Russian. It was called "ЙЦУКЕН," where the most needed letters were situated under the index finger. This helped increase the speed of writing by almost 2%. At the same time, efforts were made to economize the keyboard by removing some keys. For example, the numerals 1 and 0 were replaced by the letters l and O. In light of this interest in the speed of typing there was no room for the letter ё. Three keys needed to be pressed in order to produce this ё: first, typists would press the letter "e", then they needed to back up the carriage, and finally place quotation marks on top of the "e". The unsurprising unintended result was that this would further slow down the speed of typing.

In the 1960s, the USSR decided to produce new models of typewriters: "Moscow", "Ukraine", "Yatren." They featured both the Cyrillic and Latin alphabets. They also included the letter ё. It was situated in the lower corner near the letter ю.

Figure 4. Typewriter "Moscow" (with the letter ё)

In the 1980s the electromechanical or so-called electric typewriter "Lubava" was developed in Ryazan, Russia. It was the license version of "Erika" from the German Democratic Republic. Despite being new, the letter ё was left out on the "Lubava" but it was included in the "Erika" (Figure 5). For this reasons, foreign typewriters were popular

---

[6] Fokin A. Why in that order? Why the 19th century? // Fokin A. 2021. https://bloha.ru/interesnoe/qwerty-ycuken-pochemu-imenno-v-yetom-poryadke/

in the USSR. Alongside "ERIKA," the models "OPTIMA" and "ROBOTRON" were also produced in the GDR, also came with the letter ё, and thus proved convenient for ordinary soviet citizens.

Figure 5. Electromechanical typewriter "Erika" (with the letter ё)

Computer technology

In 1963 ASCII was created in the USA. It was a 7-bit code system for typing symbols as codes. Each letter presented on the table has its own digital code. At the time of release, it had only 128 symbols for Latin languages. Later, some modifications were implemented in order to include other national symbols. So, the number of codes available increased to 256. The capital letter Ё was located in 168th place and it was identified with the code 0xA8, meanwhile the lowercase letter ё was placed in 184th place.

Moreover, there was another modification of ASCII. It was assumed that it should have been switched up to the US-ASCII as well as switched up to the national option. Switching up was done instantly using symbol "SO," however, the most famous code was KOI-7. In the KOI-7 the lowercase letter ё was marked as 35 (23h), meanwhile the capital one ё was marked as 36 (24h).

Withe development of personal computers in the USSR the committee of computer science and IT started a process to develop a new keyboard for the IBM PC/XT computers, which was finalized in 1983. The letter ё managed to keep its place (figure 6).

Figure 6. Soviet laptop –"Electronica MK-106" (with the letter ё)

However, despite the fact that the committee demanded it, some early projects for the keyboard still ignored the letter ё. As a result, the letter ё was again absent from some home-made keyboards used for the operating system DOS (figure 7-9).

Figure 7. Soviet laptop "Electronica MS 1504" (without the letter ё)

Figure 8. "Corvette PC 8020" (without the letter ё)

Figure 9. "Electronica MS 1208" (without the letter ё)

Microsoft decided to create keyboards where ё was located left to the "1," in the second row of keys. Meanwhile, in Russia two types of keyboards were used at the same time. In the older keyboard, searching for the letter ё was still complicated. It was located in the last row, to the right of the letter ю. This placement was very uncomfortable for the fingers, and consequently caused a severe slowdown in typing.

Moreover, Microsoft Word did not support the letter ё in its spelling check system. Even today the system considers the 6th and 7th letters to be identical. So, writing the word "поёзд" would not be marked as an error. This flaw can pose some issues, for example if we were to create a dictionary or use tools to search the text.

In Microsoft Word 2010 the check-system for letters in the dictionary is disabled by default, and it has to be deliberately activated. In order to make this operation, you have to tick the box explicitly dedicated to the use of the letter ё (Figure 10).
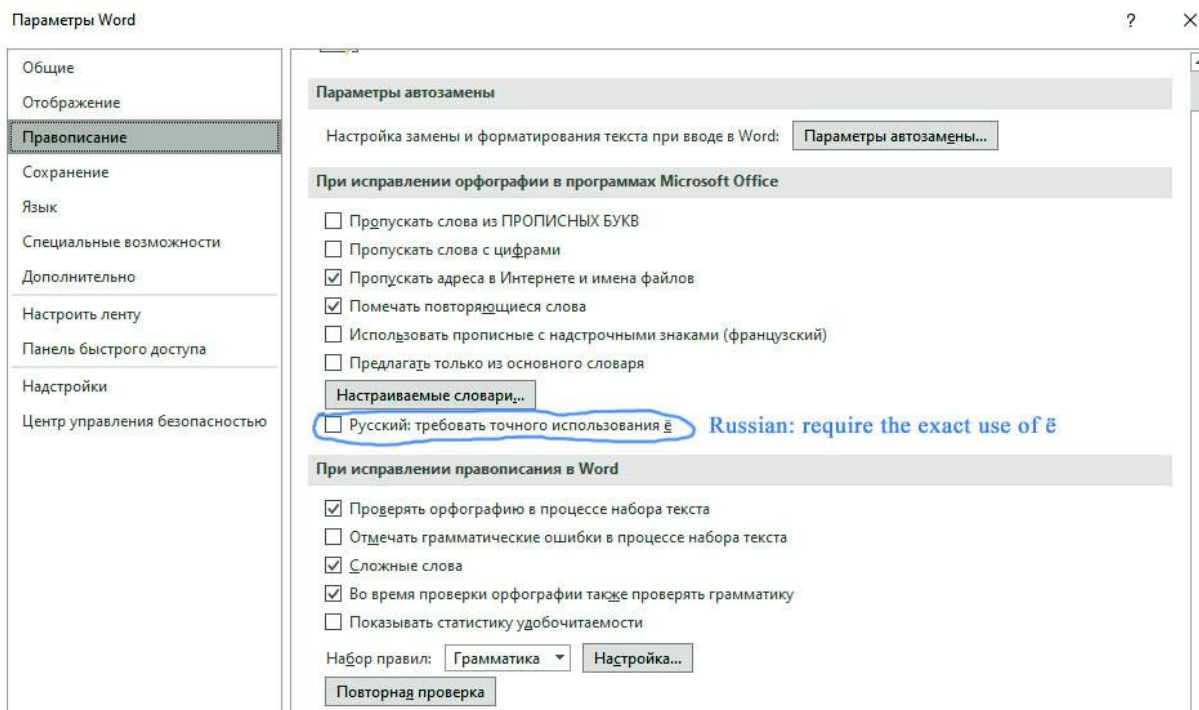
Figure 10. Word option to require the exact use of ё

