DOGS2010

DOGS2010 is the eighth conference offering an overview of current state and research directions in digital signal processing. Traditional conference topics are digital speech, image and biomedical signal processing.
The eighth DOGS will take place at the hotel Norcev at Iriški venac, one of the mountain tops of Fruška gora, from December 16th to 18th, 2010. The previous conferences took place in Novi Sad (1996), on Fruška gora (1998), in Novi Sad and Sremski Karlovci (2000), in the Dunđerski castle near Bečej (2002), in Sombor (2004), Vršac (2006) and Kelebija (2008).
This conference is unique within the boundaries of former Yugoslavia. Among its topics speech technologies traditionally stand out, being a very challenging problem highly dependent on language. Speech technologies help us to preserve our languages through application in various user services, and they are of immense importance to visually or hearing impaired people.

CONFERENCE TOPICS

Digital speech signal processing: speech generation and perception, phonology and phonetics, speech pathology, speech technologies (speech analysis and synthesis, speech recognition, speaker identification and verification), speech coding and transmission, speech cryptography techniques, noise reduction, analog and digital speech processing systems, implementation and applications (communications, computer telephony, etc.), and others.

Digital image processing: image coding and transmission, image analysis and segmentation, linear and non-linear image filtering and restauration, image modelling and representation, digital transformations, movement detection and estimation, implementation and applications (communications, multimedia, robotics, control, etc.), and others.

Digital Biomedical Signal Processing: linear and non-linear processing of 1-D biomedical (cardiovascular, neural and other) signals, decomposition, transformation, biomedical imaging, medical statistics.

CONFERENCE OVERVIEW

- Introductory invited papers
- Papers presenting an overview of current research (reviews of noteworthy papers published between two DOGSes)
- Original, previously unpublished papers
- Practical demonstrations of digital signal processing applications
- Round table discussions

DOGS is organized by the Faculty of Technical Sciences of Novi Sad, in cooperation with the Faculty of Electrical Engineering in Beograd, the Faculty of Electronics in Niš and Provincial Secretariate for Science and Technological Development of Vojvodina, under the auspices of the Ministry of Science and Technological Development of the Republic of Serbia and the IEEE Section of Serbia and Montenegro.

CONFERENCE COMMITTEE

ORGANIZATION COMMITTEE

PhD
Alpineon, Ljubljana

Milan Sečujski, PhD
FTN Novi Sad

OFFICIAL LANGUAGES

The official languages of the conference will be Serbian and English. Papers must be written and abstracts presented in one of the two.

ATTENDANCE FEE

The conference attendance fee is 5000 RSD, and it covers:

- Participation in programme activities
- Cost of proceedings (hard-copy and CD-ROM)
- Extra-curricular activities

Attendance fees should be paid into the bank account that will be announced shortly.

INSTRUCTIONS FOR WRITING PAPERS

Papers should be submitted abridged (1 to 2 pages) and/or unabridged, according to the instructions for authors (in Serbian orEnglish language), in an electronic form to the e-mail address:

secujski@uns.ac.rs

PROCEEDINGS

Accepted papers will be published in the Proceedings, available as a Book and a CD-ROM. Both will be available at the conference.

DEADLINES

31.07.2010. Submission of abstracts
15.10.2010. Submission of unabridged papers (up to 4 pages; invited up to 8 pages)
15.11.2010. Acceptance information

dogs2010

digitalna obrada govora i slike
iriški venac, decembar 2010.

