

Workshop: First step towards an interactive map of Balkan linguistic features

Date: 26th – 27th November 2018

Venue: University of Zurich

In typology and beyond it is a well-known fact that the so-called Balkan languages, such as Albanian, Aromunian, Bulgarian, Greek, Macedonian, BCS, Romani and Romanian, share a set of structural features on the morpho-syntactical, phonological and lexical level. Many descriptions of Balkan languages are based on researches on the Standard varieties while neglecting peculiarities of certain dialects, especially those ones which are spoken in the periphery of the whole Sprachbund area or in the periphery of individual language areas.

An attempt to fill this lacuna has already been made in the scope of the project *A Small Dialectological Atlas of the Balkan Languages*. Furthermore, the amount of diverse linguistic material (dialect descriptions, corpora etc.) on Balkan dialects is rich and grows larger throughout ongoing data collection. Additionally, we witness the ever-growing importance of Digital Humanities, the development and implementation of cut-edge technologies for visualizing linguistic data and, most importantly, the trend to make linguistic resources available for a broader scientific or non-scientific community. Bearing in mind both the abundance of available language material as well as the possibilities new technologies involve, we consider the time suitable to start a first attempt to synthesis the data which has been collected so far.

Creating a single interactive map or linguistic atlas of Balkan linguistic features is a grand enterprise which requires a long-term collaboration of a wide range of experts of the respective languages and dialects and which cannot be achieved during a single workshop. With this workshop we try to lay the foundation stone of such a collaboration by pursuing two main goals during two main sessions.

Day 1: Section 1 *Data presentation and evaluation*

Participants will be given twenty minutes for a presentation on their data and research, followed by 10 minutes of discussion. At the end of the day the audience should understand how you can contribute for future collaboration.

Day 2: Section 2 *Hands-on*

The participants should witness the advantages of modern data visualization by applying some methods on their own data. Therefore, Dr Peter Ranacher, the head of the GIS Group at UZH, will give an overview and training on methods of geo-linguistic data. By this, the participants will gain a better idea of the technical possibilities and the requirements which have to be fulfilled by their corpus before it can serve as material for a common data base or use it for their own research.

The major part of this section will be dedicated to the hands-on session, where the participants will be guided step-by-step through the process of preparing and processing linguistic data and visualizing the results onto a map. For this purpose, R software will be used. Ideally, the participants will be using a small portion of their own material.

By the end of the workshop the following goals should have been reached:

- Bringing together researchers from different fields of Balkan/Southeast European linguistics who have conducted or are currently conducting linguistic research on the basis of linguistic data on Balkan languages/dialects
- A better overview of current corpus linguistic research on Balkan languages and the linguistic data available
- A better understanding of the technical possibilities of geo-linguistic data visualization
- A stronger desire to collaborate and network in the Balkan linguistic community
- In a nutshell: The beginning of a long-term collaboration and a first step towards an interactive map of Balkan linguistic features

If you have any questions, feel free to contact us: Teodora Vuković: teodora.vukovic2@uzh.com
Veton Matoshi: veton.matoshi@lmu.de