Currently the letter ё is implemented in all modern keyboards. However, the problem remains that it is hard to find it. It is located in the upper left corner, to the left of the row of numbers under the "Esc" button. Also, the button for ё has two other functions, the symbol (~) and the one for reversed quotes (`) (figure 11). On Apple's keyboard the letter ё is situated in the third row, to the right of the letter э.

Figure 11. Modern keyboard (with the letter ё)

However, there are keyboards where the letter ё is not indicated on the keyboard. Since the letter is in the code system, however, knowledgable users can still produce it with the same key under the esc.

To be sure, knowledgable users have other opportunities as well. Excel Macro is a record and playback tool that simply records one's Excel steps and the macro will play it back as many times as one wants. Macros appeared in the 2010. VBA Macros save time as they automate repetitive tasks. However, one needs to have some basic knowledge of VBA in terms of making advanced modifications in the macro. Thus, by assigning the appropriate macro to a key, the letter could be generated on any computer keyboard.

Mobile Phones

Some older models of cell phones came with a small physical keyboard. Here, the letter ё was not shown. This is due to the fact that the quantity of letters in the Russian language is supposed to be 33. However, in the English language there are only 26. As a result, not all letters can be fitted into that type of layout. Again, for the sake of economizing space, some symbols were hidden and ё was among them.

In figure 12 we see an example of this old keyboard. Worthy of attention is the button 3, which allows to type the letters Д, Е, Ж and З. In the alphabetical order after the letter ё the next letter is е, so ё was hidden behind that button. If you want to write it, you have to tap three times the button 3.

Figure 12. The keyboard of a push-button phone (without the letter ё)

In modern smartphones the letter ё is also hidden. The option will be displayed if

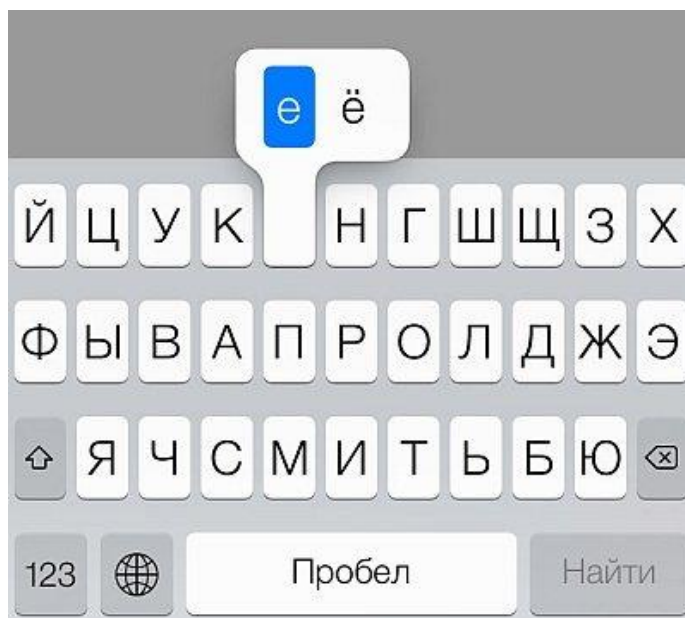you tap and press for some seconds on the letter e (figure 13).

Figure 13. The digital keyboard of a smartphone

If one opts for a typescript layout instead, the letter ё is located near the space button. On MacBook the letter ё on the keyboard is the key \. It is situated to the right of the letter ю.

CATALOGUING PERSONAL NAMES

The main issue connected with the implementation of the letter ё concerns personal names. Crispin Thurlow noticed that words with diacritics appeared incorrectly and inconsistently in library catalogues. For instance, one can notice a transliterative and diacritic mistake in the *Ragnarök* manga, both in the title and the author name (Thurlow, 2021). Most of those technical troubles derive from the tendency to use ASCII in computer coding. Since the ASCII system didn't allow to use diacritics but was nevertheless used for creating library catalogues, these troubles became widely diffused.

