

ЯЗЫК  
ТЕКСТ  
СОЦИУМ  
2023

LANGUAGE  
TEXT  
SOCIETY  
2023-CON

**4-я Международная  
научная конференция  
«Язык. Текст. Социум»**

**4th International  
Conference  
“Language. Text. Society”**

**19–20 октября 2023, Самара, Россия**

**October 19–20, 2023, Samara, Russia**

**Материалы конференции**

**Conference Proceedings**

Содержание		Table of Contents	
<b>Водоватова, Татьяна Евгеньевна</b> , Самара (Россия) Тексты комментариев в персональном общении: лингвостратегический аспект	3	<b>Tatyana Vodovatova</b> , Samara (Russia) Comments in personal communication: linguostrategic aspect	
<b>Мамилова, Светлана Абдурахметовна</b> , Москва (Россия) Испанские словари о русизмах «рубль» и «копейка»	4	<b>Svetlana Mamilova</b> , Moscow (Russia) Spanish dictionaries about Russianisms “ruble” and “kopeck”	
<b>Васина, Арина Владимировна</b> , Самара (Россия) Приём red herring как проявление ненадежного рассказчика	7	<b>Arina Vasina</b> , Samara (Russia) The red herring technique as a manifestation of unreliable narrator	
<b>Кириллов, Андрей Геннадьевич</b> , Самара (Россия) Фейк-новости как жанр и явление в медиадискурсе	11	<b>Andrey Kirillov</b> , Samara (Russia) Fake news as a genre and phenomenon in media discourse	
<b>Хазова, Анастасия Борисовна</b> , Москва (Россия) Количественное измерение распределения частей речи на материале корпуса интернет-текстов	14	<b>Anastasiia Khazova</b> , Moscow (Russia) Quantitative Measurement of the Distribution of Parts of Speech Based on the Corpus of Internet Texts	
	21	<b>Timur Akishev</b> , Almaty (Kazakhstan) Exploring the relationship between loanword frequency and productivity	
	23	<b>Luca Iezzi</b> , Pescara (Italy) Linguistic and sociolinguistic aspects of online communication: the role of Instagram in shaping language and identity	
	26	<b>Anna Faktorovich</b> , Quanah, TX (USA) The New Computational-Linguistics Author-Attribution Method Used to Reattribute the British Renaissance to Six Ghostwriters	

## Тексты комментариев в персональном общении: лингвостратегический аспект

Водоватова Татьяна Евгеньевна  
Самарский университет государственного  
управления «Международный институт  
рынка»  
Самара, Россия  
vodovatovaimi@mail.ru

**Аннотация**—Выступление посвящено анализу речевых стратегий, используемых в комментариях в англоязычном персональном интернет-дискурсе.

**Ключевые слова**—интернет-дискурс; комментарии; речевые стратегии

Современная интернет-лингвистика характеризуется значительными достижениями, однако не все ее объекты и явления изучены достаточно подробно. В имеющихся работах, посвященных изучению комментариев, представлены особенности функционирования текстов этого типа в различных видах дискурса, а также лингвистические свойства текстов комментариев. В лингвостратегическом аспекте интернет-комментарии в англоязычном персональном дискурсе специально пока не изучались.

При выделении стратегий в текстах комментариев в англоязычном персональном дискурсе в настоящем исследовании мы исходили из представляющегося бесспорным положения о том, что выбор тех или иных стратегий как способа формально-семантической организации комментария определяется коммуникативными целями автора, и, соответственно, функциями

## Comments in personal communication: linguostrategic aspect

Tatyana Vodovatova  
Samara University of Public Administration  
“International Market Institute”  
Samara, Russia  
vodovatovaimi@mail.ru

**Abstract**—The presentation is focused on the analysis of speech strategies used in comments in English-language personal Internet discourse.

**Key words**—internet discourse; comments; speech strategies

текста. Специфической, жанрообразующей коммуникативной функцией интернет-комментария является оценка. Но это не единственная функция для данных текстов, которым присущи также функции интернет-дискурса, в рамках которого функционирует жанр комментария. К ним относятся цели социализации, инкультурации, самопрезентации, развлечения.

Анализ фактического материала позволил прийти к выводам о том, что стратегиями, организующими структурно-семантическое пространство текстов комментариев в англоязычном персональном дискурсе и позволяющими достичь максимального коммуникативного эффекта, являются стратегии согласия, несогласия и развития темы текста-стимула. Все эти стратегии можно отнести к стратегиям кооперативного характера; деструктивных стратегий в анализируемом материале не обнаружено.

## Испанские словари о русизмах «рубль» и «копейка»

Мамилова Светлана Абдурахметовна  
Московский государственный университет  
Москва, Россия  
serdle@mail.ru

**Аннотация**—Доклад посвящен актуальной теме присутствия русизмов в испанских словарях разных типов. В данном докладе рассматриваются словарные дефиниции слов «рубль» и «копейка», данные в тринадцати испанских словарях от 1793 г. до 2007 г. издания (это словарь Испанской королевской академии, словарь Ñ, словарь Марии Молинер, словарь Sopena, «Словарь испанского языка, с терминами науки и искусств» Эстебана Террероса-и-Пандо в 4 томах, словарь Anaya, словари Larousse, словарь испанского языка для иностранцев, словарь Clave, школьный учебный словарь, энциклопедический иллюстрированный словарь Espasa, словарь Саламанка). Была выявлена вариативность дефиниций в зависимости от специфики словаря, а также вариативность слова «копейка» и в орфографии, и в категории рода и числа. Мы также обратились к корпусу испанского языка (Corpes XXI) с целью рассмотреть количество употреблений слов «rublo» и «kopek» за последние годы в текстах разных жанров и тем и дать статистику. Мы брали употребление этих слов во всей испаноязычной зоне.

**Ключевые слова**—русизм; рубль; копейка; заимствование; испанский язык; реалья; испанский словарь

Наш доклад посвящен двум русизмам — денежным единицам в испанских словарях — rublo и kopek. Они были взяты как примеры присутствия русизмов в испанском языке, т. к. встречаются почти во всех рассмотренных нами словарях, в отличие от многих других русизмов. Так, в словаре Испанской королевской академии

## Spanish dictionaries about Russianisms “ruble” and “kopeck”

Svetlana Mamilova  
Lomonosov Moscow State University  
Moscow, Russia  
serdle@mail.ru

**Abstract**—The report is devoted to the topical topic of the presence of Russianisms in Spanish dictionaries of various types. This report discusses the dictionary definitions of the words “ruble” and “penny” given in thirteen Spanish dictionaries from 1793 to 2007 (these are the Dictionary of the Spanish Royal Academy, the dictionary of the Spanish, the dictionary of Maria Moliner, the dictionary of Sopena, the “Dictionary of the Spanish language, with terms of Science and Arts” by Esteban Terreros y Pando in 4 volumes, the dictionary of Anaya, Larousse dictionaries, the Dictionary of the Spanish language for foreigners, the dictionary of Slave, the school educational dictionary, the encyclopedic illustrated Espasa Dictionary, Salamanca Dictionary). The variability of definitions was revealed depending on the specifics of the dictionary, as well as the variability of the word “penny” both in spelling and in the category of gender and number. We also turned to the Corpus of the Spanish language (Corpes XXI) in order to consider the number of uses of the words “rublo” and “kopek” in recent years in texts of different genres and topics and to give statistics. We took the use of these words throughout the Spanish-speaking area.

