

ЕЗИКОВИТЕ ТЕХНОЛОГИИ И РЕСУРСИ – НОВИ ПЕРСПЕКТИВИ В ОБУЧЕНИЕТО ПО БЪЛГАРСКИ ЕЗИК (БЪЛГАРСКАТА ЛЕКСИКАЛНО-СЕМАНТИЧНА МРЕЖА БУЛНЕТ И БЪЛГАРСКИЯТ НАЦИОНАЛЕН КОРПУС)¹⁾

Светла Коева, Светлозара Лесева, Ивелина Стоянова, Мария Тодорова
Институт за български език „Проф. Л. Андрейчин“, Българска академия на науките

Резюме. Статията представя подходи за прилагане на електронните езикови ресурси и технологии в обучението по български език. Активното интегриране на езиковите технологии в учебния процес допринася за повишаването на интереса към учебното съдържание, изграждането на умения за самостоятелна работа и развитието на изследователските умения на учениците. Представени са два основни езикови ресурса за български език – лексикално-семантичната мрежа БулНет и Българският национален корпус. Приложението им в обучението е демонстрирано с примерни задачи.

Keywords: education, Bulgarian language, language technologies, language resources

1. Технологиите в класната стая

В последните години обучението по български език е изправено пред предизвикателството да запази своята роля за изграждане на езиковата компетентност и книжовната култура на учениците, като отговори на динамичното развитие на технологиите, на интереса на децата и младежите към тях и на новите методи за достъп до информация и учене в дигиталната епоха. От определящо значение е да се поддържа както актуалността на преподаваните знания, така и ефективността и атрактивността на методите за тяхното представяне и усвояване, за да се осигури траен интерес към ученето и да се стимулират изследователските и аналитичните умения на учениците. Основните положителни страни на активното използване на езикови ресурси и технологии в обучението по български език са следните:

- Учебният процес може да стане по-ефективен и привлекателен.
- Предоставят се широки възможности за разработване на интерактивни уроци, задачи, упражнения, които могат да бъдат съобразени с потребностите на различни групи ученици, включително на ученици с различни увреждания.
- Ученикът има възможност да участва активно в изследователския процес на формулиране и проверка на езикови хипотези. Стимулират се сътрудничеството и работата в екип.
- Работата на учителя в клас, както и предварителната му подготовка, може да се улесни значително. Предоставя се възможност за привлекателно онагледяване на информацията и разнообразни начини за представяне на съдържанието.

В настоящата статия представяме два големи електронни езикови ресурса на Института за български език при Българската академия на науките – лексикално-семантичната мрежа БулНет и Българския национален корпус, които могат да намерят приложение в обучението по български език, като спомогнат за разширяването и обогатяването на учебното съдържание с нови знания и емпиричен материал за употребата на езика и нови възможности за внедряване на изследователските методи в обучението.

2. Приложение на БулНет в обучението по български език

2.1. Кратко представяне на БулНет

БулНет е лексикално-семантична мрежа за български език, създадена по модела на принстънския УърдНет (Милър, 1995; Фелбаум, 1998). Словното богатство на езика е представено чрез система от отношения, които го организират по подобие на начина, по който е организирана понятийната система и свързаната с нея лексикална информация. Основната единица на БулНет е синонимното множество (синонимна група, синсет, англ. *synset*), на което съответства отделно понятие, характеризирano в лексикално-семантичната мрежа посредством лексикално значение, примери и семантични отношения с други множества. Основополагаща концепция за лексикално-семантичната мрежа е, че системата от понятия е в голяма степен езиково независима, поради което могат да бъдат създадени взаимно свързани мрежи за различни езици. Синонимните множества в БулНет са съотнесени с принстънския УърдНет, а посредством него и с други езици (например балканските езици), с релации на еквивалентност между синонимните множества – преводни еквиваленти, които обозначават едно и също понятие. Това превръща БулНет в богат многоезиков ресурс, който има многобройни приложения при превода, включително автоматичния превод, езиковедските теоретични и съпоставителни изследвания и обучението както по български език, така и по други езици.

БулНет съдържа около 250 хиляди лексикални единици за български език, структурирани в повече от 121 хиляди синонимни множества, като над 134 хиляди лексикални единици са създадени и проверени от експерти, а останалите са автоматично създадени и подлежат на проверка.

Синонимните множества са свързани помежду си с над 250 хиляди връзки, които включват семантични, морфосемантични, деривационни и екстралингвистични релации. Подробно описание на типовете релации, тяхната семантика и свойства са представени от Коева (2014а: 154 – 173). По-долу са представени част от релациите.

