



Boosting the vitality of the Belarusian language

BASNET (BELARUS) (https://www.inthefieldstories.net/network/basnet/),
 EAPCONNECT (EUROPE'S EASTERN PARTNERSHIP COUNTRIES) (https://www.inthefieldstories.net/network/eapconnect/),
 EUROPE (https://www.inthefieldstories.net/network/europe/)
 ARTS & CULTURE (https://www.inthefieldstories.net/category/arts-and-culture/),
 DIGITAL HUMANITIES (https://www.inthefieldstories.net/category/digital-humanities/)

"When you lose a language and a language goes extinct, it's like dropping a bomb on the Louvre." Michael Krauss (https://era.org/en/krauss-rewarded-for-language-work/)

Languages, like endangered species, are easier to keep alive than to bring back from extinction. The survival and development of the Belarusian language has been boosted by the work of the Speech Synthesis and Recognition Laboratory (SSRL) at the United Institute of Informatics Problems of the National Academy of Sciences of Belarus (UIIP NASB), with the support of the national research and education network BASNET and the EU-funded EapConnect project. The Belarusian language is underrepresented both on the national level and on a global scale.

Yuras Hetskevich, head of SSRL, says: "It is crucial to popularise the language and underscore the importance of the use of the mother tongue in 'the land under white wings'. Paradoxically, this can be done by increasing the visibility of the language in the international community."

What determines the vitality of a language? Fifteen years ago, the main criteria were the number of students receiving education in the language, the availability of books written in it, and its presence in state institutions. Although these factors are still important, new rules have been dictated by new technologies: the presence of the language in the virtual space — in computer support, in applications, on social networks and on the Web in general — affects if and to what extent it is alive.

The current situation

One of the most important factors is how the language is valued by the younger generation. Youngsters have been growing more interested in Belarusian, partly owing to the virtual world, where they can communicate and search for information using the language. The amount of online content in Belarusian is growing, on official portals, websites, social networks and blogs. But while this helps the language to develop — for example with the appearance of new words such as slang — speech norms are often breached or intentionally butchered on the Internet. Online translators also leave much to be desired. Many use a form of translation from Russian (the other state language) that creates phrases that seem authentically Belarusian but are not. There is a lot of history in such mistakes and the history of the Belarusian language has been complicated. If one wants to understand one's own culture, one has to know his own language. Dictionaries and other tools that can be used by amateurs, professional and computational linguists exist in online platforms and websites, but many people are unaware of them.

Such problems can be gradually solved by drawing attention to the language, the tools that are already available, and the projects that are yet to be implemented. According to a 2019 decree, all national websites shall be available in both state languages, which requires official support, centralised and systematic activities.

Successful beginning

The new decree made it possible for the Speech Synthesis and Recognition Laboratory to develop websites for healthcare organisations in Belarusian as well as Russian. They also develop electronic resources for computational linguistics, such as audio guides for museums and exhibitions and various mobile applications — first of all, in the Belarusian language, to popularise it among the citizens-but also in English, Russian, Chinese, Polish, so that visitors from other countries can enjoy it. The content is always carefully and meticulously fact-checked and proofread by historians, linguists and other professionals.

In 2016 Yuras Hetskevich and his team participated in the EapConnect project's Enlighten Your Research (EYR) programme in the Eastern Partnership (EaP) region. As a result, the necessary data storage resources in BASNET cloud infrastructure were allocated and a new mobile application 'KrokApp' was successfully developed. This uses speech systems as well as audio and visual imagery in different languages to inform visitors about the history of cities in different languages. This also led to the creation of other audio guides later.

Quick steps towards worldwide recognition

With the support of EapConnect, Yuras Hetskevich and a colleague participated in the 2019 CLARIN conference held in Germany. This is the pivotal and largest event for computational linguists in Europe. CLARIN (Common Language Resources and Technology Infrastructure) is a network of organisations that host digital language resources and services from all over Europe and beyond, which are accessible in one online environment to support scholars, researchers, students and scientists, especially in the humanities and social sciences. As a result of this participation, an application to create a CLARIN knowledge centre in Belarus was submitted.

The team also attended CLARIN events in Bulgaria, the Netherlands and the UK, gaining hands-on experience of computational linguistics, learning how to create a national CLARIN consortium, and presenting 'Krokam' audio guides that showed the Belarusian culture to advantage via electronic devices. As a result, the team now cooperates on the Distant Reading project by adding Belarusian resources to the European corpus and provides Belarusian resources to the CLARIN Virtual Language Observatory. The Knowledge Centre for Belarusian text and speech processing (K-BLP) has also been recognised as a CLARIN Knowledge Centre.

"Many of these fruitful steps became possible due to the support of BASNET and the EapConnect project. By joining CLARIN, Belarus has boosted its visibility, which has already opened the doors to various international projects." Yuras Hetskevich.

Further information:

- United Institute of Informatics Problems of the National Academy of Sciences of Belarus (https://uiip.bas-net.by/eng/)
- Belarusian text and speech processing tools (https://www.corpus.by/)
- The CLARIN (https://www.clarin.eu/) infrastructure
- CLARIN Virtual Language Observatory (https://vlo.clarin.eu/search?&q=uiip) with Belarusian resources
- Krokam (https://krokam.com/) personal audioguides
- Distant Reading (https://www.distant-reading.net/) for European Literary History