Identification of the Native Language of a Person

Yana Lazorenko [1], Viktor Shevchenko[2], Igor Sinitsyn[2]

[1] *Taras Shevchenko National University of Kyiv, Bohdan Havrylyshyn St. 24, Kyiv, 02000, Ukraine*
[2] *The Institute of Software Systems of the National Academy of Sciences of Ukraine, Academician Glushkov Avenue, 40, corps. 5., Kyiv, 03187, Ukraine*

Abstract

With the great increase of population movement, caused either by temporarily needs due to travelling or by long-term ones due to work, etc., there appears a need to improve the processes of movement control and identification of groups of people. The primary need is to identify if a person belongs to a certain nationality or territory of primary residence. In addition, it can be useful for scientific social and political researches, as well as for the field of tourism and entertainment. Such identification of people by language environment can also help to find out more about the cultural environment, which will make possible to predict the preferences of entire groups of customers better.

The analysis of the audio recording of the speaker's speech was divided into 3 stages, each of which contains a defined step-by-step instruction for the processes of data preparation and handling. Firstly, the sound itself was analyzed, because voice and pronunciation are one of the fastest and most effective tools for identifying a person or a group of people. Then the audio recording was converted into text and an analysis of the lexical composition of the studied fragment of the conversation was done. At the end, the results of the program, got in the previous two stages, were compared and the complex evaluations were done according to the criterias determined in the theoretical researches. The mentioned criterias were used with weighting factors assigned to them. As a result, an assumption about the speech environment of the speaker was given by the program.

The work also describes the factors which affects the formation of pronunciation and the change of various intonations, the relationship between text and sound. Some dependencies between sound parameters and the pronunciation of sounds in the Russian and Ukrainian languages were mathematically formalized.

The topic of the work is wide, as it considers not only lexical, but also acoustic features of speech. In this way, not only changeable or situational qualities are taken into account, for example, the vocabulary of the conversation, but also those ones that are related to the person regardless of the context of the conversation, his or her emotional state or what language he or she is speaking at the very moment. The criterias considered in the work can help to make assumptions about the speaker's "natural", most familiar language, in other words the language he or she uses most of the time in life.

Keywords

Speaker identification, text analysis, sound characteristics, acoustic spectrum of sounds

1. Introduction

Every year the transport, logistics and tourism industries are actively developing, because people travel, relocate due to a change of job, move due to climate and health problems, etc.

Therefore, the automation of the processes of control of movement and identification of groups of people is becoming more and more relevant and requires the improvement of existing tools and the application of new approaches.

13th International Scientific and Practical Conference from Programming UkrPROGP'2022, October 11-12, 2022, Kyiv, Ukraine
EMAIL: yana_lazorenko@knu.ua (A.1); gii2014@ukr.net (A.2); ips2014@ukr.net (A.2)
ORCID: 0000-0002-3987-2338 (A.1); 0000-0002-4120-0784 (A.2); 0000-0002-9457-7454 (A.2)


One of the fastest means is to determine a person's belonging to a certain nationality or territory of primary residence by pronunciation. To do this, it is only necessary to analyze a short fragment of the conversation, which takes a minimum of time. At the same time, this method is quite effective, because the pronunciation is formed over a long time, and it also takes a lot of time to correct it. Some differences in pronunciation and accent are established in childhood and remain forever. Thus, it allows to determine which language was a person's first or which language he mainly speaks. In addition, pronunciation is also a marker of a person's territorial origin, because even different regions of the same country have different dialects.

Thus, the processing of human speech for the purpose of identification of his/her personality is an actual and rather broad task, because it includes the analysis of both text and sound, as well as both innate features (characteristics of the human voice itself) and acquired or situational ones (due to dialects in a certain area, or intonation, etc.). At the same time, there are currently no automated means that use a complex analysis of such factors, which makes this method of identification impossible for now. However, in order to solve this problem, it is important to carry out research and development, because language and speech are a primary and very visible marker of many common features of social groups in society.

2. Analysis of the state of research issues

Most of the applied models or expert systems for processing live human speech are focused specifically on identifying the speaker's language and recognizing his speech. In fact, the main task of these solutions is to transcribe the audio recording into text. Further, the text is mainly analyzed for its logical content, i.e. the main task is to determine what the text is about and to select the necessary data from it.