I. Семантични релации

– еквивалентност (синонимия): *кола – автомобил, лингвистика – езикознание, говоря – приказвам, вървя – ходя*;

– род – вид (хиперонимия – хипонимия): *лекар – педиатър, птица – чапла, говоря – шепна, ям – лапам*. Разновидност на хиперонимно-хипонимните отношения при съществителните е релацията между хипероним и хипоним – уникален обект, изразен със собствено име: *писател – Радичков, река – Марица*;

– цяло – част (холонимия – меронимия) – отношение между цяло и част от него: *сграда – стена, книга – корица*; отношение между множество и член на множеството: *гора – дърво, футболен отбор – футболист*; отношение между вещество, материал и под. и негова съставка: *бронз – мед, хляб – брашно*;

– противоположност (антонимия): *красив – грозен, брат – сестра, влизам – излизам*;

– причина – следствие (каузира – каузиран): *полагам – лежа, убивам – умирам*;

– събитие – темпорално включено събитие (събитие – субсъбитие): *спя – хъркам, купувам – плащам*;

– свойство – проява на свойството (има атрибут – е атрибут): *красота – красив, красота – грозен*;

– релация на подобие между глаголи (глаголна група): *пътувам – пътувам*;

– релация на подобие между прилагателни (подобен на): *красив – хубав, красив – великолепен*;

– допълнителни връзки към подобни по значение прилагателни или глаголи („виж още“): *красив – привлекателен, режа – изрязвам*.

II. Словообразователни релации (деривационно свързване):

– съществително – производно относително прилагателно: *антропология – антропологически, медицина – медицински*;

– прилагателно – производно наречие: *красив – красиво*;

- словообразователно свързани съществителни и глаголи: *пиша – писател*;
- глагол – производно причастие: *играя – игран*.

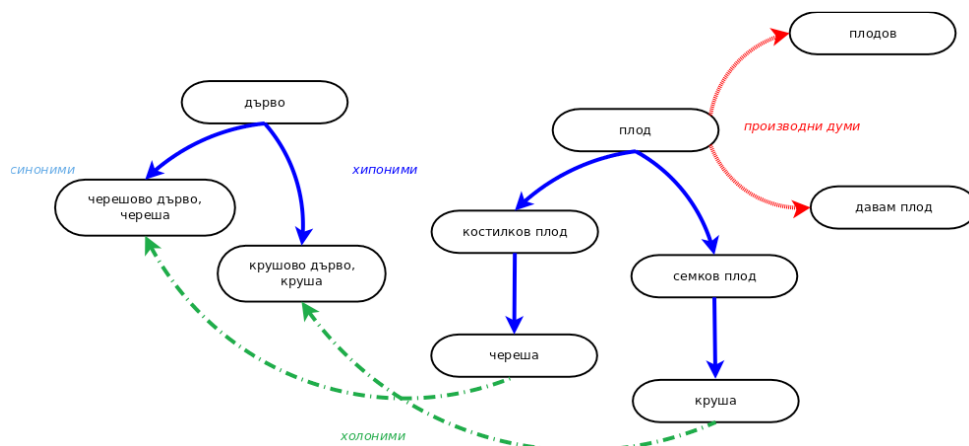
III. Морфосемантични релации

Морфосемантичните релации са семантични отношения, произтичащи от семантиката на съответната деривационна връзка (Коева, 2014: 167). Включените морфосемантични релации са между глаголи и съществителни, които изразяват различни „роли“ по отношение на глаголите:

- одушевен волеви вършител (агент): *пиша – писател, ора – орач*;
- средство, приспособление и под., с чиято помощ вършителят извършва действието (инструмент): *перфорирам – перфоратор, четкам – четка*;
- неодушевен агент или причина (с помощта на): *освежавам – освежител*;
- вещество, материал и под., посредством което се извършва действието (материал): *изолирам – изолатор*;
- превозно средство, чрез, в, от което се извършва действието (превозно средство): *бомбардирам – бомбардировач*;
- част от тялото, с която се извършва действието (агент част на тялото): *жиля – жилище*;
- обект, включително одушевен, на действието (засегнат обект): *интервюирам – интервюиран*;
- действие или събитие, което произтича от действието (събитие): *организирам – организиране*;
- свойство, произтичащо от действието, състоянието и под. (свойство): *привлека – привлекателност*;
- резултат или следствие от действието (резултат): *втечнявам – течност*;
- състояние, произтичащо от действието, състоянието и под. (състояние): *превъзхождам – превъзходство, удивя – удивление*;
- обект, чиято функция или предназначение е да се извърши действието (използване): *червя – червило, пудря – пудра*;
- мястото, където се извършва или състои действието, състоянието и под. (местоположение): *кова – ковачница, пека – пекарна*;
- мястото или обектът, който е крайна цел на действието (дестинация): *адресирам – адресат*.

Фигура 1 представя графично някои от семантичните релации в БулНет (синонимия, хиперонимия, холонимия, меронимия), както и деривационни релации. Онагледена е думата *плод* и нейни хипоними и словообразователно свързаните с нея синонимни множества *плодов* и *давам плод*. Думата *череша*, от своя страна, е представена в две от значенията си – като хипоним на *плод* и на *дърво*.