The adoption of a transliteration method that is not consistently applied by different institutions presents a variety of bureaucratic issues for all those people who have a surname that includes the letter ё. This means that the name used in everyday life and the one present in one or more documents may differ. As a consequence, during a formal technical checking of the papers for a registration, for government services, for exams, or for receiving a bank card, one may receive an automatic rejection because the surnames do not match.

In 2007 a law was enacted which obliged people and authorities to record ё in first names, last names, and patronyms. This caused further issues. In particular, the documents of persons who have documents dating before and after 2007 caused some problems. For example, Ksenia Zubacheva is a mother of three children from

71

Kaliningrad. She was refused state benefits because her surname was written differently on her children's birth certificates. In one document it was written with "е" while in another with "ё" (Zubacheva, 2018). In her story about this, the author said that she pronounces her surname as "Zubachyova" (or "Зубачёва" in Russian), but in all official documents (including the by-line of her article) it was written "Zubacheva" ("Зубачева") in order to avoid mishaps.

TECHNICAL ISSUES WITH THE LETTER Ё IN MODERN SOCIETY

There are obviously a lot of technical difficulties connected with the letter ё that an ordinary user faces in modern society. We will analyze some of those issues with the help of some fascinating jokes as well as some ridiculous stories, which could be found on the internet. A lot of those stories are connected with the difficulties of using the letter ё in everyday life. For example, in figure 14 we can see the ~ button on the keyboard. The letter ё should be printed on it, however, it is not. As a result, the user might think that the system does not support this letter as it supposed to be and as it really does.

Figure 14. The letter ё is not printed on the keyboard

The most common situation is when people do not know how to type the letter ё on their smartphones, due to the fact that the letter ё is located inconveniently in the digital keyboard. As we mentioned previously, it is hidden behind the letter e, and requires a prolonged tapping. Figure 15 shows a dialogue where one of the users confesses his amusement when seeing the use of the letter ё, because he or she is unable to find it.
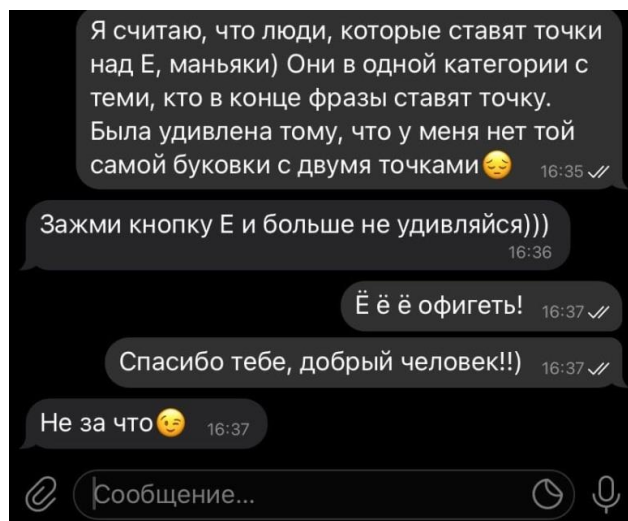
Figure 15. Dialogue with a person who doesn't
know how to write the letter ё on the smartphone

Some people just get used to the embedded systems of a text editor that automatically corrects the mistakes in words and sentences. Therefore, when typing the letter ё, the system automatically changes it for e. However, it also happens that people who know that they typed the word correctly get stuck and share this information with others. On the internet there are a lot of screenshots testifying that sometimes the letter ё is considered a mistake by the auto-correction system (Figure 16).
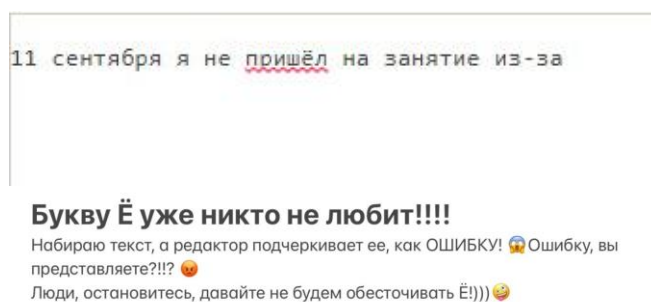

Figure 16. A message of indignation for seeing a word correctly written with the
letter ё highlighted as an incorrect one

Many users, understandably, show some frustration at not being able to use a letter of their alphabet in the grammatically correct way. But there are also a lot of people for whom the absence of the letter ё in the technical devices represents an even greater issue. We refer to the case already mentioned of people who have the letter ё in their surnames. In figure 17 we can see a screenshot which shows how a name, written correctly, is

considered an error by the system, which therefore prevents the operation from continuing.

