CODE-SWITCHING SPEECH RECOGNITION FOR CLOSELY RELATED LANGUAGES

Tetyana Lyudovyk, Valeriy Pylypenko

International Research/Training Center for Information Technologies and Systems, Kyiv, Ukraine
tetyana.lyudovyk@gmail.com, valeriy.pylypenko@gmail.com

ABSTRACT

This work presents an approach to recognition of multi-speaker conversational speech with code-switching between Ukrainian and Russian languages. Both inter-sentential and intra-sentential code-switching is handled. The approach takes into account peculiarities of phonetic systems of the closely related Russian and Ukrainian languages. A cross-lingual LVCSR system is developed. The acoustic model and pronunciation lexicon are based on Ukrainian phone set. Modeling of pronunciation variation in lexicons helps to cope not only with code-switching speech but also with accented speech. Results of code-switching speech recognition are presented. The approach is suitable especially in cases of intra-sentential code-switching where language identification is problematic.

Index Terms— mixed speech, bilingual speech, code-switching, Ukrainian, Russian

1. INTRODUCTION

Code-switching means alternation of different languages within or between sentences. In countries where more than one language is spoken by a significant number of people the code-switching speech is a common phenomenon.

In Ukraine the majority of population is bilingual. Code-switching between Ukrainian and Russian languages is a frequent event penetrating the broadcast media, meetings, parliamentary debates, court hearings etc. This complicates the automatic speech recognition, in particular the automatic subtitling and the automatic transcription of audio archives.

The following issues should be accounted for:

- Ukrainian and Russian are closely related languages;
- Russian language spoken in Ukraine differs from Russian language spoken in Russia;
- code-switching speech is often speech with an accent.

This work is dedicated to the recognition of code-switching conversational Ukrainian-Russian speech. Inter-sentential and intra-sentential code-switching is handled. It is not known in advance which language is spoken. This situation is typical for recognition of dialogues and interviews.

2. RELATED WORK

The problem of code-switching is relevant for Hong Kong [1, 2], Singapore [3, 4], Taiwan [5], and India [6, 7]. Language pairs are English/Chinese [1-4] and English/Hindi [6, 7]. In [5] the code-switching between Mandarin and Taiwanese dialects of Chinese is investigated.

For code-switching speech recognition a multi-pass and a single-pass approaches have been proposed.

The multi-pass approach consists in finding boundaries of monolingual fragments in multilingual speech, providing language identification (LID) [8], and using an appropriate language monolingual system to recognize monolingual fragments. This approach is highly dependent on the accuracy of finding the boundaries of monolingual fragments, and on the accuracy of language identification.

In case of closely related languages with similar phonetics LID becomes difficult. The pair Ukrainian/Russian is a great challenge [9]. In case of intra-sentential Ukrainian/Russian code-switching LID becomes very problematic.

Within a single-pass approach, a bilingual recognition system is used with one bilingual acoustic model (AM), one bilingual language model (LM) and one bilingual pronunciation lexicon. This approach does not imply the LID.

An AM in a single-pass approach may be built by training an AM on bilingual data [3, 4, 5]; using one of the monolingual AMs [6, 7]; pooling the existing monolingual AMs as they are; mixing monolingual AMs by sharing phonemes belonging to both languages.

A bilingual LM for recognizing sentences spoken in more than one language may be trained using the pooled text data [10].

A pronunciation lexicon accounting for accented speech may be created by extending a canonical lexicon with accented pronunciations [11].

The recognition of Russian speech spoken in Russia is addressed in [12-17]. The reported results are: 18.3% WER for Russian broadcast news [12], 50.7% WER for Russian

conversational telephone speech [13], 45.75% WER for colloquial, highly emotional and heavily accented speech [14]. As far as we know, the recognition of Russian speech spoken in Ukraine has not been investigated yet.