БулНет съчетава предимствата на лексикално-семантичните мрежи (богата мрежа от различни по семантика и свойства връзки между синоним-



Фигура 1. Представяне на семантични отношения между синонимни множества

ните множества) с тези на тълковните речници (подробна и експертна дефиниция и примери за употребата), специализираните речници (термини и друга специализирана лексика), словообразователните речници (словообразователни и морфосемантични релации), дву- и многоезичните речници (връзка с лексикално-семантичните мрежи за други езици). Това го прави ценен ресурс с редица приложения в обучението по български език²⁾.

2.2. Приложение на БулНет в обучението по български език

Приложението на БулНет в обучението по български език може да бъде насочено към обогатяването на речника на учениците, усвояването на умения за уместно използване на лексикалните средства на българския език, придобиването на знания за системните отношения между лексикалните единици, придобиването на умения и навици за използване на компютърни речници при работата в клас и самостоятелната подготовка.

На първо място, БулНет може да се използва за проверка и усвояване на непознати значения. С помощта на лексикално-семантичната мрежа значенията на думите се овладяват по-лесно и ефективно, тъй като се предлага едновременно както достъп до информацията за дадено синонимно множество – синоними, тълковна дефиниция, примери за употреба и др., така и онагледяване на множеството от семантични отношения, в които синонимното множество участва.

Пример 1 показва няколко от значенията на думата *кола*, която е представена с 11 различни значения и съответно участва в 11 различни синонимни множества. Думи, които имат повече от едно значение, са подходящи за изследователски задачи, свързани с разпознаване на отделните значения в различен контекст, а могат да бъдат използвани и в гатанки, кръстословици и игрословици.

Пример 1. Думи с повече от едно значение (многозначни думи и омоними).

кола:1 (тонизираща газирана безалкохолна напитка, подправена с екстракт от ядките на дървото *кола*)

кола:2; каруца:1; талига:1 (открит дървен фургон за превоз на товари, теплен от кон или друго товарно животно)

автомобил:1; кола:5; моторно превозно средство:1 (сухопътно безрелсово превозно средство на колела, със собствен двигател и система за управление).

Една от насоките за разширяване на знанията и комуникативните умения на учениците в областта на лексиката, които БулНет позволява, е системната полисемия в езика, възникваща чрез метафоричен или метонимичен пренос на значения. Част от примерите включват: (а) естествени обекти и артефакти с аналогично предназначение, например *крило:4* (орган за летене) и *крило:3* (на летателен апарат), *чело:2* (част от тялото) и *чело:5* (работна повърхност на инструмент); (б) артефакти от различен тип, например елементи от конструкцията на сгради и на превозни средства (*прозорец, покрив, врата* и много други); (в) организации, населени места, помещения и др. и множеството от хора в тях (*село, град*); (г) институции и сградите, в които се помещават (*университет, болница, училище*); (д) съдове, помещения и др. и тяхната вместимост (*шиши, чаша, лъжица, кола*) и др.

БулНет позволява изследването на значението и употребата на различни съставни лексикални единици като фразеологизми, съставни думи, съставни термини и др. Търсенето в лексикално-семантичната мрежа може да се основава както на цялата съставна лексикална единица, така и на някой от компонентите ѝ, което позволява намирането на фразеологизми или съставни думи, съдържащи определени думи, и запознаване със значенията им, както е илюстрирано в пример 2. Представено е едно от осемте глаголни синонимни множества, съдържащо фразеологизъм с думата *очи*.

Пример 2. Синонимно множество, съдържащо съставни лексикални единици.

следа зорко:1; не изпускам от очи:1

Дефиниция: държа под строго наблюдение

Част на речта: v (глагол)

Пример за употребата: *Не изпускайте от очи затворника.*

хипероним: внимавам:2; следя:7; оглеждам се:1; огледам се:1; озъртам се:1; озърна се:1.

БулНет предоставя примери за употребата на лексикалните единици, които могат да се използват в областта на семантиката и лексикологията за илюстриране на многозначността, омонимията и отделните значения. Пример 3 представя употребата на думата *глас* в различни примери с няколко различни значения, извлечени от лексикално-семантичната мрежа.

Пример 3. Употреба на дума в различни значения.

глас:1. *Зад нас прозвуча пронизителен глас.*

глас:3; вот:3; гласувал избирател:1. *Надявах се на повече гласове на изборите.*

глас:6. „Таймс“ не е гласът на Ню Йорк.

Множеството от релации между синонимните множества в БулНет предоставя богата информация, която може да бъде използвана при изясняването на значенията на думите и обогатяването на речника на учениците с нови думи и значения, при откриване и наблюдения върху системните отношения в лексиката на езика – синонимия, антонимия, хиперонимия, меронимия и други, както и при словообразователен анализ.

2.3. Примерни задачи, основани на БулНет

Представената в БулНет информация дава възможност за конструиране на различни типове изследователски задачи.

