

Михаил Вячеславович Копотев
(University of Helsinki)

Использование электронных ресурсов в преподавании русского языка

Одна из первых научных работ Арто Мустайоки назвалась «Опыт составления частотной грамматики русских существительных» (Мустайоки 1973). В исследовании были сформулированы принципы создания частотной грамматики с помощью вычислительной машины и поставлены задачи, которые не потеряли актуальности до сих пор. Более того, их решение в силу развития компьютерной лингвистики становится реальностью только в последнее время. При всей широте научных интересов проф. А. Мустайоки применение компьютера для решения исследовательских и педагогических задач всегда остается одной из важных составляющих его научной деятельности. Результатом этого стало создание более десяти обучающих компьютерных программ (см. список в конце книги) и целого ряда статей и монографий, основанных на компьютерном подходе к изучению языка. В настоящее время на Отделении, возглавляемом проф. А. Мустайоки, продолжается работа по созданию компьютерных систем, призванных помочь преподавателям и исследователям русского языка в их работе. В настоящей статье кратко представлены два проекта, работа над которыми ведется в настоящее время на Отделении славянских и балтийских языков и литератур Хельсинкского университета. Оба доступны в интернете и, будучи независимыми друг от друга, могут быть использованы как по отдельности, так и в качестве составных частей единого учебного процесса.

I. ХАНКО (www.slav.helsinki.fi/hanco)

Несмотря на то, что мировая корпусная лингвистика имеет уже многолетнюю историю и значительный научный и практический потенциал, корпусная русистика по разным причинам начала развиваться значи-

тельно позже¹. Одним из первых проектов такого рода стала разработка под руководством проф. А. Мустайоки корпуса ХАНКО – Хельсинкского аннотированного корпуса русского языка².

Корпус ХАНКО представляет собой журнальные тексты на русском языке, в которых каждой единице приписана вся релевантная лингвистическая информация. Поскольку корпус создается на Отделении, тесно связанном с преподаванием русского как иностранного, было сразу ясно, что ресурсом будут пользоваться не только лингвисты, но и студенты, изучающие русский язык, преподаватели русского как иностранного и даже учителя и школьники. Это, разумеется, не значит, что разработчики избегают употребления сложных лингвистических терминов, однако выбор параметров осуществляется так, что собственно лингвистическая информация подается в максимально традиционной форме. Из этого вытекает основной принцип создания: направленность на максимальный охват грамматической информации, а не на объем материала. Задачей составителей стало предоставление для широкого использования аннотированного корпуса, содержащего достаточно точную грамматическую информацию, основанную на устоявшихся теоретических концепциях, которые приняты в известных лингвистических трудах и/или учебной литературе. Уже сейчас для пользователей доступен поиск по морфологическим и синтаксическим параметрам. В перспективе планируется добавить функциональные (семантические) характеристики, разработанные в рамках теории функционального синтаксиса (см. Мустайоки 2006). Понятно, что при таком тщательном подходе к аннотированию корпус объемом в несколько миллионов словоупотреблений потребовал бы колоссальных усилий при ручной обработке материала. Поэтому объем ХАНКО в настоящий момент всего 100 000 текстоформ.

Еще одна особенность ХАНКО является следствием «расширенного» подхода к грамматике: в корпусе предусмотрена возможность более чем одной интерпретации языковых фактов. Так например, в предложении (1) невозможно однозначно определить, является словоформа *Ронибони* формой именительного или творительного падежа. В

¹ Необходимо отметить взрыв интереса к созданию корпусов в русской лингвистике в 2000-ые гг. (см. подробнее Шаров 2003; Резникова & Копотев 2005; Копотев & Мустайоки 2008). Крупнейшим из них, безусловно, является Национальный корпус русского языка (www.ruscorpora.ru), разработчики которого с самого начала работы тесно сотрудничали с хельсинкской группой (см. Плунгян 2005).

² Это название, предложенное Арто, соотносится с географическим местом в Финляндии, исторически тесно связанным с восточным соседом: именно у мыса Ханко (ранее Гангут) произошло Гангутское сражение, в XIX веке на мысе располагались дачи русской аристократии, после зимней войны 1939–40 гг. эта территория была арендована Советским Союзом.

ХАНКО реализована возможность давать таким единицам двойную квалификацию.

- (1) Причем последние по времени свидетельства относятся к 70-м годам – «очевидец» вспоминает седого, как лунь, тяжелобольного старика, которого остальные заключенные почему-то называли **«Ронибони»**.