In Ukraine the research aimed at the LVCSR recognition of Ukrainian speech began in 2007 on the basis of the Ukrainian Parliamentary Speech Corpus [18] and was continued on the basis of the Ukrainian Broadcast Speech Corpus [19, 20] which contains both read and spontaneous speech. The achieved performance is 27% WER. German researchers report 11% WER for read Ukrainian speech when OOV is 0.53% [21].

3. PECULIARITIES OF UKRAINIAN AND RUSSIAN PHONETICS

Ukrainian and Russian languages are closely related: both are members of the East Slavic group, and both of them use the Cyrillic alphabet. Ukrainian and Russian phonetics are similar. Although Cyrillic notation is more appropriate for phoneme sets of these languages, we follow the SAMPA-like notation [22-24] in this work.

The main phonetic features of Russian language are described in [13, 15-17]. Here we will underline phonetic peculiarities of Russian and Ukrainian languages.

Following the traditional phonology, there are 6 vowels and 32 consonants in Ukrainian, while there are 6 vowels and 36 consonants in Russian. At the same time in practice, namely in speech recognition and speech synthesis, a wider range of phonemes is used [13, 21, 25, 26].

The speech synthesis system for Ukrainian language [26] uses 58 phonemes. Stressed and unstressed vowels are regarded as different phonemes. Stressed vowels are denoted by uppercase letters, while unstressed vowels are denoted by lowercase letters.

All consonants but /j/ are divided into pairs formed by a palatalized phoneme and a corresponding non-palatalized phoneme.

In the speech recognition system for the Ukrainian [20] 54 phonemes are used (12 vowels and 42 consonants). Three infrequent palatalized consonants are neglected: /g'/, /dz'/, and /dZ'/.

Differentiating between stressed and unstressed phonemes is observed also in Russian speech synthesis [25]. As for speech recognition, the number of Russian phonemes is still under investigation ranging from 43 to 55 [13, 16].

The majority of phonemes in the two languages are phonetically close and are labeled with the same symbols. This concerns half of the vowels. The other vowels sound similar but are labeled with different symbols. All symbols of Russian consonants are present in the Ukrainian phoneme set.

On the contrary, 6 of Ukrainian consonants used for speech recognition are unique. Phonemes /h/, /dz/, /dZ/, /Z'/, /ts'/, and /tS/ are absent in Russian.

Ukrainian voiced pharyngeal fricative /h/ and Russian voiced velar plosive /g/ are written by the same letter and Russian /g/ is often replaced by Ukrainian /h/ in Russian speech in Ukraine. This replacement makes Russian speech appear to have a Ukrainian accent.

The second difference between Ukrainian and Russian phonetics lies in number of vowel reduction levels. There are two vowel reduction levels in Russian and no vowel reduction in Ukrainian. When Ukrainians speak Russian, they do reduce unstressed vowels, but not as markedly as Russians speaking their mother language.

The third difference is related to devoicing of consonants in some positions present in Russian and absent in Ukrainian. In fact, in real speech this devoicing is common for both languages.

As for phonotactics, the main difference concerns the frequency of the stressed mid front vowel /E/ in a position after a palatalized consonant, e.g. Russian /l' E s/ (*forest*). This vowel in this position is rather rare in Ukrainian but not in Russian.

4. PROPOSED APPROACH

The goal of this work is to use mainly the phonetic knowledge to build a system capable of recognizing code-switching Ukrainian-Russian speech without significant degradation in performance of monolingual speech recognition.

Unlike the closely related Mandarin/Taiwanese pair [5], where the difference exists only in spoken form but not in written one, in our case the orthography is mainly different in both languages while phonetically Russian and Ukrainian are close. Our main hypothesis is that people speaking Ukrainian and Russian in Ukraine use the same set of phonemes. As the Ukrainian phonetics includes all Russian phonemes it is possible to use an AM based only on Ukrainian phone set (UKR_AM) to recognize Russian and code-switching speech. Further, we assume that based only on the main Russian/Ukrainian phonotactic difference it is possible to extend the Ukrainian phone set by a single Russian stressed vowel /E/ preceded by a palatalized consonant. This vowel is denoted as /'E/. The corresponding AM is denoted as UKR+1_AM. As a reference a combined phone set consisting of 55 Ukrainian and 50 Russian phonemes is investigated (UKR+RUS_AM).