Задачите за правилно разпознаване на семантичните отношения между думите се основават на понятийните отношения, които са отразени в лексикално-семантичната мрежа посредством синонимните множества и релациите между тях.

Задача 1. Глаголите *тананикам* и *нея* се отнасят помежду си така, както:

А. *куцам* и *ходя*; Б. *влизам* и *излизам*; В. *пълзя* и *лазя*.

Отношението между *тананикам* и *нея* е между по-конкретно и по-общо значение, т.е. между хипоним и хипероним (глаголът *тананикам* означава специфичен начин, по който се извършва действието, спрямо глагола *нея*). Същото отношение се изразява от глаголите в подточка А. Отговор Б. и отговор В. са неправилни – *влизам* и *излизам* са антоними (изразяват противоположна посока на движение), а *пълзя* и *лазя* са синоними.

Задача 2. Съществителните *писател* и *биограф* се отнасят помежду си така, както:

А. *къща* и *стая*; Б. *къща* и *дом*; В. *сграда* и *къща*.

Релацията между *писател* и *биограф* е отношение между хипероним и хипоним. На това условие съответства отговорът в подточка В. – *къщата* е вид *сграда*. Другите два отговора не са верни. Отношението между *къща* и *стая* е

отношение между цяло и негова част (холонимия – меронимия), а *къща* и *дом* могат да бъдат синоними.

Задача 3. Посочете поне по един синоним и един антоним на думата:

А. бедност; Б. победа; В. късмет.

Възможни отговори:

А. мизерия, недоимък, нищета, сиромашия и др.; богатство, охолство, благосъстояние и др.

Б. успех, триумф, тържество и др.; неуспех, провал, крах и др.

В. шанс, сполука, щастие и др.; несполука, малишанс и др.

Тъй като повечето включени думи имат повече от едно значение, е възможно посочването на други, подходящи за определено значение, синоними и антоними.

Задачите за включване или изключване на дума по отношение на група от думи на базата на свързващите ги семантични отношения могат да се основават на структурата на лексикално-семантичната мрежа.

Задача 4. Коя от думите във всеки един от редовете е излишна?

А. летателен апарат, делтапланер, самолет, хеликоптер, дирижабъл;

Б. птица, бозайник, насекомо, риба, ято;

В. пейзаж, стенопис, живопис, натюрморт;

Г. предно стъкло, покрив, врата, прозорец, автомобил.

В подточка А. думите *делтапланер*, *самолет*, *хеликоптер* и *дирижабъл* представляват видове летателни апарати и са хипоними на *летателен апарат*, надредно понятие спрямо останалите и следователно е излишната дума. В подточка Б. *птица*, *бозайник*, *насекомо* и *риба* са представители на биологични класове, докато *ято* е множество от птици. В подточка В. излишната дума е *живопис*, тъй като означава вид изобразително изкуство, а *пейзаж*, *стенопис* и *натюрморт* са жанрове в живописа. В подточка Г. *предно стъкло*, *покрив*, *врата* и *прозорец* са мероними (съставни части) на автомобил (холоним). Може да се обърне внимание, че между *ято* и *птица* и *автомобил* – *покрив*, *предно стъкло*, *врата*, *прозорец* съществува холонимно-меронимна релация от различен тип (множество – член на множество в първия случай и цяло – неделима съставна част във втория).

Задача 5. Коя от думите в третата колона може да се добави към всяка от предните две колони с някое от значенията си?

тъкачница	белачка	пералня
ковачница	отварачка	горелка
книжарница	сеялка	пръскалка

Думите в първата колона означават мястото, където се извършва определен вид действие – *тъкане*, *коване*, *продаване на книги*, а думите във втората колона – инструмент за извършване на определено действие – *белене*,

отваряне, сеене. Думата *белачка* е многозначна и може да означава и *жена*, която бели. Думите *отварачка* и *сеялка* помагат да се прецени, че в колоната са включени инструменти, а не дейтели. Правилният отговор е *пералня*, тъй като е единствената дума в третата колона, която означава и мястото, и инструмента, средството (*перална машина*).

Задачите, свързани със словообразуване и значението на словообразователните средства, са насочени към познаване на морфемния строеж на думите и изследване на отделните деривационни стъпки при конструиране на дадено значение, както и към обобщения за многозначни, синонимни и омонимни морфеми.

Задача 6. В кой от редовете словообразователно свързаните думи не изразяват едно и също отношение:

А. *пиша* – *писател*, *чистя* – *чистач*, *стрелям* – *стрелец*, *издам* – *издайник*, *адресирам* – *адресант*;

Б. *уча* – *учител*, *пръскам* – *пръскалка*, *мета* – *метла*, *пера* – *пералня*, *нагрявам* – *нагревател*.