**Key words**—Russianism; rublo; kopek; borrowing; Spanish; reality; Spanish dictionary

(21 издание) их насчитывается 25. Все рассмотренные нами словари были изданы в Испании с 1793 г. по 2007 г.: словарь Испанской королевской академии, словарь Ñ, словарь Марии Молинер, словарь Sopena, «Словарь испанского языка, с терминами науки и искусств» Эстебана Террероса-и-Пандо в 4 томах, словарь Anaya, словари

Larousse, словарь испанского языка для иностранцев, словарь *Clave*, школьный учебный словарь, энциклопедический иллюстрированный словарь *Espasa*, словарь Саламанка. Надо отметить, что «рубль» — одно из самых ранних заимствований, оно относится к XVI–XVII вв. по разным источникам, а «копейка» относится к XIX в. Среди методов, применявшихся в работе, можно назвать метод лексикографических дефиниций и метод количественного анализа. Новизной является то, что русизмы рассматриваются в контексте именно словарей, а не художественных или публицистических текстов.

Русские слова попадали в испанские словари и в XVII в. (тогда их было менее 10), в словарях XX в. уже около 30. Наши исследования в корпусе показали, что в эпоху глобализации тема заимствований в естественных языках приобретает особую актуальность. Изучение заимствований является важной задачей, это помогает понять, каковы социально-культурные взаимосвязи между странами. По-видимому, первые русизмы появились в результате первых дипломатических и культурных контактов. Естественно, что при этом в испанский язык проникали слова, обозначающие русские реалии. Это и есть т. н. безэквивалентная лексика, которая всегда представляет трудности при переводе и зачастую просто транслитерируется.

В испанском языке имеется несколько десятков русизмов (т. е. заимствований из русского языка). Мы будем рассматривать лексические заимствования, обозначающие российские реалии. Если слово *rublo* во всех словарях пишется одинаково, то слово *corres* имеет множество вариантов: *corresa*, *correso*, *сóрес*, *сореск*, *кóрес*, *кóрек*, *кóреск*. Различны и определения, данные в разных словарях этим словам. Это зависит от специфики словаря. В большинстве

словарей дается этимологическая справка об этих понятиях. Существуют различные виды словарей. В своей работе испанский лексикограф Х. Фернандес-Севиля выделяет такие словари: исторические (*Diccionario de autoridades 1726–1739* гг.), этимологические (словарь Короминаса начало 1980-х), идеологические, всеобщие (или тезаурусы, где собрана вообще вся лексика языка), структурные (слово определяется как система отношений), генеративные (где определяются сочетаемость слова, его контексты, правильное употребление; основаны на учении Н. Хомского), словари синонимии и антонимии, нормативные (*Diccionario de dudas*), специальные (например, словарь морских терминов), словарь конструкций и управления, (*Diccionario de construcción y régimen* Р. Х. Куэрво), двуязычные словари. Мы рассматривали различные словари — и этимологические, и энциклопедические, и толковые словари, а также словари словоупотребления. В докладе дается краткая характеристика словарям и определения слов «рубль» и «копейка» в каждом с переводом на русский язык. Для нашего исследования оказалось очень важным обращение к диахроническому корпусу испанского языка (CNDHE), который позволяет получить объективные данные о времени и количестве употребления слов «*rublo*» и «*corres*» в испанском языке, хотя мы отдаем себе отчет в том, что не все тексты, включающие эти русизмы, были оцифрованы. Мы ограничились промежутком с XII в. до 1975 г. Оказалось, что слово «*rublo*» бытует в испанских документах с 1779 г. : всего 43 раза в 17 документах. Слово «копейка» мы проверяли в разных написаниях; «*korrek*» там встретилось 1 раз в документе 1931 г. (мн. ч. *korreks* 1 раз в документе 1980 г.), «*corres*» ни разу, как и «*корес*», и «*кореск*». Мы также обратились к корпусу испанского языка *Corpes XXI* с целью рассмотреть

количество употреблений слов «rublo» и «корек» (сорес) за последние годы в текстах разных жанров и тем. Причем мы брали употребление этих слов во всей испаноязычной зоне. По актуальному испанскому корпусу XXI в. можно сделать вывод о том, что слово «rublo» чаще всего встречается в художественной литературе, а слово «копейка» в вариантах «сорес» и «корек» — исключительно в художественной литературе (главным образом, в переводах русских авторов). При этом вариант «сорес» гораздо более частотный, хотя в словаре Испанской королевской академии определение дается на вариант «корек», т. е. такое написание признано Академией нормативным для испанского языка.

#### **Список литературы и источников:**

1. Алексеев М. П. Русская культура и романский мир. Ленинград, 1985.
2. Античные теории языка и стиля. СПб, 1996.
3. Ахманова О. С. Словарь лингвистических терминов.
4. БСЭ. <https://bse.slovaronline.com/37160-RUBL>.
5. Влахов С., Флорин С. Непереваемое в переводе. Москва, 1980.
6. Волин Б. М., Ушаков Д. Н. Толковый словарь русского языка. Т. 3. Москва, 1939.
7. Касарес Х. Введение в современную лексикографию. Москва, 1958.
8. Лингвистический энциклопедический словарь, под ред. В. Н. Ярцевой. Москва, 1990.
9. Ожегов С. И., Шведова Н. Ю. Толковый словарь русского языка. Москва, 1997.
10. Садиков А. В. Испанский язык сквозь призму лексики. Москва, 2019.
11. Corpus del Español del Siglo XXI (CORPES). URL: <https://www.rae.es/corpes/> (дата обращения 02.06.2023).
12. Diccionario de la lengua española URL: <https://dle.rae.es/> (дата обращения 25.05.2023).
13. Fernández-Sevilla J. Problemas de lexicografía actual. Bogota, 1974.
14. Haensch G. Los diccionarios del español en el umbral del siglo XXI: Problemas actuales de la lexicografía. Los distintos tipos de diccionarios, una guía para el usuario. Bibliografía de publicaciones sobre lexicografía. Salamanca, 1997.

## Приём red herring как проявление ненадежного рассказчика

Васина Арина Владимировна  
Самарский университет государственного  
управления «Международный институт  
рынка»  
Самара, Россия  
arinavasina.00@list.ru

**Аннотация**—Данная работа исследует понятие red herring и его использование в качестве приема ненадежного рассказчика. Red herring описывает ситуацию, когда информация в истории преднамеренно вводит читателя (или слушателя) в заблуждение, отклоняя его от верного пути рассуждений. В работе проанализированы техники, которые авторы задействуют для создания ненадежного рассказчика, с фокусом на том, как использование red herring может влиять на мнение читателя о надежности и достоверности повествования. Также были рассмотрены эффекты, которые эти приемы оказывают на повествование, создавая сюжетные повороты и неожиданные развязки. Работа включает анализ нескольких известных литературных произведений, где рассматривается использование red herring как средство создания ненадежного рассказчика. Полученные результаты позволяют лучше понять принципы литературных приемов и их влияние на восприятие читателя.

