# Compilation of Language Resources and On-line Dissemination of Knowledge about

# Endangered Languages and Linguistic Heritage

## Katarzyna Klessa & Nicole Nau

Adam Mickiewicz University in Poznań, The Institute of Linguistics. Contact us at: **klessa@amu.edu.pl**, **naunicol@amu.edu.pl**

## Abstract

This contribution presents selected results of two projects concerning compilation of language resources and dissemination of knowledge about endangered languages and language documentation primarily among non-academic target groups. Both projects include data from one of the Baltic languages, Latgalian. The aim of the paper is to discuss the main features of the resources as well as to show the interest of these projects from a Baltic perspective and their usefulness for further research and development.

## 1. Languages in Danger: Dissemination and Education

The interactive website available on-line at http://languagesindanger.eu/ delivers information about selected endangered languages from all over the world, and is designed for students and teachers of secondary schools as well as for young researchers with a view to increase awareness and interest in these topics [1, 2]. The implemented features include: interactive components (interactive map, interactive quizzes), teaching materials (ready-to-use classroom materials, multimedia resources), a textbook for background reading (10 chapters enriched with visual and multimedia material). The components were created using source data available from existing repositories e.g. DoBes [3] as well as with the support of new data gathered by individual researchers and language documenters.

More than 20 languages are selected as featured languages of the website. Locations of these featured languages are displayed in the interactive map together with exercises and multimedia materials. They are also referred to in more detail in other sections of the website.

Among others, these featured languages include Latgalian [4] (spoken in Latvia), Karaim (spoken in Lithuania and Poland), Wilamowicean [5] (spoken in Poland). There are also teaching materials about Finno-Ugrian languages, which could be especially interesting for Estonian schools.

An important novelty of the package is the European perspective and the availability of all contents not only in English but also in Polish, Hungarian, German and Dutch (all localized versions to be published at the end of September 2014).

## 2. Linguistic Heritage: Compilation of Language Resources and Dissemination

The main goals of the project Dziedzictwo językowe Rzeczypospolitej (The Linguistic Heritage of the Polish-Lithuanian Commonwealth, referring to the languages spoken in the territory once named 'Rzeczypospolita' in Polish) were:

a) to provide information about languages for a general audience (above 20 languages described so far, including Baltic (Latgalian; Lithuanian dialects spoken in Poland) and other languages also spoken in the Baltic States (e.g. Yiddish, Romani, Karaim, Russian dialects of Old Believers).
b) to start a scientific documentation database providing resources for researchers (4 languages featured so far: Latgalian, (Polish) Yiddish, Wilamowicean, Halcnovian).

Latgalian resources in the on-line database contain among others folklore texts collected by Stefania Ulanowska in the 1890s. The original edition contained the Latgalian texts written in the semi-standardized writing system of the time and a Polish translation. In the present database, a new transliteration based on today's orthography was added. Moreover, for all the original fairy tales (and selected songs and short tales) English translations have been provided. For some of the tales, audio recordings were made and also included in the database.

The Linguistic Heritage database can be flexibly extended by new languages and source data.