В подточка А. всички двойки представят отношение между глагола и неговия вършител, което е изразено от наставките на съществителните *-тел*, *-ач*, *-ец*, *-ник*, *-ант*. В подточка Б. първата двойка е пример за агентивно (дейтелно), а останалите 4 двойки – за инструментално значение, изразено от наставките *-(а)лка*, *-ла*, *-(а)лня*, *-тел*. Може да се посочи, че една и съща наставка може да има различно значение, например *-тел* в *писател* и *учител* в сравнение с *нагревател*. С помощта на подобни задачи се обръща внимание на морфемния строеж на думите и едновременно с това се акцентира върху значението, което носят определен тип морфеми (в случая еднаквото значение на различни наставки), и върху различните значения на една и съща морфема.

Задача 7. Допълнете всеки от редовете, така че да се получи поредица от думи, които са производни една от друга и изразяват същите отношения както думите в останалите редове. Подредете думите във всеки ред, така че всяка следваща дума да е непосредствено образувана от предходната.

А. *подвързвам*, *подвързвачница*, ...;

Б. *пека*, *пекар*, ...;

В. *тъкачница*, ...,

Ако се обединят наблюденията от А., Б. и В., се вижда, че могат да се извлекат тройки от (а) глагол, (б) съществително, означаващо лице деятел, производно от глагола, и (в) съществително, обозначаващо мястото на извършване на действието, производно от агентивното съществително. Получените редове са *подвързвам* – *подвързвач* – *подвързвачница*; *пека* – *пекар* – *пекарна* или *пекарница*; *тъка* – *тъкач* – *тъкачница*. Може да

се обърне внимание на различните наставки, с помощта на които се произвеждат аналогични значения – *-ач, -ар* (за вършител) и *-ница, -на* (за място на извършване), както и на факта, че в някои случаи по различни словообразователни модели възникват синонимни значения – *пекарна* или *пекарница*.

Задача 8. Допълнете всеки от редовете с думи, производни на посочената дума, които изразяват посочените отношения, като всяка следваща дума е образувана от предходната:

А. *бръсна*; ДЕЯТЕЛ; ПРОФЕСИЯ;

Б. *меля*; МЯСТО; ДЕЯТЕЛ;

В. *уча*; ДЕЯТЕЛ; ПРОФЕСИЯ;

Г. *бръсна*; ДЕЯТЕЛ; МЯСТО.

Получават се следните тройки *бръсна – бръснар – бръснарство*; *меля – мелница – мелничар*; *чета – учител – учителство*; *бръсна – бръснар – бръснарница*.

Те илюстрират морфемния строеж и значението на наставките и различната последователност на елементите в даден словообразователен модел. При *меля – мелница – мелничар* съществителното, означаващо деятел, се образува от съществителното, означаващо мястото, където се върши действието, а при *бръсна – бръснар – бръснарница* (и аналогичните *тъка – тъкач – тъкачница*, *пека – пекар – пекарна*) – последователността е глагол – деятел – място.

БулНет има широко приложение в задачи за редактиране на текст, които са свързани със: (а) заместване на повторени думи и изрази; (б) заместване на (стилово) неуместни думи и словосъчетания. Подходящи кандидати за заместване при повторения могат да бъдат синоними, хипероними или хипоними, холоними или мероними, антоними и др. на повторената дума или словосъчетание.

Задача 9. Намерете подходящо съответствие на подчертаните думи и изрази.

*Всеки човек има права и задължения като гражданин на обществото – действия, които са ни позволени, и действия, които **не са ни позволени**. Така например, **забранено е** да се продава алкохол на **хора** под 18 години.*

Решението на задачата може да бъде улеснено чрез проверка в БулНет за синоними или антоними на повторената дума *позволени*, съответно – *забранен* или *разрешен*. В резултат от това, тази част от изречението може да бъде редактирана по следния начин: ... и действия, които **не са разрешени** ... или ... и действия, които **са забранени**. Въз основа на синонимите в БулНет думата *забранено* от второто изречение може да бъде заменена с *незаконно*, а *хора* – с *лица*: Така например, **незаконно е** да се продава алкохол на **лица** под 18 години.

3. Българският национален корпус в езиковото обучение

3.1. Представяне на Българския национален корпус

Българският национален корпус (БНК) е най-големият корпус с фокус върху български език, в който корпусните единици са снабдени с подробни метаданни и разнообразна лингвистична анотация (Коева и др., 2011; Коева и др., 2012; Коева, 2014б). Българоезичната част съдържа над 240 хиляди текстови единици, които наброяват повече от 1,2 милиарда думи. Корпусът съдържа също така 48 паралелни корпуса с различна големина, наброяващи общо над 4,2 милиарда думи. Всички текстове на чужд език имат съответен български текст, като най-големи са паралелните корпуси за английски, румънски и полски език.

В БНК се използва унифициран модел за изграждане на корпуса и обработка на текстовите единици, независимо от езика, категорията на текста и т.н. Събирането на текстове се извършва основно автоматично, което осигурява възможност за бързо обогатяване и разширяване на корпуса (Коева и др., 2012). Корпусът осигурява представителност на езиковите данни чрез покритие на разнообразни категории текстове, представящи многообразието на българския език през XX и началото на XXI век.