Figure 17. The error says 'check the correctness of the surname' due to the letter ё

Some sites do not recognize names and surnames with ё, so that people are excluded from some services and they have to spend time and resources to find a solution. In figure 18 we can see the dialogue, which shows that a person with the letter ё in the surname cannot sign up to get a doctor's appointment using the intended site. As a result, such people have to do this in person or by way of a phone call.
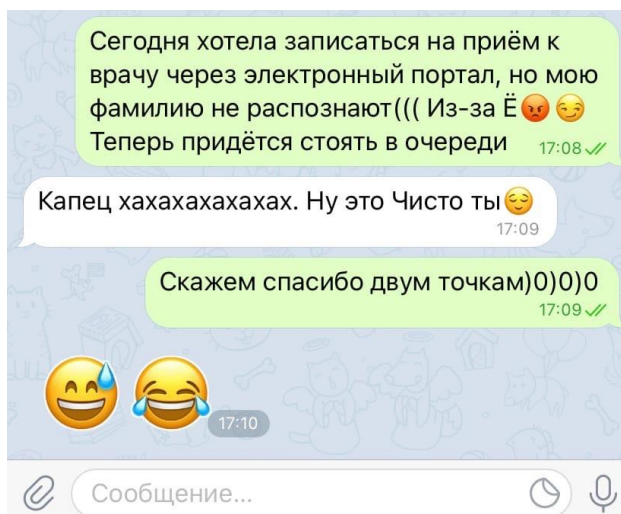

Figure 18. Inability to sign up to a doctor's appointment using the digital form for the person whose surname includes the letter ё

This issue has spread everywhere for lots of forms of the electronic registrations as well as for the payment terminals. In figure 19 we can see an outraged user whose surname includes ё, complaining about unjust discrimination. The error text that is shown below translates as: 'If your surname, name or patronym have the letter ё, the payment method in this terminal is impossible. The payment is possible only in the offices of banks'.

Меня раздражает дискриминация людей
с фамилией с Ё!!!!😡
Сегодня пошёл в миграционную службу и в
очередной раз столкнулся с проблемой:
«Если в вашей фамилии, имени, отчестве
содержится буква Ё, оплата через данный
терминал услуг не производится.
Оплата возможна только через офисы
банков...»

Figure 19. Electronic payment denied to a user with the letter ё in the surname

From a linguistic point of view, it should also be noted that, in some cases, writing e instead of ё changes the main idea of a statement. There are a lot of jokes about that theme on the internet (some manufactured, some real). For instance, there is this statement for music lovers that was found and shared on the internet: "Some think that everything goes away when you cry. I believe that everything goes away when you eat." The last word, which is written with the letter ё, means "eat." However, the word was intended to be written with the letter e, and in that case it would have meant "sing." Thus, a famous phrase about the spiritual power of music becomes advice to fight sadness with consumption.

Another technical resource that has become a gold mine for interesting cases are online translators. These kinds of services are not always accurate and frequently mistake "e" and "ё." This has become a problem especially for those who are studying Russian as a second language. In figure 21, for example, we see „everybody sleeping" translated as „still sleeping" since "все (everybody)" was mistaken for "всё."
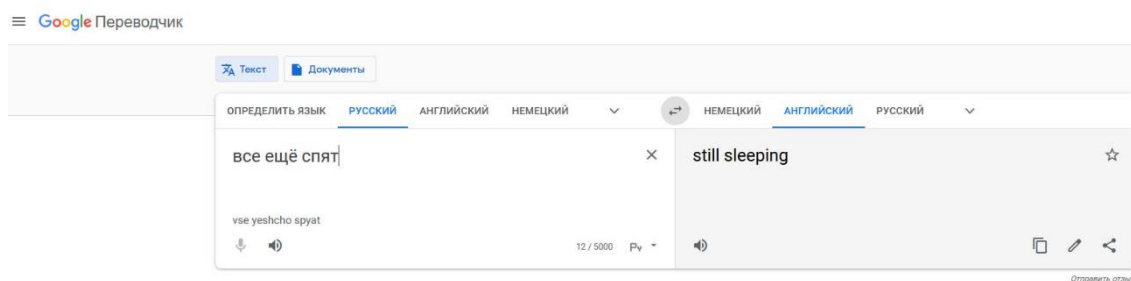
Figure 21. The word with "ё" is mistakenly translated as the word with "e"

At the same time, recently, we can see a tendency for the letter ё to be treated rightfully among the technical fields, even in those where it has no special place. When spelling the alphabet, Russians usually skip the letter ё. So "ё" is not present in all cases

where letters and figures are used alongside numbers (for instance, in the alphanumeric combination used in birth or marriage certificates, in the numbers of cars, in a degree, etc.) As a result, when people see the appearance of "ё" in a list or a call number, they can be so amused to post it on the social networks as an interesting case.
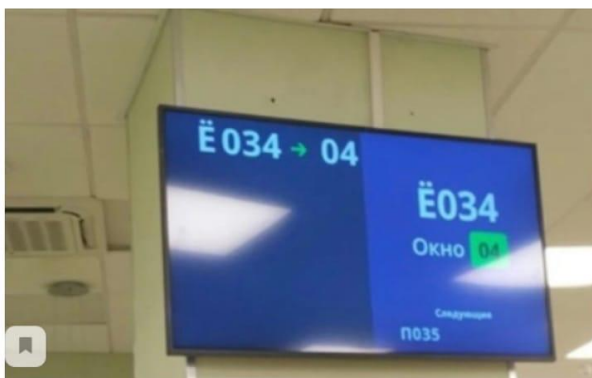
Если вам не хватает в жизни буквы Ё, то идите в Сбербанк 😂😂

Figure 22. Appearance of the letter ё in the call number for an order

LETTER Ё AND AUTOCORRECT

Modern typing devices are programmed to anticipate the words that are about to be written. Often, when we type the first letters of a word, the program automatically checks an integrated vocabulary and presents us the alternatives. One example of this kind of technology is called T9 (Text on 9 keys). It was created in 1999 by Cliff Kushler in the USA. At the time, in order to write a word as simple as "Hello" on a cell phone keyboard, the combination of keys was much more cumbersome. Given that you had to scroll letter by letter on a keyboard with only 9 keys, the result was this order of input: 4433555666. Thanks to T9, that combination decreased to 5 symbols, because T9 was capable of completing words by itself.

