



A New User-Interface Model for Albanian-Speaking Blind and Visually Impaired Science Students

Valbon Ademi¹, Vanko Cabukovski², Lindita Ademi³

¹Faculty of Natural Sciences and Mathematics, State University of Tetovo, (Republic of Macedonia),

²Faculty of Natural Sciences and Mathematics, Ss. Cyril and Methodius University, (Republic of Macedonia), ³Faculty of Philology, State University of Tetovo, (Republic of Macedonia)

¹valbon.usht@gmail.com, ²cabukv@hotmail.com, ³lindita.ademi@unite.edu.mk

Abstract

Being able to read text, find out information and know about the latest news has always been a challenge for those who cannot access the printed version, such as the blind and visually-impaired. The advent of internet and advancements in User Interfaces (UIs) as well as in assistive technologies make this challenge increasingly feasible. A lot of effort has gone into making software applications and web pages more accessible to the blind and visually impaired students. There are few Windows based stand-alone text-to-speech screen reader applications that provide application-specific speech output through script files written specifically for them. All these screen readers usually cover widely speaking languages. Albanian language is not covered with no one of the screen readers.

In this paper a new user-interface model for Albanian speaking blind and visually impaired is presented. The new model is built up on the research results an testing with different screen readers and different groups of blind and visually impaired Albanian speaking students, the screen readers analysis, text-to-speech technologies analysis, specially developed text-to-speech generator from Albanian texts, the mathematical analysis done on general Albanian texts as well as the basic principles of lexicons creations. It has been tested to blind and visually impaired science students. The new user-interface model is a very useful tool for Albanian-speaking blind and visually impaired science students giving them an equal studying opportunities as the other students.

1. Introduction

Intelligent User Interfaces (IUI), or sometimes called an interface agent is a user interface (UI), that includes some aspects of the artificial intelligence. The area of intelligent user interfaces covers various topics and deals with application of artificial intelligence and knowledge-based techniques on questions of man-computer interaction [1] [2] [3].

One of the major benefits from the rapid development of the intelligent user interfaces are the applications known as screen readers, which to a greater extent make the lives of the blind and visually impaired people easier [4][5][6]. The screen readers are text-to-speech systems that transform the text into speech.

The speech synthesis is an artificial production of human speech from written texts. The speech synthesis is the central part of the text-to-speech technology, which today has a wide range of applications with a huge significance of their application in auxiliary technologies and tools for blind and visually impaired people and dyslexia, where the screen readers belong.

Currently there are solutions that enable natural speech generation for various world languages [7][8][9][10]. However, they are not universal solutions which can be used for other languages as well, since the speech generated for other languages is not recognizable and natural [8][9].

Because of this, for each language one should look for a solution that refers to its specifics, always aiming towards a voice creation that will respond to the nature of the language.

In this paper, we describe the basic principles of designing a system for speech synthesis of the Albanian language from written texts.

2. Generating speech from written texts in Albanian language

Currently there are various technologies for converting a written text into speech. Their mutual goal is artificial generating of natural speech that is maximally comprehensive. But, unfortunately, such perfect convertors cannot be found yet.

The last reform in 1972, brought the standard Albanian literature language. There was a linguistic move to "unite" the two dialects of the Albanian language: Gheg and Tosk. While it seems that the official language



is the language of the media and the schools, it is not the language of each speaker of the Albanian language [12].

The Albanian language and its alphabet consist of 36 letters, all of which are in Latin script, except for two letters which are written with diacritical script ç and ë. The number of vowels is seven: a, e, ë, ë, i, o, u, y. Nine letters are considered as compounds, because they represent combinations of two letters: dh, ð, gj, ll, l, nj, rr, sh, j, th, ð, xh, d3 and zh 3.

The verbs can have one or two forms: *Z* active form and/or *Z-hem* passive form. Both forms have a same past participle, and the four tenses of the passive form are made from the same tense as the active form by adding the prefix "u": for example, *lava* (I washed -aorist) *u lava* (I washed myself). For some tenses the prefixes *të* and *do* are used, and for other tenses some other prefixes are used. For example, *laj, të laj, do të laj*, (active form) and *lahem, të lahem, do të lahem* (passive form) with 6 different tenses.

In the Albanian language, the nouns and pronouns change most. A large number of words of masculine gender in singular change into feminine gender in plural. Hence, the gender is not a constant propriety of a noun. The declination position is at the end of the word, but for three pronouns, it is in the middle, for example, *cilido, ciitdo, and cilindo*.