Всички текстови единици са снабдени с подробно описание на метаданни на български и английски език, което позволява лесното им класифициране, извличане на статистически данни за корпуса или на подкорпуси за специфични цели. Документите са организирани в мрежа от категории: стил (обща категория на текста), жанр, тематична област, автор, година на създаване и др., които характеризират изчерпателно текстовете.

Корпусът е анотиран на различни нива на лингвистична информация. Българските и английските текстове са токънизирани, разделени на изречения, анотирани с информация за част на речта, лема и граматични характеристики, както и със значенията от БулНет и принстънския УърдНет. Части от паралелните корпуси са съотнесени със съответните текстове на български език. Корпусът предлага няколко възможности за достъп. Някои текстове, към които няма ограничение за разпространение от авторски права, се предоставят за свободно изтегляне (<http://dcl.bas.bg/BulNC-registration/>). Основният начин на достъп е чрез подаване на заявки към онлайн системата за търсене (<http://search.dcl.bas.bg/>). Тя позволява заявки по леми или словоформи и (наредени или ненаредени) комбинации от леми или словоформи и техни граматически характеристики. Пример 4 показва заявка за търсене на наредена двойка от произволно прилагателно и която и да е форма на съществителното *решение*, заедно с примерни резултати, които включват нечленуваните форми *решение* и *решения*. Пример 5 показва допълнител-

ната семантична информация, приписана от БулНет, както и граматична информация.

Пример 4. Заявка за наредена последователност от произволно прилагателно (*{POS=A}), следвано от произволна форма на думата *решение* (*решение/F/*).

<{*{POS=A} решение/F/>³⁾

Намерени: 14689:

Не търсете прекалено лесни решения.

Изглежда, напоследък не съм в състояние да взимам правилни решения.

Естествено ще ми се противопоставя при всяко важно решение.

Няма смисъл да търсят компромисно решение с благородниците от Артоа.

Пример 5. Семантична и граматична информация за думи от резултатите. Не търсете прекалено лесни решения.

лесни:

A__0f pf_ (прилагателно име, нечленувана форма, форма за множествено число,)

- който не представя трудности и не изисква големи усилия (*лесна слава*)
- който се намира в изобилие или се придобива без голям труд (*лесни пари*)
- който може да бъде излъган, подведен, използван, нападнат и т. н. без особени усилия или без да окаже съпротива (*лесна плячка*)

3.2. Приложения на Българския национален корпус в обучението по български език

Българският национален корпус позволява извършването на наблюдения върху разпространението и функционирането на определени лексеми, техни форми, употреби. Друга посока на използване на БНК е извличането на информация за лексикалната и граматичната съчетаемост, значението и употребите на лексеми, форми, фразеологизми, изследването на синтактични модели и т.н. Не на последно място БНК има приложение като източник за подбор на автентични илюстративни примери, които да онагледяват езиковите факти и явления.

3.3. Създаване на задачи, основани на БНК

Задачите за изследване на различните значения и употреби на думи и изрази включват наблюдения върху примери от различни стилове, жанрове и тематични области на БНК. Търсенето може да става по форма на думата или по лема. Тези задачи се основават на предимството, че на всички думи в БНК автоматично са приписани съответстващите им значения от БулНет.

Задача 10. Посочете най-често срещаните значения на *лице* в административния стил и публицистичния стил. Обяснете съответните значения и ги илюстрирайте с пример.

За решението на задачата е необходимо да се зададе съответният функционален стил, в рамките на който се осъществява търсенето. Това става с помощта на бутона Мета и попълване на полето Стил – административен или публицистичен. Търсенето показва, че в административния стил *лице* най-често се използва със значение на човек (като член на обществото). Характерни съчетания са *физическо лице*, *юридическо лице*, *длъжностно лице* и др. В публицистичния стил това също е едно от често срещаните значения, но преобладава значението 'предната част на главата при човека' – *гримираното лице*, *красиво лице*, *изразително лице*.

Задача 11. Посочете значенията на изрази *подпиша смъртната присъда* и ги подкрепете с примери.

При решението на задачата може да се използват примери, извлечени от БНК с помощта на следната заявка: <подпиша/F/|подписвам/F/ [0,3] смъртната [0,3] присъда>, с която се открива последователност от произволна форма (/F/) на *подпиша* или *подписвам*, следвана от 0 до 3 произволни думи, следвани от формата *смъртната*, следвана от 0 до 3 думи, следвани от формата *присъда*.

Задачите за изследване на съчетаемостта между думите изграждат умения за анализ върху езиков материал и конструиране на изводи за езиковата употреба.

Задача 12. Попълнете многоточието, така че да се получи смислово вярно и граматически правилно изречение. Обяснете разликата в значението на глагола *вярвам* в отделните изречения.

А. А. Действително Джош не е отявлен атеист, все пак вярва Бог и ходи на църква на Великден, Коледа и още два-три пъти годишно.

