

Litera

Правильная ссылка на статью:

Соловьева А.А. — Методика применения мини-корпусов для анализа языкового варьирования этикетных выражений двуязычных языковых ситуаций в социолингвистическом аспекте // Litera. – 2023. – № 10. DOI: 10.25136/2409-8698.2023.10.39183 EDN: WGLTKE URL: https://nbpublish.com/library_read_article.php?id=39183

Методика применения мини-корпусов для анализа языкового варьирования этикетных выражений двуязычных языковых ситуаций в социолингвистическом аспекте

Соловьева Анна Андреевна

ORCID: 0000-0001-9472-5966

кандидат филологических наук

старший преподаватель, кафедра русского языка как иностранного и общетеоретических дисциплин,
Российский государственный аграрный университет - Московская сельскохозяйственная академия
имени К. А. Тимирязева

127247, Россия, г. Москва, пр-д Бескудниковский, 4-2-84

✉ yarlee@yandex.ru



[Статья из рубрики "Лингвистика"](#)

DOI:

10.25136/2409-8698.2023.10.39183

EDN:

WGLTKE

Дата направления статьи в редакцию:

16-11-2022

Дата публикации:

16-10-2023

Аннотация: Предметом рассмотрения является процесс создания и методика применения мини-корпуса для решения задач в области изучения выбора языка в сфере речевого этикета. Рассматриваются различные этапы составления мини-корпуса от особенностей сбора материала до построения теоретических моделей использования языков речевого этикета. Вводится терминологический аппарат исследования. Сравниваются особенности различных методов и обосновывается необходимость применения комплексного подхода к составлению мини-корпуса. Констатируется

нехватка корпусов для изучения языкового выбора в двуязычной ситуации. Описываются преимущества и недостатки выбранной методики. В частности, указывается на необходимость отбора аутентичных примеров спонтанной речи, взятой из видеоматериалов, на основе метода невключенного наблюдения. Показано место методики мини-корпусов в совокупности методов, применяемых в данном исследовании. Новым является применение методики мини-корпусов на материале единиц билингвальных языковых ситуаций. Результатом становится описание на основе корпуса алгоритмов использования говорящими этикетных выражений на том или ином языке под влиянием различных конфигураций факторов и параметров социального контекста. Делается вывод о том, что данный метод позволяет, с одной стороны, получить необходимый и достаточный материал для анализа, с другой стороны, не предполагает использования и обработки большого массива данных. Перспективой исследования методики является применение ее в других ситуациях двуязычия для дополнения и корректировки полученных в ходе исследования теоретических моделей. Таким образом, метод составления двуязычных мини-корпусов позволяет выходить на уровень обобщения языковых фактов, достаточный для построения универсальных моделей для интерпретации закономерностей выбора языка этикетных выражений при двуязычии.

Ключевые слова:

социолингвистика, корпусная лингвистика, двуязычие, этикетные выражения, методы, языковой выбор, контент-анализ, каталанский язык, язык хинди, мини-корпус

Введение

Лингвистические исследования предполагают работу с лингвистическими фактами, языковыми единицами разных уровней. Сбор и обработка языковых единиц должна соответствовать целям и задачам исследования. Кроме того, методы должны обладать логикой научного анализа собственно лингвистических задач. Проблематика выбора методологии лингвистического исследования является предметом активного изучения современной отечественной лингвистики [\[1\], \[6\], \[7\], \[9\]](#).

В настоящем исследовании рассматривается особая категория единиц языка – речевой этикет.

Под речевым этикетом в данном случае подразумеваются «регулирующие правила речевого поведения, система национально специфичных стереотипных, устойчивых формул общения, принятых и предписанных обществом для установления контакта собеседников, поддержания и прерывания контакта в избранной тональности» [\[8, 11\]](#). Для исследования речевого этикета необходимым оказывается работа на лексическом уровне языка. При этом важным оказывается понятие контекста, что в двуязычной языковой ситуации осложняется сосуществованием двух языков внутри одного социокультурного сообщества, и, соответственно, двуязычным контекстом.

В парадигматике билингвальная языковая ситуация ставит говорящего перед выбором из единиц на двух языках. В данной ситуации говорящий делает выбор, опираясь на социолингвистический контекст, на характеристики типа адресата и самой языковой ситуации [\[4, 10\]](#).

С точки зрения синтагматики в некоторых случаях необходимо учитывать также языковой

контекст. При этом немаловажно учитывать особенности самих этикетных выражений.