In front of the most adjectives, some nouns and some pronouns there is a particle called article. These articles have declinations: their four forms can be *i, e, të, së*. These declensions vary according to the place of the articulated adjective or articulated noun in nominal syntagm.

From the conducted analysis of the Albanian written texts, it can be noticed that the usage of the letters is not the same. Around 450 most often used words in the Albanian language cover more than 50% of the texts in Albanian. At generating the speech, first are provided the acoustic files of the most often used words, that cover significant percentage of the written texts, while for other words the generating of the speech should be done through single letters.

In a similar way as it was the case with words, it is the case of generating the speech from syllables. The idea is based on the fact that the number of syllables can be smaller and through them to cover all written texts. Theoretically, in the Albanian language there are $36 \times 36 = 1296$ possible diphones [11]. On the basis of the calculation of the frequency of the diphones in the Albanian language only 45 of them can cover around 50 % of the written texts. Since their total number is small, it is reasonable that all diphones should be included in the database of the acoustic files.

The equivalent mass consists of 200 units which are considered to be necessary for the Albanian language, but also, additional units can be added to it. All words that are not in the equivalent mass, will be interpreted letter by letter.

From all these approaches to the speech generating the following main conclusions are drawn:

- Generating speech letter by letter is comprehensive but it is not natural and connected to difficult discontinuities in the speech.
- Generating word with quality speech results – we get comprehensive and natural speech. However, the only disadvantage is that not all the words can be included in the acoustic database.
- The use of diphones as basic units is characterized with the fact that it provides homogenous and stable solution, as any text can be interpreted through a combination of acoustic added files to 892 identified diphones.
- A possibility which offers compromised solutions - considerable number of the most used words are uttered according to the record of the original, while others are comprised of diphones, and several are uttered through special letters.



3. A new model of user interface for Albanian speaking language

In the group of existing intelligent interfaces for accessibility, with certain exceptions, the English language ones dominate. The choice of the option for intelligent interface for blind and visually impaired persons from non-English speaking regions comes to choice of two options – creation of a new product or localization of some existing solution.

The Albanian language is not included in the set of languages for which there is a final solution for this subject. We are dedicated to find a model in which texts written in Albanian language can be transformed into speech.

In this case we need a preliminary language preparation. This preparation covers the contents of the basic databases of a vocabulary or terminology book, including special IT terminology, as well as their vocalization. The speech is a vocalization form of human communication. Every uttered word is created by phonetic combination of limited set of vocal, sound units – vowels and consonants.

The quality of the text conversion into speech is measured by two main parameters: comprehensiveness and naturalness. The comprehensiveness should be done with clarity of the heard speech, and naturalness refers to the similarity of the speech with the artificially generated common speech.

According to the characteristics of the Albanian language, the process of conversion the sentences into words is a process of segmentation and classification of the written form of the sentence. Since the classification and segmentation are mutual here, it is not possible to do it one by one. On the contrary, our approach is first to make *temporary* segmentation in potential written forms called tokens, and then the next step will be to examine each token and solve any ambiguity.

In the phase of pre-processing the text, it is usually first divided into sentences, then divided into tokens which are separated by characters of blank space. The module for pre-processing the text is also responsible for management of non-standard words such as the abbreviations and numbers. In the phase of morphological analysis, the input text is analysed in order to find morphemes in each word.

The next phase is prosodic modelling, which in the linguistic science includes intonation, rhythm and lexic accentuation in the speech. The prosodic characteristics of the speech unit can be grouped in the characteristics of syllable, word, phrase or sentence.

The perceived quality of the synthesized speech is largely determined by the naturalness of the prosody generated during the synthesis.

After finding the optimal solutions for adaptation of the specifics and characteristics of the Albanian language, on the basis of the above said for the way of work on multilingual text into speech system, we continue with the following stages: text normalization, creation of acoustic database and at last, text conversion with the help of applications for speech generating from text written in Albanian language. The whole process is given in Fig. 1.

