



<https://doi.org/10.70728/b.series.fil.v08.i09.140>

УДК 81’322:81’27

ЛИНГВОКОГНИТИВНЫЕ И СОЦИОЛИНГВИСТИЧЕСКИЕ ПАРАМЕТРЫ РЕЧЕВОГО ЖАНРА «ПОЗДРАВЛЕНИЕ» И ИХ КОРПУСНАЯ РАЗМЕТКА

Юлдашева Сабина Юсуфовна

*преподаватель кафедры русского языка и литературы, Кокандский
государственный университет; e-mail: yuldashevasabina18@gmail.com; ORCID:
<https://orcid.org/0009-0008-4685-4884>*

Аннотация. В статье предлагается многоуровневая схема разметки текстов речевого жанра «Поздравление», в которой лингвокогнитивные и социолингвистические параметры жанра представлены как формальные признаки корпуса. Материалом служит специализированный корпус поздравительных текстов на узбекском и русском языках. Схема включает три уровня: метаданные коммуникативной ситуации, композиционно-прагматическую разметку (речевые ходы) и лингвокогнитивную разметку (фреймовые слоты и концепты пожеланий). Схема позволяет количественно исследовать зависимость речевых средств от социальных факторов и может использоваться как обучающие данные для систем автоматической классификации и генерации этикетных текстов.

Ключевые слова: речевой жанр, поздравление, корпусная разметка, фрейм, концепт, социолингвистические параметры, узбекский язык, компьютерная лингвистика.

«TABRIK» NUTQIY JANRINING LINGVOKOGNITIV VA SOTSIOLINGVISTIK PARAMETRLARI HAMDA ULARNING KORPUSDA ANNOTATSIYALANISHI

Annotatsiya. Maqolada «Tabrik» nutqiy janri matnlarini ko‘p bosqichli annotatsiyalash sxemasi taklif etiladi. Unda janrning lingvokognitiv va sotsiolingvistik parametrlari korpusning formal belgilari sifatida ifodalanadi. Tadqiqot materiali o‘zbek va rus tillaridagi tabrik matnlarining maxsus korpusidir. Sxema uch bosqichdan iborat: kommunikativ vaziyat metama’lumotlari, kompozitsion-pragmatik annotatsiya (nutqiy yurishlar) va lingvokognitiv annotatsiya (freym slotlari va tilak konseptlari). Sxema ijtimoiy omillarning til vositalari tanloviga ta’sirini miqdoriy jihatdan o‘rganish imkonini beradi hamda etiket matnlarini avtomatik tasniflash va generatsiya qilish tizimlari uchun o‘quv ma’lumotlari bo‘lib xizmat qilishi mumkin.

Kalit so‘zlar: nutqiy janr, tabrik, korpus annotatsiyasi, freym, konsept, sotsiolingvistik parametrlar, o‘zbek tili, kompyuter lingvistikasi.

COGNITIVE AND SOCIOLINGUISTIC PARAMETERS OF THE SPEECH GENRE OF CONGRATULATION AND THEIR CORPUS ANNOTATION

Abstract. The article proposes a multi-layer annotation scheme for texts of the speech genre of congratulation, in which the cognitive and sociolinguistic parameters of the genre are encoded as formal corpus features. The material is a specialized corpus of congratulatory texts in Uzbek and Russian. The scheme comprises three layers: situational metadata, compositional-pragmatic annotation (rhetorical moves), and cognitive annotation (frame slots and wish concepts). The scheme enables a quantitative study of how social factors shape linguistic choices and can serve as training data for systems of automatic classification and generation of etiquette texts.

Keywords: speech genre, congratulation, corpus annotation, frame, concept, sociolinguistic parameters, Uzbek, computational linguistics.

Введение

Поздравление — один из самых частотных этикетных жанров повседневного и институционального общения. Оно сопровождает семейные праздники, профессиональные события, религиозные и государственные торжества. С распространением мессенджеров и социальных сетей объём поздравительных текстов значительно вырос, а сами тексты стали объектом автоматической обработки: их создают чат-боты, шаблонные сервисы и языковые модели.

Для прикладной и компьютерной лингвистики это ставит новую задачу. Чтобы система могла определить повод и степень официальности поздравления или создать текст, уместный для конкретного адресата, ей необходимы размеченные данные. Создание таких данных требует лингвистически обоснованной схемы разметки, которая учитывает как содержательную сторону жанра, так и социальные условия его употребления.

Актуальность исследования определяется тем, что этикетные жанры практически не представлены в корпусных ресурсах узбекского языка, а описания жанра «Поздравление» в отечественной и зарубежной лингвистике носят преимущественно качественный характер. Развитие национальных корпусов и лингвистического обеспечения для узбекского языка относится к приоритетным направлениям прикладной лингвистики в Республике Узбекистан.

Цель статьи — разработать схему корпусной разметки, которая переводит лингвокогнитивные и социолингвистические параметры жанра «Поздравление» в формальные, воспроизводимые признаки. **Задачи** исследования: определить параметры жанра, значимые для его описания и автоматической обработки; построить многоуровневую схему разметки и инструкцию для разметчиков; оценить надёжность схемы; показать возможности количественного анализа, которые она открывает.