**Ключевые слова**—нарратология; литературный прием; ненадежный рассказчик; отвлекающий маневр; red herring; повествование

### Введение

В нарративной литературе понятие *red herring* относится к технике, используемой авторами для введения читателя (или слушателя) в заблуждение. Red herring (или отвлекающий маневр) представляет собой информацию, улику или событие, которые намеренно

## The red herring technique as a manifestation of unreliable narrator

Arina Vasina  
Samara University of Public Administration  
“International Market Institute”  
Samara, Russia  
arinavasina.00@list.ru

**Abstract**—This paper explores the concept of red herring and its use as an unreliable narrator technique. “Red herring” describes a situation where information in a story deliberately misleads the reader (or listener), diverting from the correct path of reasoning. The paper analyzes the techniques that authors used to create unreliable narration, focusing on how the use of “red herring” can influence the reader’s opinion of the reliability and credibility of the narrative. The effects that these techniques have on the narrative, creating plot twists and unexpected endings, were also examined. The work includes an analysis of several famous works of literature that examine the use of “red herring” as a means of creating an unreliable narrator. The results obtained allow us to better understand the principles of literary techniques and their impact on the reader’s perception.

**Key words**—narratology; literary device; unreliable narrator; red herring; narration

привлекают внимание и отводят от верного пути рассуждений или от главной сюжетной линии.

Цель использования red herring состоит в том, чтобы вызвать у читателя эмоциональную или интеллектуальную реакцию, сделать его более вовлеченным в историю и заставить задумываться о возможных развитиях сюжета. Однако,

когда red herring используется в качестве приема ненадежного рассказчика, он может вводить читателя в заблуждение и нарушать его доверие к надежности повествования. Читатель может быть уверен в одной версии событий, но по мере развития истории сталкивается с противоречиями и новыми информационными фрагментами, которые переворачивают его представления с ног на голову.

В целом, red herring является мощным литературным инструментом, который помогает создавать интригу и напряжение, но он также требует осторожного использования, чтобы избежать чрезмерного разочарования у читателя. Это может быть источником дискуссий и интерпретаций, делая литературное произведение более глубоким и многогранным.

## Результаты и обсуждение

Термин red herring происходит из области охоты, где копченая селедка использовалась для отвлечения собак от прочих запахов. Государственные структуры и представители могут использовать red herring с целью отвлечь внимание мировой общественности от истинных виновников какого-либо серьёзного происшествия или для того чтобы дезинформировать противника — СМИ называют это «отвлекающим маневром». В литературе red herring может быть использован для создания интриги, напряжения и сюжетных поворотов. Он может вводить читателя в заблуждение, создавать ложные улики и подозрения, чтобы сохранить тайну или обеспечить неожиданный разворот событий.

Уже в 1836 в выпуске «Таймс» от 12 августа можно было увидеть использование red herring в качестве его изначального смысла, но как метафоры,

которая впоследствии вошла в употребление как «отвлекающий маневр»:

*Mr. Rice called Lord Lyndhurst's intimation that the Government might take off the whole of the stamp duty 'a false drag—the scent of a red-herring to draw off the hounds'.*

В 1884 в Liverpool Daily Post от 11 июля видно использование red herring именно как «отвлекающего маневра»:

*The talk of revolutionary dangers is a mere red-herring.*

Во вторых президентских дебатах в США в 2021 году между Миттом Ромни и Барак Обамой одна женщина спросила, как кандидаты «ограничат доступность штурмового оружия». Ни один из них не ответил на этот конкретный вопрос. Вместо этого Барак Обама произнес речь о предотвращении насилия до того, как оно выйдет из-под контроля, а Ромни сосредоточился на хороших школах и воспитании детей в семьях с двумя родителями. Они оба отвлекли внимание от первоначального вопроса, используя отвлекающие маневры. Этот пример наглядно демонстрирует, как red herring используется в политике для «сохранения лица», когда высокопоставленным лицам задают каверзные и неловкие вопросы или когда нельзя дать однозначный ответ.

Отвлекающий маневр может использоваться в учебных целях. Red herring присутствует в экзаменационных задачах для студентов-юристов, чтобы ввести их в заблуждение и отвлечь от правильного вывода по юридическому вопросу. Это предназначено для проверки понимания студентами основного права и их способности правильно различать существенные фактические обстоятельства.



В художественной литературе «отвлекающий маневр» используется автором для того, чтобы подбросить ложную зацепку, которая приводит читателей к неправильным выводам. Например, персонаж епископа Мануэля Арингаросы в «Коде Да Винчи» Дэна Брауна на протяжении большей части романа представлен так, как будто он находится в центре церковных заговоров, но позже выясняется, что он просто жертва, обманутая и подставленная настоящим антагонистом истории. В имени этого епископа Браун использовал игру слов: *aringa rosa* с итальянского переводится как «розовая селёдка».

В первой повести о Шерлоке Холмсе под названием «Этюд в багровых тонах» автор применяет тактику отвлечения внимания. Злодей, чтобы сбить с толку полицию и читателя, оставляет на месте преступления надпись на немецком “*Rache*” (что означает «месть»), заставляя персонажей и читателей ошибочно подумать, что в происшествии причастен немецкий человек.

В другом произведении Артура Конан Дойла, «Собака Баскервиль» (“*The Hound of the Baskervilles*”): В этом детективе Шерлок Холмса автор использовал *red herring* в виде «проклятия Баскервиль». Читатели были введены в заблуждение и направлены на мысль, что загадочное создание связано с легендарным проклятием, тогда как истинная причина злодеяния была совсем иной.

В комиксах, которые считаются «сплавом» литературы и изобразительного искусства, также можно найти примеры отвлекающего маневра. В «Хранителях» Алана Мура начинается с загадки того, кто убил персонажа по прозвищу Комик. На первых страницах художник дает подсказку — пару рук в коричневом свитере, тянущихся, чтобы убить Комика. Позже в книге появляется Холлис Мейсон, другой

персонаж в коричневом свитере, а так же читателю становится известно, что он ненавидит Комика. Однако это оказывается отвлекающим маневром, и Холлис Мейсон не убийца.

Компьютерные игры используют комбинацию из визуального, текстового и аудио-составляющего, чтобы произвести впечатление на аудиторию. С годами головоломки в видеоиграх становятся все более и более предсказуемыми, поскольку игроки улавливают различные приемы, используемые дизайнерами. В ответ на эту тенденцию дизайнеры начали использовать все больше отвлекающих факторов. Например, в игре “*Flight of the Amazon Queen*” игрокам противостоит горилла, преграждающая путь. В тексте сообщается, что гориллу нельзя ни отвлечь, ни победить. После некоторого исследования игрок обнаруживает неподалеку банан, но он ничего не делает; это просто отвлекающий маневр.

Прием *red herring* как проявление ненадежного рассказчика является важным и вызывающим интерес аспектом литературного искусства. Через использование этого приема, авторы могут вводить читателя в заблуждение, создавать интригу и неожиданные сюжетные повороты. Он позволяет повысить напряжение и удивить аудиторию, при этом подчеркивая сложность и неоднозначность литературного процесса. Отвлекающий маневр призывает к более глубокому чтению, анализу и исследованию произведений, что позволяет лучше понять сложность и многообразие творчества.

