

КОРПУС ИЛИ ЭКСПЕРИМЕНТ?¹

Шмелев А. Д. (shmelev.alexei@gmail.com)

Московский педагогический государственный университет, Москва, Россия; Институт русского языка им. В. В. Виноградова РАН, Москва, Россия

В статье обсуждается общенаучное разграничение наблюдения и эксперимента применительно к лингвистическим исследованиям.

Утверждается, что некоторые лингвистические задачи должны решаться на основе наблюдения, т. е. посредством обращения к корпусным данным, тогда как для решения других необходим лингвистический эксперимент; при этом исследователь может проводить эксперимент как с опорой на собственную языковую компетенцию, так и путем опроса информантов.

Описаны проблемы, связанные с проведением опроса информантов для выявления региональных языковых норм на фоне общеязыковых норм с одной стороны и регионального узуса - с другой.

Ключевые слова: наблюдение, эксперимент, корпус, узус, языковая норма, региональная норма

¹ Статья написана при финансовой поддержке РФНФ в рамках научно-исследовательского проекта РФНФ («Кодифицированная норма и региональный узус»), проект № 11-04-00501а.

CORPUS OR EXPERIMENT?

Shmelev A. D. (shmelev.alexei@gmail.com)

Moscow Pedagogical State University, Moscow, Russia;
Vinogradov Institute of Russian Language, Russian Academy
of Sciences, Moscow, Russia

The paper discusses general distinction between observation and experiment as applied to linguistic research. It claims that the resolution of certain issues has to be based on observation, that is, on corpus data (e. g., studies of ancient languages) while other issues require experiment (in particular, issues concerning the difference between common usage and linguistic standards). Two main varieties of experiment may be distinguished, namely, experiment on the researcher's linguistic competence and interrogation of informants. The latter is the only accessible experimental method if the researcher has no linguistic competence of the language or dialect under investigation. To illustrate the point, the paper uses the example of regional linguistic standards (in particular, norms of pronunciation). It discusses various pitfalls on the way to an accurate account of regional linguistic standards. To avoid falling into one of those pitfalls, the researcher should be clear in his/her mind about the ultimate objective of the investigation and word questions to the informants in a clear form if s/he is going to use a questionnaire.

Key words: observation, experiment, corpus, usage, linguistic standard, regional linguistic standard

1. Наблюдение и эксперимент: общенаучное разграничение

Дискуссия на конференции «Диалог-2011», посвященная использованию корпусов в лингвистических исследованиях, показала, что ряд участников дискуссии был готов характеризовать исследования, проведенные на основе «Национального корпуса русского языка», как разновидность экспериментальных исследований. Представляется, что такое словоупотребление находится в противоречии с общенаучным разграничением двух основных способов получения данных: наблюдения и эксперимента. Различие между ними может быть описано следующим образом. Наблюдение имеет своим объектом явления

и события, происходящие независимо от исследователя, который тем самым изучает их *in vivo*. Напротив того, эксперимент предполагает целенаправленную деятельность, направленную на получение данных, которые без эксперимента просто не существовали бы; тем самым производится исследование *in vitro*. Существенно при этом иметь в виду ограниченность экспериментального метода. Прежде всего, существуют вещи, которые в принципе нельзя исследовать экспериментально, так что для их изучения приходится ограничиваться наблюдением². Так, климатические условия в той или иной местности, по-видимому, не подлежат экспериментальному изучению. Едва ли имеет научный смысл экспериментальная проверка таких гипотез, как «если я забываю зонтик, идет дождь, а если я беру его с собою, светит солнце», а более содержательные эксперименты с климатом едва ли возможны (имеются в виду научные эксперименты, а не воздействие на климат как побочный эффект человеческой деятельности). Точно так же экспериментальные методы в собственном смысле слова не могут использоваться в астрономии. Кроме того, чрезвычайно важно, что если часто приходится ограничиваться наблюдением, не прибегая к экспериментам, то эксперимент без наблюдения вообще невозможен. Во-первых, эксперименту, как правило, предшествует наблюдение, на основании которого и выдвигаются экспериментально проверяемые гипотезы. Во-вторых, что важнее, наблюдение над полученными экспериментальными данными составляет неотъемлемую часть всякого эксперимента; в противном случае проведение эксперимента лишилось бы смысла. Если же оказывается, что исследователь не нуждается в экспериментальных данных (либо проведение эксперимента невозможно в принципе) и ограничивается данными, полученными в ходе наблюдений, он прибегает к их обработке, но такая обработка не должна называться «экспериментом».