PROGRAM KONFERENCIJE

Četvrtak, 16. decembar

19.00 Registracija učesnika

20.00 Koktel dobrodošlice

Petak, 17. decembar

9.00 Sala A – Sesija A1 (Govor)

A1.1 ROLE OF PROSODY IN HUMAN-COMPUTER INTERACTION – Milana Bojanić, Vlado Delić

A1.2 PROSODIC CONSTITUENTS IN SERBIAN – Maja Marković, Tanja Milićev

A1.3 TRAJANJE GLASOVA I NAJUTICAJNIJI FAKTORI U SRPSKOM JEZIKU – Sandra Sovilj-Nikić

A1.4 ANALIZA TRAJANJA I INTENZITETA ZVUČNIH GLASOVA /dž, ž, r, l/ U TIPIČNOJ I ATIPIČNOJ PRODUKCIJI – Silvana Punišić, Slobodan Jovičić, Zorka Kašić, Slavica Golubović

A1.5 MODELOVANJE IZGOVORA AFRIKATA /c/ – Milan Vojnović, Miško Subotić

A1.6 MODELOVANJE ATIPIČNOG IZGOVORA AFRIKATA /c/ – Milan Vojnović, Silvana Punišić

9.00 Sala B – Sesija B1 (Biomedicinski signali)

B1.1 IDENTIFIKACIJA GOVORNIH MOŽDANIH ZONA FUNKCIONALNOM MAGNETNOM REZONANCOM – Olivera Šveljo, Katarina Koprivšek, Miloš Lučić, Mladen Prvulović, Branimir Reljin, Milka Ćulić

B1.2 ROBUST FEATURE-BASED REGISTRATION FOR CT-MR IMAGES – Nemir Ahmed Al-Azzawi, Wan Ahmed K. Wan Abdullah

B3.4 KORIŠĆENJE GENETSKOG PROGRAMIRANJA ZA DETEKCIJU POPLAVLJENOG POLJOPRIVREDNOG ZEMLJIŠTA – Predrag Lugonja, Nemanja Petrović, Dubravko Ćulibrk, Vladimir Crnojević

B3.5 KLASIFIKACIJA SLIKA ZASNOVANA NA ADABOOST ALGORITMU I STABLIMA ODLUKE SA HOG I LBP OBELEŽJIMA – Marko Panić, Predrag Lugonja, Dragan Letić, Dubravko Ćulibrk, Vladimir Crnojević

B3.6 PREPOZNAVANJE PEŠAKA ZASNOVANO NA HOG I LBP OBELEŽJIMA PRIMENOM RANDOM FOREST ALGORITMA – Dubravko Ćulibrk, Marko Panić, Dragan Letić, Predrag Lugonja, Vladimir Crnojević

B3.7 AUTOMATIZOVANI VIDEO NADZOR SAOBRAĆAJA KORIŠĆENJEM DETEKCIJE I PRAĆENJA POKRETNIH OBJEKATA – Dragan Letić, Branko Brkljač, Predrag Lugonja, Dubravko Ćulibrk, Vladimir Crnojević

B3.8 EYE LOCALIZATION USING CORRELATION FILTERS – Vitomir Štruc, Jerneja Žganec Gros, Nikola Pavešić

B3.9 APPLICATION OF THE PROGRESSIVE WAVELET CORRELATION TO CONTENT-BASED IMAGE RETRIEVING – Igor Stojanović, Ivan Kraljevski, Slavčo Čungurski

18.00 Sala A – Sesija A4 (Govor)

A4.1 KONSTRUKCIJA DEO PO DEO UNIFORMNOG KVANTIZERA I PRIMENA U KODOVANJU GOVORNOG SIGNALA – Zoran Perić, Jelena Nikolić, Aleksandra Jovanović

A4.2 KOMPRESIJA SA GUBICIMA I BEZ GUBITAKA GOVORNOG SIGNALA VISOKOG KVALITETA – Zoran Perić, Milan Savić, Milan Dinčić

A4.3 UTICAJ KARAKTERISTIKA AMBIJENTA NA KVALITET GOVORNOG SIGNALA – Petar Prokić, Slobodan Jovičić

A4.4 UTICAJ NAČINA UPOTREBE MOBILNOG TELEFONA NA FORMANTNE FREKVENCIJE – Nikola Jovanović, Slobodan Jovičić