## Заключение

Прием *red herring* в контексте ненадежного рассказчика играет значительную роль в развитии сюжета и создании эффекта неожиданности в литературе. Этот хитроумный прием

позволяет авторам вводить читателя в заблуждение, вызывать сомнения и создавать напряжение. Он служит инструментом, с помощью которого рассказчик может управлять восприятием и мнением читателя, создавая уникальный литературный опыт.

Проявление отвлекающего маневра также обращает внимание на роль доверия в литературном процессе. Читатель должен быть осведомлен о возможности обмана и не всегда принимать информацию от рассказчика бесспорной правдой. Это призывает к развитию искусства критического мышления и глубокого анализа произведений, чтобы раскрыть скрытые мотивы и истинные намерения рассказчика.

Отвлекающие маневры помогают продлить тайну и напряжение, лежащие в основе истории. Как и любой поворот сюжета, они удерживают внимание читателя, удивляя его, но в конечном итоге не раскрывая тайны сюжета. Как только читатель попадет на отвлекающий маневр, он будет более склонен не доверять своим собственным инстинктам, и ему будет труднее составить мнение о том, что на самом деле происходит в истории. Это очень эффективно для создания и поддержания напряжения.

В целом, прием red herring в контексте ненадежного рассказчика представляет интересный элемент в литературе, который вызывает вопросы о правде, обмане и способности литературы манипулировать нашими ожиданиями и восприятием. Это позволяет нам лучше понять хитрость литературного ремесла и наслаждаться фрагментами, в которых лежит ключ к истине, заставляя нас возвращаться к произведениям с новым опытом и переосмысливать свои представления о повествовании.

В заключение также важно отметить, что данное исследование включает в себя лишь часть существующих примеров без составления какой-либо статистики или тенденций. Возможно продолжение изучения red herring в различных областях помимо художественной литературы, а также подробное рассмотрение влияния отвлекающего маневра и эмоциональной реакции читателей/слушателей.

### Список литературы:

1. Тимохин А. Что такое отвлекающий маневр? // recture.ru. URL: <https://recture.ru/common/что-такое-отвлекаяushchiy-manevr/>.
2. Dupriez B. M. 1991. A Dictionary of Literary Devices: Gradus, A-Z. Toronto; Buffalo: University of Toronto Press, 1991.
3. Lieb M., Mason E., Roberts J., Rowland C. The Oxford Handbook of the Reception History of the Bible. Oxford University Press, 2011. <https://doi.org/10.1093/oxfordhb/9780199204540.001.0001>
4. Literary Devices. <https://www.literarydevices.com/>.
5. Moore A. The Watchmen (comic book). Archived from the original on 24 June 2016. <https://archive.org/details/WatchmenByAlanMooreAndDaveGibbons/mode/2up>.
6. Quinion M. 2008. The Lure of the Red Herring // World Wide Words. <https://worldwidewords.org/articles/herri-ng.htm>.
7. Red Herring: Definition and Examples // Literary Terms. <https://literaryterms.net/red-herring/>.
8. red herring // Oxford English Dictionary. OED Third Edition, September 2009. Archived from the original on 24 September 2015.
9. Tindale C. W. Fallacies and Argument Appraisal. Cambridge University Press, 2007.
10. Turco L. The Book of Literary Terms: The Genres of Fiction, Drama, Nonfiction, Literary Criticism and Scholarship. Hanover, NH: University Press of New England, 1999.

## Фейк-новости как жанр и явление в медиадискурсе

Кириллов Андрей Геннадьевич  
Самарский университет государственного  
управления «Международный институт  
рынка»  
Самара, Россия  
andreygkirillov@yahoo.com

**Аннотация**—В докладе рассматривается сосуществование жанра фейк-новостей и фейк-новостей как части информационной войны или пропаганды в медийном дискурсе. Делается попытка выделить некоторые общие основания, которые однако не позволяют объединять два вида ложных новостей в один жанр.

**Ключевые слова**—медиадискурс; жанр; фейк-новость; пропаганда

Одним из наиболее интересных и обсуждаемых медийных явлений являются так называемые фейк-новости (fake news). Данное понятие, однако, может обозначать довольно различные сущности, что требует уточнения термина «фейк-новость». В научном мире лживые и недостоверные произведения получили название фактоидов. В. И. Карасик считает, что по своей семантике фактоиды не отличаются от фактов, так как принципиально не верифицируемы в текстах новостей, и главную роль занимает прагматический аспект, цель публикации или подача сообщения [Карасик 2017, 25]. Однако мы сталкиваемся с различными по своему характеру видами фейк-новостей.

С одной стороны под фейк-новостями понимаются сатирические произведения, часто публикующиеся в специализированных СМИ («Панорама» в России и The Onion в США). В научной литературе такие фейк-новости могут называться юмористическими [Шилихина

## Fake news as a genre and phenomenon in media discourse

Andrey Kirillov  
Samara University of Public Administration  
“International Market Institute”  
Samara, Russia  
andreygkirillov@yahoo.com

**Abstract**—The presentation examines the coexistence of the genre of fake news and fake news as part of info-war/propaganda in media discourse. An attempt is made to identify some common grounds, which, however, do not allow combining the two types of fake news into one genre.

**Key words**—media discourse; genre; fake news; propaganda

20196]. Цель подобных произведений – развлечение, акцентирование внимания на острых общественно-политических и иных событиях и явлениях, а также юмористическое изобличение журналистских и пропагандистских приёмов, штампов. Будем называть этот тип фейк-новостями<sup>1</sup>.

Когда термин стал использоваться Дональдом Трампом и для описания информационных войн, под фейк-новостями стали понимать откровенно искажённые и лживые новостные сообщения, порой граничащие с пропагандой и дезинформацией, лишённые сатирической составляющей. Цель таких сообщений заключается в устрашении, дезинформации, искажении представлений о каких-либо явлениях, странах или людях. Подобные фейк-новости стали оружием информационных войн, задействующих мощь СМИ и социальных сетей, то есть, продвигая подобные ложные сообщения как через официальные каналы, так и через

лидеров общественного мнения и широкую публику. Обозначим данный тип фейк-новостями<sub>2</sub>.

Структура обоих видов фейк-новостей не различается, поскольку они создаются по модели традиционных новостей. Единственным отличием может быть приписка мелким шрифтом в конце фейк-новости<sub>1</sub> о том, что данная информация представляет собой сатирическую ложь. Фейк-новости<sub>2</sub> никаких пометок не несут, поскольку их задача как раз в том, чтобы пройти неотличимо от достоверной информации.

Существуют попытки классификации фейк-новостей, в которых зачастую включаются оба вида [Суходолов 2017]. Развести их в таком случае можно только при выделении признака целеполагания. В некоторых исследованиях две разновидности фейк-новостей предлагается объединять в качестве «жанров-двойников» [Шилихина 2019а].

Основой для разграничения между двумя видами фейк-новостей выступает определение цели сообщений. Если цель — создать сатирическое, пародийное произведение, не скрывающее данного намерения, — мы имеем дело с фейк-новостями<sub>1</sub>. Если же целью становится преднамеренное манипулирование общественным мнением, дезинформация, то перед нами фейк-новости<sub>2</sub>. Единственным эксплицитным показателем выступает маркирование фейк-новостей<sub>1</sub> или целых сайтов, на которых они размещаются.