2. Наблюдение и эксперимент: применение общенаучного разграничения к лингвистическим исследованиям

Все сказанное имеет непосредственное отношение и к лингвистическим исследованиям. Многие виды лингвистических исследований заведомо не предполагают и не могут предполагать эксперимента. Так, при изучении древних языков приходится ограничиваться наблюдением над данными, извлекаемыми из корпуса имеющихся в распоряжении исследователя текстов на этих языках. Разумеется, в какой-то момент могут быть обнаружены новые тексты, и они дадут материал для проверки описания, созданного на основе предшествующего корпуса; однако это не будет экспериментальной проверкой, поскольку имеет место именно обнаружение, а не искусственное создание новых данных. Точно так же и астрономические наблюдения могут дать

² Для простоты я отвлекаюсь от возможности косвенных экспериментов с объектами, моделирующими изучаемые условия. В любом случае таким образом нельзя получить непосредственные данные, касающиеся интересующего нас объекта.

материал для проверки существующих теорий, однако это не делает астрономические наблюдения экспериментом³.

Изучение живой устной речи также в норме должно вестись *in vivo*, т. е. посредством наблюдения, а не эксперимента. Возможно, впрочем, экспериментальное исследование спонтанного речевого поведения носителей языка в некоторой искусственно созданной ситуации.

Однако следует понимать, что есть языковые данные, которые можно получить только экспериментально, поскольку непосредственному наблюдению они недоступны. (Из сказанного вытекает, что для древних языков эти данные нельзя получить в принципе, хотя иногда их можно реконструировать на основе косвенных соображений.) Так, важнейшим элементом языковой способности носителей языка является их умение отличить «правильное» высказывание от «неправильного». Необходимо добавить, что во многих случаях бинарное противопоставление «правильное»– «неправильное» оказывается недостаточным, так что приходится вводить какие-то промежуточные точки шкалы. При этом следует иметь в виду, что «промежуточный» статус на шкале «правильности» того или иного высказывания может быть обусловлен двумя различными обстоятельствами. Во-первых, сами информанты могут колебаться в оценке степени правильности высказывания, характеризовать его как «допустимое, но странное», «сомнительное» и т. п. Во-вторых, иногда обнаруживается разнобой во мнениях информантов: одни считают высказывание правильным, другие — нет. В обоих случаях одно наблюдение над корпусом не позволит выявить такие высказывания. В самом деле, может случиться так, что информанты, считающие, что высказывание правильное, будут употреблять его, тогда как мнение других, возможно, даже более многочисленных информантов, отвергающих высказывание как неправильное, выявлено не будет. Сомнительные высказывания также могут встретиться в корпусе, и их статус не будет отличаться от прочих, несомненно правильных, высказываний.

Здесь мы сталкиваемся с более общей проблемой, касающейся наблюдения над готовыми текстами. Исследователи, имеющие дело с древними текстами, знают, что в них могут встречаться ошибки разного рода, но исходят из того, что ошибка не может быть частотной. Однако, имея дело с современными языками, мы нередко обнаруживаем, что носители языка пользуются понятием «распространенная ошибка». Понятно, что одно наблюдение над текстами не позволит отличить такую ошибку от «правильных» (т. е. не вызывающих ни малейшего сомнения у носителей языка) высказываний.

В силу сказанного для древних языков оказывается затрудненным разграничение нормы и узуса. Возможны лишь разного рода ухищрения, позволяющие в каких-то случаях проводить такое разграничение на основе косвенных данных (так, Клайв Льюис отмечал, что если в старых текстах делается

³ Само собою разумеется, что не является экспериментом ситуация, когда исследователь конструирует некоторое языковое выражение, а потом проверяет, обнаружится ли оно в корпусе древних текстов.

предостережение против некоторой ошибки, то это значит, что соответствующее речевое явление было довольно распространено в узусе).

В отношении современных языков многие из указанных трудностей могут быть преодолены при помощи эксперимента. Самый примитивный эксперимент состоит в опросе информантов, которым предъявляется высказывание и делается предложение оценить это высказывание с точки зрения языковой правильности («можно ли так сказать?»). Однако следует иметь в виду, что результаты такого эксперимента вовсе не обязательно будут отражать лингвистическую реальность (т. е. в данном случае подлинное мнение информантов о правильности предъявленных высказываний). Подводные камни, связанные с проведением опросов, известны специалистам, работа которых во многом опирается на опрос как метод исследования (не только полевым лингвистам, но и, напр., социологам), и я не буду здесь на них останавливаться.