A4.5 PRIMENA GOVORNIH TEHNOLOGIJA U ADAPTACIJI RAČUNARSKE IGRE LUGRAM ZA SLEPU I SLABOVIDU DECU – Branko Lučić, Nataša Vujnović Sedlar

A4.6 MODELING MACEDONIAN INTONATION FOR TEXT-TO-SPEECH SYNTHESIS – Branislav Gerazov, Zoran Ivanovski, Ružica Bilibajkić

A4.7 TIME ENCODED SIGNAL PROCESSING FOR SPEECH QUALITY ASSESSMENT – Ivan Kraljevski, Igor Stojanović, Slavčo Čungurski, Sime Arsenovski

A4.8 SPEECH SYNTHESIS OF DISSIMILAR LANGUAGES USING THEIR PHONETIC SUPERSET – Slavčo Čungurski, Ivan Kraljevski, Igor Stojanović, Blerta Prevalla

SPEECH SYNTHESIS OF DISSIMILAR LANGUAGES USING THEIR PHONETIC SUPERSET

Slavčo Čungurski[1], Ivan Kraljevski[2], Igor Stojanović[3], Blerta Prevalla[4]

[1]Faculty of ICT, FON University, Skopje, Macedonia; e-mail: chungurski@fon.edu.mk
[2]Faculty of ICT, FON University, Skopje, Macedonia; e-mail: ivan.kraljevski@fon.edu.mk
[3]Faculty of Computer Sciences, Goce Delchev University, Stip, Macedonia; e-mail: igor.stojanovik@ugd.edu.mk
[4]Faculty of ICT, FON University, Skopje, Macedonia; e-mail: blerta.prevalla@fon.edu.mk

ABSTRACT

The two most common languages in Macedonia are Macedonian and Albanian language. The main idea of this paper is to upgrade the phonetic inventory of Macedonian language to its superset – phonetic inventory of the Albanian language, and use it for speech synthesis in both languages.

The structure of general speech synthesis system for the Macedonian language is used [1] for its upgrade to speech synthesis system, which will produce Albanian language also. The modification of the system is described and there are details about the specific features of the languages important for the speech synthesis, like sound and vowel system and prosodic elements.

Because concatenative speech synthesizers are based on a collection of speech segments, the creation of common speech corpus for these languages is described.

1. INTRODUCTION

The systems which synthesize speech with connection of previously recorded speech segments take significant place among the systems for text to speech conversion (TTS systems). These TTS systems are called concatenative speech synthesizers. These systems are simple and they do not need deep knowledge of phonetic transitions and co-articulation effects, which is case with other kinds of speech synthesizers based on rules defined by linguists.

Concatenative speech synthesizers require setting of serious task for definition and recording of speech and its processing for extracting convenient speech segments. Consequently, this paper presents proper definition and development of common speech corpus for Macedonian and Albanian language.

2. GENERAL STRUCTURE OF SPEECH SYNTHESIZER AND ITS UPGRADE

The general functional organization of speech synthesis system is shown in Figure 1, and it consists of two main parts:

- Natural Language Processing (NLP) module, which gets text as input, makes analysis of the text, create its transcription into phones and recognize the prosodic elements of the input. The output of this module is symbolic information for the phones and the prosody for the input text.

- Digital Signal Processing (DSP) module, which gets the symbolic information for phones and prosody from the NLP module and after certain processing of the input information, gives synthetic speech as output.