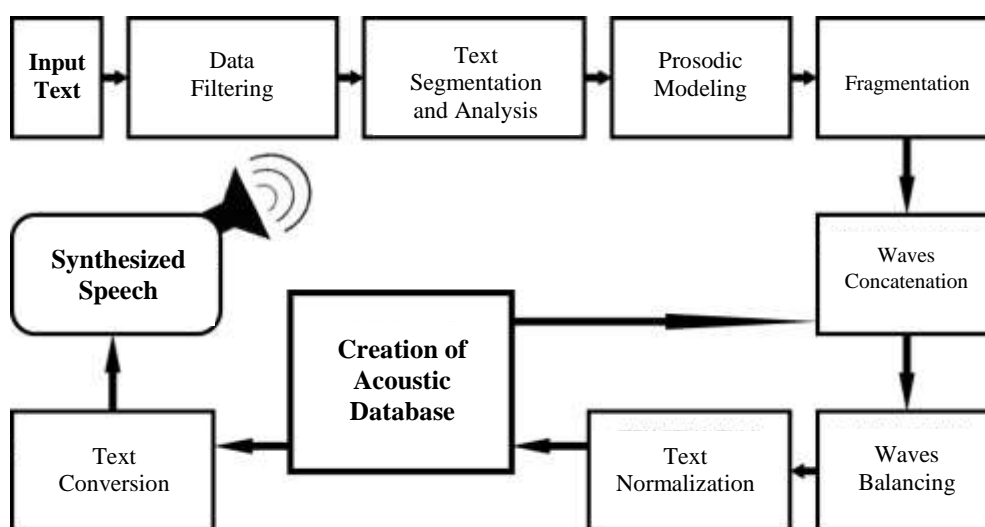


Fig. 1: The user interface model



4. Conclusion

In the group of existing intelligent interfaces for accessibility, with certain exceptions, dominate those in English. The choice of option for intelligent interface for blind people from non-English speaking regions comes down to two options – creating a new product or localization of some existing solution. The proposed new user-interface model for Albanian speaking blind and visually impaired is based on a optimal solutions for adaptation of the specifics and characteristics of the Albanian language for work on multilingual text into speech systems. It has been successfully tested on Albanian speaking blind and visually impaired science students with WebAnywhere screen reader. The speech generated provides homogenous and stable speech quality fully understandable. This new user-interface model is a very useful tool for Albanian-speaking blind and visually impaired science students giving them an equal studying opportunities as the other students.

References

- [1] Galitz O. Wilbert, *The Essential Guide to User Interface Design: An Introduction to GUI Design Principles and Techniques (CourseSmart)*, Wiley Publishing 2007.
- [2] Mayhew J. Deborah, *Principles and Guidelines in Software User Interface Design*, Prentice Hall, 2008.
- [3] Johnson Jeff, *Designing with the Mind in Mind: Simple Guide to Understanding User Interface Design Rules*, Elsevier, 2010.
- [4] Bigham J. Jeffrey P., Lau Tessa, Nichols Jeffrey, TrailBlazer: Enabling blind users to blaze trails through the web. In *Proceedings of the 12th International Conference on Intelligent User Interfaces (IUI 2009)*, Sanibel Island, FL, USA, 2009.
- [5] Asakawa C., What's the web like if you can't see it? In *Proceedings of the 2005 International Cross-Disciplinary Workshop on Web Accessibility (W4A)*. pp. 1–8. 2005.
- [6] Bigham P. Jeffrey, Prince M. Craig, Ladner E. Richard, *WebAnywhere: Enabling a Screen Reading Interface for the Web on Any Computer*, In *Proceedings of the 2008 international cross-disciplinary conference on Web accessibility (W4A)*, pp. 132-133, ACM New York, 2008.
- [7] Feldman R., Sanger J., *The text mining handbook - advanced approach in analyzing unstructured data*, Cambridge Press 2007.
- [8] Gormez Zeliha, Orhan Zeynep, TTTS:Turkish text-to-speech system, 12th WSEAS International Conference on Computers, Heraklion Greece, pp. 977-981, 2008.
- [9] Parssinen Kimmo, *Multilingual text-to-speech system for mobile devices: development and applications*, Tampere University, Tampere 2007.
- [10] Taylor Paul, *Text-to-Speech Synthesis*, Cambridge University Press, 2009.
- [11] Hamiti Mentor, Dika Agni, *Speech Generation for Albanian Written Texts*. In *Proceedings of the ITI 2010, 32nd International Conference on INFORMATION TECHNOLOGY INTERFACES*, pp. 85-90. Cavtat/Dubrovnik, Croatia, 2010.
- [12] Piton Odile, Lagji Klara, Përnaska Remzi, *Electronic Dictionaries and Transducers for Automatic Processing of the Albanian Language*, *Natural Language Processing and Information Systems, Lecture Notes in Computer Science Volume 4592*, pp 407-413, 2007.