Возможны и другие типы опросов. Так, можно моделировать какую-то ситуацию и наблюдать за речевым поведением испытуемых в данной ситуации. При этом прямолинейный способ получения ответа состоит в том, чтобы эксплицировать параметры ситуации, предъявить их респонденту, и спросить, что он сказал бы в такой ситуации. Более изощренный способ заключается в том, чтобы искусственно поставить испытуемого в некоторые условия и наблюдать за его речевым поведением, — это приближает искусственно созданные условия к естественным и тем самым эксперимент оказывается похож на наблюдение *in vivo*, но все же отличается от него.

Здесь, впрочем, важны формулировки вопросов. Если мы спрашиваем респондента, что бы он сказал в некоторой ситуации, то предполагаем получить данные о том, как представляет себе респондент собственное речевое поведение (излишне говорить, что не всегда эти представления оказываются адекватными). Можно спросить, что респондент считает правильным сказать в данной ситуации или что обычно говорят в такой ситуации, и получить другие ответы.

Помимо этого, в применении к живым языкам возможно непосредственное наблюдение над повседневной речевой деятельностью носителей языка, а не только наблюдение над результатами такой деятельности — готовыми текстами. Известны трудности, связанные с таким наблюдением: поведение людей, ставших объектом наблюдения, может меняться в силу того, что им становится известно, что за ними наблюдают⁴, а также разные приемы, направленные на преодоление этих трудностей (так, для изучения речи некоторых социальных групп может использоваться так называемое «включенное наблюдение» и т. п.).

Заметим, что то, что говорилось о соотношении наблюдения и эксперимента вообще, в целом применимо и к лингвистике. В частности, и в лингвистике, прежде чем проводить опрос информантов, конструировать искусственные примеры, придумывать воображаемые ситуации, исследователь обычно

⁴ Подобные трудности иногда возникают и в естественных науках: так, при наблюдении за поведением животных в естественных условиях важно «не спугнуть» животное.

обращает внимание на какие-то явления, с которыми он столкнулся, наблюдая повседневную речевую деятельность, а любой эксперимент включает в себя наблюдение в качестве составной части. Наконец, можно заметить, что если в результате эксперимента был получен значительный массив данных, эти данные могут составить некий корпус, над которым другие исследователи могут проводить наблюдения, уже не занимаясь экспериментами.

3. Роль интроспекции

Одним из методов лингвистического эксперимента является интроспекция. Внутренняя форма этого слова как будто подсказывает, что интроспекция — это разновидность наблюдения (наблюдение над самим собою). Однако интроспекция как чистое наблюдение встречается довольно редко. Конечно, бывает так, что, сказав или написав нечто без целей лингвистически исследовать соответствующее языковое явление, лингвист про себя отмечает, что именно он сказал или написал и в дальнейшем использует это наблюдение в своей работе. Однако гораздо чаще интроспекция включает эксперимент: лингвист искусственно конструирует некоторое высказывание (часто переставляя или заменяя какие-то языковые единицы, которые он уже наблюдал в речевой деятельности) и оценивает его правильность. Другой тип эксперимента в рамках интроспекции — попытка лингвиста представить себе, что бы он сказал в некоторой воображаемой ситуации.

Интроспекция обладает целым рядом важных преимуществ по сравнению с работой с информантами. Однако необходимо понимать, что, ограничиваясь интроспекцией, лингвист замыкается на собственном идиолекте, который может по каким-то характеристикам отличаться от идиолектов других носителей языка. Помимо этого, эксперимент, сводящийся к интроспекции, в большей степени ориентирован на норму, а не на узус; конечно, относительно каких-то высказываний, которые представляются ему неправильными, лингвист может понимать, что неоднократно слышал или читал (а возможно, и сам произносил) их, однако в точности оценить степень их распространенности в узусе путем интроспекции невозможно. Поэтому, если целью исследования является изучение узуса (напр., количественное), интроспекция мало что дает.