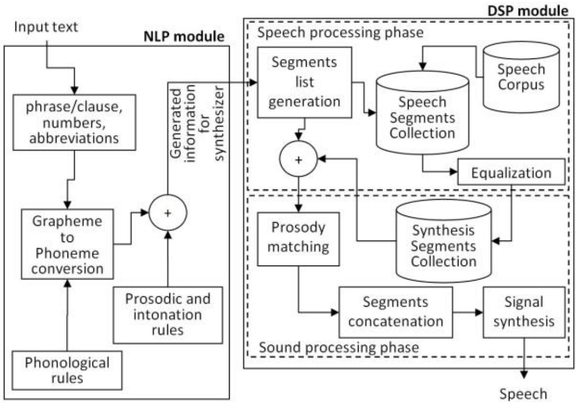

Figure 1: Structure of general TTS system

This structure is used for development of TTS-MK speech synthesizer for Macedonian language [1]. This synthesizer is evaluated and it gives good results from the aspect of its usability [3]. The idea of this paper is to use the development of its DSP module and modified instance of its NLP module for some another language. The chosen language is the standard Albanian language because it is widely present in Macedonia and its neighbor countries.

The modification of the general TTS system is shown in Figure 2 and it is consisted in upgrade of the system with additional NLP module, upgraded speech corpus and introducing language recognizer.

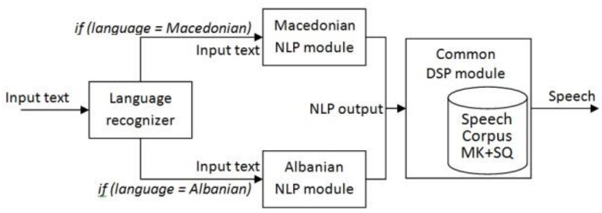

Figure 2: Modification of general TTS system

2.1. Language Recognizer

Language recognizer is a module introduced in this paper for recognition of the language of the input text. The input text can be consisted of several sentences, and these sentences can be in different languages.

This module creates a queue of sentences and identifies the language for each element in the queue. Then it passes the queue to the appropriate NLP module.

Because this paper represents the startup of this project, this module opens a new field for research. In this phase, the main identification of Macedonian and Albanian language is based on a writing system for the input text. If the writing system is Cyrillic then the language is identified as Macedonian, otherwise it is identified as Albanian.

But, the usage of the Latin letters in Macedonian language is not very rare, so the mentioned research in this field is more than necessary.

2.2. NLP Module

This module consists of three main parts: text analyzer, grapheme-to-phoneme unit and prosodic generator.

Text analyzer is made of:

- Pre-processing unit where input sentences are transformed into word arrays. This unit also identifies the numbers and abbreviations and transforms them into words.

- Module for morphological analysis which executes morphological analysis of the text for recognition of the affixes (prefixes and suffixes) which are added to the basic forms of the words. This module also confirms the correct accent of the words.

This module of TTS systems for the Macedonian language seems very complicated because in the Macedonian language, in opposite of other languages, a plural of the words, as well as words with an articles, adds whole syllables at their end. This can look very difficult and it points to some complications in morphological analysis of the text, but thanks to antepenultimate word stress in Macedonian language and the fact that this analysis is strictly for speech synthesis, this procedure is very simple. This is supported with the fact from [5], that in the Macedonian literature language there is no reduction of the vocals, which means that the vowels sounds very similar in their accented and unaccented form. As a result from this fact, this considerably simplifies the whole speech synthesis process for Macedonian language, because the vowels have only one sound variant.

Antepenultimate word stress means that the accent falls on the third last syllable in words with three or more syllables, and on the first or only syllable in other words. Of course (as in [6]), there are some exceptions in cases of accent aggregates (for example: Kiselà voda, Gazì-baba), loanwords (for example: ateljè, genèza) and words with specific ends (for example: ìzam, ìst – alpinizam, alpinist), but there is a solution for these cases. This solution implements an exception dictionary for accentuation rule.

The accent in the Albanian language is quite different from the one in the Macedonian language. It may fall on different syllables, so there are ultimate, penultimate and antepenultimate word stresses. In the case of ultimate word stress the accent falls on the first syllable from the end (for example: lirí, barí, atdhé), but they are rare and their accent comes out of the context, without the need for any particular sign. This group are includes the words with Turkish origin. This can be solved with a dictionary. The penultimate word stress is when the accent falls on the second syllable from the end (for example: dórë, bímë, shkóllë) which is the most often. Antepenultimate word stress is when the accent falls on the third syllable from the end (for example: flútura, kúmbulla, pópulli, plúhuri).

