Information Technology for Entering Text Based on Tools of the Special Virtual Keyboard Mobile and Auxiliary Devices

Olexander Barmak[1[0000-0003-0739-9678]], Ruslan Bahrii[1[0000-0001-5219-1185]], Iurii Krak[2,3[0000-0002-8043-0785]], Veda Kasianiuk[2[0000-0003-3268-303X]]

[1]National University of Khmelnytsky, 11, Institutes str., 29016, Ukraine
barmakov@khnu.km.ua, bagrijro@khnu.km.ua
[2]Taras Shevchenko National University of Kyiv, 64/13 Volodymyrska str., 01601, Ukraine
krak@univ.kiev.ua, veda.kasianiuk@gmail.com
[3]Glushkov Cybernetics Institute, Kyiv, 40 Glushkov avenue, 03187, Ukraine
krak@nas.gov.ua

Abstract. Information technology for the realization of human communication using residual human capabilities, obtained by organizing text entry using mobile and auxiliary devices is proposed. The components of the proposed technology are described in detail: the method for entering text information to realize the possibility of introducing a limited number of controls and the method of predicting words that are most often encountered after words al-ready entered in the sentence. A generalized representation of the process of entering text is described with the aid of an ambiguous virtual keyboard and the representation of control signals for the selection of control elements. The approaches to finding the optimal distribution of the set of alphabet characters for different numbers of control signals are given. The method of word prediction is generalized and improved, the "back-off" statistical language model is used, and the approach to the formation of the training corpus of the spoken Ukrainian language is proposed.

Keywords: virtual keyboard, distribution of symbol, text prediction, corpus of words, statistical language model, N-gram.

1 Introduction

The article presents the results of research and creation of alternative communication technology for people who have temporarily lost their ability to verbal communication. It is proposed to organize communication by entering text and following voice reading it using standard mobile devices. Propose approaches that allow entering text with a limited number of controls, for example, using the four keys of the virtual keyboard.

In modern society communication is a vital necessity for a person, one of its main needs. A large layer of people with speech disorders needs additional means of alternative communication for communication. Alternative communication is all communication techniques that supplement or replace ordinary speech for people who do not have the opportunity to communicate as a result of congenital or acquired disease [1].

In the world there are various systems for the implementation of non-verbal communication. People with hearing disabilities use sign language to communicate. The authors proposed approaches for implementing communication in sign languages using virtual reality systems [2]. Many alternative means of communication are limited to devices that can be used only in stationary conditions. Significant constraints on existing systems, what using text entry are the low input speed and, as result, communication that associated with the use of slow methods of selecting controls [3].

The main requirement for an alternative communication system is the speed of text entry considering individual human features and quick adaptation without additional training. The problem is the small amount of available control signals that can be used to generate messages. One possible solution to this problem is the use of virtual ambiguous keyboards, which requires research of optimization methods to effectively solving the problem of ambiguous selection [4]. Another direction is to speeds up the text entry process by reducing additional user actions. This requires conducting research using Natural Language Processing methods (NLP) and Statistical Language Models (SLM) to improve prediction of words [5].

Research objective was to develop information technology for the implementation of communication for people who have temporarily lost their ability to speak. As a result of the realization of this objective: analyzed alternative information channels what suitable for communication and suggested ways to use them, developed mechanisms for rapid text input in Ukrainian with a limited number of control signals, developed system of prediction of the text, developed information technology for the implementation of alternative communication using standard mobile devices.

2 Information technology of alternative communication

Information technology (IT) what based on mobile devices was created for solving set tasks. It implements communication by replacing verbal communication to communication by voice reading of text messages, which entering by a limited number of controls (Fig. 1) [6].

It is proposed to use the characters of verbal communication of the Ukrainian language (letters) for the transmission of text messages. This is due to the fact that people with temporary speech disorders usually prefer to use the language they know, and not learn new communication paradigms.

The input information for the proposed IT is information obtained from alternative human information channels. To provide alternative communication was necessary to intellectualize the process of entering text information for a limited number of control signals. Acceleration of the text input process is possible through the use of redundancy of natural language, which involves using the virtual keyboard, the keys of which contain grouped letters of the alphabet.

