

Tamás Váradi and the International Lexicography

Lipp Veronika¹

¹ Nyelvtudományi Intézet
lipp.veronika@nytud.hu

1. European Network of e-Lexicography (2013–2017)

In the last 10 years a clear need has arisen for a common approach to e-lexicography that forms the basis for a new type of lexicography fully embracing the pan-European nature of many of the vocabularies of languages spoken in Europe.

The aim of the COST European Network of e-Lexicography (ENeL, 2013–2017) was to increase, co-ordinate and harmonize European research in the field of e-lexicography and to make authoritative information on the languages of Europe easily accessible.

Tamás Váradi joined the project in 2013 as the leader of the Department of Language Technology and Applied Linguistics of the Research Institute for Linguistics. He became one of the most active members of the Management Committee of the network.

The aim of the network was to make lexical knowledge of small and large languages available in a European dictionary portal. The portal was finalized in 2017 and it serves as the central reference point for all dictionary users who look for reliable, authoritative dictionary information on the languages of Europe and their histories.¹

The network enabled cooperation and the exchange of resources, technologies and experience in e-lexicography and provided support for dictionaries which were not online yet.

There were numerous meetings in which we discussed establishing standards for innovative e-dictionaries that fully utilize the possibilities of the digital medium and tried to establish new ways of representing the common heritage of European languages by developing shared editorial practices and by interconnecting already existing information.

Tamás Váradi was involved mainly in two working groups' tasks: one was engaged in the integrated interface for European dictionary content,

¹ <http://www.dictionaryportal.eu/hu/>

whereas the other investigated how the pan-European nature of the vocabularies of the languages of Europe could be represented in monolingual dictionaries and within the European dictionary portal.

The first working group investigated how authoritative dictionary information on the languages of Europe can be made accessible to both the general and academic public. A European dictionary portal was set up, which gives information on scholarly dictionaries of the languages of Europe and provides access to these dictionaries. It investigated the possibilities of interlinking the contents of European dictionaries, explored user requirements with respect to the presentation of dictionary content, and mapped the possible involvement of users in the creation of dictionary content.

The other working group investigated how the pan-European nature of the vocabularies of the languages of Europe could be represented in monolingual dictionaries and within the European dictionary portal.

Tamás organized the 6th action meeting in Budapest with great success in February, 2017.

2. European Lexicographic Infrastructure

Another lexicographical project called ELEXIS started in 2018. It was a continuation of the ENel project, and it aimed at developing lexicographical tools. The lexicographical landscape in Europe is currently rather heterogeneous. There are standalone lexicographical resources, which are typically encoded in incompatible data formats due to the isolation of efforts, disabling the reuse of data in natural language processing, linked open data and the Semantic Web, or in digital humanities. ELEXIS will introduce common standards, develop conversion tools, and most importantly, it will interconnect the existing resources so that they can be used to develop new modern data which can be used in ways that new digital technologies require.

The main aim of the ELEXIS project is to introduce common standards, develop conversion tools, and most importantly, it will interconnect the existing resources so that they can be used to develop new modern data which can be used in ways that new digital technologies need.

The Research Institute for Linguistics is actively involved in this project under Tamás Váradi's leadership. Aligning senses across resources and languages is a challenging task with beneficial applications in the field of natural language processing and electronic lexicography. In this project the goal was to align word senses in monolingual dictionaries.

The alignment is carried out at sense-level for various resources in Hungarian and in 14 other languages (Ahmadi et al., 2020). Another task is to create parallel sense-annotated datasets in many European languages to be used for evaluation purposes in word-sense disambiguation.

The main merit of Tamás Váradi is that he has placed Hungarian lexicographical research into the international context. He has established extensive relationships with international lexicographical institutions and prominent lexicographers. I am very grateful for having been able to work with him over the past few years and I hope it will continue long into the future.

Dear Tamás, I wish you a happy birthday and much success in your further career.

References

- Ahmadi, S., McCrae, J.P., Nimb, S., Khan, F., Monachini, M., Pedersen, B. S. Declerck, T., Krek, S., Lipp, V., Váradi, T. et al.: A Multilingual Evaluation Dataset for Monolingual Word Sense Alignment. In: Calzolari, N. et al. (eds.) Proceedings of the 12th Language Resources and Evaluation Conference. pp 3232–3242. ELRA (2020)