On the other hand, there are very few programs or applications that would focus specifically on the features of language constructions used in the text or on specific features of pronunciation. And again, they solve one applied problem in general - text recognition. However, it does not allow to correlate the received data with the personality of the speaker and does not provide any characteristics about him, except, in fact, what language he is currently speaking. In reality, appearance and speech skills carry the majority of information both about a person individually and about groups of people, depending on the characteristic chosen for classification.

Therefore, the first thing we have is a voice. It contains features that are unique to a specific person. Therefore, quite a lot of research is currently being conducted and systems are being designed for identifying a person by voice. M.V. Tkachenko, R.M. Fedorenko, Y.V. Kondratenko, and I.H. Zotova considered in detail the characteristics of the voice, the principles of their formation, and the possibility of using this data to determine the speaker [1]. There are already ready-made systems that are used in banking, medical systems and part of US government bodies, for example, developed by Nuance Communications Inc. [2] and Biocatch [3]. These materials are suitable for emphasizing the relationship between unique voice parameters and speaker identification. At the same time, the results of these works are also useful for the classification of groups of people based on their speech, where it is necessary to generalize their common speech features and at the same time neglect the specified unique features. In addition, a person's voice partially changes during life depending on age, diseases, emotional state, etc. Therefore, it is also necessary to take into account and "smooth out" such errors when analyzing the speech of a whole group of people. This will predict a large number of errors and avoid the effect of "overtraining" of a model based on neural networks.

Y.S.Yamnenko and O.V.Smirnov worked on more general characteristics of the voice and patterns of their changes. In the joint work of the authors, the principles of changing the voice depending on the emotions and psychological state of the speaker are given [4]. This is one of the most important aspects that must be taken into account when trying to classify people into groups based on speech, because atypical sounding fragments of an audio recording may not indicate the difference of a person from a certain group, but simply the peculiarities of his mental state at a specific moment. If some similar emotionality is present on the audio recordings of most of the people in the group, but it is known that they are in different psychological states, then it is possible to talk about the intonation of speech. This is one of the important tools for identifying groups of people, as

each language differs in its intonation and level of its expression. At the same time, people usually leave the intonation of their native language in their speech, even when they switch to foreign languages. And it is quite difficult to correct such a "habit". In this way, it is possible to classify people according to their native language or the language in which they communicate most of the time, distinguish the specificity of their speech and, accordingly, characterize their environment and the territory of their preferred stay.

Alla Bagmut, doctor of philological sciences, has been researching the intonation features of language theoretically for many years. In her writings, she not only provides thorough theoretical materials on intonation as a phenomenon of voice change, but also analyzes in great detail the rules of building intonation and emotions in speech according to the syntactic and semantic features of sentences. Thus, some monographs of philology are devoted to the analysis of only one type of sentence, for example, a narrative sentence with equal intonation [5], [6]. At the same time, Ruslan Minnigalimov managed to formalize some of these regularities in the work "Analysis and Synthesis of Ukrainian Speech" [7], where he mathematically described the regularities of changes in the frequencies of the main tone for narrative affirmative and negative as well as interrogative sentences. However, this is not enough for a comprehensive analysis of human speech.

Therefore, the goal of the work is to identify patterns of changes in voice characteristics inherent in certain languages; to compare the differences between the speech in different languages and their formalization; to analyze both sound and lexical components of speech for a comprehensive analysis of human speech; to determine the sequence of speech analysis steps by voice and text when classifying speakers by the language they mainly use. A comparison of speakers of the Ukrainian and Russian languages was chosen for the research, but the general principles of the approach are universal and can be used to study other languages.

3. Available theoretical approaches

The vast majority of methods of speaker identification by voice and text analysis are based on machine learning algorithms. This approach is quite effective, but at the same time it is quite flexible and dependent on the input training data. Therefore, the effectiveness of the program can vary significantly depending on the selected sample for training. Incorrectly calculated statistical data can also affect the operation of such a system. Instead, combining machine learning with well-defined rules and mathematically formalized regularities always improves the quality of the solution and makes it more stable.

Many speech recognition programs use a dynamic programming method that works by comparing audio or text samples with a standard, and a number of additional methods for separating individual elements (sounds or phonemes) that make up syllables and morphemes. Neural networks, Markov models and methods of discriminant analysis have become widely used.

