AUTOMATIC GENERATION OF RIGHT INTONATION MARKS AND SPEECH FOR MEDICAL DOMAIN TEXTS FOR BELARUSIAN NOOJ MODULE

Dz. Dzenisiuk, Yu. Hetsevich, N. Drahun, A. Bakunovich, J. Shynkevich
United Institute of Informatics Problems of the NAS of Belarus, Minsk
e-mail: d.dzenisiuk@gmail.com, yury.hetsevich@gmail.com, ndrahun@gmail.com, bakunovich.andrei@gmail.com, silenoschestra@gmail.com

The main problem in this paper is automatic generation of prosodic transcription of long sentences for Belarusian. Due to difficulties of generation of intonationnaly coloured (expressive) synthetic speech from long sentences without punctuation or any visible intonational marks this problem is actual.

To resolve this problem, we collected the special corpus of medical domain texts. The text corpus was created by getting news from the website of medical organization.

Analyzed text corpus is about 270 texts, 71897 word forms, more than 6000 sentences.

Because of the absence of punctuation marks between 4^{th} and 10^{th} words in most of analyzed long sentences of the corpus, we decided to develop some syntactic grammars for Belarusian NooJ module that could add additional right intonation markers into the long sentences by dividing sentences into syntactic phrases (sequence of word forms or phonetic words).

After the implementation of the developed NooJ grammars in Belarusian text to speech system, we obtain invisible punctuation marks in long sentences taking into account the nearest context in a text.

As a result, the quality of synthetic speech generated by Belarusian text to speech system will be improved. Because of the additional pauses (after processing with NooJ grammars) the level of perception of such speech be the end user will increase.

References

1. Hetsevich Y., Okrut T., Lobanov B. (2016) Grammars for Sentence into Phrase Segmentation: Punctuation Level. In: Okrut T., Hetsevich Y., Silberztein M., Stanislavenka H. (eds) Automatic Processing of Natural-Language Electronic Texts with NooJ. NooJ 2015. Communications in Computer and Information Science, vol 607. Springer, Cham

2. Pitrelli, J.F., et al.: The IBM expressive text-to-speech synthesis system for American English. IEEE Trans. Audio Speech Lang. Process. 14(4), 1099–1108 (2006)

3. Froehlich, P., Hammer, F.: Expressive text-to-speech: a user-centred approach to sound design in voice-enabled mobile applications. Telecommunications Research Centre Vienna (FTW). (http://userver.ftw.at/~froehlich/papers/JDS2004_ExpressiveTTS.pdf)

4. Leskovec, L.: Lexical Stress Assignment and Pronunciation Formalization in Expressive TTS/Trinity College Dublin. (http://www.cs.tcd.ie/research_groups/clg/COST2102.IS2009/content/program/node61.html)

5. Kawanami, H., et al.: Designing Speech Database with Prosodic Variety for Expressive TTS system, Nara Institute of Science and Technology, Takayama-cho. (http://gandalf.aksis.uib.no/lrec2002/pdf/337.pdf)

6. Lobanov, B.M., Tsirulnik, L.I.: Computer speech synthesis and cloning. Belarussian science, p. 342 (2008). (in Russian)