Б. Ако се вярваше Грегорович, кражбата е била извършена преди много години.

В. Той трябва да бъде уверен в собствените си способности. Трябва да бъде с устойчива психика и да вярва себе си.

Г. Вярвам истинността на разказа ти.

В А. глаголът *вярвам* изисква предлога *в*, в Б. – предлога *на*, а във В. и Г. – предлога *в*. Значението, с което е употребено съчетанието в А., е 'изповядвам определени религиозни или идеологически убеждения'. Съчетанието с на (Б.) означава 'имам вяра в достоверността или истинността на нещо, уповавам се на някого за нещо'. Във В. значението е: 'убеден съм в способностите, възможностите и под. на някого или в полезността или ефикасността на нещо', а в Г. – 'приемам нещо за истинно'. При обсъждането

на задачата може да се обърне внимание на близостта на значенията и на ролята на предлога.

БНК може да се използва и за задачи, свързани с извличане на граматични обобщения въз основа на примери от корпуса.

Задача 13. Кои глаголи не образуват сегашно деятелно причастие и деепричастие?

- А. непреходните глаголи;
- Б. преходните глаголи;
- В. глаголите от свършен вид;
- Г. глаголите от несвършен вид.

Отговорът на задачата може да бъде основан на анализ на извадка от корпуса при заявка за сегашно деятелно причастие или деепричастие: $*\{IVF=y\} | * \{IVF=z\}$:

Има само хора, **убиващи** хора. (V IM _ T)

Съвсем тихо, **ходейки** на четири крака, аз се приближих. (V IM _ IN)

Протегна ръка, **опипвайки** въздуха. (V IM _ T)

Пресичащи деца! (V IM _ IN)

Друг вариант на задачата е учениците самостоятелно да направят заявка, например за търсене на сегашно деятелно причастие при глаголи от свършен вид $*\{VA=PE\ IVF=y\}$ (0 срещания) и от несвършен вид $*\{VA=IM\ IVF=y\}$ (140 554 срещания), за да достигнат до съответните изводи. Търсенето по граматични характеристики може да бъде подпомогнато чрез избор на част на речта и граматични характеристики от бутона Асистент.

4. Вместо заключение: други приложения на езиковите ресурси и технологии в езиковото обучение

Комбинирането на знанията, представени в БНК и БулНет, дава още по-широки възможности за приложение в езиковото обучение. БулНет предоставя информация и за преводните еквиваленти на думите (пример 6) и може да бъде използван като двуезичен речник за всички двойки между български, английски и румънски. Наличието на примери за употребата, както и допълнителни семантични отношения дава по-подробна информация от обикновен двуезичен речник, с което БулНет става подходящ като учебно помагало в чуждоезиковото обучение, както и за съставяне на учебни задачи.

Пример 6. Преводни еквиваленти на многозначна дума в различни значения.

bg - n: глас:1 **en** - n: voice:4 **ro** - n: voce:4

bg - n: глас:3; вот:3; гласувал избирател:1 **en** - n: vote:9; voter turnout:1 **ro** - n: vot:4

bg - n: глас:6 **en** - n: voice:1 **ro** - n: glas:2; voce:5

В паралелните корпуси на БНК могат да се проверяват преводните съответствия на конкретно значение в реален контекст и да се направи справка с тълковната дефиниция в БулНет.

Пример 7. Преводни съответствия в реален контекст.

bg: Такова единодушие в едно демократично и плуралистично общество е ясен показател за пълната обвързаност на Албания с тази кауза.

en: The cohesion of the government, opposition and the majority of the Albanian **public** on the Iraqi crisis illustrates the desire for Albanian foreign policy and diplomacy ...

БНК и БулНет позволяват разработването на задачи с изследователски характер, които да провокират интереса на учениците към наблюдението на езиковите явления. Илюстрация за такава задача може да бъде изследването на употребата и значението на непознати фразеологизми.

Задача 14. Митко играе с топка и без да иска, счупва прозорец. От къщата излиза един дядо и извиква: *Ей, веднага си обирай крушите, хлапе!* Какво иска да каже възрастният човек? Използвайте БулНет и БНК, за да проучите употребата и значението на съчетанието *обирам си крушите*.

БНК може да се използва, за да се провери употребата на израза. Намерените 144 срещания насочват към това, че вероятно изразът е фразеологично съчетание. От тях 69 (48%) са с повелителната форма на глагола – *обирай, обирайте*. 143 срещания са от текстове от художествената литература и едно – от журналистически текст. Анализът на контекста, в който се употребява изразът, подчертава неофициалния характер на съчетанието. За да видим значението му, можем да го потърсим в БулНет, където ще намерим синоними като *изчезвам, пръждосвам се*, както и дефиницията: *‘заминавам (от някъде) незабавно, обикновено по принуда, за да избегна неприятности или защото не съм желан’*.