Also, to divide the audio recording into syntactically complete fragments according to the content, the division by pauses is used, which usually indicates the punctuation marks in the text. For example, a comma is followed by a minimal pause, and a dash or period is followed by a longer one [8]. Based on these pauses, their frequency and duration, assumptions are made about the emotional state of the speaker, his speech problems, etc., but at the same time, they do not take into account the extent to which pauses are generally used in a particular language. Therefore, it remains impossible to characterize the speaker in the context of society - only his personal flaws or experiences at a specific moment.

Methods of identifying a person by text are based on identifying the most frequently used words and constructions by the author. Thus, just as with audio recordings, it makes it possible to characterize a specific person personally, but not him in society. That is because the particular words that a particular person prefers or his personal style of speech do not mean that the traits of that person are inherent in the entire group of people with whom he interacts. To identify the characteristics of a group it is necessary to focus not on the uniqueness of certain speakers, but on the contrary, on the similarity of their speech within a group formed, for example, by a national characteristic. By summarizing the acquired knowledge, it is possible to deduce the features of uniqueness for the entire

group, which already makes it possible to identify this group among other groups of society and further correlate individual speakers to one or another group.

Therefore, the main contradiction between the described approaches and practical needs is that the principles of identification and characterization of speakers are based exclusively on the personal characteristics of a person's speech and consider the person himself only as a separate entity, independent of others. However, a person's speech is greatly influenced by his environment and his place of residence, which already suggests that several people can have identical speech characteristics. In addition, a person is subconsciously able to recognize different pronunciations, dialects, accents, etc. literally by a few phrases, but there are actually no working software tools for this. It can be concluded that the speech of a group of people can be generalized and formalized according to some sound and textual characteristics. And the found dependencies between these linguistic characteristics and the linguistic environment of a person, the territory of residence, occupation, etc., can be used for practical tasks, where the personal qualities of people can be neglected, but at the same time, their classification as groups according to social, political or geographical criteria is important. Such automation can be integrated into various fields of human life, into the processes of statistical sociological research, etc.

4. Linguistic criteria for the classification by speech

In order to recognize the speaker's belonging to a certain group by his speech in various fields and for various applied tasks, it is important to analyze everyday conversation, the kind that a person speaks in life under casual circumstances [9]. It is worth noting that it is more effective to recognize the language of everyday conversation by sound than by text, especially in the context of the languages chosen for this work. This is due to the fact that currently Ukrainian, Russian and partially Belarusian are quite strongly mixed on the territory of Ukraine. That is, in fact, most people speak "surzhyk", which is closer to Ukrainian, sometimes to Russian, and in the northern regions to Belarusian. In addition, the border regions of Russia have dialects where the vocabulary is somewhat "ukrainized", that is, they also speak "surzhyk". Therefore, a simple analysis of the text can give significant errors. Rather, it is appropriate to use it as one of the criteria for determining a language with a certain weighting factor.

Accent can be a vivid marker of language - in Russian and Ukrainian, accents differ greatly and in many words that are used precisely in everyday conversations [10]. So accent is such an intermediate element between text and sound analysis, because it is necessary to recognize on which syllable the sound is amplified in speech, and then compare it with the syllable on which it is usually amplified in another language. Of course, this is also not an absolute indicator, because due to a long assimilation, we got used to emphasizing some words in the Russian way. But still, emphasis should be included in the set of criteria, and moreover with a fairly high coefficient. It actually determines the linguistic environment in which a person grew up, and therefore, over time, this feature as a marker is firmly fixed in a person's speech.

By breaking down language into smaller fragments we eventually arrive at sounds. There are many differences here, and they are significant. Therefore, actually, many specifically Ukrainian words are really impossible to pronounce in Russian, and those who did not have the Ukrainian-speaking environment simply do not know how to pronounce, for ex., the same sound " и", but in Ukrainian. Therefore, the analysis of the sound differences of the language is more complicated, but on the other hand, it is more unambiguous and effective, because even when a person speaks in "surzhyk" and uses the words of another language, the pronunciation still remains the same for all words.

Actually, the complexity of the proposed method lies primarily in the fact that it is more difficult to recognize and catch individual sounds in speech than whole words. In addition, in fact, there is little "technical" literature on philology, few ready-made libraries for language identification by individual sounds, or software algorithms for such a task. All the rules of pronunciation of sounds and the differences between sounds are explained by philologists mostly intuitively: for example, if we hear a lot of soft "л", then it is the dialect of Poltava region, if "в" softens to "у", then it is Western Ukraine.

