

## Многоязычный ГРАЛИС-КОРПУС в системе параллельных южнославянских корпусов

Бранко Тошович (Branko Tošović)  
Институт славистики Университета им. Карла и Франца в г. Грац  
branko.tosovic@uni-graz.at

Арно Вонисш (Arno Wonisch)  
Институт славистики Университета им. Карла и Франца в г. Грац  
arno.wonisch@uni-graz.at

О степени взаимной близости / отдаленности южнославянских языков существуют различные мнения, особенно по отношению к очень близким языкам, какими являются (а) сербский, хорватский, боснийский/бошняцкий и черногорский, (б) македонский и болгарский. Различия во взглядах особенно сильно проявляются, когда речь идет об определении их социолингвистического статуса — являются ли они отдельными языками, вариантами одного (полицентрического) языка, диалектами одного национального языка, одним языком с различными (политическими) названиями и т.п. Отсутствие исследований, которые на базе электронных корпусов объективно представили бы соотношения южнославянских языков и предоставили бы релевантный материал для разъяснений всех связанных с этим существенных вопросов, создает благоприятные условия для субъективных, тенденциозных толкований и политизации. Данная проблема сильно обострилась распадом Югославии и военными конфликтами в Юго-Восточной Европе в 90-ые годы XX столетия. Ситуацию значительно усложнили расхождения во мнениях относительно оправданности / неоправданности официального провозглашения и кодификации некоторых южнославянских языков. На усложнение межславянских языковых отношений повлияло и изменение роли некоторых из них в обществе. Статус языка взаимного общения на уровне одного государства потерял сербохорватский язык, словенский язык превратился из языка республики в язык нового государства, некоторые южнославянские языки стали языками национальных меньшинств (например, сербский в Хорватии и хорватский в Сербии). Непростая и неоднородная языковая ситуация на территории распространения южнославянских языков требует объективных исследований интралингвистических, социолингвистических и психолингвистических аспектов их соотношения на репрезентативном корпусном материале.

В рамках проекта «Die Unterschiede zwischen dem Bosnischen / Bosniakischen, Kroatischen und Serbischen» / «Различия между боснийским, бошняцким, хорватским и сербским языками» (Грац, Австрия, 2006–2010 гг.) создан ГРАЛИС-КОРПУС (нем. GRALIS-KORPUS) — комплексный и многоязычный корпус письменных и устных текстов для изучения славянских языков и преподавания того или иного славянского языка в качестве родного или иностранного (<http://www-gewi.kfunigraz.ac.at/gralis/index.html>). Разработчики корпуса преследовали следующие цели: (а) создать возможности для системнолингвистического, социолингвистического и психолингвистического изучения отношений между славянскими (в первую очередь южнославянскими) языками для того, чтобы ответить на вопрос, чем различаются славянские литературные языки между собой и насколько они далеки друг от друга, (б) предоставить разнообразный материал для изучения межславянских сходств и различий, (в) показать, как славянские языки функционируют и взаимодействуют на конкретном материале и на всех языковых уровнях (фонетико-фонологическом, орфоэпическом, грамматическом и стилистическом). В настоящее время пользователями ГРАЛИС-КОРПУСА являются слависты (лингвисты и литературоведы) из различных стран,

которые используют данные корпуса в рамках коллективных и индивидуальных проектов, а также в процессе обучения.

Концепция корпуса предполагает, что его можно использовать как одноязычный и как многоязычный (параллельный) корпус, но он в первую очередь предназначен для анализа соотношений между близкородственными (разными славянскими) и неблизкородственными (например славянскими и германскими) языками. Таким образом ГРАЛИС-КОРПУС является одним из немногих параллельных и мультилингвальных корпусов двух или более современных славянских языков (кроме — по нашим данным — Польско-русского и Польско-украинского корпусов Варшавского университета<sup>1</sup>, Болгарско-русского параллельного корпуса университета «Св. св. Кирил и Методий» в Великом Тырнове<sup>2</sup>, Параллельного корпуса славянских и других языков ParaSol<sup>3</sup>, Корпуса параллельных текстов Национального корпуса русского языка<sup>4</sup>).

Инфраструктура ГРАЛИС-КОРПУСА предоставляет возможность поиска в следующих языковых парах: (а) любой из славянских языков и немецкий язык, (б) славянские языки, принадлежащие к разным группам (южнославянские, восточнославянские, западнославянские), (в) языки только из одной группы (например, южнославянские). Морфосинтаксическая аннотация и поиск по ее параметрам проводится при помощи особой программы — ГРАЛИС-МОРФОГЕНЕРАТОРА, служащей для автоматического анализа морфологической структуры боснийского / бошняцкого, сербского и хорватского языков. В рамках ГРАЛИС-КОРПУСА существуют и индивидуальные подкорпуса (корпуса текстов писателей). К настоящему времени разработаны четыре: корпус лауреата Нобелевского премии по литературе Иво Андрича (1892–1975; корпус текстов на сербском / сербохорватском, болгарском, македонском, словенском, русском и немецком языках), корпус величайшего повествователя, юмориста и сатирика БРАНКО ЧОПИЧА (1915–1984; корпус текстов на сербском / сербохорватском, русском и немецком языках), корпус ЗОРАНА ЖИВКОВИЧА, в настоящее время самого переводимого писателя бывшей Югославии (род. 1948; корпус текстов на сербском, хорватском, русском и немецком языках) и корпус виднейшего македонского писателя, поэта и филолога БЛАЖЕ КОНЕСКОГО (1921–1993; корпус текстов на македонском, сербском, русском, украинском и немецком языках). Все подкорпуса письменных текстов ГРАЛИС-КОРПУСА содержат около 5 300 000 словоформ.

В центре внимания в докладе находится (а) глобальное представление ГРАЛИС-КОРПУСА со всеми его подкорпусами, (б) сопоставление этого параллельного корпуса с другими многоязычными (южно)славянскими корпусами и (в) обсуждение и демонстрация возможностей проведения сопоставительных лингвистических (конкретно: морфологических) исследований южнославянских языков на материале ГРАЛИС-КОРПУСА.

---

<sup>1</sup> <http://pol-ros.polon.uw.edu.pl/index.php?id=04&lang=pl>.

<sup>2</sup> <http://rbcorpus.com/index2.php>.

<sup>3</sup> <http://www.slavist.de>.

<sup>4</sup> <http://www.ruscorpora.ru/search-para-en.html>.