This phonetically based approach to AM creation requires careful development of a bilingual pronunciation lexicon. Three variants of the lexicon are created to model the pronunciation of Ukrainian and Russian speech as well as Ukrainian speech with a Russian accent and vice versa.

A bilingual LM is trained on a mix of Ukrainian and Russian texts to model inter-sentential and intra-sentential code switching.

5. SPEECH DATA

5.1. Ukrainian Broadcast Speech Corpus

In our study we used speech data of the Ukrainian Broadcast Speech Corpus (AKUEM) [19]. This corpus consists of broadcast recordings in Ukrainian and Russian. It represents mainly broadcast conversational speech. More than 4000 speakers are presented. There are 160 speakers who speak both Ukrainian and Russian. Speech data were transcribed manually. Such events as background noise and music, overlapping speech, breathing, laughing, cough, hesitations, mispronunciations, self-corrections etc. were marked. Recordings were divided into speech segments approximately corresponding to phrases and lasting up to 15 seconds. Some quantitative characteristics of the corpus are shown in Table 1.

Table 1. Characteristics of the Ukrainian Broadcast Speech Corpus.

Speech Corpus Characteristics	Ukrainian speech	Russian speech
Speech length (in hours)	116	190
Total number of word-forms	962 504	1 721 606
Number of unique word-forms	69 500	83 500
Number of speakers	1 723	2 781

5.2. Test Data

Test data consisted in speech material from AKUEM corpus. 12 bilingual speakers were chosen. They are public people, mainly politicians. Test data were selected from 38 talk-shows. The total duration of test speech was 4.7 hours (2.4 hours of Ukrainian and 2.3 hours of Russian speech). In total, the test set consisted of 18,000 Ukrainian and 19,000 Russian words. Recordings with background noise and speech errors (mispronunciations etc.) were not excluded. Hesitations and breath made up 2.1% and 3.5% of all test data.

There were 102 code-switchings between Ukrainian and Russian speech in test data, 24 of them were intra-sentential switchings. Here is an example of intra-sentential switching from Russian to Ukrainian: “Мы продолжим этот разговор | після невеликої перерви” (“*We'll continue this conversation* (in Russian) | *after a short break* (in Ukrainian)).

6. UKRAINIAN SPEECH RECOGNITION SYSTEM

The study is based on a LVCSR system for Ukrainian language [20] which is capable of recognizing broadcast conversational speech. The system is developed using HTK toolkit [27].

The AM is trained on Ukrainian speech from AKUEM corpus. In addition to 55 Ukrainian phonemes, special models for pause, hesitations, breathing, smacking etc. are used.

LM is a bigram model trained on all texts of AKUEM corpus (20 MB) and texts from the Internet (400 MB).

The pronunciation lexicon contains 116k word-forms. On average, one word-form has 1.5 phoneme transcriptions reflecting the canonical and the spontaneous pronunciation. Frequent word-forms, including numbers and names of persons, count up to 10 phoneme transcriptions. Canonical transcriptions were generated automatically by grapheme-to-phoneme rules. Spontaneous transcriptions were generated from canonical ones by additional rules elaborated on the base of phonetic knowledge and recognition experiments.

Recognition performance of this system on this study test set is about of 77% word accuracy for Ukrainian speech segments.

7. EXPERIMENTAL SYSTEMS

Two monolingual and three bilingual systems were designed based on speech and texts from AKUEM corpus and HTK decoder. Bilingual systems differ in acoustic models and pronunciation lexicons but share the same language model.

7.1. Phone sets

Three phone sets were used in this work for recognition of code-switching speech:

- only all Ukrainian phonemes (12 vowels and 43 consonants);
- all Ukrainian phonemes + Russian /’E/;
- all Ukrainian phonemes + all Russian phonemes (105 phonemes).