Все сказанное касается и проблемы соотношения нормы и узуса, а также более специальных вопросов, напр. соотношения региональной нормы и регионального узуса с общезыковой нормой и общезыковым узусом. Здесь исследователь, даже живущий в соответствующем регионе, не может ставить во главу угла интроспекцию. Региональный узус (и его соотношение с общезыковым) в принципе может изучаться посредством наблюдения (хотя, напр., для изучения произносительного узуса нужен был бы представительный корпус устной речи), но региональная норма в ее соотношении с общезыковой нормой требует обращения к работе с информантами. В этом отношении первостепенную роль играет корректная методика опросов и интерпретация результатов.

4. Отклонение регионального узуса от общеязыковой нормы: проблемы интерпретации и способы исследования

При изучении соотношения общеязыковой нормы и регионального узуса в фокус внимания исследователя должны попасть случаи, когда региональный узус отличается от того, что предписывается кодифицированной общеязыковой нормой. Для простоты будем исходить из того, что словари и справочники, кодифицирующие общеязыковую норму, во всех случаях содержат адекватное описание (хотя на самом деле это, конечно, не так). Тогда отличия регионального узуса от того, что предписывается нормативными словарями и справочниками, могут характеризоваться как отклонения регионального узуса от общеязыковой нормы.

Такие отклонения могут интерпретироваться различным образом в зависимости от того, каковы в данном случае общеязыковой узус и региональная норма. Логически мыслимы четыре ситуации:

1. Как общеязыковой узус, так и региональная норма соответствует общеязыковой норме. В этом случае мы можем говорить о специфических явлениях в региональном узусе, которые не становятся нормативными даже в соответствующем регионе.
2. Общеязыковой узус соответствует общеязыковой норме, а региональная норма допускает употребление, характерное для регионального узуса. В этом случае можно говорить о специфической региональной норме, отличной от общеязыковой нормы.
3. Региональная норма соответствует общеязыковой норме, а в общеязыковом узусе обнаруживаются то же отклонение, что и в региональном узусе. В этом случае можно говорить об отсутствии региональной специфики: соответствующее отклонение от нормы обнаруживается не только в рассматриваемом регионе, но и распространено более или менее повсеместно.
4. Отклонение от общеязыковой нормы обнаруживаются не только в региональном, но и в общеязыковом узусе; при этом данное отклонение допускается региональной нормой. В этом случае некоторое распространенное повсеместно отклонение от (общеязыковой) нормы «повышает» свой статус в региональной разновидности языка.

Замечание. Понятно, что языковое выражение, употребительное или даже известное лишь в некотором регионе, является отклонением от общеязыковой нормы лишь при наличии альтернативного и при этом общеязыкового способа выражения. Очевидный пример — региональная топонимика. Вероятно, наименование *Столбы* известно не только жителям Красноярска; но даже если бы его в других регионах не знали, оно все равно оставалось бы нормативным с точки зрения общеязыковой нормы (разумеется, это верно не для всех региональных топонимов). Особо следует упомянуть случаи, когда та или иная реалья получает разные наименования в различных регионах либо, наряду

с общеязыковым наименованием, у него имеется особое «региональное» наименование. Часто бывает так, что региональное наименование является общеязыковой нормой, когда используется для обозначения данной реалии в соответствующем регионе (напр., для придания местного колорита), тем более и реалии в разных регионах часто не тождественны. Так, сельский дом на севере России именуется *изба*, а на юге — *хата*. Соответственно, северные бревенчатые дома уместно называть *избами*, а южные дома (как правило, выполненные не из цельных, а из половинных бревен, снаружи обмазанные глиной и побеленные) — *хатами*. Слова *арык* или *шак* (вместо слов *канал* и *осел*) уместны, когда речь идет о Средней Азии; так же, если речь идет о Средней Азии, вместо «общеязыкового» слова *абрикос* (в применении к дереву) может использоваться *урюк* (при этом, поскольку носители русского языка часто именно в Средней Азии сталкиваются с цветущими абрикосовыми деревьями, сочетание *цветущий урюк* оказывается едва ли не более частотным, нежели *цветущий абрикос*)⁵.