For example, when typing some letters, T9 automatically tries to create the word starting from that set and shows it on screen. If it matches with the word that you wanted to type, you can tap the space button and go directly to the next one. Otherwise, T9 continues to make suggestions with every new letter that is added to the set.

There is a different Russian vocabulary for different keyboards (see Figure 23). In this picture we can see registered words for different systems. Moreover, there is a separate vocabulary where some words are located having letter ё (words with the letter ё will be corrected automatically if T9 recognizes them) (Figure 24).

Yandex Keyword	Smart Keyboard	Google Keyword
..	а-	
не	а-а	абажур
и	аа-а	абажура
в	а-аа	абажуром
а	а-а-а	аборигена
ты	а-ааа	абразив
что	а-аа-а	абрека
на	а-а-а-а	абхазец
как	а-аааа	абхазски
с	а-а-ааа	ав
да	а-а-а-а-а	авансцену
тебя	а-а-а-а-а-----	авоське
у	а-ааааа	авоськи
ну	а-а-а-а-а-а	авоськой
так	а-ааааааа	авоську
это	а-а-а-ааааа	аврал
все	а-а-а-а-а-а-а-а	авт
меня	аааааааыыы-радуга	авто
мне	аааа-выходит	автобазе
то	аааа-тогда	автоматная
тебе	аааа-х	автоматной
нет	а-а-а-во	автоматные
уже	аааай-выгнулся	автоматными
но	аааамяуррр-опять	автоматчика
еще	ааа-понятно	автомобиля

Figure 23. Russian vocabulary for different keyboards

Yandex Keyword	Smart Keyboard	Google Keyword
всё	автоплатёж	No words with ё
ещё	автоприём	
её	адьё	
чё	аёт	
моё	аётам	
котёнок	аксёнов	
неё	аксёнова	
твоё	аксёновой	
чём	актёр	
идёт	актёра	
пошёл	актёрам	
своё	актёрами	
днём	актёров	
пришёл	актёрской	
пойдём	актёры	
придётся	алгёбру	

Figure 24. Words with letter ё in vocabulary for different keyboards

In modern smartphones, T9 was substituted by a much more sophisticated autocorrection tool, due to the fact that digital QWERTY keyboards are now the standard. Unlike T9, today's autocorrector does not write the word during the user's typing, but corrects it afterwards if it recognizes a mistake. This is because the keyboard with its own vocabulary is built-in in the middleware. The system automatically scans saved words in the dictionary and offers them by default during typing.

There are two types of dictionaries: systematic ones and custom ones. In the first case we can see the original set of words inserted (Figure 23 and Figure 24). The second list is more dynamic. New words are constantly added by monitoring the users, or they can be manually entered by them. At the beginning there are no words in the dynamic dictionary and it has frequently a poor range of words (it is assumed to have 500-1000 words). Due to the allocated memory reaching maximum capacity, older words are deleted automatically.

The autocorrection of the words with the letter ё depends on that letter being in the vocabulary base of the keyboard. For example, if we take the word "ёж" ("hedgehog" in Russian), we can see how it is treated by different keyboards. In three of the best established keyboards, the word "еж" was not changed for "ёж", although in the line of autocorrection there was an option for "ёж."
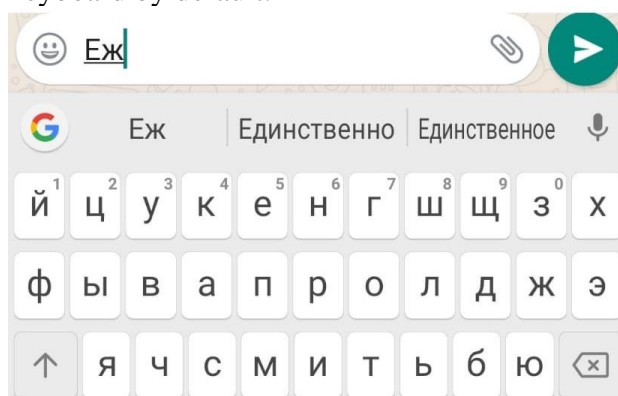
1. Samsung Keyboard by default.