7.2. Grapheme-to-phoneme converters

For creating Russian and Ukrainian lexicons, automatic grapheme-to-phoneme (g2p) rule-based converters are usually used [13, 15].

In this study we used one g2p converter for Ukrainian language [26] and two different g2p converters for Russian language.

The first g2p converter for Russian is based on g2p rules reflecting pronunciation inherent to Russians living in Russia [25] (no-accent-style converter). This converter uses 50 Russian phonemes.

The second converter (accent-style) for Russian lexicon is based on Ukrainian phonemes with determined correspondence between Ukrainian and Russian phone sets. This converter developed by us for Russian Text-To-Speech

synthesis reflects to some extent the pronunciation by Ukrainian people while they speak Russian.

Thus, in no-accent-style converter three levels of vowel reduction are taken into account: zero level (no reduction of stressed vowels), first level (e.g. first pre-stressed vowel in a word) and second level (e.g. second pre-stressed vowel counting from right to left). Reduced vowels of the first level are denoted by the following symbols: /a/, /o/, /e/, /u/, /i/, /y/. Reduced vowels of the second level are denoted by /@/ (for reduced vowels preceded by non-palatalized consonants) and /$/ (for reduced vowels preceded by palatalized consonants).

In accent-style converter only two levels of reduction are incorporated because vowel reduction is not as strong in Ukrainian as in Russian. The second difference between two converters concerns the phoneme /g/. Accent-style converter generates two variants of phoneme transcription for words containing this consonant: one variant with a plosive /g/ and one with a fricative /h/. Other g2p rules of both converters are identical. Table 2 shows examples of phoneme transcriptions of Russian word-forms generated by two g2p converters.

Table 2. Examples of word-forms and corresponding phoneme transcriptions generated by different grapheme-to-phoneme converters for Russian language.

Examples of word-forms	**Phoneme transcriptions generated by unaccented converter**	**Phoneme transcriptions generated by accented converter**
думаю	d U m @ j u	d U m a j u
испуганный	i s p U g @ n @ j	i s p U g a n y j i s p U h a n y j
получается	p @ l u tS' A j $ ts a	p a l u tS' A j e ts a

7.3. Bilingual pronunciation lexicons

Three pronunciation lexicons for code-switching speech recognition were created based on three phone sets. They are named: UKR_LEX, UKR+1_LEX, and UKR+RUS_LEX. Ukrainian word-form transcriptions are the same in three lexicons. Russian word-forms are transcribed differently. Russian phoneme transcriptions are generated by:

- accent-style converter based on Ukrainian phonemes;
- accent-style converter based on Ukrainian phonemes + Russian phoneme /'E/;
- no-accent-style converter based on Russian phonemes. All phonemes in Russian transcriptions receive prefix "r_" for distinguishing between Russian and Ukrainian phonemes.

Word-forms in which different vowels may be stressed have more than one phoneme transcription, e.g. /s t r A n y/ ("*countries*") and /s t r a n Y/ ("*country*").

Lexicons contain 4980 word-forms written equally in Ukrainian and Russian and 840 word-forms not only written but also pronounced equally, e.g. "наш" ("*our*"), "народ" ("*people*").

All lexicons have the same number of word-forms (110k), but they differ in number of phoneme transcriptions per one word-form. The most compact UKR+RUS_LEX contains 1.33 phoneme transcriptions per one word-form while UKR_LEX and UKR_LEX+1 contain 1.35 transcriptions per one word-form. In these two lexicons, accented Russian pronunciation is modeled by additional phoneme transcriptions.

7.4. Acoustic models

In addition to two monolingual AMs, three bilingual AMs were trained corresponding to three phone sets and three lexicons: UKR_AM, UKR+1_AM, and UKR+RUS_AM.

All AMs include 19 extra language-independent acoustic models for silence, breath and different types of hesitations.

7.5. Language model and ASR decoder

As our focus of interest was acoustic and pronunciation modeling, we decided to exclude the influence of language model on recognition results. A bilingual bigram LM was trained on all texts of AKUEM corpus comprising 2.7 million word-forms.