From everything considered above we can form a basic sequence of speech processing steps:

1. preparation of an audio recording, including filtering of extraneous noise;
2. detection and rejection of an emotion, since it speaks about the mental state of the speaker, and not about the peculiarities of speech;
3. transcribing the audio recording into text and analyzing it:
    3.1. determination of the language of the text;
    3.2. identifying words that have the same or similar spelling in both languages;
    3.3. identifying words that are not common for a given language, which speaks of "surzhyk" or dialects;
4. comparing the accent of the speaker in words that are similar in spelling in both languages with the correct accent of these words in one and another language;
5. extracting several fragments or samples of individual sounds (especially distinct in Ukrainian and Russian vowels) from the audio recording, smoothing their sound characteristics and comparing the obtained data with the "reference" characteristics of these sounds in both languages;
6. assumptions about the language environment or the territory of preferred stay based on the algorithm of analysis of the received data according to criteria, taking into account certain weighting coefficients for them;
7. comparison the result of this classification with the classification of the speaker using machine learning methods, errors estimation and assumptions adjusting.

5. General audio recording processing

A wide range of human emotions can be summarized into three main groups: positive, negative and neutral. The first ones are used to indicate happiness, satisfaction, good mood. This is achieved by increasing the tone of the voice, sometimes its volume and rhythm. At the same time, the intonation of phrases goes up smoothly, without jumps.

Negative emotions, on the contrary, express dissatisfaction, unpleasantness and bad feeling. The emotions of this group differ more from each other in the change of sound parameters, since both a person in a state of depression and a person in a state of passion can feel negative emotions, but they will be transmitted differently. However, in general, the main difference from the positive ones is the lack of smoothness in the sound, the gradual transitions between fragments on the audio recording. In addition, for many negative emotions there is a noticeable decrease in pitch [4], while the volume may fluctuate depending on the intensity of the emotion and the context of its use.

Signs of emotions of the first two groups must be identified and rejected, smoothing the recording, so that they do not distort the results of further analysis of the speaker's speech. In other words, the audio recording should be expressively close to neutral emotions that convey the calm, relaxed and confident state of the speaker. In such an audio recording, there will be no jumps in sound characteristics, and the overall smoothness will be preserved.

In order to filter emotions, the work analyzed entire sentences, because most often they contain only one meaningful syntagm, that is, a fragment that is semantically and emotionally indivisible. Therefore, the arrays of values of sound characteristics were presented as generalized functions of lines of polynomial approximation of the third order, because the cubic function showed optimal results during the work: the necessary smoothing of the graphs was provided, but at the same time the largest differences in values were preserved. Then the resulting dependencies were compared with the patterns of changes in sound characteristics for positive and negative emotions [11]. If the emotion was defined as positive or negative, then the record was smoothed with a moving average, otherwise it was left unchanged and assumed to be emotionally neutral.

Continuing the work directly with the sound we will move on to the division of the sound series into small phonetic units. It was mentioned earlier that vowel sounds are the most distinctive in Russian and Ukrainian pronunciation. Therefore, we separate them from the signal fragments. To do this, we divide the audio recording into syllables, because a syllable always contains one and only one vowel sound. You simply cannot divide the recording into intervals of the same duration, because the syllables differ in duration. Therefore, we will use the method of discarding consonants, that is, we will first find syllables after consonants. Then we will select from a piece of audio recording, where

there is only one syllable, a section where the sound characteristics have greater values, which indicates that the voice itself is used to pronounce vowel sounds.

On a vowel sound all parameters increase, including volume, while on a consonant they decrease and become almost zero on sibilants and whistles. This suggests that it is possible to divide the sound track into segments by syllables, using the value of a certain threshold, when we assume that all values below it will correspond to consonant sounds, while above - to vowels. A minus 30-40dB (dBFS) threshold of 40-50ms worked best, but this of course depends on the original volume of the audio recording. Let's demonstrate this graphically on the volume array of some recording compared to the text version of the phrase. Let's divide it into intervals and choose one of the compounds as an example, as shown in Fig. 1.