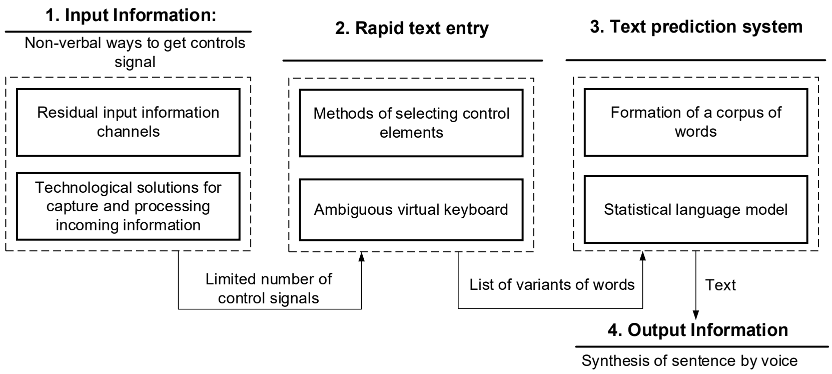

Fig. 1. Diagram of Information Technology

To improve the rapid of entering information and minimize user interaction with the IT device, a prediction system was proposed that automatically suggests the following words that are most commonly encountered after the words already entered in the sentence. For the text prediction system proposed language model and formed training corpus of words adapted to the required type of communication.

Consider the components of information technology in more detail.

3 Method of entering text information

The main difficulty of any means of alternative communication is that a large set of language symbols must be associated with a very limited set of controls. Selection techniques, which are used for the alternative communication, have significant limitations and disadvantages [7], so to speed up the selection process it is proposed to group the letters of the alphabet into control elements.

In general, the method for entering of text information a limited number of control signals is presented as follows (see Fig. 2):

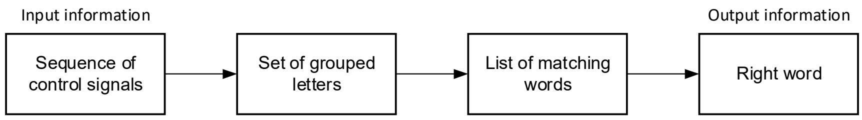

Fig. 2. Method of entering of text information

1) *input information*: sequence of control signals to enter the desired word;
2) representation of control signals in the form of a sequence of associated sets containing groups of letters;
3) list of words-candidates existing in the language (corpus) is possible with this sequence;
4) *output information*: choosing the right word.

In addition to non-traditional ways of transferring control signals for a significant number of modern people, it is already typical to manage the process of entering text information using various types of keyboards (both physical and virtual). Without loss of generality, we will consider this method of transferring the control signals in the proposed information technology the main one.

4 Ambiguous virtual keyboard

The virtual keyboard displays the character layout on the device screen. Ambiguous virtual keyboards have several characters per key that make them more productive than normal ones. The purpose of such a keyboard is to reduce the effort at entering text.

In general, the components of the ambiguous keyboard can be represented with 5 parameters:

$$K_{AMB} = \langle P_{sel}, k_{abc}, k_{sec}, A_{ord}, P_{dis} \rangle, \tag{1}$$

where $P_{sel} = \{direct\ selection, scanning, encoding\}$ – selection techniques; $k_{abc} = \{4...8\}$ – number of keys with grouped letters; $k_{sec} = \{1...5\}$ – number of auxiliary keys; $A_{ord} = \{alphabetic, keyboard, statistical\}$ – order of following the letters in the grouped keys; $P_{dis} = \{multi\text{-}press, two\text{-}key,\ scanning, T9\}$ – techniques disambiguate selection.

In Figure 3 shows a generalized diagram of the components ambiguous keyboard.

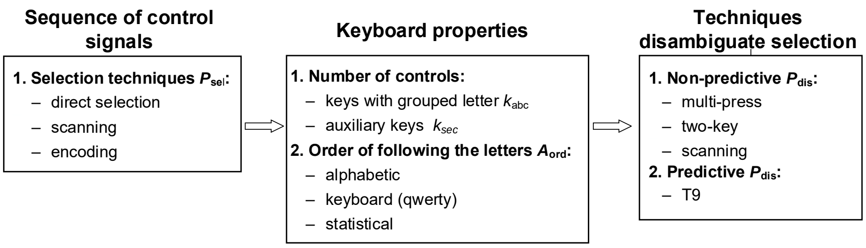

Fig. 3. Components of the ambiguous keyboard

With these components you can describe the known types of entering text information by ambiguous keyboards. For example, according to the above description, the ambiguous keyboard used to enter text on mobile devices with numeric buttons can be represented as: $K_{AMB} = \langle P_{sel} = direct\ selection,\ k_{abc} = 8,\ k_{sec} = 1,\ A_{ord} = alphabetic,\ P_{dis} = T9 \rangle$. That is, a direct selection method is used, 8 keys with grouped letters and 1 auxiliary, alphabetic order of letters and prediction method T9 with disambiguate.