The same HTK decoder was used in all experiments.

8. EXPERIMENTS AND RESULTS

A series of experiments were conducted to evaluate the performance of the proposed approach to code-switching speech recognition.

Ukrainian and Russian monolingual systems trained only on AKUEM corpus were regarded as base systems (BASE_UKR and BASE_RUS). Russian monolingual lexicon was created using the no-accent-style grapheme-to-phoneme converter modeling Russian pronunciation peculiar to Russia.

The following bilingual systems were created for experiments:

UKR: UKR_AM, UKR_LEX;
UKR+1: UKR+1_AM, UKR+1_LEX;
UKR+RUS: UKR+RUS_AM, UKR+RUS_LEX.

The results are summarized in Table 3.

9. DISCUSSION

The experimental results show that three bilingual system configurations are capable of recognizing conversational code-switching speech with a word accuracy of 60-65%. The bilingual system based only on Ukrainian phonemes (UKR) outperforms the bilingual system based on combination of all Ukrainian and all Russian phonemes (UKR+RUS). This can be explained by our hypothesis that speakers living in Ukraine use the same phonemes while speaking Ukrainian and Russian. This is reflected in pronunciation lexicon UKR_LEX where phoneme transcriptions of Russian words are created in accent style. On the contrary, UKR+RUS_LEX contains phoneme transcriptions created in no-accent style modeling Russian pronunciation in Russia. Another reason is a significantly larger phoneme inventory of UKR+RUS system.

Table 3. Monolingual and code-switching speech recognition results (Acc, %).

Experimental Systems	**Ukrainian Speech (2.4 hours)**	**Russian Speech (2.3 hours)**	**Mixed Speech (4.7 hours)**
BASE_UKR	73.29	-	31.18
BASE_RUS	-	69.71	25.89
UKR	70.86	60.08	65.36
UKR+1	71.02	59.64	65.21
UKR+RUS	66.79	53.79	60.15

Comparing the recognition results of different systems only on Ukrainian speech segments revealed that the performance of bilingual systems based on Ukrainian phonemes did not degrade significantly compared to the monolingual Ukrainian system. The absolute degradation of 2-3% can be explained by a very similar, and often identical, pronunciation of frequent Ukrainian and Russian words, e.g. “і”/“и” (“*and*”), “ви”/“вы” (“*you*”), “ми”/“мы” (“*we*”) etc. Besides, some speakers speak Ukrainian with a Russian accent. This leads to a pronunciation with highly reduced unstressed vowels, which is peculiar to Russian language. That is why Ukrainian words like “політики” (“*politicians*”) and “євросоюз” (“*European Union*”) pronounced as /p a l' I t' i k' i/ instead of /p o l' I t y k y/ and /j e v r a s a j U s/ instead of /j e v r o s o j U z / can be recognized as Russian words “политики” and “евросоюз”.

Interestingly, adding one Russian phoneme to Ukrainian phone set gives a slight increase in accuracy for Ukrainian speech. This may be due to some very frequently encountered Ukrainian and Russian words spelled equally but differing in pronunciation namely by this phoneme, e.g. “не” (“*not*”, /n e/ in Ukrainian and /n' e/ in Russian, “те” (“*that*”, /t e/ in Ukrainian and “*those*” /t' e/ in Russian) etc.

10. FUTURE WORK

The analysis of code-switching speech recognition results suggests that they can be improved by post-processing. To smooth the recognition results, first a simple text-based language identification can be applied and then a bilingual dictionary can be used, which contains Ukrainian-Russian homophone word-forms, e.g. “лінія”/“линия” (“*line*”), “ера”/“эра” (“*era*”). Using the bilingual homophone dictionary means a simple substitution because, in general, the closely related syntactic systems of Ukrainian and Russian allow doing so.

The question remains what g2p converter should be used to model Russian language spoken in Russia and how large a phone set needs to be in that case.

We also plan to handle the challenging problem of intra-word Ukrainian-Russian code-switching which is highly characteristic of Surzhyk dialect.