There are also a limited number of words in the Albanian language which have the same phonetic form (identical) and are distinguished only by the accent. In these cases, to avoid misunderstandings, signs for the accent can be also used, for example: bári (shepherd) - barí (grass), minístri (ministry) - ministrí (minister), stóli (chair) - stolí (jewelry).

- Grapheme-to-phoneme unit. One of the main common features of both languages is their simple rule for grapheme to phoneme transcription. It means that in Macedonian and Albanian language, each letter means one phone. So, the phonetization task for both languages is reduced to trivial checks and solutions based on dictionaries and morphophonemic rules, like in English or French language for example, are not necessary. This unit also makes contextual checks of each phoneme, which means phoneme appearance in context with others. This is convenient because concatenation of longer speech segments, if available in the speech corpus, will produce more natural intonation.

The task of the prosodic generator is to supply naturalness to the synthesized speech. Apart from the fact that synthesized speech with natural intonation is more pleasant for listening, it is easier for understanding. Also, during a speech, which is practically uninterrupted, the listener can easily recognize the bounds between the words. The problem of the synthesizers without prosodic elements is a problem that occurs to the listener during the thought flow from the one who transmits the message, because he has difficulties to detect the end of one and the beginning of other word, as shown in [8]. For speech synthesizer it is very important to cover prosodic elements and proper intonation.

2.3. DSP Module

The execution in DSP module is divided into two phases: speech processing and sound processing.

Speech processing is one of the most important phases in TTS synthesis in general. It means that in this phase it is mandatory to make definition and record a universal speech corpus for Macedonian and Albanian language, as well as processing of the speech segments. In this phase, the symbolic information for the phones and the prosody for the input text are applied on the recorded speech corpus. The segments from the corpus and their concatenation order are established also in this phase.

The other phases of the DSP module are standard for general speech synthesis engines and they are discussed in details in [4].

3. DEFINITION OF THE SPEECH CORPUS

Several continuous steps are needed for definition and creation of the speech corpus for this TTS system:

- Selection of speech segment types needed for the synthesis in both languages
- Definition of the set of the phonemes which covers all sound occurrences in both languages
- Selection of the set of segments, which covers the whole phoneme set from the previous step
- Text selection which will cover whole speech segment set from the previous step

3.1. Speech segment types

The speech segments should satisfy several basic conditions. Segments should be fluently connectable and in the same time their number should be kept as low as possible. On the other side, longer segments are increasing the density of concatenation points, which brings to better quality of the synthetic speech. This is opposite to the constraints of speech generating time, as shown in [1], in situations when the speech corpus is big, which brings to longer searches for appropriate segment in the corpus.

Anyway, diphones are the most used basic segments for speech synthesis. They are very appropriate for concatenation because they always start and end in areas where the wave shapes of the phonemes are phonologically independent of the environment. Diphones are not too numerous and they have roles in most phonetic transitions, but frequent occurrence of connections of basic segments leads to problems with fluency of the connections and decreasing the quality of the synthesized speech. These occurrences represent the biggest problem of speech synthesizers based on diphone database. Other problem with this kind of synthesis is reproduction of co-articulations, which appear in the speech and often have an influence on the phoneme. The solution for this problem is based on implementation of disyllables which start in the middle of one vowel and end in the middle of the next vowel. Such segments are easier for extraction from the speech, as well as their concatenation, but their number is much bigger than the number of diphones.

These facts lead to the best solution for generating a speech corpus for this system, which will include mostly diphones and include certain disyllables that appear very often in the languages themselves. In the frame of this research a text database, with whole printed media published in Macedonia in period of one year (as in [2]), is used for statistics on occurrences of diphones and disyllables in Macedonian and Albanian language.