The improvement of the method for entering text information with a limited number of control signals (Figure 4) within the framework of the above generalized components of the keyboard (1) is given in the following steps:

1) representation of the sequence of control signals by any of the indicated selection techniques P_{sel};
2) reducing the number of controls k_{abc}, k_{sec} and optimizing the distribution of the alphabet set A_{ord};
3) solving the problem of ambiguous selection P_{dis} by the predictive method for increasing the efficiency of text entry.

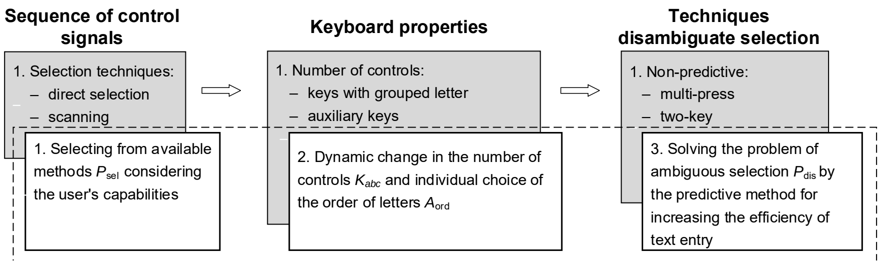

Fig. 4. Diagram the improvement of the method of entering text information with a limited number of control signals

5 Representation of the control signal for selection

In today's development of IT devices, there is a tendency to minimize their size, including the perspective development of chip devices that can be implanted in the human body. As a result, it became possible to use these devices not only under fixed conditions, but also with free movement. The use of such devices by people with disabilities makes it possible to implement alternative communication for them, which will not be tied to a particular place.

Known techniques for selecting symbols can be assigned to three main categories: *direct selection*, *scanning* and *encoding* [8]. The speed of selection directly affects the speed of communication and therefore it should be measured to determine the fastest method, which will be used in the future. In the framework of IT, the representation of a continuous and discrete control signal for the selection is proposed. The result of processing the input signal is to determine its type and the fixation of the methods of selecting control elements.

For a *continuous* control signal, the selection is carried out by a cursor, which can be moved using special devices that capture the movement of the eyes, heads, hands, etc. To make a selection, hold the cursor over the desired control. If this is the keys of the virtual keyboard, then for each time interval holding the cursor over them, the next letter that belongs to the key will be selected. This mode allows you to implement a direct selection of letters and can be used in the absence of words in the dictionary.

For a *discrete* control signal, each of its states is associated with a separate function or control element. If these states are only two, then the scanning selection mode is applied. In this mode, the internal timer cyclically "highlights" the controls and awaits confirmation of the choice from the user. Of course, the speed of text entry in this mode is quite small, but this mode allows you to use it to people who other ways of communication are not available due to physical constraints.

In the case where the number of states is not sufficient to assign each control element unique control state, a combination of selection modes is used. First of all, it is proposed to bind the keys of the virtual keyboard, as most operations are performed when selecting letters. Other interface features can be selected using encoding. Then the sequence of signal states is used to select them. This allows the maximum use of all available states of the discrete signal coming from a special device.

You can also use the usual physical keyboard and touch type control using the direct selection method to text entry. Touch control allows the use of alternative communication on different mobile devices, which is optimal for people with temporary communication disorders.

6 Research of variants distribution of set of the letters Ukrainian alphabet for different order followings

Decreasing the number of control signals will improve the efficiency of text entry for people with disabilities. In addition, it expands the list of alternative means of communication that can be used to generate control signals. If the scanning selection method is used, decreasing the number of keys should significantly reduce the search time for each character.

Determining the minimum number of keys for ambiguous keypads is an optimization problem, as with a decrease in the number of keys, the number of letters in them increases, which leads to an increase interpretations error of user actions. This requires additional refinement of choice and reduces the efficiency of text entry. So, for a keyboard with 9 keys (T9 standard), the number of letters located on each of them is on average 4th. In this case, the standard algorithm T9 successfully solves the problem of ambiguous selection [9]. Reducing the number of keys to 4 requires placing on each key, on average, about 8 letters, which leads to an increase in the list of possible words. This much complicates the selection of the user's expected words, and requires optimization of the distribution of the letters on the keys and the use of additional prediction algorithms.