Figure 25. Keyboard Samsung

2. Google Keyboard
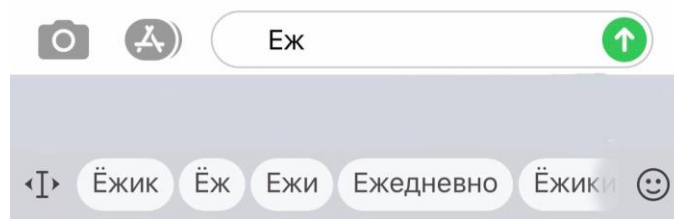

Figure 26. Keyboard Google

3. Yandex Keyboard

Figure 27. Keyboard Yandex

A similar situation arises with the word "ёжик" (уменьшительно-ласкательное от слова "еж," that is, "hedgehog"). The word "ежик" was changed automatically to было "ёжик" in all keyboards.
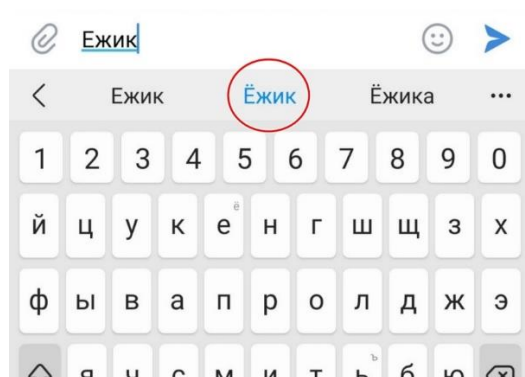
1. Samsung Keyboard by default.

Figure 28. Keyboard Samsung
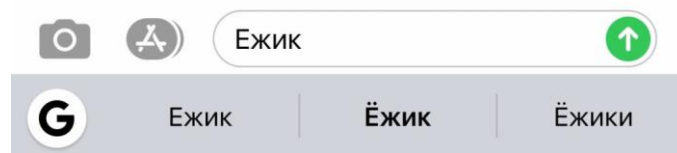
2. Google Keyboard
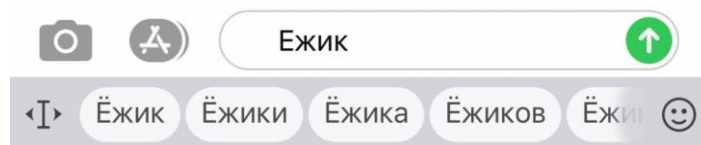

Figure 29. Keyboard Google

3. Yandex Keyboard

Figure 30. Keyboard Yandex

According to these examples we can conclude that the autocorrect intervention in different keyboards depends on certain factors. Firstly, how the software works depends on the type and producer of the keyboard. As a result, all systems work independently of each other. Obviously, every artificial intelligence was taught differently in every part of its production as well as receiving a different programming by IT-engineers. The second condition, we assume depends on the users. According to their choices, the correction is set in a certain way.

In Table 1, 40 words are represented. Twenty of them have the letter ё and are used in the vocabulary of T9, meanwhile, the other 20 words are not used. We have tested different keyboards and checked them by writing some of those words. The results are shown in the table.

Table 1. Autocorrect words with the letter ё

Word	word with letter ё in the T9 base vocabulary	Autocorrection of words		
		Yandex Keyboard	Smart Keyboard	Google Keyboard
The replacement of ё in all sentences				
Пойдём	+	Пойдём	Пойдём	Пойдём
Ждём	+	Ждём	Ждём	Ждём

Даёшь	+	Даёшь	Даёшь	Даёшь
Приём	+	Приём	Приём	Приём
Шёл	+	Шёл	Шёл	Шёл
Везёт	+	Везёт	Везёт	Везёт
Мёд	+	Мёд	Мёд	Мёд
Зелёное	+	Зелёное	Зелёное	Зелёное
Тётя	+	Тётя	Тётя	Тётя
Смеётся	+	Смеётся	Смеётся	Смеётся
Путёвка	+	Путёвка	Путёвка	Путёвка
Плетёная	+	Плетёная	Плетёная	Плетёная
Перебьёт	+	Перебьёт	Перебьёт	Перебьёт
Пьёшь	-	Пьёшь	Пьёшь	Пьёшь
Врёшь	-	Врёшь	Врёшь	Врёшь
Флёр	-	Флёр	Флёр	Флёр
Мёртвый	-	Мёртвый	Мёртвый	Мёртвый
Ввёл	-	Ввёл	Ввёл	Ввёл
Вертолёт	-	Вертолёт	Вертолёт	Вертолёт
Перевёл	-	Перевёл	Перевёл	Перевёл
Самолётик	-	Самолётик	Самолётик	Самолётик
Пришёл	-	Пришёл	Пришёл	Пришёл
Рулём	-	Рулём	Рулём	Рулём
Насчёт	-	Насчёт	Насчёт	Насчёт
Вдвоём	-	Вдвоём	Вдвоём	Вдвоём
Орёт	-	Орёт	Орёт	Орёт
The replacement of ё in some Keyboards				
Бельё	+	Бельё	Белье	Белье
Живёт	+	Живет	Живет	Живёт
Днём	+	Днем	Днем	Днем
Ребёнок	+	Ребенок	Ребёнок	Ребёнок
Самолёт	+	Самолет	Самолет	Самолет
Отчёт	+	Отчёт	Отчет	Отчет
Объём	-	Объём	Объем	Объем
Тушёнка	-	Тушёнка	Тушенка	Тушёнка
Счёт	-	Счёт	Счёт	Счет
Приём	-	Приём	Прием	Прием
Отдаёшься	-	Отдаёшься	Отдаешься	Отдаешься
Лёд	-	Лёд	Лед	Лед
There is no replacement of ё in sentences				
Рёв	-	Рев	Рев	Рев
Перетёр	-	Перетер	Перетер	Перетер
Зёрнышко	-	Зернышко	Зернышко	Зернышко
Отвёртка	-	Отвертка	Отвертка	Отвертка
Позёмка	-	Поземка	Поземка	Поземка
Клёпка	-	Клепка	Клепка	Клепка

So, some words can be written with the letter e. The example can be given of the words "еж" and "ежик" ("hedgehog"), mentioned above (figure 25 and figure 30). Those are two different forms of the same word, with the same meaning. By this it could be said that autocorrection creates its own rules in writing words with the letter ё, based on a combination of the built-in dictionary and the individual vocabulary of the user. By this it could be confirmed that whatever word and first letter are used, the dictionary creates its own base of words, known as well as unknown ones, if it is not done by default.

