

Светла Коева, Диана Благоева, Сия Колковска  
Институт за български език „Проф. Л. Андрейчин“  
при Българската академия на науките  
София, България

## ПРОЕКТЪТ *БЪЛГАРСКИ НАЦИОНАЛЕН КОРПУС*: РЕЗУЛТАТИ И ПЕРСПЕКТИВИ

(Резюме)

В статията се представя проектът Български национален корпус (БНК), насочен към изграждане на голям по обем репрезентативен корпус на българския език, достъпен за свободно ползване в интернет. Корпусът е изцяло морфосинтактично (и отчасти семантично) анотиран и е снабден с детайлно описание на метаданните, което улеснява обработката му и извличането на информация. В сегашния момент БНК се състои от над 400 милиона графични думи и съдържа повече от 10 000 текста. В статията се описва накратко структурата на Корпуса и се дискутира въпросът за приетите критерии за представителност и балансираност на Корпуса. Представят се функциите на системата за разширено търсене в Корпуса и се привеждат примери за извличането на конкорданси и колокации с помощта на тази система. БНК дава възможност за провеждането на разнообразни по характер изследвания в различни области като лексикографията, компютърната лингвистика, теоретичната лингвистика, обучението по български език и пр. В статията се представят някои предварителни наблюдения върху възможностите за използване на Корпуса при проучването на системата на българските местоимения.

✉ Светла Коева  
*svetla@dcl.bas.bg*  
✉ Диана Благоева  
*diablag@mail.bg*  
✉ Сия Колковска  
*sia\_btb@yahoo.com*

Публикувано: 30 септември 2011