Alphabet and keyboard («QWERTY») alphabetic ordering are well-known. For both of these cases, it is not possible to change the order of their sequence, so a study with a uniform distribution of letters in groups and the best variants were determined [10].

The frequency order of the letters is a list of letters of the Ukrainian alphabet in order of decreasing the frequency of their use. To find the optimal letter distributions a set of letters divided into classes according to the frequency of their use. Within these classes, it is possible to change the order of following letters, keeping to the distribution of letters with similar frequencies in different groups. Based on these conditions, an optimization problem has been formed:

1) *Constraint 1*: letters with similar frequencies should be placed in different groups;
2) *Constraint 2*: letters in groups should be evenly distributed;
3) *The criterion of the optimization problem*: the total number of words with the same code should be the smallest.

To solve the optimization problem, an algorithm for actions that considers the distribution constraints is proposed:

1) *input information:* a set of letters in a certain order of follow;
2) set of letters is divided into classes so that in one class there are letters with similar frequencies;
3) for each class, a random letter is selected and placed in a group that does not already have a letter from this class;
4) *output information*: formed groups.

For the Ukrainian alphabet, the number of classes is defined in the range from 4 to 8, depending on the number of groups. Thus, each letter that belongs to a particular class belongs to a unique group that enforces the *constraint 1.* Compliance with such an order of distribution of letters ensures the fulfillment of the constraint 2 - the letters are distributed evenly.

7 Solving the problem of ambiguous selection

In order to minimize the interaction between the user and the virtual keyboard, it is proposed to use an algorithm for eliminating ambiguity, similar to T9 (TNKey algorithm), since this method demonstrates better performance than methods that use a several actions to select a letter [11].

To solve the problem of ambiguous selection an algorithm is proposed which allows obtaining a set of words that correspond to a sequence of user actions with a virtual ambiguous keyboard. The algorithm consists of the following steps:

1) encoding of all words in the dictionary (corpus) for a given distribution of letters on the keys;
2) obtaining the code of the current word corresponding to the sequence of actions (keystrokes) of the user;
3) searching the word dictionary corresponding to the code.

In the process of text entry the current word often does not refers to the words of the Ukrainian language, but only reflects the sequence of actions. The code obtained when encoding this sequence is compared to an internal dictionary and generates a list of words corresponding to it.

The use of ambiguous virtual keyboards leads to ambiguous selection. When encoding all dictionary words for a given letter distribution, words that have the same code can be mistakenly offered to the user as an expected word. Such situations interrupt the text entry process and require a refinement of the choice. The quantitative definition of such situations is the frequency of errors, that is, the percentage of words that are misinterpreted.

Such a quantitative estimate of the error rate is determined based on the frequency of word use in the texts. Assume that in a set of words with the same code, the word having the highest frequency of meetings corresponds to the expected one, and other words are not offered correctly. Thus, the quantitative estimation of the error rate is defined as the total number of mistakenly proposed words, normalized to the total number of words, consideration their frequency.

In [10], variants of the distribution of the set of letters of the Ukrainian alphabet into groups in a certain order of succession were investigated. Using the results obtained and the proposed approach, a quantitative estimate of the error rate for a different number of groups and different orders of letters are calculated (Fig. 5).

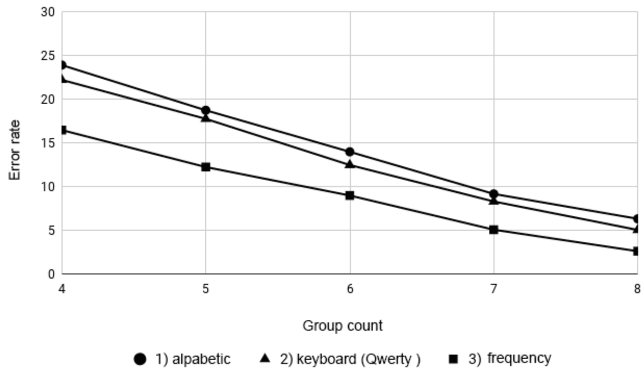

Fig. 5. Error rate for different sequence of letters

Thus, the error rate for 4 groups was 16.43% for the frequency order of letters, which is the best option compared to the keyboard (22.16%) or alphabetic (23.85%) orders. Similar relationships have also been confirmed for another number of groups.

The use of recommended letter distributions for the Ukrainian alphabet provides a more convenient way to enter text for people with different experience with digital devices and ensures the efficiency of the set with decreasing controls.