Однако, для того чтобы отнести конкретную ситуацию к тому или иному типу необходимо уметь устанавливать в каждом конкретном случае общеязыковую и региональную норму и общеязыковой и региональный узус. Как уже говорилось, общеязыковую норму можно условно приравнять к рекомендациям нормативных словарей и справочников. Как общеязыковой, так и региональный узус, вообще говоря, следовало бы изучать методом наблюдения. Поскольку речь идет о современной языковой ситуации, возможно не только наблюдение над готовыми результатами речевой деятельности (корпусом текстов), но и непосредственное наблюдение над речевой деятельностью в прямом режиме (последнее — если исследуется устный узус). Именно наблюдение над речевой деятельностью в прямом режиме часто является исходным стимулом для последующих наблюдений над корпусом текстов и лингвистическими экспериментами. Однако, как уже говорилось, непосредственное наблюдение в прямом режиме сопряжено с риском, что сам факт наблюдения будет воздействовать на речевое поведение носителей языка; кроме того, оно плохо применимо к созданию письменных текстов. В любом случае результаты непосредственного наблюдения в прямом режиме затруднительно использовать для решения многих важных исследовательских задач — в частности, сопряженных со статистическим анализом полученных результатов (так, оценка частотности того или иного произносительного варианта при непосредственном наблюдении может делаться лишь «на глазок»). По отношению к письменным текстам часто удобным оказывается «сегментно-статистический анализ» [Беликов 2011], использующий в качестве источника корпуса текстов различные сегменты русского Интернета. Однако для исследования особенностей регионального узуса, специфичных для устной речи, данные Интернета мало что дают. Попытка получить материал для наблюдения методами полевой лингвистики (в духе проекта петербургских лингвистов «Один речевой день» [Богданова 2008]) во многих случаях также не дает и едва ли может дать в обозримом будущем достаточный материал для статистически достоверных выводов,

⁵ Ср. обсуждение этого вопроса в статье [Shmelev, Shmeleva 2009, 340].

касающихся особенностей общезыкового и регионального узуса. Здесь более эффективным может оказаться «скрытый» эксперимент, когда искусственно создается ситуация, в которой велика вероятность употребления информантом языкового выражения, интересующего исследователя. Что касается до региональной нормы, то в отношении письменного языка она может изучаться методом наблюдения — если принять положение, согласно которому определенные типы текстов заведомо написаны нормативным языком (напр., официальные постановления местных органов власти). Однако допущения такого рода в отношении устной нормы сильно исказило бы картину. В самом деле, такая методика предполагала бы, что мы а priori выделяем некоторое заданное множество «образованных» носителей языка данного региона, речь которых может считаться «образцовой», и постулируем, что любое языковое явление, которое встретилось в их речи, является принадлежащим норме (хотя бы региональной). Однако такой подход был бы неоправданным упрощением: норма была бы приравнена к узусу «образованных» носителей языка, при этом случаи, когда некоторое языковое явление, встретившееся в речи одних «образованных» носителей языка, решительно отвергается другими, вообще не нашли бы отражения в описании.

5. Изучение соотношения нормы и узуса путем опроса информантов: подводные камни

Сказанное подводит нас к выводу, что здесь, как и в других подобных случаях, не обойтись без эксперимента. При этом, как уже говорилось, чрезвычайно важна корректная методика эксперимента и интерпретации его результатов. Иллюстрируем это на примере произносительных норм, отклонения от которых наблюдаются в узусе.

Предположим, исследователь в ходе наблюдений обнаружил, что некоторая словоформа в региональном узусе часто произносится не с тем ударением, которое предписывается общезыковой нормой (определяемой, на основе нормативных словарей и справочников). Правильная интерпретация этого факта зависит от того, насколько частотно такое отклоняющееся от нормы ударение в региональном узусе, встречается ли оно в региональном узусе чаще, нежели в общезыковом узусе, осознается ли оно как отклонение от нормы «образованными» носителями языка, живущими в данном регионе.

Если для данного региона исследователь располагает представительным корпусом устной речи, в котором интересующая его словоформа встречается достаточно часто, чтобы можно было сделать какие-то статистически значимые наблюдения, то дополнительных экспериментов для изучения регионального узуса не потребуется. В противном случае, чтобы получить информацию о том, как в данном регионе произносят интересующую нас словоформу, требуется провести эксперимент. Методологически предпочтителен «скрытый» эксперимент, когда создается ситуация, которая естественным образом будет

располагать информанта к произнесению данной словоформы. Однако иногда оказывается, что экономнее просто дать информантам прочесть вслух текст, уже содержащий данную словоформу. Разумеется, в последнем случае произношение информантов может быть несколько иным, нежели в естественных условиях, в большей мере ориентированным на их представления о норме, но этим часто можно пренебречь. К более существенным искажениям может привести то, что при проведении такого эксперимента способ произнесения каждого информанта считается один раз, независимо от того, использует ли он соответствующее слово в своей повседневной речи. Некоторая поправка могла бы состоять в том, чтобы не считать ответы информантов, очевидным образом испытывающих затруднение при произнесении данной словоформы (об этом может свидетельствовать заминка при чтении). Впрочем, на практике при проведении такого эксперимента все равно редко достигается репрезентативность отбора информантов, так что возможными сдвигами можно пренебречь.