11. CONCLUSIONS

This paper proposes an approach to recognition of code-switching speech for closely related Ukrainian and Russian languages. The approach allows handling both inter-sentential and intra-sentential code-switching. Language identification which is complicated because of close relation and high degree of phonetic interference between languages is avoided.

The approach is based on phonetic similarities between Ukrainian and Russian. Bilingual acoustic models and lexicons are created considering these similarities.

The usefulness of the approach was explored by comparing the recognition performance of two monolingual systems with the performance of three bilingual systems.

Bilingual AMs and lexicons do not degrade significantly the recognition performance on Ukrainian speech segments compared to monolingual Ukrainian AM and lexicon. Modeling of pronunciation variation in lexicons helps to cope not only with code-switching speech but also with speech with accent. Borrowing Ukrainian phoneme models outperforms putting together the two language-dependent sets of phoneme models by 5%.

A multi-speaker conversational code-switching speech is recognized with a word accuracy of 65%.

12. ACKNOWLEDGEMENTS

The authors would like to thank the Special Recording Systems, Ltd for the opportunity to use the Ukrainian Broadcast Speech Corpus (AKUEM).

13. REFERENCES

[1] Joyce Y. C. Chan, Houwei Cao, P. C. Ching, and Tan Lee, “Automatic recognition of Cantonese-English code-mixing speech”, *International Journal of Computational Linguistics and Chinese Language Processing*, 14(3):281-304, 2009.

[2] Ying Li, Yue Yu and Pascale Fung, "A Mandarin-English Code-switching Corpus", *Proceedings of the Eighth International Conference on Language Resources and Evaluation*, 2012, pp. 2515-2519.

[3] Weiner, J., Vu, N. T., Telaar, D., Metze, F., Schultz, T., Lyu, D.-C., Chng, E.-S. and Li, H., "Integration of Language Identification into a Recognition System for Spoken Conversations Containing Code-Switches", *Proceedings of the Third International Workshop on Spoken Languages Technologies for Under-resourced Languages*, Cape Town, South Africa, 2012.

[4] Vu, N.T., Lyu, D., Weiner, J., Telaar, D., Schlippe, T., Blaicher, F., Chng, E., Schultz, T., Li, H., "A First Speech Recognition System for Mandarin-English Code-switch Conversational Speech", *ICASSP, 2012*, pp. 4889-4892.

[5] Dau-cheng Lyu, Ren-yuan Lyu, Yuang-chin Chiang, Chun-nan Hsu, "Speech Recognition On Code-Switching Among The Chinese Dialects", *ICASSP, 2006*, vol. 1, pp. 1105-1108.

[6] Kiran Kumar Bhuvanagiri and Sunil Kumar Kopparapu, "An Approach to Mixed Language Automatic Speech Recognition", *Proceedings of the Oriental COCOSDA*, Kathmandu, Nepal, 2010.

[7] Kiran Bhuvanagirir, Sunil Kumar Kopparapu, "Mixed Language Speech Recognition without Explicit Identification of Language", *American Journal of Signal Processing*, 2(5): 92-97, 2012.

[8] Niesler, T., Willett, D., "Language identification and multilingual speech recognition using discriminatively trained acoustic models", *ISCA Tutorial and Research Workshop on Multilingual Speech and Language Processing*, 2006, paper 004.

[9] Martin, A. and Greenberg, C., "The 2009 NIST Language Recognition Evaluation," *Proc. of Odyssey 2010: The Speaker and Language Recognition Workshop, 2010*, paper 030, pp. 165–171.

[10] Weng, F., Bratt, H., Neumeyer, L., Stolcke, A., "A study of multilingual speech recognition", *Proceedings of the EUROSPEECH, 199*7, Rhodes, Greece, pp. 359-362.

[11] Kim, M., Oh, Y.R., Kim, H.K., "Non-native pronunciation variation modeling using indirect data driven method", in Noam R. Shabtai [Ed], *Advances in Speech Recognition*, pp. 83-104, 2010.