Наконец, последний принцип связан с последовательным выделением многокомпонентных единиц. Известно, что при машинной обработке текстов за исходную единицу часто принимается текстоформа, или набор знаков от «от пробела до пробела». Однако очевидно, что такой «орфографический» критерий не может служить надежным основанием для выделения словоформы (см., напр., Мельчук 1997, 198–199). При таком подходе определенная часть важной лингвистической информации теряется (сослагательное наклонение и сложное будущее время глагола, аналитические формы прилагательных и наречий, составные числительные и др.), что признается неизбежным допущением при автоматическом или полуавтоматическом аннотировании. Кроме того, обычно не учитываются и так называемые «эквиваленты слов», или фраземы типа «в течение», «так как», «несмотря на», «друг друга» и др. Все эти единицы получили отражение в создаваемом корпусе.

Известно, что ни одна грамматика не в состоянии дать ответы на все возникающие вопросы и никакой, даже самый хороший учебник не заменит работы с реальным языковым материалом. ХАНКО, как и другие аннотированные корпуса, предоставляет преподавателю и исследователю инструмент для решения бесчисленного количества конкретных учебных или исследовательских задач (Добрушина 2005; Национальный корпус 2005, 2007). Учитель с помощью корпуса может быстро собрать необходимый, а главное живой материал по определенной теме.

Корпус ХАНКО в этом смысле представляет собой удобный инструмент для работы особенно в иностранной аудитории в силу своей компактности, аккуратности аннотирования, актуальности текстов и не в последнюю очередь наличием интерфейса на английском языке, который для многих иностранных студентов является более известным чем русский.

Приведем несколько примеров использования корпуса при составлении заданий для работы в аудитории.

1. Для иностранного студента представляют определенные сложности русские приставочные глаголы, обозначающие разные фазы движения. С помощью корпуса изучение этой темы можно сделать существенно более интересным и динамичным. Так, преподаватель может быстро задать запрос «найти все глаголы, заканчивающиеся в инфинитиве

на *-ехать*», что даст обозримый список примеров (43 глагола), которые можно сразу использовать как материал для анализа разницы в приставочных образованиях: *поехать, приехать, уехать, выехать* и др. Более того, следуя идее Н.Р. Добрушиной (2005, 321–324), на основе полученного списка предложений можно просто и эффективно создать задания для контрольных или самостоятельных работ. Для этого необходимо всего лишь скопировать найденные примеры и удалить глагольные приставки с тем, чтобы студенты сами восстановили пропуски. См. ниже пример задания:

В следующих предложениях добавьте к глагольным формам подходящие по смыслу приставки:

1. Когда он **...ехал** из аэропорта в гостиницу, то обнаружил, что в чемодане явно не его вещи.
2. Он последовательно сбрасывал с политического Олимпа своих перспективных политических союзников, подозревая их в желании до срока **...ехать** в Кремль.
3. Я пытался убедить его скрыться, **...ехать** из Венгрии.

2. Корпус представляет собой прекрасный инструмент для так называемого «проблемного обучения» (*problem-based learning*; см. Кудрявцев 1991, Mathews-Aydinli 2007), суть которого состоит в том, что студенты в небольших группах самостоятельно решают конкретные мини-задачи, помогающие осваивать материал. Примером такого рода «проблемы» может стать тема «конкуренция творительного/именительного падежа в составном именном сказуемом». Запросы типа «именительный/творительный падеж существительного/прилагательного в составном именном сказуемом» выдают студентам обозримое количество примеров, которые можно анализировать, определяя условия употребления конкурирующих падежных форм. Например:

- (2) Другая **команда** – питерские чекисты, **лидером** которых является секретарь Совета безопасности Сергей Иванов.
- (3) Правда, **отечественными** остаются только пошив и моделирование, сырье преимущественно **импортное**.

3. Многоуровневая разметка корпуса открывает возможность сочетать поиск синтаксического, морфологического и лексического материала для изучения сложного, плохо описанного реального употребления конкретных морфологических форм. Так например, анализ глагольной пары «говорить-сказать» демонстрирует, что наиболее типичное употребление глагола *сказать* – форма *скажем*, используемая в

качестве вводной конструкции; при этом, однако, активно используются и формы прошедшего времени. Наиболее типичное употребление глагола *говорить* – форма настоящего времени в позиции сказуемого полного двусоставного предложения: *X говорит*. В составе вводной конструкции (*как говорят*) этот глагол употребляется гораздо реже. Ср. ниже типичные примеры:

- (4) А вот, **скажем**, КВН одинаково хорошо идет хоть по-русски, хоть по-азербайджански, хоть на смеси языков.
- (5) Она **говорит**: «Валленберг был для нас как мессия».