Можно выделить несколько параметров, по которым пытаться разграничивать анализируемые явления: ложность содержания (оба вида имеют ложное содержание), маркированность ложности (только у фейк-новостей<sub>1</sub>), сатира/пародия против политики (фейк-новости<sub>1</sub> против фейк-новостей<sub>2</sub>), манипулятивность (фейк-новости<sub>2</sub>),

формат (новость для обоих видов). Как видно, почти все эти параметры требуют наличия фоновых знаний, проверки фактов и особого знания, что в совокупности представляется не собственно языковым инструментарием. Чётких лингвистических характеристик каждого вида, отличающих бы их друг от друга, не обнаруживается. Наличие в фейк-новостях<sub>1</sub> иронии и сарказма не является универсальным, а считывание данных приёмов не представляется всеобщим.

Единственными объединяющими чертами выступают ложность содержания и формат новостей. Поскольку целью несатирических фейк-новостей является преднамеренное введение в заблуждение, манипуляция, пропаганда и формирование общественного мнения на ложных основаниях, нельзя автоматически считать их новостями или просто каким-то зеркальным отражением сатирических фейк-новостей. Поэтому объединение фейк-новостей<sub>1</sub> и фейк-новостей<sub>2</sub> в гипержанр медиадискурса представляется невозможным.

Фейк-новости<sub>1</sub> представляют собой сатирический жанр медиадискурса. Фейк-новости<sub>2</sub> являются не отдельным жанром, а лишь искажённой формой жанра новостного сообщения.

### Список литературы:

1. Карасик В. И. Фактоиды как лингвокультурное явление // Политическая лингвистика. 2017. №3. С. 21-30.
2. Суходолов А. П. Феномен «Фейковых новостей» в современном медиaprостранстве // Евразийское сотрудничество: гуманитарные аспекты. 2017. №1. С. 87-106.
3. Шилихина К. М. Жанры-«двойники»: пародийные жанры в компьютерно-опосредованном дискурсе // Коммуникативные исследования. 2019а. №3. С. 810-826.

[https://doi.org/10.25513/2413-6182.2019.6\(3\).810-826](https://doi.org/10.25513/2413-6182.2019.6(3).810-826).

4. Шилихина К. М. Юмористические фейковые новости как

коммуникативная технология и жанр онлайн-коммуникации // Известия ВГПУ. 2019б. №5 (138). С. 139-143.

## Количественное измерение распределения частей речи на материале корпуса интернет- текстов

Хазова Анастасия Борисовна  
Национальный исследовательский  
университет «Высшая школа экономики»  
Москва, Россия  
ahazova@hse.ru

**Аннотация**—В статье содержится обзор подходов к изучению корпусов интернет-текстов: приводится история создания таких коллекций, перечисляются главные корпуса, собранные отечественными и зарубежными исследователями и повлиявшие на подход к сбору данных для них. Рассматриваются плюсы и минусы использования интернет-корпусов для исследования языковой системы: с одной стороны, подчеркивается простота сбора данных и их машиночитаемость, а также возможность получения социолингвистически и жанрово сбалансированной репрезентативной выборки. С другой стороны, если эти параметры будут нарушены, статические данные о присутствии языкового явления в пространстве интернет-коммуникации окажутся недостоверными, то есть будут вводить в заблуждение. Кроме того, описываются статистические меры, зарекомендовавшие себя в области анализа корпусов текстов. Специфической целью данного исследования является определение способности количественного измерения частеречного распределения единиц в корпусе интернет-текстов указывать на значимые колебания, мотивированные изменениями средств ввода информации. Для ее реализации был проведен эксперимент по выявлению частеречного распределения на материале корпуса текстов, собранных в социальной сети Vkontakte. Он позволил обнаружить особенности функционирования в них местоимений-существительных и глаголов. Было установлено, что эти части речи демонстрируют ответ на особенности канала коммуникации и воздействие ситуативных экстралингвистических факторов. Особенно выделяется поведение

## Quantitative Measurement of the Distribution of Parts of Speech Based on the Corpus of Internet Texts

Anastasiia Khazova  
National Research University Higher School of  
Economics  
Moscow, Russia  
ahazova@hse.ru

**Abstract**—The article provides an overview of approaches to the study of Internet text corpora: the history of the creation, the main corpora collected by domestic and foreign researchers and influenced the approach to the collection of this type of corpus are listed. The pros and cons of using Internet corpora for the study of the language system are considered: on the one hand, the simplicity of data collection and their machine-readability are emphasized, as well as the possibility of obtaining a sociodemographically and genre-balanced representative sample. On the other hand, if this parameter is violated, it is possible to obtain unreliable static data about the presence of a linguistic phenomenon in the Internet communication space. In addition, statistical measures that have proven themselves in the field of text corpus analysis are described. The objective of this study is to determine the ability of quantitative measurement of the partial distribution of units in the corpus of Internet texts to indicate significant fluctuations motivated by changes in the means of information input. An experiment was conducted to establish a partial distribution on the material of the corpus of texts collected in the Vkontakte social network. The peculiarities of the functioning of subject pronouns and verbs were found on the corpus material. In the course of the study, it was found that these parts of speech potentially demonstrate a response to the peculiarities of the communication channel and the effects of extralinguistic factors. In this regard, the behavior of personal pronouns is particularly distinguished – there is a sharp drop in the use of "I" and the increasing use of "we" adjacent to it. The use of personal pronouns, on the one hand, is one of the

личных местоимений – наблюдается резкое падение употребления «я» и возрастающее употребление «мы», выступающего в качестве предпочтительной альтернативы. Использование личных местоимений, с одной стороны, является одним из ярких маркеров устной речи и потому наследуется компьютерно-опосредованной коммуникацией. С другой, оно демонстрирует социолингвистические тенденции, происходящие в интернет-пространстве и изменяющие жанровые характеристики производства текстов. В заключении делается вывод о возможности использования частеречного измерения корпусов интернет-текстов для отслеживания влияния канала коммуникации и экстралингвистических факторов на язык, а также обсуждаются дальнейшие перспективы исследований в этой области.

**Ключевые слова**—частеречное распределение; корпус текстов; коммуникация; социальные сети; интернет

## Основы сбора корпусов интернет-текстов

Социальные сети, появившиеся более 25 лет назад<sup>1</sup>, стали одним из быстро прижившихся способов опосредованной коммуникации между людьми. Они открыли для исследователей быстрый доступ к текстам, авторами которых становятся люди разных возрастов, геокультурных пространств и круга интересов. Это позволяет объединять множество текстов в лингвистические корпуса, отражающие языковые особенности интернет-коммуникации и позволяющие отследить их эволюцию.

За сравнительно небольшой срок существования лингвистических корпусов сложилась устойчивая практика их описания: исследователи фиксируют в метаданных такие показатели как объем, тип разметки, источники, из которых были

brightest markers of oral speech and is inherited by computer-mediated communication. On the other hand, it demonstrates the trends taking place in the Internet space and changing the nature of writing texts. In conclusion, it is concluded that it is possible to use partial measurement of Internet text corpora to track the influence of the communication channel and extralinguistic factors on the language, and further prospects for research in this area are discussed.