Такие исходные данные требуют от исследователя разработать методику фиксации единиц речевого этикета с сохранением характеристик речевой ситуации и социального контекста.

Разработка методики подобного рода исследований является актуальной, так как при большом числе одноязычных автоматизированных корпусов существует проблема составления корпусов, включающих в себя тексты одновременно на двух и более языках для одной языковой ситуации. Данное исследование методики предполагает создание таких корпусов как инструмента исследования полиязыковой среды.

Целью данного исследования является построение поэтапной методики создания мини-корпуса, необходимого для проведения исследования этикетных выражений в ситуациях двуязычия.

В задачи исследования входит: 1) продемонстрировать принцип работы методики составления мини-корпуса; 2) доказать эффективность данной методики на примере анализа данных языковых ситуаций в Каталонии и Северной Индии; 3) очертить сферу применения и перспективу использования данного метода в лингвистических исследованиях.

Методы

Этикетные выражения являются особыми единицами языка, так как они выполняют ряд функций, отличающих их от других единиц в языке. Р. О. Якобсон считал одной из главных функций этикета в языке фатическую, отмечая, что «фатическая функция, осуществляется посредством обмена ритуальными формулами или даже целыми диалогами, единственная цель которых – поддержание коммуникации» [\[10, 67\]](#). Основное назначение сообщений, обладающих этой функцией, состоит в том, чтобы «установить, продолжить или прервать коммуникацию, проверить, работает ли канал связи, привлечь внимание собеседника или убедиться, что он слушает внимательно» [\[10, 67\]](#).

Для исследования были выбраны две ситуации двуязычия – Северная Индия и Каталония (Испания).

Параметры данных языковых ситуаций соответствуют требуемым в исследовании. Так, в обеих ситуациях существует языковое варьирование. На них также влияют схожие параметры [\[4, 10\]](#) и факторы социального контекста, что позволяет провести их сопоставление.

Для анализа языкового варьирования применяется ряд социолингвистических методов, в том числе наблюдение, анкетирование, количественных и качественный контент-анализ [\[3\]](#).

Методологическим ядром при исследовании языковых единиц в ситуации двуязычия выступает контент-анализ, а именно его качественная разновидность. «Качественный контент-анализ направлен на понимание изучаемых явлений; на анализ взаимосвязей и процессов между этими явлениями; он ориентирован на охват всей совокупности и сложности изучаемых явлений и направлен на изучение единичных случаев» [\[2, 70\]](#).

В данном исследовании метод контент-анализа и разработанная на его основе методика логически продолжают перечисленные У. Лабовым методы социолингвистических

исследований, так как основывается на языковых фактах, полученных в ходе невключенного наблюдения.

Существенным преимуществом применения в социолингвистике качественного контент-анализа на основе корпуса является то, что исследование позволяет рассматривать единицы как в языковом, так и в социальном контексте одновременно, что необходимо для более точного понимания закономерностей выбора языка.

Материалом для контент-анализа послужили единицы речевого этикета на двух языках, собранные в мини-корпусы.

Под мини-корпусом понимается корпус с числом вхождений исследуемых единиц в количестве 100. Необходимость составления собственного мини-корпуса была вызвана отсутствием специализированных двуязычных корпусов этикетных выражений исследуемых регионов на материале спонтанной речи говорящих. Кроме того, данный вид корпуса был выбран, так как он статистически соответствует критериям для проведения качественного контент-анализа, направленного на изучение единичных явлений языка.

Составление мини-корпуса состояло из нескольких этапов.

На первом этапе необходимо было обнаружить, отобрать и зафиксировать видеоматериал. В основу корпуса легли текстовые версии видеоматериалов. Видеоматериалы должны были содержать элементы спонтанной речи, в связи с чем преимущество отдавалось видеороликам следующих жанров: розыгрыши, социальные эксперименты, разговорные видеоблоги, документальные съемки о жизни простых людей и съемки скрытой камерой. Тип видеоролика фиксировался в таблице мини-корпуса.

Отобранные видеоматериалы прослушивались с целью выявления в них исследуемых речевых ситуаций и единиц речевого этикета. Далее, при выявлении интересующей единицы она заносилась в специальную таблицу, пример которой представлен ниже. Также в таблице отражался языковой контекст, описывалась ситуация, указывались ссылки на источник, таймкод, место, группа этикетных выражений и перевод на русский язык.