8 Words prediction method

Word prediction speeds up the text entry process by reducing additional user actions. To do this, it is necessary that as many as possible input words are predicted as the words "by default", that is, they correspond to the word that the user expects.

In general, the method of words prediction, which most often occurs after the words entered in the sentence, can be represented as follows (Figure 6):

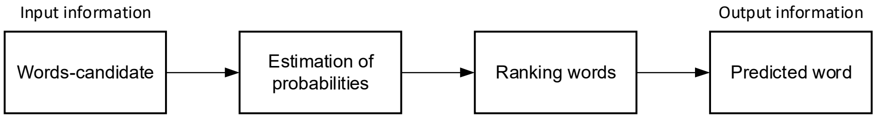

Fig. 6. Method of words prediction

1) *input information*: list of words-candidate that correspond to the sequence of user actions;
2) the evaluation of the probabilities of words-candidate, considering the previous words of the sentence is presented in the form of different language models;
3) the ranking of words for probabilities and the definition of the word "by default";
4) *output information*: predicted word.

The purpose of the prediction method is to provide a list of words sorted by probability values. For each candidate word, its probability is estimated considering the language model that uses the statistics of the corpus of the words.

The improvement of the prediction method is submitted by implementing the following two steps:

1) prediction of the most probable words using statistical language model, which would allow to realize a problem with acceptable computational complexity (for the possibility of implementation on mobile devices);
2) formation of the corpus of words of the spoken Ukrainian language to improve the quality of prediction and reduce user interaction with the device.

Statistical models of the language are used for predictive text. In the field of communication AAC for word prediction is used N-gram model that calculates the probability of the last word, as the probability of a sequence of words in a certain corpus [12].

The maximum likelihood estimation (MLE) method uses to estimate these probabilities. It consists in determining the parameters that maximize the probability of this similarity for given words. Thus, the MLE estimate for the parameters of an N-gram model by getting as normalized counts from a corpus that is statistically representative for language model [13]. For example, we can estimate the probability bigram word w_n, considering the previous word w_{n-1}, calculating the entry (count) of bigrams $C(w_n, w_{n-1})$ and normalize by unigram count for word w_{n-1}:

$$P(w_n|w_{n-1}) = \frac{C(w_{n-1}, w_n)}{C(w_{n-1})} \tag{2}$$

One of the most important problems of N-gram models is the problem of data sparse, which grows rapidly with increasing model order. In fact, the MLE provides zero probability for any sequence of words that is missing from the corpus. To solve the problem of sparse data and improve the overall quality of the prediction without increasing the computational complexity, it is suggested to use a model with a "back-off" with appropriate optimization of the parameters. In the case of a limited corpus of words, the statistical model language, what allows satisfying these requirements is the

Katz's backoff model [14]. The main idea behind Katz's backoff model is to evaluate the conditional probability of a word by a "backoff " to a N-gram of lesser order in the case where N-grams of higher order are not found in the training corpus. So, the model with the most complete information is used to provide the best results.

In particular, for the highest order N, it is proposed to use the trigram model and perform recursively backoff to the bigrams and unigrams:

$$P_{BO}(w_n \mid w_{n-2}, w_{n-1}) = \begin{cases} P^*(w_n \mid w_{n-2}, w_{n-1}) & if\ \ C(w_{n-2}, w_{n-1}, w_n) > 0 \\ \alpha(w_{n-2}, w_{n-1}) P_{BO}(w_n \mid w_{n-1}) & if\ \ C(w_{n-2}, w_{n-1}) > 0 \\ P_{BO}(w_n \mid w_{n-1}) & else \end{cases}, \quad (3)$$

where:

$$P_{BO}(w_n \mid w_{n-1}) = \begin{cases} P^*(w_n \mid w_{n-1}) & if\ \ C(w_{n-1}, w_n) > 0 \\ \alpha(w_{n-1}) P^*(w_n) & else \end{cases}, \quad (4)$$

$$P^*(w_n \mid w_{n-2}, w_{n-1}) = \frac{C^*(w_{n-2}, w_{n-1}, w_n)}{C(w_{n-2}, w_{n-1})}, \quad (5)$$

where P_{BO} – probability calculated by Katz's backoff model, P^* – smoothed probability, estimated by Good-Turing, C – count of N-gram, C^* – smoothed count of N-grams, w_n – predicted word, w_{n-2}, w_{n-1} – previous words, α – backoff coefficients.