Для установления региональной нормы в рассматриваемой области трудно представить себе более эффективный способ, нежели опрос информантов. При установлении орфоэпической нормы следует ориентироваться не на то, как носители языка произносят единицу, для которой необходимо выработать орфоэпическую рекомендацию, а на их сознательную оценку степени нормативности существующих орфоэпических вариантов [Шмелев 2004], и это касается как общезыковой, так и региональной нормы. Иными словами, тот факт, что какое-то ударение, отвергаемое общезыковой нормой, широко распространено в данном регионе, в том числе в речи радиодикторов или университетских преподавателей, еще не означает, что его следует счесть региональной нормой: другие дикторы и преподаватели (или даже большинство таковых) могут это ударение решительно отвергать. Решение вопроса, имеем ли мы дело с региональным узусом, отклоняющимся от нормы, или с особой региональной нормой, не может быть принято только на основе фактов узуса: необходимо учитывать оценку этих фактов говорящими. Так, в Америке почти все русские говорят (или, по крайней мере говорили несколько лет тому назад) *на Интернетe*, а не *в Интернетe*; необходимо специальное исследование, чтобы решить, как трактовать это: как ошибку, вызванную отрывом от родной речи, или как особую черту русского языка в Америке.

При проведении опроса важно найти оптимальную формулировку вопросов и правильно интерпретировать полученные ответы. Так, может ввести в заблуждение формулировка вопроса: «Какое ударение представляется вам наиболее привычным?» Респонденту может быть не вполне ясно, спрашивают ли его о том, какое ударение он привык слышать, или о том, как он сам привык произносить данное слово. Целесообразно разграничить три типа вопросов:

- (1) С каким ударением обычно произносят это слово окружающие вас люди?
- (2) С каким ударением произносите это слово вы?
- (3) Как правильно произносить данное слово?

Желательно каждому респонденту дать все три вопроса, сделав тем самым различие между ними максимально явным. Помимо этого полезно предусмотреть возможность «уклончивых» ответов: на первый вопрос — «не обращал внимания» или «произносят по-разному», на второй вопрос — «не использую этого слова», на третий вопрос — «не знаю, надо посмотреть в словаре». Наличие возможности дать «уклончивый» ответ отчасти снижает роль личностных характеристик информантов⁶, которые в противном случае могли бы привести к снижению достоверности интерпретации полученных результатов. Кроме того, респонденты должны понимать, что возможен ответ «и так, и так», «два варианта»⁷ и т. п. на первый и третий (а может быть, и на второй) вопросы.

При этом необходимо понимать, что ответы информантов не следует интерпретировать буквально: если респондент ответил на второй вопрос, что он произносит слово так-то и так-то, из этого вовсе не обязательно следует, что он действительно его так произносит. Здесь могут играть роль и ложные представления о собственной речи (в отношении ударения это случается относительно редко, но для других языковых явлений редкостью не является), и влияние представлений о норме или о «престижности» того или иного варианта⁸.

После того как получены ответы, можно переходить к интерпретации результатов. К сожалению, обсуждение того, как на основе этих результатов сделать вывод о региональной норме, находится за пределами данной статьи (некоторые схематичные указания, по необходимости краткие, содержатся в уже цитированной заметке [Шмелев 2004]).

6. Заключительное замечание

В заключение повторим, что методика исследования должна быть адекватна поставленным задачам. Из того неоспоримого факта, что есть задачи, которые нельзя решить иначе, как обращением к корпусу текстов, никак не следует, что такое обращение является оптимальным способом решения любой лингвистической задачи и дело лишь в том, какой корпус выбрать и какой метод его анализа применить. Для многих осмысленных задач наблюдений над реальными текстами недостаточно и необходим лингвистический эксперимент, будь то работа с информантами или интроспекция. Этот вывод мог бы казаться вполне тривиальным, если бы энтузиазм, связанный с использованием корпусных методов иногда не приводил к тому, что его сторонники начинали

⁶ Так, стереотипы уверенных, неуверенных и нерелефлирующих носителей языка и влияние указанных характеристик на оценку этими носителями правильности языковых выражений обсуждаются в статье [Беликов 2009].