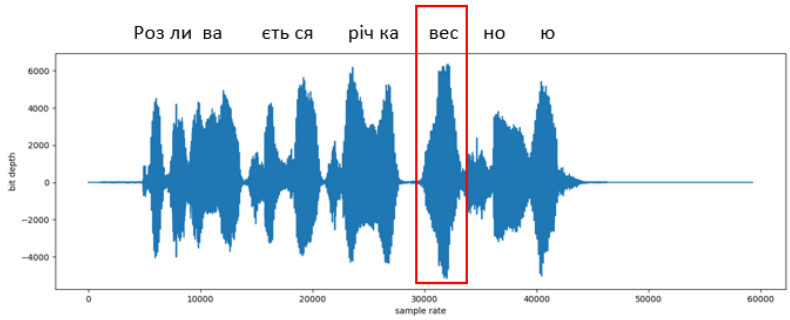

Figure 1: Graphic correspondence of text and voice

Next we work out the syllables separately, similarly to the phrase, but we set a slightly higher volume threshold and shorten the duration of the intervals. Fig. 2 shows that in this way we select the central part of the fragment-syllable, and discard the pieces shaded in gray in the image, because they are consonant sounds. As a result, after dividing the record into syllables and discarding "conditionally voiceless" particles, we get an array consisting exclusively of vowel sounds.

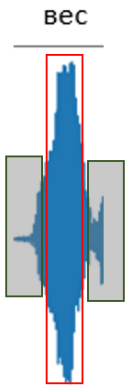

Figure 2: Graphic correspondence of letters (sounds) of a syllable and a voice fragment

6. Comparison of vowel sounds on an audio recording

In order to distinguish sounds from each other, as well as their sounding in the Ukrainian and Russian languages, it is necessary to approximate them according to the 6th order (because for the analysis of sounds, more precise detail is needed, therefore the 3rd order is no longer enough) and generalize the result, and in addition, determine what to take as a standard. To do this, we will consider the system of sounds of the Ukrainian and Russian languages.

So, Ukrainian contains 6 vowel phonemes: а, о, у, е, и, і. Russian - also 6: а, о, у, е, и, ы. Therefore, you can compare sounds in pairs. Let's consider how the sound characteristics change for each of them from the point of acoustics. Oleksandr Ishchenko in the study "Acoustic Classification

of Ukrainian Vowel Sounds" provides graphs of changes in the volume of sounds by formants (overtones that give the sound of a musical instrument or voice a characteristic color - timbre). Similar studies were conducted by S.V. Knyazev and S.K. Pozharytska. Approximating the graphic dependencies found by the authors, we get the following functions of changing the intensity of sound characteristics, which are presented in Fig. 3-8.

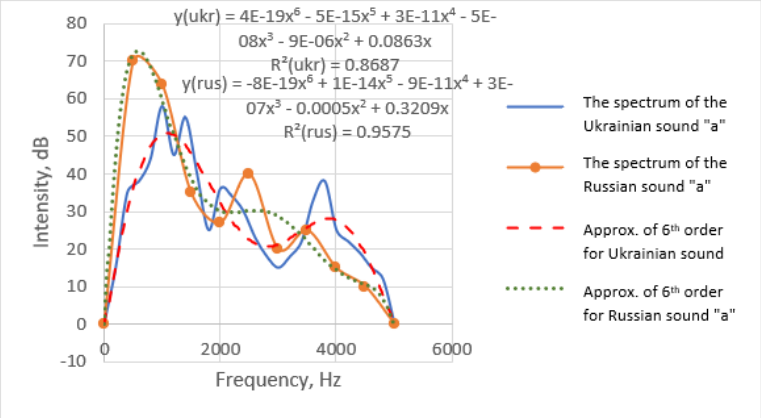

Figure 3: Comparison of the Ukrainian sound 'a' and the Russian sound 'a'

From these graphs, we can see that the Russian pronunciation of the sound 'a' has a much higher intensity in the initial interval.