3.2. Phoneme set definition

In the phase of selection of the phonemes, which will be available in the corpus, the most important constraint is that the phoneme set should comprise all possible speech occurrences in both languages. The phonologic inventory of the Albanian language has characteristics of a superset of the phonologic inventory of the Macedonian language. This means that speech corpus of TTS-MK could be upgraded to reach the speech corpus needed for speech synthesis in the Albanian language also.

Therefore, as in [1] the phoneme set of 33 phonemes for the Macedonian language is used with its 6 vowels and 27 consonants, and additional non-Macedonian consonants *dh*, *r* and *th*, and vowel *y*.

The result is a phonologic inventory of the Albanian language needed for speech synthesis, consisted of 7 vowels and 30 consonants.

The 7 vowels, shown in Table 1, include the 5 standard vowels from Macedonian alphabet, a, e, i, o, u, plus the schwa /ə/ semivowel which appears when the consonant r takes the role of vowel (for example in the word brza). The schwa semivowel is a standard phoneme in the Albanian language, so the only upgrade of vowels set is with the vowel *y* /y/ which is not present in the Macedonian language.

	Front rounded	Front unrounded	Central	Back
Close	y /y/	и /i/		у /u/
Mid		е /ɛ/	ъ /ə/	о /ɔ/
Open			а /a/	

Table 1. Vowels subset of diphones by IPA

The 30 consonants include the 26 standard consonants from the Macedonian alphabet. The additional consonant in the phoneme set is the allophone *ŋ* of the consonant n, thanks to its specifics. It is so specific, for a difference of the allophones of the other consonants, because the consonant *n* changes a lot in its articulation domain, when it appears before the consonants *k* or *g*. Therefore, this allophone also appears in the phoneme sets for other languages.

The upgrade of the consonants set is with the three Albanian consonants *dh* /ð/, *r* /ɾ / and *th* /θ/.

The Table 2 shows all the consonants for the Macedonian language with the Albanian upgrade, including the allophone /ŋ/ of the consonant n, divided in categories by International Phonetic Association and with International Phonetic Alphabet (IPA) symbols.

	Bilabial		Labio-Dental		Dental		Alveolar		Post-Alveolar		Palatal		Velar	
Nasal	м /m/						н /n/				њ /ŋ/		ŋ /ŋ/	
Plosive	п /p/	б /b/					т /t/	д /d/			ќ /c/	ѓ /ɟ/	к /k/	г /g/
Affricate							ц /ts/	ѕ /dz/	ч /tʃ/	џ /dʒ/				
Fricative			ф /f/	в /v/	th /θ/	dh /ð/	с /s/	з /z/	ш /ʃ/	ж /ʒ/			х /x/	
Approximant											j /j/			
Trill							р /r/							
Flap							r /ɾ/							
Lateral							љ /ɫ/	л /l/						

Table 2. Consonant subset of diphones by IPA

3.3. Set of speech segments

It is a fact that quality speech synthesizer should produce speech on any input text, meaningful or not. To achieve this, the best way is to include all theoretically possible diphones in both languages in the speech corpus, because these diphones can appear on the boundaries of two neighboring words in a sentence.

Disyllables which are good candidates for membership in the speech corpus are those which contain one consonant between two vowels and which will meet problems in concatenation of the two diphones which creates the disyllable. Also, the corpus should contain disyllables which cover more frequent groups of two or more consonants, which will not lade the system and which will improve the quality.

It is also important for the corpus to contain diphones and disyllables, which contain beginnings and ends of isolated speech units. These speech segments contain transitions from silence to phone and vice versa. This leads to adding an additional element to the phoneme set, while generating the diphone database. This element is silence (marked as "_" in this system). The transition from silence to silence is irrelevant, so the final number of theoretically possible diphones for this system is 38*38-1=1443.