Анализ литературы

Теоретической основой исследования является теория речевых жанров. М. М. Бахтин определил речевой жанр как относительно устойчивый тип высказывания, характеризующийся тематическим содержанием, стилем и композицией [1]. Т. В. Шмелёва предложила модель речевого жанра, включающую коммуникативную цель, образ автора, образ адресата, событийное содержание, фактор прошлого и будущего и языковое воплощение [4]. Каждый из этих параметров соотносится с наблюдаемыми языковыми признаками, что делает модель удобной основой для разметки.

В теории речевых актов поздравление относится к экспрессивам по классификации Дж. Сёрла [12]. Однако в реальных текстах экспрессив сочетается с пожеланием, комплиментом, благодарностью. Для описания такой последовательности продуктивно понятие речевого хода (move), разработанное Дж. Суэйлзом в рамках жанрового анализа [13]: ход представляет собой функционально выделяемый фрагмент текста, решающий определённую коммуникативную задачу.

Лингвокогнитивный аспект исследования опирается на фреймовую семантику Ч. Филлмора [9] и на концептологические работы В. И. Карасика, рассматривающего этикетные тексты как источник сведений о ценностной картине мира [2]. Социолингвистический аспект связан с теорией вежливости П. Браун и С. Левинсона, согласно которой выбор речевых средств определяется социальной дистанцией, соотношением власти участников и степенью официальности [8].

Речевой этикет как система формул детально описан Н. И. Формановской [3]. Для узбекской лингвистики важны исследования Ш. Сафарова и М. Хакимова, посвящённые прагматическим нормам узбекского общения, в том числе особой роли уважения к старшим [5; 6]. Методологическую основу корпусной части работы составляют труды Т. МакЭнери и Э. Харди [10], а также принципы разметки, сформулированные Дж. Пустейовски и Э. Стаббс [11]. Оценка надёжности разметки опирается на рекомендации Р. Артштейна и М. Поэзио [7].

Анализ литературы показывает, что жанр «Поздравление» хорошо описан в качественном аспекте, однако схем его корпусной разметки, объединяющих когнитивные и социальные параметры, не существует. Особенно заметен этот пробел для узбекского языка.

Методология

Материал. Корпус включает два сопоставимых подкорпуса — узбекский и русский. Источники: поздравительные открытки, официальные поздравления организаций и государственных органов, публичные публикации в социальных сетях, сообщения в мессенджерах (с согласия авторов), записи устных поздравлений. Объём корпуса: [N] текстов, [N] словоупотреблений. Все персональные данные обезличены.

Методы. В работе использованы методы жанрового и фреймового анализа, корпусные методы (частотный анализ, анализ коллокаций, конкордансы), метод лингвистической разметки, а также статистические методы. Разметка выполнялась в инструменте [INCEpTION / BRAT / другое] двумя независимыми разметчиками по

единой инструкции. Согласованность оценивалась коэффициентом к Коэна; значимость различий проверялась критерием χ^2 . Обработка данных проводилась средствами [Python, AntConc — уточнить].

Схема разметки. Схема включает три уровня: уровень 1 размечается на весь текст, уровни 2 и 3 — на его фрагменты (таблица 1).

Таблица 1. Схема разметки поздравительных текстов

Уровень	Признак	Значения
1. Метаданные	Повод	день рождения, юбилей, религиозный праздник, государственный праздник, свадьба, рождение ребёнка, профессиональное достижение
1. Метаданные	Канал	открытка, официальное письмо, социальная сеть, мессенджер, устная речь
1. Метаданные	Официальность	официальное, полуофициальное, неофициальное
1. Метаданные	Отношения участников	родство, дружба, коллеги, иерархия «старший — младший», организация — лицо
1. Метаданные	Возраст, пол, статус	для адресанта и адресата отдельно
2. Речевые ходы	Обращение	имя, термин родства, должность, обращение с формулой уважения
2. Речевые ходы	Поздравление	экспрессивная формула (tabriklayman, поздравляю)
2. Речевые ходы	Оценка	комплимент, оценка заслуг или события
2. Речевые ходы	Пожелание	одно или несколько пожеланий
2. Речевые ходы	Завершение	формула завершения, подпись
3. Когнитивный	Концепт пожелания	здоровье, счастье, долголетие, благополучие, успех, семья, мир, вера, любовь
3. Когнитивный	Тип выражения	клише, модифицированное клише, авторская формулировка
3. Когнитивный	Образные средства	метафора, сравнение, гипербола, цитата, религиозная формула

Для каждого признака составлена инструкция с определением, примерами и пограничными случаями. Так, формула *Alloh rozi bo'lsin* размечается как религиозная формула и не относится к ходу «Пожелание», если стоит после подписи как ритуальное завершение. Пожелание с несколькими концептами (*здоровья и счастья*) размечается как один ход с несколькими концептами. Модифицированным клише считается устойчивая формула с авторским распространением (*крепкого, как сталь, здоровья*).