# www.languagesindanger.eu


Figure 1. The website www.languagesindanger.eu - welcome page


Figure 2. An example exercise on Latgalian at the Interactive Map at www.languagesindanger.eu.
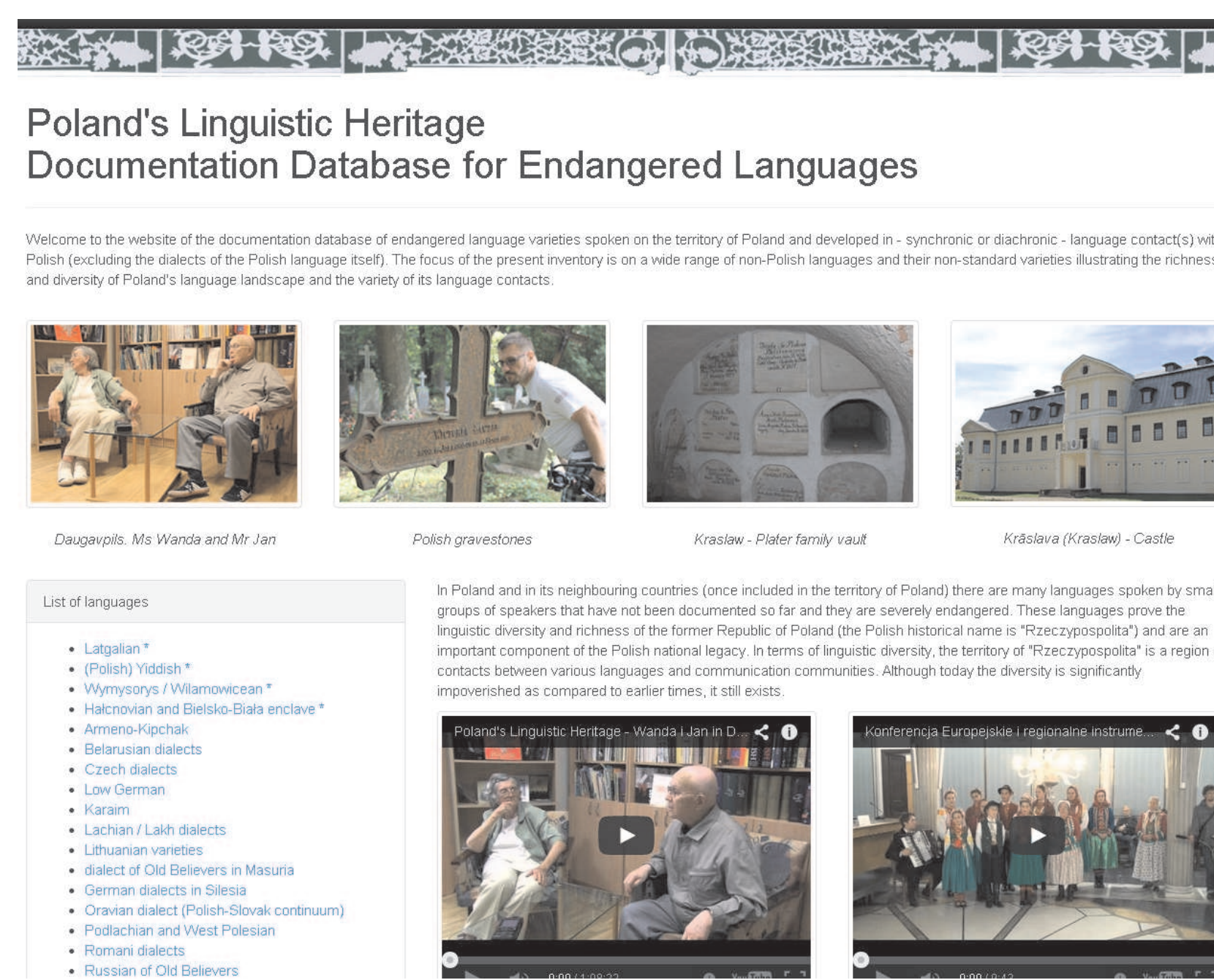
# www.inne-jezyki.amu.edu.pl