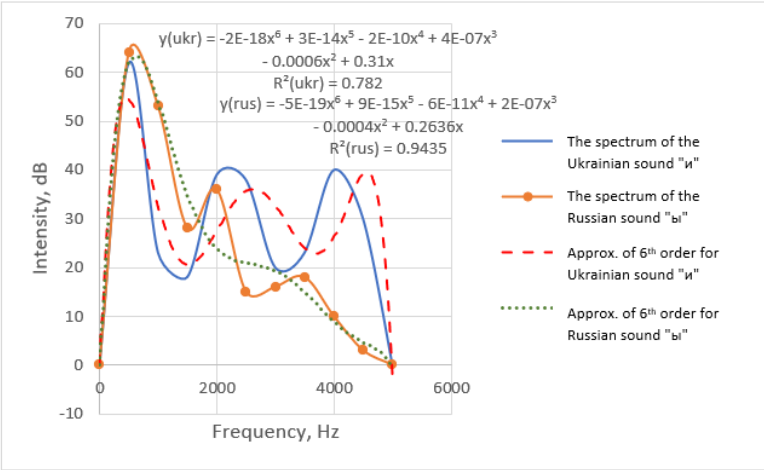

Figure 4: Comparison of the Ukrainian sound 'и' and the Russian sound 'ы'

Here again we see a significant difference - the Ukrainian sound is more uniform throughout the interval, while the Russian counterpart has a pronounced "vowel" part only at the beginning.

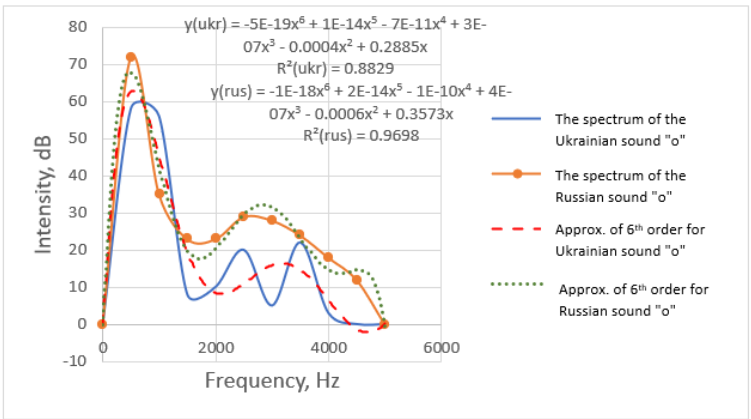

Figure 5: Comparison of the Ukrainian sound 'o' and the Russian sound 'o'

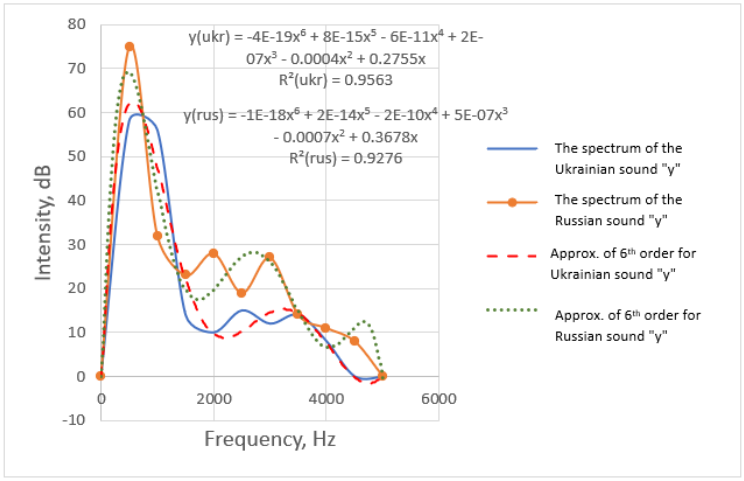

Figure 6: Comparison of the Ukrainian sound 'y' and the Russian sound 'y'

The Ukrainian pronunciation of the sounds "o" and "y" is more truncated at the end, the volume drops more sharply.

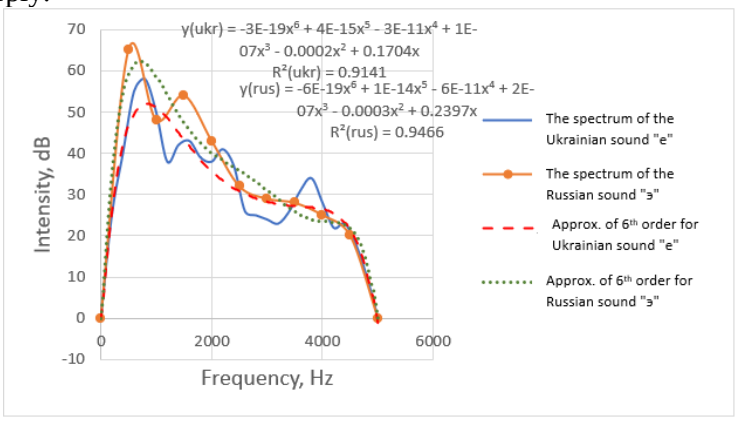

Figure 7: Comparison of the Ukrainian sound 'e' and the Russian sound 'э'