**Key words**—partial distribution; corpus of texts; communication; social networks; Internet

получены тексты, а также информацию о периоде их создания и геокультурной среде. Опционально также указывается информация о лицензии, релевантных публикациях и условиях использования.

Целью нашего исследования является определение способности количественного измерения частеречного распределения единиц в корпусе интернет-текстов указывать на значимые колебания, мотивированные изменениями средств ввода информации. Как мы предполагаем, количественное измерение частеречного распределения в корпусах, собранных на основе социальных сетей, способно показать переход от активного использования персональных компьютеров на мобильные устройства с точки зрения изменения способа ввода информации, а также установить элементы текста, которые оказались наиболее чувствительны к этому процессу. С точки зрения лингвистики здесь становится

1 Первая социальная сеть Classmates появилась в 1995 г.

важен сопутствующий фактор: ограничение допустимой длины текста, принуждающее к максимальной экономии средств выражения мысли.

К достоинствам интернет-корпусов как объектов исследования текста можно отнести простоту и скорость автоматической обработки больших массивов данных, в которые можно включать тексты с различными социалингвистическими и стилистическими параметрами, отражающие разнообразные особенности речевого поведения. Но в этом случае критически важно обеспечить ее репрезентативность, иначе результаты анализа могут оказаться несостоятельными.

Важно также обеспечить точность лингвистической разметки. В работе морфологических и синтаксических анализаторов встречаются ошибки, единого подхода к аннотированию семантических характеристик еще нет, в любом типе разметки может присутствовать субъективность. Также стоит упомянуть и требования к интерпретации данных (результатов), полученных на исследовательской выборке. Их нельзя экстраполировать на весь язык без дополнительной проверки.

Источником текстов для нашего исследования стала социальная сеть Vk. Мы намеренно не выбрали популярные платформы, ведущим материалом которых являются видеоролики (Youtube, TikTok), поскольку информации, интересной с точки зрения полнотекстовой коммуникации, там не так много.

Объем корпуса составил 81485414 словоупотреблений (3570918 постов). В выборку включены тексты пользователей в возрасте от 17 до 55 лет за 2012-2016 г. Такой подход к сбору данных позволяет учесть различия в пользовательских привычках использования сети Интернет

для создания текстов, так как опыт использования персональных компьютеров и мобильных телефонов, сформированный у людей с разными социалингвистическими характеристиками, значительно отличается. Такой охват разных платформ позволяет создать репрезентативный корпус, который объединяет тексты разных периодов (даты и времени написания сообщения) и стилистического своеобразия. Положительной чертой некоторых социальных сетей является наличие информации об устройстве ввода, а также данных о географии создания постов, но в случае с Vk такая информация предоставляется не всегда.

Часть характерных особенностей корпуса интернет-текстов диктуется социодемографическим составом пользователей и тематической направленностью сети. Так, лексика сообщений (т.е. характер речевого поведения) может изменяться в зависимости от среднего возраста авторов, поскольку пользователи используют Vk для не только для целей общения, но и ведения бизнеса, деловой переписки. Лингвистические эксперименты, проводимые на корпусах интернет-текстов, зачастую имеют ограниченный список управляемых факторов. Связано это, в первую очередь, с характером выборки - данные таких корпусов стараются подбирать по заранее известным параметрам, например, на узком круге платформ за определенный период, ограничивая поиск тематической группой. Но заранее предсказать, попадет ли интересующее нас явление в достаточном количестве в набор текстов интернет-коммуникации, нельзя.

Кроме того, нельзя исключать влияние на ее состав средств ввода информации -- компьютеров и мобильных телефонов -- некоторые конструкции удобнее вводить с клавиатуры компьютера.



Например, в начале 2010-х годов было популярно написание текста с переменным использованием символов верхнего и нижнего регистра.

В последние годы наблюдается также преднамеренное смешение русской и английской лексики, кириллицы и латиницы, использование большого количества специальных символов, для которых требуется переключение раскладки. Однако вводить короткие тексты, сопровождаемые смайликами, значительно удобнее на мобильных телефонах. Эргономика этих устройств также влияет на длину сообщений, поэтому интерфейс многих социальных сетей изначально разрабатывался под экраны мобильных телефонов, в то время как более старые сети создавались для использования на компьютерах и адаптировали свои приложения и сайты (Bodomo 2010, 129).

Самой очевидной технологической чертой, которая неизбежно накладывает отпечаток на лингвистические характеристики интернет-текстов, создаваемых в мобильных приложениях, является ограничение их длины, измеряемой числом символов. Например, на платформе Twitter в момент появления оно было равно 140 символам, после 2017 года поднято до 280. Расхождения между платформами по параметру длины текста затрудняют сравнение лингвистических особенностей текстов, поскольку малый объем вынуждает пользователей искусственно упрощать (искажать)

стандартный лингвистический код. К тому же в ряде социальных сетей используется формат лонгрида – информационного текста, допускающий 1500 символов. Справедливо это и для социальной сети Vk, долгое время максимальное количество символов для заметок составляет 15985 знаков.

### **Количественное измерение распределения частей речи в корпусе интернет-текстов**

Первым этапом работы стало написание скрипта на Python, позволяющего вычленив из датасета информацию о дате написания поста, которая служит основанием для принятия программой решения о добавлении текста к одному из подкорпусов.

После этого с помощью библиотеки `rumystem3` была наложена морфологическая разметка. Выбор морфологического анализатора продиктован наличием возможности разрешения омонимии без привлечения дополнительных средств. Для сокращения времени последующей обработки было принято решение фиксировать только показатель части речи, а всю остальную информацию удалять. Следующим этапом стал подсчет числа употреблений каждой части речи в каждом корпусе. Все данные были объединены в таблицу 1, подсчитан также процент употребления каждой части речи относительно общего объема корпуса.

Таблица 1. Процентное соотношение распределения частей речи в корпусе интернет-текстов по годам

Table 1. Percentage ratio of the distribution of parts of speech in the corpus of Internet texts by year

	<b>2012</b>	<b>2013</b>	<b>2014</b>	<b>2015</b>	<b>2016</b>
<b>S</b>	<b>60,85%</b>	<b>63,04%</b>	<b>63,37%</b>	<b>63,41%</b>	<b>63,73%</b>
<b>A</b>	0,48%	0,52%	0,52%	0,50%	0,51%
<b>ADV</b>	1,76%	1,78%	1,77%	1,79%	1,79%
<b>ADVPRO</b>	0,00%	0,00%	0,00%	0,00%	0,00%
<b>ANUM</b>	0,00%	0,00%	0,00%	0,00%	0,00%
<b>APRO</b>	0,01%	0,01%	0,01%	0,01%	0,01%
<b>COM</b>	0,00%	0,00%	0,00%	0,00%	0,00%
<b>CONJ</b>	0,01%	0,01%	0,01%	0,01%	0,01%
<b>INTJ</b>	0,01%	0,00%	0,00%	0,00%	0,00%
<b>NUM</b>	0,00%	0,00%	0,00%	0,00%	0,00%
<b>PART</b>	0,02%	0,02%	0,02%	0,02%	0,02%
<b>PR</b>	0,00%	0,00%	0,00%	0,00%	0,00%
<b>SPRO</b>	<b>14,26%</b>	<b>13,10%</b>	<b>13,01%</b>	<b>12,94%</b>	<b>12,80%</b>
<b>V</b>	<b>22,61%</b>	<b>21,52%</b>	<b>21,28%</b>	<b>21,32%</b>	<b>21,12%</b>

В исследуемом материале не было обнаружено значительных различий в частеречном употреблении, но присутствуют некоторые отклонения, заслуживающие внимания.