[12] Lamel, L., Courcinous, S., Despres, J., Gauvain, J-L., Josse, Y., Kilgour, K., Kraft, F., Bac, L.V., Ney, H., Nussbaum-Thom, M., Oparin, I., Schlippe, T., Schlüter, R., Schultz, T., Fraga Da Silva, T., Stüker, S., Sundermeyer, M., Vieru, B., Vu, N.T., Waibel, A., Woehrling, C., "Speech Recognition for Machine Translation in Quaero", *Proceedings of the International Workshop on Spoken Language Translation*, 2011.

[13] Lamel, L., Courcinous, S., Gauvain, J.-L., Josse, Y., Le, V.-B., "Transcription of Russian conversational speech", *Proceedings of Third International Workshop on Spoken Languages Technologies for Under-resourced Languages*, pp. 162–167, 2012.

[14] Psutka, J., Ircing, P., Psutka, J.V., Hajic, J., Byrne, W.J. and Mirovsky, J., "Automatic Transcription of Czech, Russian, and Slovak Spontaneous Speech in the MALACH Project", *Proceedings of the 9th European Conference on Speech Communciation and Technology*, pp. 1349-1352, 2005.

[15] Karpov, A., Kipyatkova, I. and Ronzhin, A., "Very large vocabulary ASR for spoken Russian with syntactic and morphemic analysis", *Proceedings of INTERSPEECH*, pp. 3161-3164, 2011.

[16] Vazhenina, D. and Markov, K., "Phoneme Set Selection for Russian Speech Recognition", *Proceedings of 7th International Conference on Natural Language Processing and Knowledge Engineering*, pp. 475-478, 2011.

[17] Karpov, A., Kipyatkova, I. and Ronzhin, A., "Speech Recognition for East Slavic Languages: the Case of Russian", *Proc. of the 3rd International Workshop on Spoken Languages Technologies for Under-resourced Languages*, pp. 84-89, 2012.

[18] Pylypenko V.V., Robeiko, V.V., "Experimental System of Computerized Stenographer for Ukrainian Speech", Proc. 13-th International Conference "Speech and Computer", pp. 67-70, 2009.

[19] Robeiko, V., Pylypenko, V., Sazhok, M., Vasylieva, N., Radoutsky, O., "Ukrainian Broadcast Speech Corpus Development", *Proceedings of the 14th International Conference "Speech and Computer: SPECOM'2011"*, pp. 435-440, 2011.

[20] Lyudovyk T.V., Pylypenko V.V., Robeiko V.V., "Automatic Recognition of Spontaneous Ukrainian Speech Based on the Ukrainian Broadcast Speech Corpus", *Proc. of the International Conference DIALOG-2011*, pp. 478–488, 2011 (in Russian).

[21] Schlippe, T., Volovyk, M., Yurchenko, K., Schultz, T., "Rapid Bootstrapping of a Ukrainian Large Vocabulary Continuous Speech Recognition System", *ICASSP 2013, pp.* 7329-7333.

[22] http://www.phon.ucl.ac.uk/home/sampa/russian.htm.

[23] http://fr.wikipedia.org/wiki/Alphabet_ukrainien.

[24]http://www.acapela-vaas.com/Includes/language_manuals/Manual_Russian.pdf.

[25] Krivnova, O., Zinovieva, N., Zakharov, L., Strokin, G., Babkin, A., "TTS Synthesis for Russian Language", *Web Journal of Formal, Computational and Cognitive Linguistics*, issue 1, 1997. Online: http://fccl.ksu.ru/ar2.htm.

[26] Lyudovyk T., Brozinski S., Noner M., Robeiko V., Sazhok M. Speech Synthesis Applied to SMS reading, *Proc. of the 13th International Conference "Speech and Computer: SPECOM'2009"*. – St. Petersburg: 2009. Pp. 300-305.

[27] Young S. et al. *The HTK Book (for HTK Version 3.4)*. Cambridge, UK, 2009.