Currently, we have an interesting situation as one uses different keyboards and

applications in the smartphone. It is undeniable that users can correct the word by themselves if they do not like how it is written. However, is also a fact that not everyone would waste their time on this.

The main issue of autocorrect is said to be that it misunderstands the context of a sentence. Autocorrection is able to change only one word, but it does not understand the context. As a result, it would lead to losing the main idea of the text. It is nevertheless essential to point out that autocorrection is developing every day. The vocabulary of the system is becoming bigger and bigger, as well as developing alongside artificial intelligence. If artificial intelligence were capable of creating a list that includes all of a user's vocabulary and recognizes its most frequent contexts of use, the system would be able to foresee all phrases and sentences in the future.

Благодаря своим жестким колючкам ежики защищаются от хищников, а также перетаскивают провизию прямо у себя на спине. Они схожи с иголками елки. Хищникам, искущенным желанием полакомиться ежатиной, придётся применить смекалку и терпение. Лесные жители, живущие по соседству с ежами, стараются не натыкаться на них. Медведи научились использовать такой приём, как зимняя спячка, в том числе, чтобы реже встречаться с ежами, выходят они из неё только весной.

Figure 31. Text typed in the smartphone

The text in figure 31, for example, reveals the working of the keyboard-system concerning the letter ё. The smartphone replaces the letter е to ё only in some words. This is illustrated once again by the word "ежик" ("hedgehog", phonetically correct form "ёжик"). Likewise, we can see the same happening with the words "жестким", "елки" and others. However, by typing it again on the same keyboard, one can get a different result. Having learned about the user's habits and preferences (if a user does not correct back words written with "ё," this is considered a preference for it), the autocorrection system replaces more occurrences of "е" by "ё"   ("ёжики", "ёлки" and "неё" – see Figure 32).
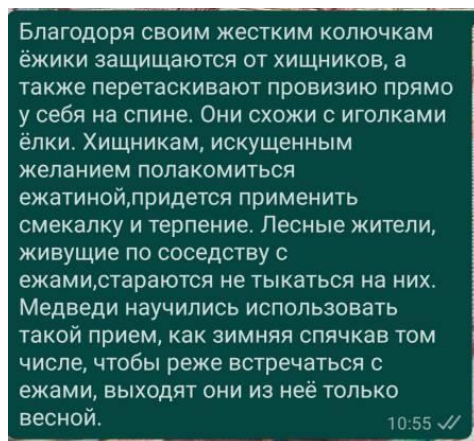
Figure 32. Changes occurring when typing the same text multiple times

SUMMARY AND CONCLUSION

In 1783 the sign ё was created. Only in 1917, however, it received the formal status of letter and its use was mandatory only during the ten years from 1942 to 1952. The ordinary use was only for avoiding ambiguity, beyond that, readers should understand by of themselves when a particular word with the phoneme [io] or ё, should be read and pronounced that way even it is written as e.

This peculiar status led to the disappearance of the letter ё in many devices. The first typewriter in the USSR was invented in 1929 but did not bring "ё" to typewriting. The special key appeared only in the mid-1950s. In the 1980s the first electric typewriter was invented, but there was no letter ё even in that machine. The letter was absent also in the first computer keyboards. All of this presents difficulties for personal names which have the letter ё, when during document-processing the letter is not supported.

Despite all the factors which were mentioned above, there was always the option, exercised by some, to create macros to support the letter ё. Also, a lot of personal computer keyboards do not indicate the letter but support it nonetheless.

There are two possible options of using the letter ё according to the rules of the Russian language. The first one is the selective one. We use "e" or "ё" to disambiguate cases when the main idea would appear to be different and when the words have different meanings. The second option is the solid one. Here one uses it when it is customarily written and pronounced. These rules have been adopted by modern computer programs. Moreover, Microsoft Word allows society to choose solid writing as an option. However, it is true that there are a lot of problems concerning writing "ё" in different text, forums, electronic registrations, payment terminals, where the letter ё cannot be supported and the error is shown on the monitor. A fascinating situation arises with smartboard keyboards which change some words automatically (including, in some cases, changes from "e" to "ё"). This, of course, is not in keeping with the modern grammar of the Russian language which tends to limit these changes to cases of disambiguation. These automatic changes result from a technical decision which draws on an analysis of historical and users'

preferences. It would seem, however, that the use of "e" or "ё" is a question of future writing, that is, it concerns the ways in which we choose to relate sounds and spellings. In a dynamic language this should not be left to personal settings and automated technical decisions, even self-learning algorithms. Instead, it is akin to voting for the future of our language and the society who struggles with these decisions every day.