За да запази своята привлекателност като учебен предмет и за да изгражда положително отношение към ученето, както и общи умения за изследване и анализ, обучението по български език може да се възползва от широките възможности, предоставени от езиковите ресурси и технологии. Настоящата статия представя примерни приложения на лексикално-семантичната мрежа БулНет и Българския национален корпус в задачи за класна и извънкласна работа и поставя въпрос за използването на електронни езикови ресурси в обучението по български език, което открива нови перспективи пред съвременното образование.

БЕЛЕЖКИ

1. Текстът е представен на Националния форум „Предизвикателства пред развитието и разпространението на българския език, литература и култура“;

12.04.2016 г., ЮЗУ „Неофит Рилски“ – Благоевград. Съорганизатори: ЮЗУ „Неофит Рилски“ – Филологически факултет, НИОН „Аз Буки“, Институт за български език – БАН, Асоциация на българските училища в чужбина.

2. БулНет е достъпен на адрес: <http://dcl.bas.bg/bulnet/>.

3. Подробна инструкция за използването на БНК и езика за заявки е на следния адрес: <http://dcl.bas.bg/bulnc/dostap/konkordansi/>

REFERENCES / ЛИТЕРАТУРА

Fellbaum, Christiane (1998, ed.) *WordNet: An Electronic Lexical Database*. Cambridge, MA: MIT Press.

Koeva, Sv. (2014a). WordNet i BulNet. V: Svetla Koeva (red. i sast.) *Ezikovi resursi i tehnologii za balgarski ezik*. Akad. izd. „Marin Drinov“, Sofia, 2014, 312 s., str. 154 – 173 [Коева, Св. (2014a). WordNet и БулНет. В: Светла Коева (ред. и съст.) *Езикови ресурси и технологии за български език*. Акад. изд. „Марин Дринов“, София, 2014, 312 с., стр. 154 – 173].

Koeva, Sv. (2014b). Balgarskiyat natsionalen korpus v konteksta na svetovnata teoriya i praktika. V: Svetla Koeva (red. i sast.) *Ezikovi resursi i tehnologii za balgarski ezik*. Akad. izd. „Marin Drinov“, Sofia, 2014, 312 s., str. 29 – 52 [Коева, Св. (2014b). Българският национален корпус в контекста на световната теория и практика. В: Светла Коева (ред. и съст.) *Езикови ресурси и технологии за български език*. Акад. изд. „Марин Дринов“, София, 2014, 312 с., стр. 29 – 52].

Koeva, Sv., D. Blagoeva & S. Kolkovska (2011). Proektat Balgarski natsionalen korpus – rezultati i perspektivi. – *Balgarski ezik*, 58 (2011), 3, str. 34 – 53 [Коева, Св., Д. Благоева & С. Колковска (2011). Проектът Български национален корпус – резултати и перспективи. – *Български език*, 58 (2011), 3, стр. 34 – 53].

Koeva, Sv., Iv. Stoyanova, Tsv. Dimitrova & Sv. Leseva (2012). Traditsii i novatorstvo v korpusnata lingvistika: Balgarskiyat natsionalen korpus. – *Spisanie na Balgarskata akademiya na naukite*, 2012, 3 [Коева, Св., Ив. Стоянова, Цв. Димитрова & Св. Лесева (2012). Традиции и новаторство в корпусната лингвистика: Българският национален корпус. – *Списание на Българската академия на науките*, 2012, 3].

Miller, George A. (1995). *WordNet: A Lexical Database for English*. Communications of the ACM, Vol. 38, No. 11: 39 – 4.

**LANGUAGE TECHNOLOGIES AND RESOURCES –
NEW ADVANCES IN BULGARIAN LANGUAGE TEACHING
(THE BULGARIAN LEXICAL SEMANTIC NET BULNET
AND THE BULGARIAN NATIONAL CORPUS)**

Abstract. The paper proposes approaches for the application of electronic language resources and technologies in Bulgarian language teaching. The active integration of language technologies in the teaching process leads to an increased interest in the subject, the formation of skills for individual work, and the development of research skills. We present in more detail two language resources for Bulgarian: BulNet – a lexical semantic network for Bulgarian, and the Bulgarian National Corpus. The application of these resources in language teaching is demonstrated through several types of linguistic problems.

✉ ¹⁾Prof. Dr. Svetla Koeva, ²⁾Dr. Svetlozara Leseva, Assist. Prof.,
³⁾Dr. Ivelina Stoyanova Assist. Prof., ⁴⁾Dr. Maria Todorova

Department of Computational Linguistics
Institute for Bulgarian Language “Prof. Lyubomir Andreychin”

Bulgarian Academy of Sciences
52, Shipchenski prohod Blvd., building 17,
Sofia, Bulgaria

E-mail: ¹⁾svetla@dcl.bas.bg, ²⁾zarka@dcl.bas.bg, ³⁾iva@dcl.bas.bg, ⁴⁾maria@dcl.bas.bg