Приведенные примеры – лишь некоторые иллюстрации тех возможностей, которые открывает перед исследователями, преподавателями и студентами ХАНКО. Уже они показывают, каким удобным и многофункциональным инструментом является корпус.

II. KARTTU (www.slav.helsinki.fi/general/karttu)

Еще одним большим компьютерным проектом, тесно связанным с преподаванием русского языка в иностранной аудитории, является система прогрессивной проверки языковой компетенции KARTTU³. Тесты разрабатывались на Отделении несколько лет и с 2005 года введены в систему преподавания (Мустайоки 2003)⁴. В своей методической части они восходит к так называемым прогрессивным тестам (*progressive tests*), давно применяющимся, например, в учебных программах подготовки врачей. Основные особенности теста KARTTU, которые выделяют его из многочисленных способов тестирования языковых знаний, сводятся к следующему.

Задания для студентов выбираются из объемной базы данных, содержащей в настоящий момент около 3 тыс. заданий по разным темам: от лексической сочетаемости до места ударения и синтаксических конструкций. Для конкретного теста («сессии») преподаватель определяет, какое количество заданий по той или иной теме следует выбрать, контролирует общее количество заданий в сессии, время, отведенное на ответы, и другие параметры. При этом выбор конкретных вопросов из базы компьютер осуществляет автоматически.

Все задания в тесте представляют собой вопросы с несколькими вариантами ответов. Итоговый результат подсчитывается по специальной

³ Старое финское слово *karttu* значит ‘валёк, ребристый деревянный брус, используемый для стирки’. Однако в названии тестов оно является сокращением от *KARTTUvan kielitaidon testi* ‘тест прогресса языковых навыков’. Надо ли говорить, что это название тоже предложил Арто?

⁴ В разработке тестов принимали участие М. Мононен, М. Копотев, Е. Никкиля, А. Леонтьев, А. Никунласси, а также студенты Отделения.

формуле, учитывающей количество предложенных вариантов ответа. Так, задание на выбор одного из двух вариантов ответа имеет вес в два раза меньший, чем задание на выбор одного из четырех вариантов, поскольку даже при случайном ответе вероятность «правильного попадания» в первом случае равна 50%, а во втором – 25%.

Учащиеся проходят тестирование («сдают сессии») несколько раз в год, при этом все полученные результаты сохраняются в базе данных, что позволяет студенту уже во время второй сессии получить сведения о прогрессе своих языковых навыков. При этом результат содержит не только обобщенную оценку – балл по 100-бальной шкале, но и детальную картину ответов по всем темам, включенным в сессию. Кроме этого, каждый студент видит средний результат учащихся своего уровня, что позволяет ему сравнить собственные успехи с достижениями товарищей. Преподаватель, работающий со студентами определенного уровня обучения, получает возможность контролировать результаты своей деятельности как в целом по группе, так и по отдельным учащимся, что в конечном итоге позволяет оперативно корректировать учебную программу на основе анализа студенческих ответов. Приведу примеры существующих заданий по русскому языку.

1. Можно ли так сказать?

Там были и русские, и армяны.

- Да
- Нет

2. Выберите вариант с правильным ударением.

- Все сд́али экзамены.
- Все сдали́ экзамены.

3. Как **нельзя** сказать по-русски?

- блестящая победа
- одержать победу
- быть в победе
- победа над врагом

На рис. 1 можно видеть интерфейс программы с заданием на выбор правильного сочетания слов.

HELSENGIN YLIOPISTO

Slavistiikan ja baltologian laitos
Humanistinen tiedekunta

77 0 sek.

Valitse yhteyteen sopiva verbi.

Девушка ... на улице.

- перебежала
- отбежала
- вбежала
- выбежала

[Department main page](#) | [University homepage](#)
[Contact information](#) | [Feedback](#)
Copyright © 2000-2006 University of Helsinki. [All rights reserved.](#)

Рис. 1

На рис. 2 приведен пример графика, иллюстрирующий общий прогресс языковых навыков одного из наших студентов.