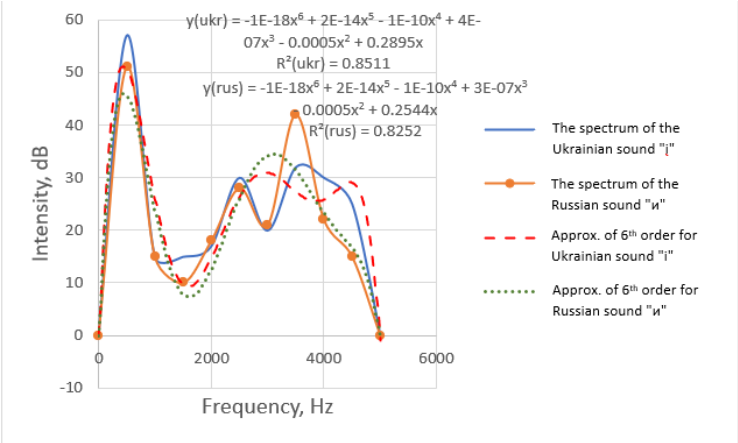

Figure 8: Comparison of the Ukrainian sound 'i' and the Russian sound 'и'

Unlike the previous sounds, the sounds 'e' and 'э' or 'i' and 'и' are practically indistinguishable, so in further analysis these sounds should not have a large weighting factor.

So, at this stage, we can already make assumptions about which pronunciation is closer to the speaker. Another significant difference between Russian and Ukrainian is that it has significantly more sounds "a" and "i", so even the quantitative characteristics of the obtained arrays must be taken into account as one of the criteria for a comprehensive analysis of speech. In the future, in order to identify the features of the speaker's speech, we will move on to text processing and analysis.

7. Lexical text processing

In order to analyze a person's speech from the point of vocabulary, it is necessary to convert the audio recording into text. After that, we divide it into an array of words and compile a dictionary of "word - number of uses" pairs, taking into account not only full-meaning words, but also functional ones, since conjunctions, prepositions and participles also differ in Russian and Ukrainian languages. At the same time, people make mistakes in them very often, using stilted words, assimilating them to how they sound in their native language.

Now let's analyze the resulting dictionary of words and divide it into words that are typical for both Ukrainian and Russian, specifically Ukrainian and specifically Russian, as well as atypical for any of the languages according to the classical dictionary, that are dialects. And here it is important to pay attention not only to quantitative characteristics of groups of words, but to qualitative ones too. A few incorrect conjunctions or dialects are enough to make an assumption about the dominant language of communication of the speaker, even if he is currently speaking another language, in a heap with the obtained results of the analysis of the audio recording.

In addition, it was noted at the beginning that accents can serve as excellent markers for speaker identification. Here, a previously prepared dictionary of full-meaning words was used, which have the same spelling in both languages, but at the same time, the stress falls on different syllables. In this way, we got a collection of "word stress array" pairs, where the stress array is an array of two elements, where the first is the ordinal number of the stressed syllable in Ukrainian, and the second is in Russian. These words were then reversed on the audio recording and divided into syllables according to the algorithm described in the previous sections. The one on which the sound characteristics had a higher intensity was chosen as the accented syllable.

The summary data obtained by the program during all stages of sound and text processing must be collected and certain assumptions about the speaker must be made. For this purpose, a separate module was created, which assigns each of the considered criteria a certain numerical weight value, makes the necessary transformations with the resulting values, multiplies them and thus approximately determines which language is more natural for the speaker - by default, for example, we take that the more final value, the higher the probability that a person mainly uses Ukrainian in daily communication.

However, of course, the results of the assumption fluctuate depending on the weighting factors and the algorithm of matching the criteria as a whole. Therefore, the program is still at the stage of improvement, in particular, the study of the degree of influence of the criteria taken into account on the accuracy of speaker identification.

