

Секция «Лингвистика: Современные лингвистические исследования: фонетика, грамматика, лексика»

**Корпусные технологии в исследовании лексических подсистем языка (на примере этнонимии русского языка)**

**Научный руководитель – Рычкова Людмила Васильевна**

***Байраммырадов Гелди***

*Аспирант*

Гродненский государственный университет имени Янки Купалы, Факультет истории, коммуникации и туризма, Кафедра перевода и межкультурной коммуникации, Гродно, Беларусь

*E-mail: Bajrammyradov\_g@grsu.by*

Бурное развитие корпусных технологий в последние десятилетия привело к коренным изменениям в области лингвистики, за которыми последовали изменения и в исследованиях лексических подсистем языка.

Национальный корпус русского языка (далее - НКРЯ) [2] предоставляет опцию поиска по лексико-семантическим признакам, основанную на семантической разметке текстов (особенности семантической разметки НКРЯ наиболее подробно описаны в [Кустова 2005: 155]). Такая разметка предназначена для того, чтобы обеспечить возможность поиска языкового материала, включающего лексические единицы, объединенные общностью определенных семантических признаков.

Например, для исследования такой лексической подсистемы русского языка, как этнонимика, целесообразно воспользоваться семантическим признаком «этнонимы» (см. рисунок 1).

Следует помнить, что разметка в корпусе осуществляется автоматически, поэтому ошибки в выдачах неизбежны, например, фамилии, омонимичные этнонимам: *Хант*, *Росс* и др. (см. рисунок 2).

Очевидно, тем не менее, что объем выдачи - 15357 вхождений в 3470 документах (см. рисунок 3) - достаточен, чтобы не только выявить этнонимы, функционирующие в современной русской речи (*англичанин, араб, венгр, викинг, германец, грек, грузин, грузинка, датчанин, еврей, еврейка, индеец, ирландец, ирландка, испанец, итальянец, калмык, литовец, мексиканец, немец, немка, норвежец, полинезиец, поляк, румын, русский, серб, славянин, сомалиец, татарин, турок, фламандец, француз, эскимос, эстонец, японец* и др.), но и выявить особенности их функционирования непосредственно в контекстах.

Даже небольшой список лексем, приведенной выше, отражает системность этнонимов: противопоставленность по семантике пола (*грузин - грузинка, еврей - еврейка, ирландец - ирландка, немец - немка*), отношения типа «род-вид» (*славянин - поляк, русский, серб; германец - англичанин, датчанин, немец, норвежец*).

Богатство контекстов и возможность получения статистических данных, которую предоставляет НКРЯ, дают дополнительные возможности для исследования этнонимии русского языка. Таким образом, корпусные технологии, позволяющие получать репрезентативные и объективные данные, могут и должны применяться в лингвистических исследованиях лексических подсистем языка.

**Источники и литература**

- 1) Кустова Г.И., Ляшевская О.Н., Падучева Е.В., Рахилина Е.В. Семантическая разметка лексики в Национальном корпусе русского языка: принципы, проблемы, перспективы // Национальный корпус русского языка: 2003-2005. Результаты и перспективы. М., 2005, С.155–174.
- 2) Национальный корпус русского языка: <http://www.ruscorpora.ru>