	Sessions	Points (max. 100)
1	Осень 2005	34.00
2	Зима 2006	54.00
3	Осень 2006	67.00
4	Зима 2007	77.00
5	Весна 2007	68.00
6	Осень 2007	73.00

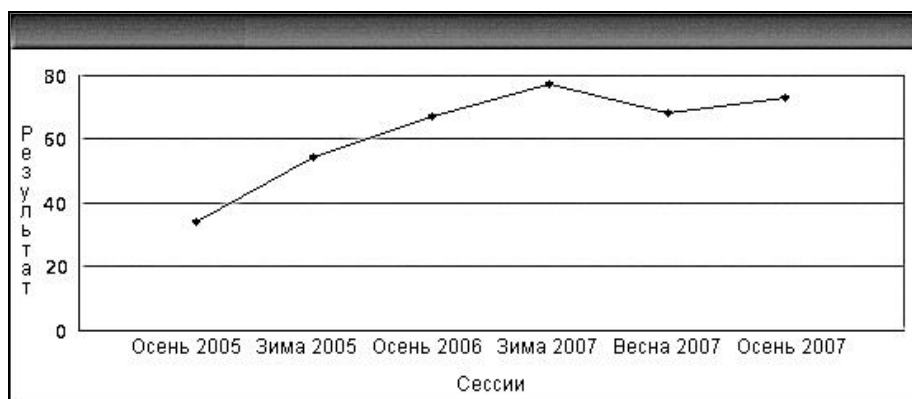


Рис. 2

Созданная на Отделении система тестирования KARTTU позволяет использовать ее как основу для создания самых разных тестов. В ней организовано удобное добавление заданий в существующий блок по русскому языку, а также возможность создания тестов по другим областям знаний. Тесты могут быть использованы как дистанционное средство проверки знаний и преподавателями вузов, и школьными учителями. При подготовке новых заданий по разным темам они могут стать прекрасной основой для создания систем проверки знаний в самых разных областях. На сайте проекта можно ознакомиться с короткой демонстрационной версией тестов.

Характерной чертой деятельности проф. А. Мустайоки является соединение научных исследований и практических нужд преподавания языка. Два проекта, описанные в настоящей статье, являются лишь иллюстрацией множества задач, решенных или успешно решаемых на Отделении славянских и балтийских языков и литератур под руководством проф. А. Мустайоки.

Адрес электронной почты автора: mihail.kopotev@helsinki.fi

Литература

- Mathews-Aydinli, J. 2007. Problem-Based Learning and Adult English Language Learners. In: *ESL Resources of CAELA*.
www.cal.org/caela/esl%5Fresources/briefs/problembased.html. (25.05.2008).
- Добрушина, Н.Р. 2005. Как использовать Национальный корпус русского языка в образовании. В кн. В.А. Плунгян (ред.), *Национальный корпус русского языка 2003–2005*, 308–329. Москва: Индрик.
- Копотев М.,В., Мустайоки А. 2003. Принципы создания Хельсинкского аннотированного корпуса русских текстов (ХАНКО) в сети Интернет. В кн. *Научно-техническая информация Сер. 2, № 6*, 33–37.
- Копотев, М.,В., Мустайоки, А. 2008. Современная корпусная русистика. В кн. *Инструментарий русистики: корпусные подходы* (Slavica Helsingiensia 34), 5–24. Helsinki: Yliopistopaino.
- Кудрявцев, В.Т. 1991. *Проблемное обучение: истоки, сущность, перспективы*. Москва: Знание.
- Мельчук, И.А. 1997. *Курс общей морфологии*, т. 1. Москва: Языки русской культуры.
- Мустайоки, А. 1973. *Опыт составления частотной грамматики русских существительных*. Хельсинки: Хельсинкский университет (магистерская диссертация).
- Мустайоки А., 2006. *Теория функционального синтаксиса: от семантических структур к языковым средствам*. Москва: Языки русской культуры.
- Мустайоки А. 2001. Упражнения по русской грамматике: таксономия и база данных. В кн. *Русский язык на рубеже тысячелетий*, 122–138. Санкт-Петербург: Издательство Санкт-Петербургского университета.
- Национальный корпус русского языка и проблемы гуманитарного образования*. Материалы международной научной конференции. 2007. Москва: ВШЭ.

- Плунгян, В.А. (ред.). *Национальный корпус русского языка 2003–2005*. 2005 Москва: Индрик.
- Резникова, Т.И., Копотев М.В. 2005. Лингвистически аннотированные корпуса русского языка (обзор общедоступных ресурсов). В кн. *Национальный корпус русского языка 2003–2005*, 31–61. Москва: Индрик..
- Шаров, С.А. 2003. Представительный корпус русского языка в контексте мирового опыта. В кн. *Научно-техническая информация*, Сер.2, № 6, 9–18.