8. Used software tools

The software algorithm based on the solutions and formalized patterns described above is written in the C# language on the .NET platform. The main classes of the application are:

1. a sound that contains a set of characteristics such as volume, duration, frequency, etc.;
2. a composition that has both general characteristics, including a stress-unstressed mark, and an array of objects of the Sound type;
3. an emotion, containing arrays of coefficients for the main characteristics of the voice;
4. an audio fragment consisting of an array of Composition objects, general characteristics, as well as an Emotion object;
5. and other auxiliary handler classes for all these entities.

C# libraries in the form of Nuget packages, such as NAudio, SoundFingerprinting, SoundTouch.Net, were used for partial remastering of audio recordings, and custom algorithms were used to perform the algorithms themselves with arrays, where Math.NET Numerics library methods were used for basic operations instead of loops.

Some tools from the Lucene.Net and SimMetrics libraries were used to edit the text, create dictionaries of words used by the speaker, and the like.

All other algorithmic operations, step-by-step described in the previous sections, were performed using custom methods.

9. Conclusions

The presented work examined vocabulary and pronunciation as components of human speech, considered the main factors by which a speaker or a group of speakers can be identified in the context of the language environment and the territory of primary residence.

On the basis of the obtained regularities and formalizations, step-by-step speech processing algorithms were created, both in terms of sound and lexical diversity. The result of the work is a software solution that can be used both individually and together with other components as a whole.

In further researches it is planned to determine the extent of the influence of the described criteria on the correctness of assumptions about the speaker's language environment, to increase the accuracy of the program's assessment of the speaker, as well as to investigate the possibility of expanding and universalizing the approach to other languages.

10. References

[1] M.V. Tkachenko, R.M. Fedorenko, Y.V. Kondratenko, I.H. Zotova (2019) Methods of automatic identification of the announcer by voice. Collection of scientific works of the Center for Military and Strategic Studies of the National Defense University of Ukraine named after Ivan Chernyakhovskyi. (3) P. 131-135. (in Ukrainian)

[2] Official website of Nuance Communications inc**.:** https://www.nuance.com/index.html

[3] Official website of Biocatch: https://www.biocatch.com

[4] Y.S.Yamnenko, O.V.Smirnov (2018) Determination of a person's emotional state based on the analysis of voice characteristics. Master's thesis: Department of Industrial Electronics, Faculty of Electronics, NTU " Igor Sikorsky KPI". 124 p. (in Ukrainian)

[5] A.Y. Bagmut (1970) The intonation structure of a simple narrative sentence in Slavic languages. Kyiv: Naukova dumka. 173 p. (in Ukrainian)

[6] A.Y. Bagmut (1980) The structure and functional-semantic aspect of the intonation of a simple narrative sentence in Slavic languages. Academy of Sciences of the USSR Institute of Slavic and Balkan Studies. 48 p. (in Ukrainian)

[7] R.T.Minnigalimov (2015) Analysis and synthesis of Ukrainian speech. Master's thesis: Department of Radio Design and Radio Equipment of the Radio Engineering Faculty, NTU " Igor Sikorsky KPI". 107 p. (in Ukrainian)

[8] L.A. Blyznychenko (1968) Speech intonation. Kyiv: Naukova dumka. 203 p. (in Ukrainian)

[9] I.R. Glyuk (2019) Study of acoustic features of the integral characteristics of the Ukrainian language. Master's thesis: Department of Acoustics and Acoustic Electronics of the Faculty of Electronics, NTU " Igor Sikorsky KPI". 91 p. (in Ukrainian)

[10] S. Y. Yermolenko (1973) Ukrainian literary pronunciation and accent. Dictionary-reference. Academy of Sciences of the Ukrainian SSR Institute of Linguistics named after O. O. Potebny. Kyiv: Naukova dumka. 724 p. (in Ukrainian)

[11] V.L. Shevchenko, Y.S. Lazorenko, O.M. Borovska (2021) Intonational expressiveness of the text at program sounding. Scientific magazine "Problems of programming". (2) P. 76-84. (in Ukrainian)

[12] S.V. Knyazev, S.K. Pozharytska (2005) Modern Russian literary language. Phonetics, graphics, spelling. Moscow: Academic Project. 426 p. (in Russian)