Необходимо отметить, что особую сложность представляет перевод, так как для понимания и анализа языковых ситуаций необходимо осуществить наиболее точный перевод с обоих языков в рамках каждого из исследуемых билингвальных регионов. Кроме того, необходимо также произвести разметку по языковой принадлежности языковых единиц этикета и контекста, чтобы исключить влияние языкового контекста на выбор языка и отделить его влияние от влияния собственно социальных факторов и параметров. Все вышеуказанные этапы трудоемки, необходимы для получения качественных результатов анализа и построения на их основе теоретических моделей использования языков речевого этикета при двуязычии.

Таким образом, были созданы мини-корпусы единиц речевого этикета на базе видеоматериалов.

С точки зрения корпусной лингвистики корпусы можно охарактеризовать следующим образом: – устные, специфические, синхронические. Корпусы включают около 1000 слов и 100 вхождений этикетных выражений в каждой их стран.

После фиксации единиц и сопутствующего контекста необходимо было разделить контексты на языковой и социальный. В данном случае задачей было доказать, что язык

этикета и язык контекста не связаны напрямую, что языковой контекст не полностью определяет язык этикетных выражений. Это требовало обнаружения ситуаций, где язык единиц речевого этикета и язык контекста не совпадают. Для этого каталанский и хинди были условно обозначены как L1, а испанский и английский как L2.

Таблица 1. Соотношение языка контекста и языка этикетных выражений (ЭВ) в примерах из видеоматериалов.

	Языки ЭВ и контекста совпадают		Язык ЭВ и контекста не совпадают
	L1	L2	L1+L2
Индия (кол-во примеров)	26	27	47
Каталония (кол-во примеров)	48	7	45

Данная таблица демонстрирует, что в Индии почти половина наблюдаемых примеров содержали смешанный языковой контекст. Такая же ситуация отмечается и в Каталонии. Ситуации отличаются тем, какая роль отводится в них контекстам L2. Так, в Индии роль контекстов L2 выше, чем в Каталонии.

Однако оба случая подтверждают предположение о непрямой зависимости языка этикетных от языка контекста. Так, для всех исследуемых групп этикетных выражений в Каталонии и для большинства таких групп в Индии смешанный языковой контекст – это нормальная среда использования этикетных выражений, такая же, как и монолингвальный контекст.

Таким образом, исходя из данных таблицы можно предположить, что несовпадение языкового контекста и языка этикетных выражений в большинстве групп – свидетельство того, что выбор языка ЭВ определяется не только языком контекста. Логично предположить, что на такой выбор также могут влиять факторы и параметры социального контекста. И данный вывод справедлив как для Индии, так и для Каталонии, что говорит о похожих общих принципах выбора языка этикетных выражений.

При этом необходимо было также доказать, что язык ЭВ не определяется в полной мере также группой этикетных выражений. Для этого нужно было установить, что в каждой из групп этикетных есть единицы двух языков.

Таблица 2. Соотношение групп этикетных выражений и языков контекста.

	L1		L2		L1+L2	
	Индия	Каталония	Индия	Каталония	Индия	Каталония
Приветствие	5	15	4	1	10	5
Прощание	0	6	4	3	0	7
Извинение	0	7	4	1	7	1
Благодарность	2	6	5	0	8	9
Обращение	20	13	9	2	22	23

В Таблице 2 можно обнаружить разное соотношение количества единиц на первом и втором языке, однако, не удастся проследить последовательной связи между конкретным языком и группой этикетных выражений.

Следующим закономерным этапом стал поиск факторов и параметров внутри социального контекста, влияющих на выбор языка этикетных выражений.

Анализ социального контекста включал последовательное рассмотрение каждого из четырех параметров («Свой» / «Чужой», социальный статус, пол и количество) в их влиянии на выбор ЭВ к релевантным типам адресата в каждом из регионов.

Для этого было подсчитано количество ситуаций использования L1 и L2 для типов адресата по каждому из параметров.

Анализ также учитывал степень влияния социальных факторов престижа и языковой лояльности в зависимости от языковой ситуации и типа адресата.

Далее были детально рассмотрены примеры использования этикетных выражений с непосредственным применением метода качественного контент-анализа. Было проанализировано соотношение социального и языкового контекстов, а также высказано предположение о причинах языкового выбора в каждом случае.

Итогом анализа видеоматериалов стало сравнение роли компонентов контекста при выборе языка в исследуемых регионах.

Результаты

Результатом применения данных, полученных в результате применения данной методики, является построение теоретических моделей [\[5, 98-103\]](#), объясняющих закономерности выбора одного из двух языков в анализируемых речевых ситуациях.