### Иллюстрации

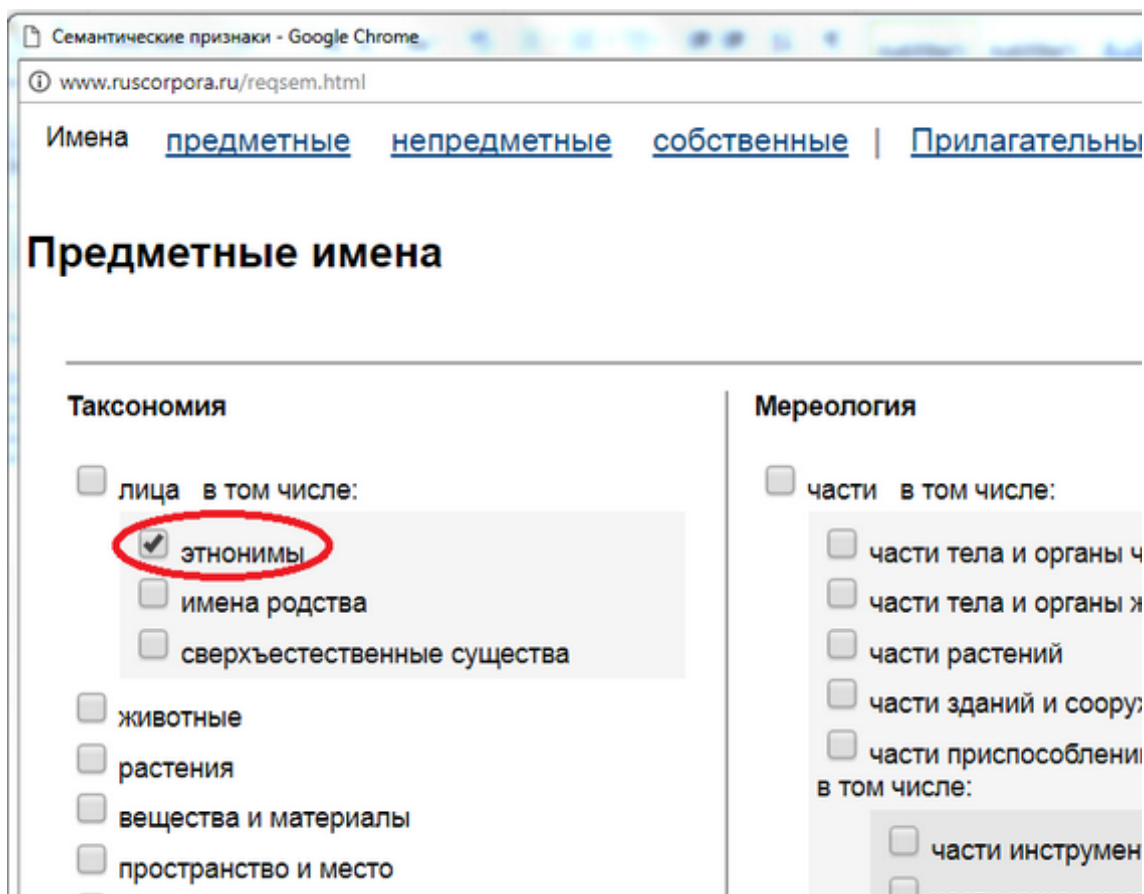


Рис. 1. Окно выбора семантических признаков

Думай, как хакер // «Computerworld», 2004 [омонимия снята] [Все примеры \(1\)](#)

По словам генерального менеджера Dimension Data Стива **Росса**, задача курса — научить его слушателей "думать, как хакеры".  
[омонимия снята] ←...→

Рис. 2. Окно выбора семантических признаков

Найдено 3 470 документов, 15 357 вхождений.

[Распределение по годам](#) [Статистика](#)

Поискать в других корпусах: [акцентологическом](#), [газетном](#), [диалектном](#), [мультимедийном](#), [обучающем](#), [параллельном](#), [поэтическом](#), [синтаксическом](#), [устном](#).

Страницы: [1](#) [2](#) [3](#) [4](#) [5](#) [6](#) [7](#) [8](#) [9](#) [10](#) [11](#) [следующая страница](#)

1. коллективный. Форум: Были вы в стране преподаваемого языка? (2008-2011) [\[омонимия снята\]](#) [Все примеры \(12\)](#)

[веса, nick] Но 4 года жила в Греции и на Кипре и говорила только на английском и часто с [англичанами](#), затем год работала в американской компании, корпоративным языком был английский, соответственно говорила только на английском, затем 2,5 года работала в производственной компании, использовала английский и немецкий, когда прилетали иностранцы и на выставках. [коллективный. Форум: Были вы в стране преподаваемого языка? (2008-2011)] [\[омонимия снята\]](#) [←...→](#)

[patrol, nick] В любом городе, да и в любой деревне Германии все мойщики туалетов – [русские](#) (и не простые, а с дипломами). [коллективный. Форум: Были вы в стране преподаваемого языка? (2008-2011)] [\[омонимия снята\]](#) [←...→](#)

[patrol, nick] Ну и естественно, как и у всех [русских](#) в Германии — мойка туалетов по ночам. [коллективный. Форум: Были вы в стране преподаваемого языка? (2008-2011)] [\[омонимия снята\]](#) [←...→](#)

[patrol, nick] Дополню, что на «работу» с зарплатой 100 Евро/ Месяц выпускника российского ВУЗа (пусть хоть с красным дипломом), если и возьмут, то это будет преподавание на каких-нибудь курсах для иностранцев ([арабы](#), [поляки](#), [эстонцы](#)...). [коллективный. Форум: Были вы в стране преподаваемого языка? (2008-2011)] [\[омонимия снята\]](#) [←...→](#)

[patrol, nick] Кстати, не подумайте, что в Германии легко найти ночную работу уборщика (мойщика), существует жёсткая конкуренция со стороны, например, [литовцев](#), [эстонцев](#), [арабов](#) и т. д. [коллективный. Форум: Были вы в стране преподаваемого языка? (2008-2011)] [\[омонимия снята\]](#) [←...→](#)

Рис. 3. Фрагмент результата поиска