Результаты

Применение схемы покажем на двух условных фрагментах. Узбекское поздравление: [Обращение: термин родства + уважение] *Hurmatli onajon!* [Поздравление] *Tug‘ilgan kuningiz bilan chin qalbindan tabriklayman!* [Пожелание: здоровье; клише] *Sizga sihat-salomatlik,* [Пожелание: долголетие; клише] *uzoq umr tilayman.* Русское поздравление: [Обращение: имя] *Дорогая Анна!* [Поздравление] *Поздравляю тебя с днём рождения!* [Оценка: комплимент] *Ты удивительный человек.* [Пожелание: счастье, успех; клише] *Желаю счастья и успехов во всём!*

Уже эти примеры показывают различия, которые схема делает измеримыми: форму обращения (термин родства с формулой уважения и вежливое *Siz* против имени и *ты*), набор концептов (здоровье и долголетие против счастья и успеха), наличие оценочного хода.

[Далее приводятся данные корпуса.]

Надёжность схемы. Значения к по признакам: [данные]. Метаданные и речевые ходы размечаются надёжнее, чем концепты пожеланий и тип выражения, требующие интерпретации.

Композиция жанра. Частотность речевых ходов в подкорпусах: [данные]. Обязательными являются ходы [данные], факультативными — [данные].

Концепты пожеланий. Распределение концептов по подкорпусам: [данные, диаграмма]. Доля концептов «здоровье», «долголетие» и «вера» в узбекском подкорпусе составляет [данные], в русском — [данные].

Социальные факторы. Зависимость длины текста, типа обращения и числа пожеланий от официальности, возраста и статуса участников: [таблица, χ^2 , p].

Обсуждение

Полученные результаты подтверждают, что качественные наблюдения о жанре поздравления можно перевести в проверяемые количественные данные. Это согласуется с положением Т. МакЭнери и Э. Харди о том, что корпусный подход дополняет интроспекцию исследователя эмпирическими свидетельствами [10]. Различия в форме обращения и наборе концептов между узбекским и русским подкорпусами соответствуют выводам М. Хакимова о центральной роли уважения к старшим в узбекском речевом поведении [6].

Наиболее сложным для разметки оказывается когнитивный уровень: границы между концептами «счастье» и «благополучие» размыты. Решением может служить опора на ключевую лексику: пожелание с лексемами *baxt*, *счастье* относится к концепту «счастье», с лексемами *farovonlik*, *достаток*, *благополучие* — к концепту «благополучие».

Практическое значение результатов состоит в том, что размеченный корпус может служить обучающей выборкой для классификации поздравлений и основой для оценки текстов, сгенерированных языковыми моделями. Схема также применима к другим этикетным жанрам (соболезнование, приглашение, благодарность) и в преподавании русского языка в узбекской аудитории.

Заключение

В статье предложена трёхуровневая схема корпусной разметки жанра «Поздравление», объединяющая теорию речевых жанров, фреймовую семантику и социолингвистику. Схема обеспечивает воспроизводимое описание жанра и служит основой лингвистического обеспечения прикладных систем обработки этикетных текстов. Перспективы исследования связаны с расширением корпуса, автоматической разметкой речевых ходов методами машинного обучения и сопоставлением узбекского материала с другими тюркскими языками.

ЛИТЕРАТУРА

1. Бахтин М. М. Проблема речевых жанров // Бахтин М. М. Эстетика словесного творчества. — М.: Искусство, 1979. — С. 237–280.
2. Карасик В. И. Языковой круг: личность, концепты, дискурс. — Волгоград: Перемена, 2002.
3. Формановская Н. И. Речевой этикет и культура общения. — М.: Высшая школа, 1989.
4. Шмелёва Т. В. Модель речевого жанра // Жанры речи. — Саратов, 1997. — Вып. 1. — С. 88–98.
5. Сафаров Ш. Прагмалингвистика. — Тошкент: Ўзбекистон миллий энциклопедияси, 2008.
6. Ҳақимов М. Ўзбек прагмалингвистикаси асослари. — Тошкент: Akademnashr, 2013.
7. Artstein R., Poesio M. Inter-coder agreement for computational linguistics // Computational Linguistics. — 2008. — Vol. 34, № 4. — P. 555–596.
8. Brown P., Levinson S. C. Politeness: Some Universals in Language Usage. — Cambridge: Cambridge University Press, 1987.
9. Fillmore C. J. Frame semantics // Linguistics in the Morning Calm. — Seoul: Hanshin, 1982. — P. 111–137.
10. McEnery T., Hardie A. Corpus Linguistics: Method, Theory and Practice. — Cambridge: Cambridge University Press, 2012.
11. Pustejovsky J., Stubbs A. Natural Language Annotation for Machine Learning. — Sebastopol: O'Reilly, 2012.
12. Searle J. R. A classification of illocutionary acts // Language in Society. — 1976. — Vol. 5, № 1. — P. 1–23.
13. Swales J. M. Genre Analysis: English in Academic and Research Settings. — Cambridge: Cambridge University Press, 1990.