REFERENCES

Abulhab, S. D. (2004). The Mutamathil Type Style: Towards Free, Technology-Friendly, Arabetic Types. *Visible language, 28*(3), 306–333.

Adali, K., & Eryiğit, G. (2015). Vowel and Diacritic Restoration for Social Media Texts. *In Proceedings of the 5th Workshop on Language Analysis for Social Media (LASM)* (pp. 53–61). Association for Computational Linguistics. https://doi.org/10.3115/v1/w14-1307

Alkhatlan, A., Kateb, F., & Kalita, J. (2020). Attention-Based Sequence Learning Model for Arabic Diacritic Restoration. In *Proceedings – 2020 6th Conference on Data Science and Machine Learning Applications, CDMA 2020* (pp. 7–12). IEEE. https://doi.org/10.1109/CDMA47397.2020.00007

Bozhko, N. N. (2011). *Russkiy Yazyk. Trudnyye Voprosy Orfografii* [Russian language. Complicated questions of orthography]. Uchitel.

Chumakov, V. T. (2010). Bukva jo v sovremennom russkom yazyke [The letter Ё in the Modern Russian Language]. *Public education*, 6, 12-34.

Cocks, J., & Keegan, T. T. (2011). A Word-based Approach for Diacritic Restoration in Māori. In: *Proceedings of Australasian Language Technology Association* (pp. 126–130). Australasian Language Technology Association. https://aclanthology.org/U11-1016.pdf

Rehr, D. (1997). Antique Typewriters and Office Collectibles: Identification & Value Guide. Collector's Books.

Ezeani, I., Hepple, M., & Onyenwe, I. (2016). Automatic restoration of diacritics for Igbo language. In P. Sojka, A. Horák, I. Kopeček, & K. Pala (Eds.), *Text, Speech, and Dialogue. TSD 2016. Lecture Notes in Computer Science* (vol. 9924, pp. 198–205). Springer. https://doi.org/10.1007/978-3-319-45510-5_23

Grozea, C. (2012). Experiments and results with diacritics restoration in Romanian. In P. Sojka, A. Horák, I. Kopeček, & K. Pala (Eds.), *Text, Speech and Dialogue. TSD 2012. Lecture Notes in Computer Science* (vol. 7499, pp. 199–206). https://doi.org/10.1007/978-3-642-32790-2_24

Hung, B. T. (2018). Vietnamese Diacritics Restoration Using Deep Learning Approach. In *Proceedings of 2018 10th International Conference on Knowledge and Systems Engineering, KSE 2018* (pp. 347–351). IEEE. https://doi.org/10.1109/KSE.2018.8573427

Korbut, A. Yu. (2017). Bukva, Porozhdayushchaya Problemy [The letter that Creates Problems]. *Pedagogical Image, 3*(36), 71-80.

Ljubešić, N. L., Erjavec, T., & Fišer, D. (2016). Corpus-Based Diacritic Restoration for South Slavic Languages. In *Proceedings of the Tenth International Conference on*

Language Resources and Evaluation (LREC'16) (pp. 3612–3616). European Language Resources Association.

Michael, A. (1997). *Antique typewriters: from Creed to QUARTEY*. Schiffer.

Novák, A., & Siklósi, B. (2015). Automatic Diacritics Restoration for Hungarian. In *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing* (pp. 2286–2291). Association for Computational Linguistics. https://doi.org/10.18653/v1/D15-1275

Orife, I. F. Ò., Adélaní, D. I., Fasubaa, T., Williamson, V., Oyewusi, W. F., Wahab, O., & Túbòsún, K. (2020). Improving Yorùbá Diacritic Restoration. *ArXiv,* 2003.10564. https://doi.org/10.48550/arXiv.2003.10564

Ruseti, S., Cotet, T.-M., & Dascalu, M. (2020). Romanian Diacritics Restoration Using Recurrent Neural Networks. *ArXiv*, 2009.02743. https://doi.org/10.48550/arXiv.2009.02743

Russo, T. A. (2002). *Mechanical Typewriters: Their History, Value, and Legacy*. Schiffer.

Thurlow, C. (2021). When Globalese Meets Localese: Transformational Tactics in the Typographic Landscape – a Bernese Case Study. *Social Semiotics, 31*(1), 88–107. https://doi.org/10.1080/10350330.2020.1810544

Zitouni, I., & Sarikaya, R. (2009). Arabic Diacritic Restoration Approach Based on Maximum Entropy Models. *Computer Speech & Language, 23*(3), 257–276. https://doi.org/10.1016/j.csl.2008.06.001

Zubacheva, K. (2018, July 31). Why 'Ё' is the Weirdest Letter in the Russian Alphabet. *Russia Beyond*. https://www.rbth.com/education/328862-yo-russian-alphabet

СВЕДЕНИЯ ОБ АВТОРАХ / THE AUTHORS

Александр Дорошев. doroshevam@icloud.com
ORCID 0000-0001-5449-4745

Анна Полякова, polyakova2.as@edu.spbstu.ru
ORCID 0000-0002-1510-3702

Alexander Doroshev, doroshevam@icloud.com
ORCID 0000-0001-5449-4745

Anna Polyakova, polyakova2.as@edu.spbstu.ru
ORCID 0000-0002-1510-3702