⁷ В принципе вариантов может быть и больше: так, слово *роженица* имеет в узусе три варианта ударения.

⁸ Я, напр., на вопрос о том, с каким ударением я произношу слово *йогурт*, отвечаю, что ударение ставлю на втором слоге; однако методом самонаблюдения я установил, что на самом деле более чем в половине случаев ставлю ударение на первом слоге.

отвергать методы, не опирающиеся на корпус, и, более того, даже именовать корпусные методы «экспериментальными».

Литература

1. *Беликов В. И.* Стереотипы в понимании литературной нормы // Стереотипы в языке, коммуникации и культуре. М.: 2009.
2. *Беликов В. И.* Чего не хватает в «оцифрованном мире» лексикографу и социолингвисту // Компьютерная лингвистика и интеллектуальные технологии: По материалам ежегодной Международной конференции «Диалог» (2011). Вып. 10 (17). М.: 2011. С. 60–67.
3. *Богданова Н. В., Асиновский А. С., Русакова М. В., Степанова С. Б., Шерстнинова Т. Ю.* Звуковой корпус русского языка повседневного общения «Один речевой день»: Концепция и состояние формирования // Компьютерная лингвистика и интеллектуальные технологии: По материалам ежегодной международной конференции «Диалог» (2008). Вып. 7 (14). М.: 2008. С. 488–494.
4. *Шмелев А. Д.* Проблема кодификации в сфере орфоэпии: личный вкус или объективные данные? // Культура русской звучащей речи: традиции и современность. Тезисы докладов международной научной конференции (2004). М.: 2004. С. 116–118.
5. *Shmelev A., Shmeleva E.* Russian Botanical Terms: Towards their Lexicographic Description. Proceedings of the 4th International Conference on Meaning-Text Theory (MTT '09), Montreal, 2009, pp. 339–348.

References

1. *Belikov V. I.* Standard Preconceptions about Russian Literary Language Norm [Stereotipy v ponimanii literaturnoi normy]. Stereotipy v iazyke, kommunikatsii i kul'ture [Stereotypes in language, communication, and culture]. М.: 2009.
2. *Belikov V. I.* What are sociolinguists and lexicographers lacking in a digitized world? [Chego ne khvataet v «otsifrovannom mire» leksikografu i sotsiolingvistu]. Komp'yuternaia lingvistika i intellektual'nye tekhnologii: Po materialam ezhegodnoi Mezhdunarodnoi konferentsii «Dialog» (2011) [Computational Linguistics and Intellectual Technologies: Papers from the Annual Conference “Dialog” (2011)], Issue 10. Moscow, 2011, pp. 60–67.
3. *Bogdanova N. V., Asinovskii A. S., Rusakova M. V., Stepanova S. B., Sherstnina T. Iu.* The Corpus of Spoken Russian: Design Principles and Approaches to Data Analysis [Zvukovoi korpus russkogo iazyka povsednevnogo obshcheniia «Odin rechevoi den'»: Kontseptsiia i sostoianie formirovaniia]. Komp'yuternaia lingvistika i intellektual'nye tekhnologii: Po materialam ezhegodnoi mezhdunarodnoi konferentsii «Dialog» (2008) [Computational Linguistics and

Intellectual Technologies: Papers from the Annual Conference “Dialog” (2008)], Issue 7 (14). Moscow, 2008, pp. 488–494.

4. *Shmelev A. D.* The Problem of Codification of the Norms of Pronunciation: Personal Preferences or Objective Data? [Problema kodifikatsii v sfere orfoepii: lichnyi vkus ili ob”ektivnye dannye?]. *Kul’tura russkoi zvuchashchei rechi: traditsii i sovremennost’. Tezisy dokladov mezhdunarodnoi nauchnoi konferentsii (2004)* [Cultivation of Russian Oral Speech: Traditions and the Present Times. Abstracts of the International Conference (2004)], Moscow, 2004, pp. 116–118.
5. *Shmelev A., Shmeleva E.* Russian Botanical Terms: Towards their Lexicographic Description. *Proceedings of the 4th International Conference on Meaning-Text Theory (MTT ’09)*, Montreal, 2009, pp. 339–348.