In order to simplify the calculation of the coefficient α for the formula (3) offered to be taken equal 0.4, as such, which was heuristically received by Google specialists for the Stupid backoff (SBO) algorithm. It is context-independent and shows approximation to other methods of prediction quality for a limited corpus of words [15].

With each backoff level r the value of the backoff coefficient is calculated according to the formula α^r, where $r = 0$ for trigrams, $r = 1$ for bigrams and $r = 2$ for unigrams. Thus, the probability value calculated by the simplified calculation (S_{Simple}) takes the form:

$$S_{Simple} = \alpha^r P, \quad (6)$$

where:

$$P\ \ (w_n \mid w_{n-2}, w_{n-1}) = \begin{cases} \dfrac{C(w_{n-2}, w_{n-1}, w_n)}{C(w_{n-2}, w_{n-1})}, & if\ \ C(w_{n-2}, w_{n-1}, w_n) > 0 \\ \dfrac{C(w_{n-1}, w_n)}{C(w_{n-1})}, & if\ \ C(w_{n-1}, w_n) > 0 \\ \dfrac{C(w_n)}{N}, & else \end{cases}, \quad (7)$$

α – backoff coefficient, r – value of backoff level, P – probability calculated by backoff model, C – count of N-gram, w_n – predicted word, w_{n-2}, w_{n-1} – previous words, N – count unigrams.

In this way, word w_n with the highest probability S_{Simple} will be defined as the word "by default".

9 Formation of the corpus of words

The choice of the training corpus is an important stage in the development of any system of predicting the text. To obtain reliable estimates of the probability, statistical language models must be trained in a large set of texts. Also, the more a training corpus is such to this type of communication, the more accurate the probability estimates.

The study modeled the kind of communication that is used in everyday communication. Target users are people with special needs who want to express their thoughts, needs or feelings through alternative communication. Since there are no special corpuses for this communication environment, other areas of communication with similar characteristics need to be considered.

To solve this problem, it is suggested to use dialogues on common topics used in vocabulary for studying foreign languages. Such dialogues simulate conversations between people who see each other, cover the most possible common situations and use a limited set of words and phrases.

For the implementation of the communication system of people with disabilities communication was created a limited corpus of spoken Ukrainian on the basis of common topics that simulate conversations between people in similar situations and use a limited set of words and phrases. To create the corps of Ukrainian language dialogues, a set of texts has been collected consisting of more than 400 dialogues on various subjects, the total volume of which was about 20,000 phrases and 100,000 words. The resulting dialogues, for the further formation of the model, were divided into basic and test sets and experimental studies were conducted to determine the enough filling of the corpus for the task of predicting words and phrases when entering the text [16].

The language model is formed from the text corpus by its partition into N-grams - unigrams, bigrams and trigrams, where each record in N-gram retains the count of words and phrases in a text corpus.

10 Efficiency of information technology

Estimated overall efficiency IT consists of the following properties of alternative means of communication: the speed of text entry a limited number of controls, quantification of the frequency of interpretations error in the process of input and quality predictive value.

The speed of text entry depends on the individual characteristics of the person - his remaining communication capabilities, the experience with digital devices and the time spent on adaptation. Therefore, for its estimation is used not the indicators of time expenditures, but the productivity of text input, that is, the minimization of the interaction of the device with the user.

The method of entering text information with a limited number of controls, which is an integral part of IT, allows you to implement the possibility of text entry with the help of 4 to 8 keys. The proposed variants of the distribution of letters for the Ukrainian alphabet on the keyboard keys allow you to individually consider the features of the user.

Using the 4 control keys is the fastest way to enter text, because the minimum number of control signals is used, but it requires high IT efficiency to reduce the number of additional actions to refine the choice. In order to compare the effectiveness of the components of the IT research was also conducted using the 6-key controls.

Using an ambiguous input method allows a certain percentage of interpretation errors. The lower the level of interpretation errors, the higher the input performance, since the time spent on correcting an erroneous selection significantly exceeds directly when entering the letters of a word. The distribution of the set of letters of the Ukrainian alphabet for different order of order greatly influences the frequency of interpretation errors.

Word quality prediction is a final component of evaluating the effectiveness of IT and accumulates the impact of all its components. To determine the predictive quality of the any text, a number of experiments were conducted using various statistical language models.

For the 4 keys of the frequency distribution, the prediction quality (Figure 7) for the backoff model was 89.2% for the words known as the N-gram model. Using such a model significantly improves the quality of prediction compared to the use of conventional probability estimates.