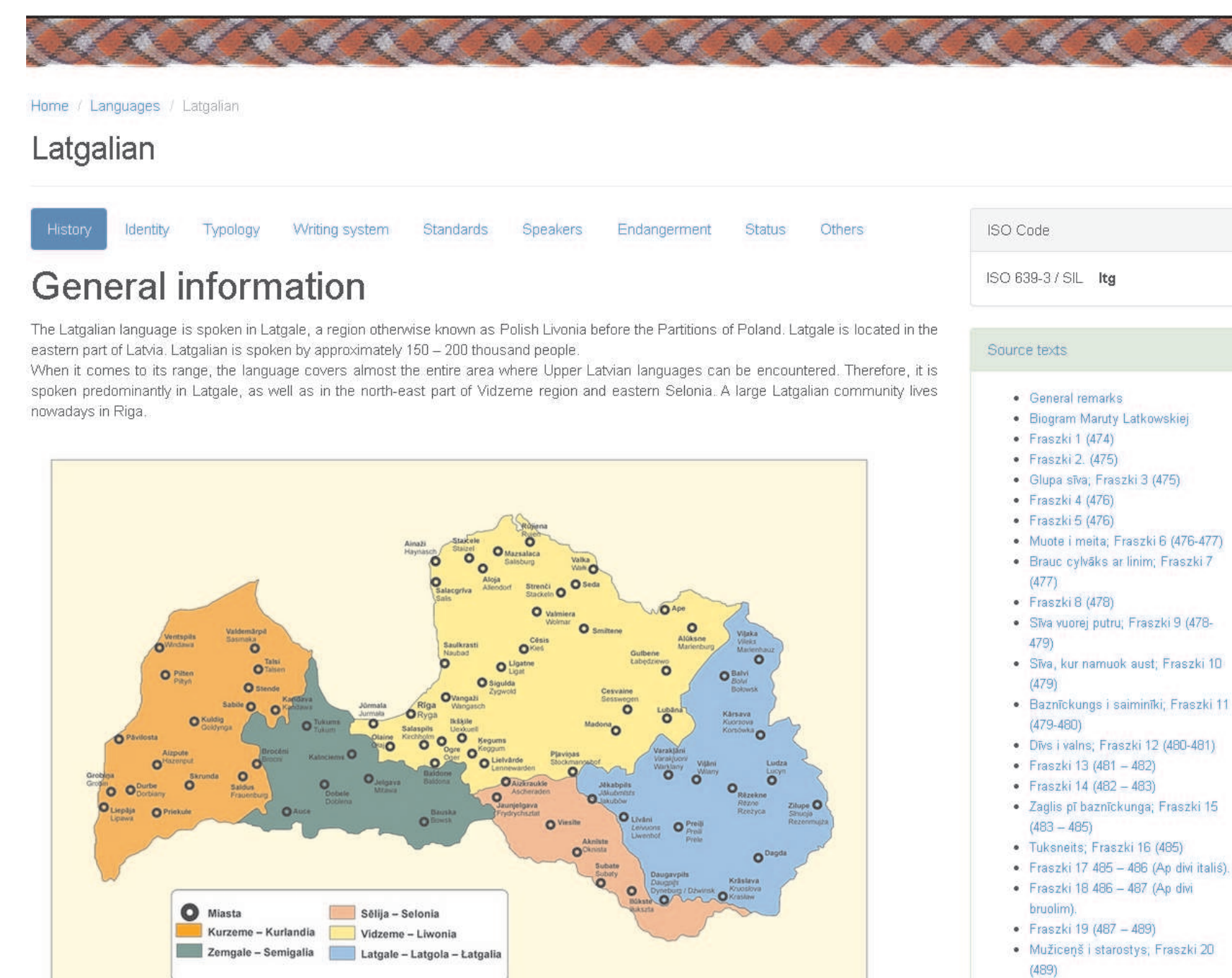Figure 3. The welcome page at www.inne-jezyki.amu.edu.pl.


Figure 4. One of sub-pages for Latgalian at www.inne-jezyki.amu.edu.pl.

## Conclusions

The architecture of both of the described websites enables further extensions by other language versions (adaptations) and materials, e.g. with the participation of a potential new partner(s) from Baltic States. For the Linguistic Heritage Database tools have to be developed that meet the needs of linguists and ethnologists who want to use the data in their research. It is important to realize that language documentation does not end when data has been collected and stored in digital archives. These archives are more than museums of dying languages, they provide powerful resources for further research.

## References

[1] Wójtowicz, R. 2014. Language endangerment in European secondary schools: Challenges and perspectives. In: Proceedings of the International Conference The Future of Education IV, Florence. On-line at: http://conference.pixel-online.net/FOE/conferenceproceedings.php]. [2] Jung, D., Klessa, K., Duray, Z., Oszkó, B., Sipos, M., Szeverényi, S., Várnai, Z., Trilsbeek P., Váradi, T. 2014. Languagesindanger.eu - including multimedia language resources to disseminate knowledge and create educational material on less-resourced languages. In: Proceedings of the 9th Language Resources and Evaluation Conference (LREC), Reykjavik. [3] DOBES (Documentation Of Endangered Languages). Available on-line at: http://dobes.mpi.nl/ [4] Nau, N. 2011. A short grammar of Latgalian. München: LINCOM Europa. [5] Wicherkiewicz, T. 2003. The Making of a Language. The case of the idiom of Wilamowice, Southern Poland. Berlin-New York: Mouton de Gruyter.

## Acknowledgements