According to the fact that the number of all possible diphones and disyllables in any language is really big and according to the upgradeability of the speech corpus, the mentioned text corpus from [2] is used for revealing the most used diphones and disyllables. This led to definition of the order for storing speech segments in the corpus and its continuous upgrades.

3.4. Text corpus for speech segments extraction

For identification of the text (words) that will be used for speech segments extraction, selection of existing Macedonian and Albanian words, which contain all possible diphones, is made, as it is shown in [1]. This selection is made in several phases.

First, the text corpus from [2] is used again in automatic selection of words, which contain all possible diphones. For each diphone there were multiple words-candidates, which contain the certain phoneme in the middle of the word, if possible. Since the number of all theoretically possible diphones is not very big, the choice of the word from the candidates for each diphone is made manually.

Of course, there were 273 diphones for which an appropriate existing word was not found. This led to the second phase of text selection. This phase comprises generation of logatoms – meaningless words of type abebibobuba, ugiga etc. These logatoms are used as words that contain certain diphones.

Finally, for extraction of diphones of type silence-phone and phone-silence, certain words were generated. These words were words which start and end with the same phone, like bab, viv etc, because each of these words is used for extraction of two diphones, starting one and ending one, as in [7].

In the first phase of the speech corpus of Macedonian language it is considered to include a small set of disyllables, but as it is mentioned above, the corpus is upgradeable and the order of disyllables that should be found in the corpus is known. So, further improvements of this synthesizer should be mainly in upgrading of the speech corpus with disyllables.

For the launching version of this system, the text corpus is ready for recording.

The text corpus for this recording session consists of exactly 543 existing Macedonian and Albanian words and 282 meaningless words for every theoretically possible diphone in the set.

Additional steps in upgrading the diphone database to speech corpus were performed with additional non-elementary speech segments. The top 180 from the most used disyllables in Macedonian language, and the top 143 from the most used disyllables in Albanian language, with at least 1000 occurrences in the text corpus were found. Their phonetic intersection counted 47, so (180+143)-47=276 is the total number of included disyllables in the corpus.

4. CONCLUSION

This paper shows a research in the field of speech synthesis for the Macedonian and the Albanian language. Elements from the general structure of TTS systems that apply to the Macedonian language were implemented and efforts for its upgrade to its phonetic superset – the Albanian language are shown.

With definition of the text that will be used for extraction of speech segments, a fulfillment of the main prerequisite for the key phase of the project, speech recording, is achieved.

Further developments in this field will be in usage of the speech corpus in the next phases of speech synthesis and continuous improvements of the speech corpus itself.

5. REFERENCES

[1] Chungurski S., Kraljevski I., Mihajlov D., Arsenovski S., Concatenative Speech Synthesizers and Speech Corpus for Macedonian Language, 30th International Conference ITI Cavtat/Dubrovnik, Croatia, June 23-26, 2008, p.669-674

[2] Chungurski S., Kraljevski I., Arsenovski S., Samardziska S., Information system for newspaper clipping, Infoteh, Jahorina, BiH 2007

[3] Chungurski S. Kraljevski I., Mihajlov D., Arsenovski S., Evaluation of TTS-MK system for speech synthesis in Macedonian language, IE3-3, ETAI 2009, Ohrid, Macedonia, 26-29 September, 2009

[4] Dutoit T., High Quality TTS Synthesis of the French Language, Ph. D. dissertation, Facultè Politechnique de Mons, France 1993

[5] Koneski B., Gramatika na makedonskiot literaturen jazik, Kultura, Skopje, 1987

[6] Panoska R., Sovremen makedonski jazik, Univerzitet Sv. Kiril i Metodij, Skopje, 1994

[7] Secujski M, Obradovic R, Pekar D, Jovanov Lj, Delic V, Alfanum system za sintezu govora na osnovu teksta na srpskom jeziku, DOGS, Becej, Srbija 2002