For the 6 keys of the keyboard distribution, the prediction quality (Figure 8) exceeded 90% compared to the distribution for 4 keys. This is due to the decrease in the number of candidate words, from which the word "by default".

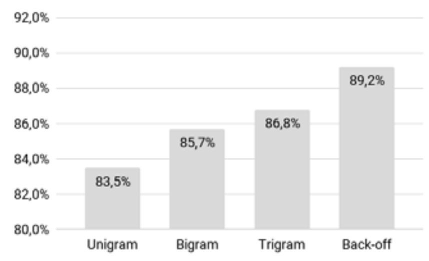

Fig. 7. The quality of prediction of any text on different models for 4 keys of distribution

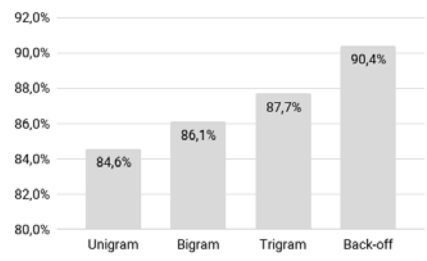

Fig. 8. The quality prediction of any text on different models for the 6-key of distribution

The study of the characteristics of the developed IT made it possible to evaluate the overall effectiveness of the text entry with a limited number of controls and the impact of its components on the final quality of predictive. In the course of the conducted experiments we obtained the performance indicators of the applied methods of input and prediction of text information for different numbers of control elements and the distribution of letters on them (Figure 9). Under the well-known approach is meant a method of entering text information using an ambiguous keyboard with alphabetical order of letters for a specified number of controls.

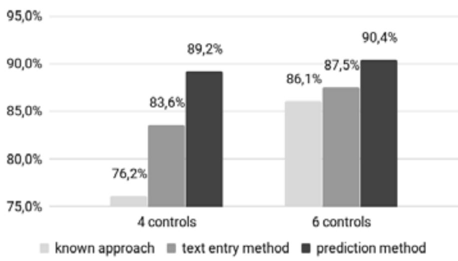

Fig. 9. Effectiveness of the components of information technology

So, the results of an experimental study of the effectiveness of the developed IT alternative communication using methods of input and prediction of text information show that the proposed IT allows people to communicate with the use of residual human capabilities by organizing text input with efficiency exceeding the well-known approaches of 4-13% depending from the number of operating controls.

11 Implementation of information technology alternative communication

A human-computer interaction model was developed for the IT under consideration (Figure 10), which enabled the implementation of software for text entry into digital devices to provide communication for people with disabilities (Figure 11).

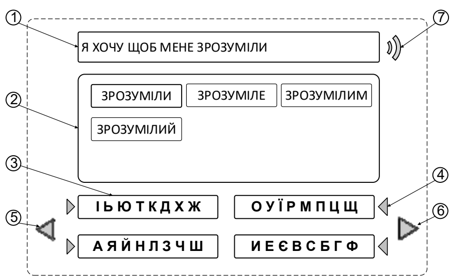

Fig. 10. User interaction model (Ukrainian alphabet)

Fig. 11. Alternative communication system interface (Ukrainian alphabet)

There is also the possibility of implementing information technology of alternative communication for other languages, such as English (Figures 12, 13).

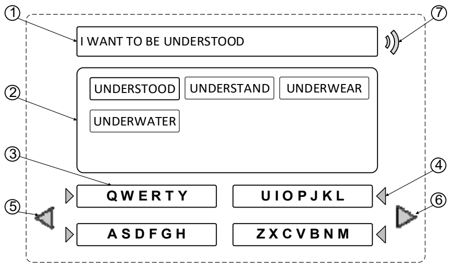

Fig. 12. User interaction model (English alphabet)

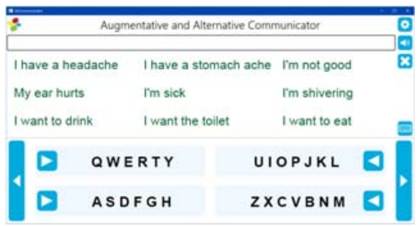

Fig. 13. Alternative communication system interface (English alphabet)

The user interaction model with the system for the organization of alternative communication considers individual human constraints (Figure 12, 13). The operating area of the model consists of the following controls: 1 – zone for displaying the entered text; 2 – a zone for displaying suggested words that correspond to the current code of the word entered; 3 – a zone that displays a virtual keyboard with the selected order of letters followed; 4 – control that allows you to select the key containing the required letter; 5 – control that allows you to undo the false choice; 6 – control element that moves to the next word; 7 – a control that activates the function of voice playback of text.