Модель 1 показала, что параметр «Свой» / «Чужой» связан с уровнем языковой лояльности региона. При высоком уровне языковой лояльности в общении со «своими» адресатами говорящие с большой вероятностью выберут этикетное выражение на L1. Если уровень языковой лояльности низкий, и L2 относится к престижным языкам, то вероятнее всего ожидать использования этикетных выражений как на L1, так и на L2.

Модель 2 демонстрирует выбор языка этикетных выражений с типами адресата, обладающими разным социальным статусом. При равном социальном статусе собеседников языковая принадлежность этикетных выражений не является значимой. Однако при неравном статусе на выбор языка влияет уровень языковой лояльности. Так, при высоком уровне языковой лояльности этикетные единицы как правило соответствуют языку контекста, а говорящие стараются избегать смешения языков. При невысоком уровне языковой лояльности и одновременно престижном L2 для адресата с неравным статусом используются единицы этикета на L2.

Модель 3 отражает закономерности выбора языка для типа адресата «Друзья» и показывает, что разница в предпочтении какого-либо языка в данном случае отсутствует. Подобное языковое поведение характерно и в случае, если общение происходит с незнакомцем-мужчиной. При разговоре с незнакомой женщиной L1 выбирается, если в билингвальной системе уже существуют собственные этикетные единицы на этом языке. В противном случае их заменяют этикетные выражения на L2.

В модели 4 о последовательном выборе языка, зависящем от количества адресатов, можно говорить только по отношению к типу адресата «Незнакомец»: при высоком уровне языковой лояльности чаще используется L1, но если L2 в данном регионе престижен, то более вероятно обнаружить выражения на L2. Однако в целом можно заключить, что язык этикетных выражений для нескольких адресатов равен языку,

используемому в данной ситуации по отношению к одному адресату.

Сфера применения данной методики может включать в себя языковые ситуации, где говорящими используется речевой этикет на более чем одном языке.

Перспективой исследования методики является применение ее в других ситуациях двуязычия для дополнения и корректировки полученных в ходе исследования теоретических моделей.

Таким образом, метод составления двуязычных мини-корпусов позволяет выходить на уровень обобщения языковых фактов, достаточный для построения универсальных моделей для интерпретации закономерностей выбора языка этикетных выражений при двуязычии.

Обсуждение

Описываемая методика обладает преимуществом перед другими методами исследования языкового выбора, в частности, перед количественным методом контент-анализа, тем, что позволяет получать объективные данные на ограниченном количестве языковых фактов. В сравнении с методом анкетирования, часто используемого в социолингвистике, методика с применением данных мини-корпуса предполагает устранение саморефлексии говорящего и, следовательно, влияние его самовосприятия на результаты анализа, а также существенно снизить влияние наблюдателя на говорящего.

Рассматриваемая методика продолжает традиции метода невключенного наблюдения и позволяет вывести данный метод на новый уровень, используя в качестве базы данных видеоматериалы спонтанной речи, что существенно облегчает и ускоряет работу исследователя.

Кроме того, предложенная методологическая разработка позволила выявить перспективное направление для исследований языкового выбора – неполное соответствие данных анкетирования и анализа видеоматериалов.

Недостатком является трудоемкость сбора и составления базы данных корпуса, что на данном этапе сопряжено с проблемами его автоматизации, так как каждая единица должна быть рассмотрена и описана по нескольким параметрам (время упоминания, место, характеристика говорящего и т. д.).

Библиография

1. Голубкова Е. Е. Использование лингвистических корпусов при решении семантических проблем // Методы когнитивного анализа семантики слова: компьютерно-корпусный подход, М.: Издательский дом ЯСК, 2015, с. 39-80.
2. Кирпиков А. Р. Качественный контент-анализ как метод исследования / А. Р. Кирпиков // XXI Международная конференция памяти профессора Л. Н. Когана «Культура, личность, общество в современном мире: методология, опыт эмпирического исследования», 22-23 марта 2018 г., Екатеринбург: УрФУ, 2018, с. 67-74., с. 70.
3. Лабов У. О механизме языковых изменений // Новое в лингвистике, Выпуск VII, М.: «Прогресс», 1975, 486 с.
4. Соловьева, А. А. Степень влияния параметра социального статуса на выбор речевого этикета при двуязычии / А. А. Соловьева // Мир науки. Социология,

филология, культурология, 2019, Т. 10, № 1, с. 10. <https://sfk-mn.ru/PDF/10FLSK119.pdf> Дата обращения: 30.10.2022.

