

# **‘Ill-bred sons’, family and friends: tracing the multiple affiliations of Balkan Slavic**

---

## **Summary of the research plan**

The South Slavic dialect continuum is characterised by an intricate encounter of affiliations: genealogically, it is intersected by an old bundle of isoglosses differentiating West and East South Slavic, areally, parts of it – the ‘ill-bred’ sons, as Schleicher (1850) called them – share a number of morpho-syntactic innovations with their neighbouring non-Slavic languages. The resulting variation becomes visible most distinctly in the Torlak dialects. Being located at the peripheries of contemporary Serbian, Bulgarian and Macedonian, they are transitional between West and East South Slavic. Yet, spoken at the outskirts of the spread zones of Balkan innovations, they are transitional between Balkan Slavic (BS) and non-Balkan South Slavic as well. A linguistic description and analysis of Torlak thus calls for bringing together insight from dialectology and areal-typology in order to study the interaction of innovations diffused through language contacts with the inherited genealogical features.

The project focuses on a set of morpho-syntactic BS innovations and their diffusion and integration into the South Slavic system from a diatopic and diachronic perspective by contrasting Torlak with the surrounding varieties and by drawing on evidence from pre-standardised vernacular sources. Systematically analysing the co-occurrence of areal innovations with inherited features will facilitate a principled description and analysis of Torlak, provide new insight into the history of Balkan Slavic and contribute to a better understanding of dialectal and areal contact and the concomitant processes of convergence and divergence.

The combination of a diachronic and diatopic perspective necessitates the comparison of non-standard data along two dimensions: (i) eastern BS with western BS varieties, and both with non-BS Serbian in order to trace the geographical extension of the relevant structures in their formal and functional aspects, (ii) contemporary data with earlier stages displayed in 18<sup>th</sup>/19<sup>th</sup> c. literary sources in order to gain insight into the diachronic diffusion of BS innovations. The focus will be on three features: (1) clitic doubling, (2) postponed definiteness marking, (3) existence of analytic and synthetic past tenses. These features are also found in languages outside the BS and Balkan area, which places the project in the larger context of (1’) conditions on clitic doubling in neighbouring non-Slavic languages, (2’) postponed definiteness markers in North Russian, (3’) development of (former) perfects in Slavic. This in turn opens up a broader contact linguistic and areal-typological perspective, concerning in particular (1’’) Slavic – Albanian/Romance, (2’’) Slavic – Baltic – Finno-Uralic, and (3’’) Slavic within the linguistic Europe.

The analysis will be based on annotated corpora for each of the varieties to be compared. Providing more fine-grained data, it becomes possible to establish correlations between features and structures and hence reveal usage conditions, illustrate converging and diverging developments, in particular as concerns their functions, and to map the data in time and space by geo-referencing and visualising them with GIScience technology. To this end, existing processing tools will be improved in such a manner that they can be applied to these still under-resourced languages. This methodological aspect adds a further dimension to the project, beyond its contribution to (Slavic) dialect syntax and the linking of dialectology and areal typology.