To process the reduced number of states of the input signal, an algorithm for linking them with virtual keyboard keys with ambiguous selection and with separate functional elements of control is proposed.

Object-oriented approach is chosen for information technology design. A general diagram of classes with the description of attributes and methods is developed that allows realizing the information technology on different platforms.

12 Conclusion

The article presents the results of the solution of the actual problem of the implementation of alternative communication for people in whom the channel of verbal communication is temporarily absent. The following basic scientific and applied results are obtained:

1) developed method for entering text with a limited number of controls;
2) developed a method of prediction the words that most often occur after the words entered in the sentences;
3) Predictable text entry system based on mobile devices for implementing communication of people with temporarily speech disorders has been developed.

Using the proposed information system of alternative communication significantly increases the level of socialization of people with special needs, improves the quality of their lives, develops self-esteem and gives them the opportunity to feel like a full-fledged personality.

References

1. Augmentative and Alternative Communication (AAC) [Internet]. Available from: http://www.asha.org/public/speech/disorders/AAC/ [Accessed: 2018-09-06]
2. Kryvonos, I.G., Krak, I.V., Barmak, O.V., Kulias, A.I.: Methods to Create Systems for the Analysis and Synthesis of Communicative Information. Cybernetics and Systems Analysis 53(6): 847–856 (2017)
3. Koester, H. H., Simpson, R. C.: Method for enhancing text entry rate with single-switch scanning. Journal of Rehabilitation Research and Development 51(6): 995–1012 (2014)
4. MacKenzie, I. S., Felzer, T.: SAK: Scanning ambiguous keyboard for efficient one-key text entry. ACM Transactions on Computer-Human Interaction 17(11): 1–39 (2010)
5. Higginbotham, J., Lesher, G. W., Moulton, B. J., Roark, B.: The Application of Natural Language Processing to Augmentative and Alternative Communication. Assistive Technology, vol. 24, pp. 14–24 (2012)
6. Kryvonos, Iu.G., Krak, Iu.V., Barmak, O.V., Bagriy, R.O.: New tools of alternative communication for persons with verbal communication disorders. Cybernetics and Systems Analysis 52(5): 665-673 (2016)
7. Cook, A.M., Cook, P.J.M.: Hussey's Assistive Technologies: Principles And Practice. 4th ed. St. Louis, Missouri Mosby, Elsevier (2015)
8. Dowden, P., Cook, A., Reichle, J., Beukelman, D., Light, J.: Choosing effective selection techniques for beginning communicators. Implementing an augmentative communication system: exemplary strategies for beginning communicators, Baltimore, MD: Paul H. Brookes Publishing Co, pp. 395-429 (2002)
9. Silfverberg, M., MacKenzie, I.S., Korhonen, P.: Predicting text entry speed on mobile phones. In: Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI '00); Hague, 01-06 April 2000, Netherlands: ACM, pp. 9-16 (2000)
10. Krak, Y.V., Barmak, A.V., Bagriy, R.A., Stelya, I.O.: Text entry system for alternative speech communications. Journal of Automation and Information Sciences 49(1): 65-75 (2017)
11. Grover, D., King, M., Kuschler, C.: Patent No. US5818437. Reduced keyboard disambiguating computer. Tegic Communications, Inc., Seattle (1998)
12. Ghayoomi, M., Momtazi, S.: An overview on the existing language models for prediction systems as writing assistant tools. In Proc. IEEE Int. Conf. Syst., Man Cybern., San Antonio, TX, USA, pp. 5083–5087 (2009)
13. Jurafsky, D., Martin, J. H.: Speech and language processing. 2nd ed. Prentice Hall Inc, New Jersey (2015)
14. Wikipedia: Katz's Back-off Model [Internet]. Available from: https://en.wikipedia.org/wiki/Katz's_back-off_model [Accessed: 2019-01-19]
15. Brants, T., Popat, A. C., Xu, P., Och, F., Dean, J.: Large language models in machine translation. In: Proceedings Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning. Prague, Czech Republic, 28-30 June 2007, pp. 858-867 (2007)
16. Kryvonos, Iu.G., Krak, Iu.V., Barmak, O.V., Bagriy, R.O.: Predictive text typing system for the Ukrainian language. Cybernetics and systems analysis 53(4): 495-502 (2017)