В период с 2012 по 2016 годы частотность существительных подросла с 60,85% до 63,73% (отметим, что корпус 2016 года сопоставим по размеру с 2013, что связано с выборкой, по которой проводился парсинг - многие пользователи перестали оставлять текстовые записи). Вместе с тем, доля местоимения-существительных снизилась с 14,26% до 12,80%. Как известно, среди местоимений-существительных особую роль в опосредованной коммуникации играют личные местоимения. Это связано с тем,

что в интернет-коммуникации доминирует тенденция устной речи к использованию личных местоимений для выражения мнения по различным вопросам (Карпов 2020, 14).

Мы решили подробнее рассмотреть те вхождения, которые были отнесены морфологическим анализатором к SPRO. Для каждого корпуса был выделен список местоимений-существительных, посчитано количество употреблений каждого из них по годам, а результат сведен в общую таблицу. Сравнивать абсолютную частоту было бы некорректно, поэтому, чтобы проследить динамику изменения употребления конкретного местоимения, для каждого местоимения посчитан ipm, а также разница в процентах :

Таблица 2. Количество употреблений личных местоимений в корпусе интернет-текстов  
 Table 2. The number of uses of personal pronouns in the corpus of Internet texts

	2012	2013	2014	2015	2016	% разницы
я	25530	18273	17473	17044	16436	36
ты	11392	6947	6336	6299	6437	43
мы	<b>9565</b>	8306	<b>9210</b>	<b>9934</b>	<b>9787</b>	-2
вы	13886	10848	10888	10881	10856	22
они	9735	8191	8247	8212	8132	16
она	9395	6616	6506	6513	6221	34
он	14853	12489	11883	11636	11377	23
оно	587	493	466	481	485	17
кто	5248	3744	3617	3389	3266	38
что	23143	19567	19502	19341	19290	17
кто-то	1164	850	845	850	776	33
что-то	1268	1030	1047	1094	1100	13
никто	1195	866	863	806	765	36
ничто	188	140	141	145	138	26

Особенно обращает на себя внимание динамика личных местоимений, а именно падение употребления “я” - разница между 2012 и 2016 годами составила 36%, причем резкий спад пришелся на 2013 г. Аналогичная динамика прослеживается в употреблении других личных местоимений: “ты” (43%), “вы” (22%), “они” (16%), “она” (34%), “он” (23%); “оно” (17%).

Мы связываем эти изменения массовым отказом пользователей от личных записей, отдавая предпочтение общению в различных группах, механизмы создания и поддержки которых, а также возможности комментирования и работы с различными аудиовизуальными материалами претерпели в 2012-2013 г. множество изменений. Для личных записей стали использоваться другие платформы на базе мобильных приложений (Twitter, Instagram).

Отдельно стоит обратить внимание на местоимение “мы”, употребление которого

в рамках общей тенденции снизилось в 2013 году, а затем показало заметный рост к 2016 году, превысив аналогичное значение 2012 года. Объяснением этому может быть переосмысление Vk как социальной сети не столько для личных записей, сколько как платформы для рабочей коммуникации: продвижения товаров и услуг, а также приглашения на различные мероприятия.

Тексты интернет-пользователей, возникшие в процессе компьютерно-опосредованной коммуникации, сделали возможным применение множества количественных методов к большим массивам данных. Несмотря на некоторые недостатки, такой способ изучения языка позволяет отслеживать разнообразные изменения на различных уровнях языковой системы.

Рассматривая корпус текстов, составленный из постов пользователей

социальной сети Vk, мы провели подсчет распределения частей речи.

На материале проведенного исследования можно сделать вывод о том, что количественное измерение частеречного распределения в корпусах текстов способно отображать как переход от использования в социальных сетях персональных компьютеров к мобильным телефонам, так и более частные случаи - нововведения на платформе и пользовательские тенденции в написании личных текстов. Обращает на себя внимание чувствительность местоимения-существительных, наиболее резко проявившаяся в частотности их употребления в период 2012 и 2013 годов, но которая однако сгладилась с 2013 по 2016 годы.

В частности, обратили на себя внимание личные местоимения, являющиеся одним из ярких маркеров устной речи и унаследованные компьютерно-опосредованной коммуникацией.

Также было обнаружено, что с течением времени постепенно увеличивается доля существительных при

сохранении среднего достаточно низкого количества употребления глаголов. Такое поведение последних может проявляться в связи с активным использованием в электронной коммуникации эллипсиса (Карпов 2020, 16). Эти изменения могут быть связаны с переосмыслением функции социальной сети и требуют дополнительного сопоставительного анализа на материале других платформ.

В дальнейшем мы планируем собрать корпус текстов 2017-2021 г., чтобы проследить аналогичные тенденции в этом периоде и выяснить, какие изменения в языке возможно выявить при помощи количественных измерений различных уровней языка.

#### Список литературы:

1. Карпов И. А., Крылова Т. В., Тимошенко С. П. Комплексный лингвистический анализ электронного комментария в социальных сетях // *Русский язык в научном освещении*, 2020. 14, 16.
2. Bodomo A. *CMC for Linguistics and Literacy*, Hong Kong: IGI Global, 2010: 129.

# Exploring the relationship between loanword frequency and productivity

Timur Akishev  
KIMEP University  
Almaty, Kazakhstan  
t.akishev@kimep.kz

**Abstract**—This study focuses on two questions related to loanword adaptation, emphasizing the semantic, structural, and productivity-based features of the borrowed linguistic content. Two perspectives that appear to be contradictory are discussed, elucidating their contribution to the adaptation of loanwords in a recipient language.

**Key words**—loanwords; loanword frequency; morphological productivity; derivation

The current study seeks to determine whether frequency engenders the productive nature of loanword material, or if its productive capacity leads to higher frequency values. The following question is the first one discussed, tentatively termed the *frequency-productivity perspective*:

1. Does the frequent nature, or popularity, of a loanword make it more possible for that word to engage in derivational relations with the recipient language's means of word formation?

First of all, it is important to pay attention to the semantic aspect of the process of loanword adaptation, specifically, the original meaning represented by a given loanword and its relevance or novelty to the recipient sociolinguistic environment. If a loanword's embedded meaning is popular with the speakers of a certain language and novel enough to represent a culturally enriching concept, the frequency of that loanword should be expected to increase, as determined by corpus-based means of reflecting the language in use. This approach is largely based on speaker perceptions of frequent items and corpus-elicited data patterns.

Loanwords whose meanings denote concepts that already exist in the adapting language can also be characterized by relatively high frequency values. Corpus-based

analyses should focus on determining the differences between the frequential properties of cultural loanwords representing a novel concept and those of their core counterparts that encapsulate existing meanings. While cultural loans are frequent in corpora, and, presumably, the language as such, due to their popularity with diverse groups of speakers of the language, core loanwords acquire increased patterns of use because they are per se more conceptually familiar to the language users and easier for them to apply in a variety of contexts and constructions.