5. Соловьева, А. А. Этикетные выражения при двуязычии: социолингвистический анализ / А. А. Соловьева, М.: Общество с ограниченной ответственностью "Русайнс", 2022, 174 с., с. 98-103.
6. Сулейманова, О. А. Принципы и методы лингвистических исследований / О. А. Сулейманова, М. А. Фомина, И. В. Тивьяева. – 2-е издание, переработанное и дополненное, М.: Общество с ограниченной ответственностью "Языки Народов Мира", 2020, 352 с.
7. Чернявская В. Е. Дискурсивный анализ и корпусные методы: Необходимое доказательное звено? Объяснительные возможности качественного и количественного подходов // Вопросы когнитивной лингвистики, 2018, № 2, с. 31-37.
8. Формановская Н. И. Русский речевой этикет: Лингвистические и методические аспекты, М.: «URSS», 2008, 160 с.
9. Фрумкина Р. М. Лингвистическая гипотеза и эксперимент (о специфике гипотез в психолингвистике) // Гипотеза в современной лингвистике, М.: Наука, 1980, с. 183-216.
10. Jakobson R. Closing Statement: Linguistics and Poetics. Style in Language / Ed. by Thomas A. Sebeok. – Cambridge, Mass.; New York; London: The Technology Press of Massachusetts Institute of Technology; John Wiley & Sons, Inc., 1960, 470 p.

Результаты процедуры рецензирования статьи

В связи с политикой двойного слепого рецензирования личность рецензента не раскрывается.

Со списком рецензентов издательства можно ознакомиться [здесь](#).

Языковые единицы разных уровней предусматривают и разность из анализа, как отмечает автор рецензируемого труда, «сбор и обработка языковых единиц должна соответствовать целям и задачам исследования. Кроме того, методы должны обладать логикой научного анализа собственно лингвистических задач». В работе точно рассматривается особая категория единиц языка – речевой этикет, на мой взгляд, выбор вполне оправдан, при этом «целью данного исследования является построение поэтапной методики создания мини-корпуса, необходимого для проведения исследования этикетных выражений в ситуациях двуязычия». Материал обладает выверенной структурой, определив устойчивый и объективный сформулированный комплекс задач – автор приступает к точечному анализу проблемы. Все этапы исследования комментируются, в первую очередь вводный блок ориентирует на полновесное объяснение сути оценки: «для исследования были выбраны две ситуации двуязычия – Северная Индия и Каталония (Испания). Параметры данных языковых ситуаций соответствуют требуемым. Так, в обеих ситуациях существует языковое варьирование. На них также влияют схожие параметры и факторы социального контекста, что позволяет провести их сопоставление». На мой взгляд, выбранная методология – контент-анализ – вполне подходит для возможной верификации проблемы, «в данном исследовании метод контент-анализа и разработанная на его основе методика логически продолжают перечисленные У. Лабовым методы социолингвистических исследований, так как основывается на языковых фактах, полученных в ходе невключенного наблюдения». Рецензируемый материал имеет ярко выраженный практический характер, его удобно использовать при формировании /

проведении тематически смежных исследований. Фактически на протяжении всего сочинения номинированный стиль соотносится с научным типом, серьезных сбивов и нарушений не выявлено. Основной блок статьи являет собой аналитический срез оценки мини-корпусов единиц речевого этикета на базе видеоматериалов. Думаю, что вполне удачно систематизации данных представлена в виде таблиц, графическое / визуальное обобщение наиболее продуктивно для языковых исследований. Автор аргументировано вводит в работу т.н. промежуточные выводы, ибо это позволяет потенциально заинтересованному читателю следить за ходом развития мыслей: например, «таким образом, исходя из данных таблицы можно предположить, что несовпадение языкового контекста и языка этикетных выражений в большинстве групп – свидетельство того, что выбор языка ЭВ определяется не только языком контекста. Логично предположить, что на такой выбор также могут влиять факторы и параметры социального контекста. И данный вывод справедлив как для Индии, так и для Каталонии, что говорит о похожих общих принципах выбора языка этикетных выражений», или «таким образом, метод составления двуязычных мини-корпусов позволяет выходить на уровень обобщения языковых фактов, достаточный для построения универсальных моделей для интерпретации закономерностей выбора языка этикетных выражений при двуязычии» и т.д. Считаю, что исследование достигло своей цели, поставленные задачи решены; материал интересен, актуален, нов в рамках выбранной языковой сферы. Рекомендую статью «Методика применения мини-корпусов для анализа языкового варьирования этикетных выражений двуязычных языковых ситуаций в социолингвистическом аспекте» к открытой публикации в журнале «Litera».