Thus, the meaning type should be a crucial notion to focus on when discussing the loanword adaptation process. The frequency of a loanword can be attributed to the novelty of the denoted concept to the recipient culture and the popularity of that concept among the speakers of the recipient language. Already existing notions embedded in loanwords, on the other hand, are easier to use due to their familiar nature. The increased application of the meaning or concept represented by a loanword in various sociocultural contexts in the recipient language undoubtedly leads to the diversification of the structural scenarios in which newly introduced forms are to be gradually adapted and used.

Therefore, the frequency-productivity perspective focuses entirely on the notion of

frequency represented by corpora and speaker-based data as the main contributor to the increased development of derivational patterns for loanwords to interact with and be adapted within the recipient language. The second approach discussed in this paper, the *productivity-frequency perspective*, is a terminologically opposite one, focusing on the contribution of the presence and development of productive forms to the increased frequency characteristic of the loanword material.

2. Does the presence of certain inherent characteristics of the loanwords facilitating the formation of a wide range of derivational patterns make the word more popular and frequently encountered in corpus-based environments and within language in use?

This approach is based almost entirely on the description of the morphological and semantic characteristics that facilitate the acquisition of the productive features by a given loanword item in the recipient language. It assumes that morphologically productive loanwords are inherently more popular and frequently encountered in the corpora or the recipient language as such due to the richness of the derived forms and derivational patterns responsible for the generation of novel lexical content in the language.

However, such a predominantly structural analysis does not include non-

productive loanwords that can still be frequent items in the corpora or popular among language users. Another limitation of this approach is connected with the selection of specific criteria or features related to the adaptational levels. According to the productivity-frequency perspective, certain cross-parametric combinations of semantic and structural types of loanwords are considered as generally more morphologically productive than others. It can be inferred that such a rule does not imply the presence of exceptions.

It is possible to conclude that the two approaches, however terminologically antithetic, may not necessarily focus on the same aspects of loanword adaptation. The first one emphasizes corpus-based and speaker-provided data and predominantly focuses on frequent items, while the second one elucidates the intrinsic linguistic features of the loanwords that can be analyzed across different parameters. Needless to say, both hypothetical approaches require further consideration, development, and testing to determine their applicability to all loanword types and parameters. They may also need to be considered as equally important to any loanword adaptation analysis, with neither excluding nor prized over the other.

# Linguistic and sociolinguistic aspects of online communication: the role of Instagram in shaping language and identity

Luca Iezzi

Università degli Studi “G. d’Annunzio” di Chieti-Pescara

Pescara, Italy

luca.iezzi@unich.it

**Abstract**—Code-switching, broadly defined here as the use of two or more language varieties in the same conversational event, is often looked at in spoken communication. However, in the past two decades, extra attention has been paid to language usage online, which finds itself at the boundary between written and oral language, going beyond the traditional borders separating the diamesic dimension and the diaphasic dimension. Furthermore, virtual communication may potentially inform us about linguistic aspects of contemporary societies. This contribution specifically looks at what role *bahasa gaul* and English play in code-switched captions on Instagram, performed by a qualitative sample of Indonesian young adults living in Bandung (Java). After analysing the aforesaid captions from a structural perspective, pointing out both extralinguistic and intralinguistic innovations, I will try to provide an overview of their functions, under the light of the English language imperialism, its place on language-in-education policy, and its importance in shaping people’s identity. Although the data come specifically from Indonesia, I will try to broaden the outcomes and give some general conclusions on online communication.

**Key words**—online communication; code-switching; English; *bahasa gaul*

## Introduction

Over the past decade, there has been increasing interest in studies about the web, especially social media (such as Facebook, Instagram, Twitter, but more recently even TikTok). Various sociolinguists, from their parts, have collected data to describe and categorise contact phenomena online, bringing about both qualitative and quantitative analyses from macro- and micro-sociolinguistic perspectives, because the language used online may potentially inform us about linguistic aspects of contemporary societies.

The present contribution concerns virtual communication in Indonesia, specifically story captions on Instagram written by young adults born in the first years of the 1990s. The analysis focuses both on a structural and functional perspective. The cornerstone of the study, so, is language practice online, characterised by dynamicity,

ephemeral forms, presence of ludic or cryptic colloquial elements, mixing between local and international traits. Studies on contact phenomena in Indonesia are scarce, and the few that are available are not always organic. Some researchers have tried to classify switching phenomena in captions, often making use of the traditional categorisations proposed by Muysken (2000) or Hoffman (1991). However, the majority of these studies have relied on captions written by celebrities or acquaintances on Instagram, but they do not seem to have a solid methodological foundation, and some results are incomplete or inaccurate. For this reason, the present contribution aims at giving a clearer and more organic overview of how online language is evolving.

## Linguistic innovations on Instagram

The data presented in the following sections show a very small picture of the wide corpus of over 3000 screenshots, containing different contact phenomena, concerning both specific Indonesian grammatical structures and the overall organisation of the sentence.

#### *N/Agg + -nya*

The Indonesian suffix *-nya* has multiple functions when attached to the preceding word (that could be a noun, an adjective or even a verb). The aforesaid functions can be summarised as follows: possessive adjective, determiner, agent, deverbal nominaliser, deadjectival nominaliser.

- |     |                    |                |       |
|-----|--------------------|----------------|-------|
| (1) | <u>Travel</u> -nya | <u>delay</u> , | [...] |
|     | travel-DET         | ritardo        |       |

The journey is delayed, [...]

- |     |       |     |         |                    |
|-----|-------|-----|---------|--------------------|
| (2) | Nanti | aku | foto-in | <u>detail</u> -nya |
|     | then  | I   | photo-V | detail-DET         |

Then I'll photograph the details

- |     |        |                    |
|-----|--------|--------------------|
| (3) | Samcan | <u>crispy</u> -nya |
|     | samcan | crispy-NOM         |

How crispy is the samcan (samcan crispness)

#### *me-/ber-/di- + V*

The Indonesian prefix *me-* is used to form transitive verbs from the root; in *bahasa gaul* this structure is substituted by *V + -in*. The prefix *ber-*, instead, is used to transform nouns into verbs or adjectives, emphasising the

function expressed by the noun root. The prefix *di-* is used to make the passive form, it is simply prefixed to the root, inverting the order of the constituents and indicating the agent through the use of the preposition *oleh*.

- |     |        |                     |         |          |
|-----|--------|---------------------|---------|----------|
| (4) | Sambil | men- <u>support</u> | sekolah | anak2?   |
|     | while  | PRFX-support        | school  | child-PL |

At the same time do they support the children's school?

- |     |       |      |          |      |                   |       |
|-----|-------|------|----------|------|-------------------|-------|
| (5) | [...] | coba | hasilnya | di   | <u>crosscheck</u> | dulu  |
|     |       | try  | product  | PRFX | crosscheck        | first |

[...] I first try the cross-checked product



### *Overall structure of the sentence*

(6)	<u>Thank you</u>	tante @-	meski	yg	diundang	<u>party</u> @-	tp @-
	thank you	aunt	although	that	invite-PASS	party	but
	masih	dapat	<u>after</u>	<u>party</u>	<u>unboxing</u>	kado	yang
	still	can	after	party	unboxing	gift	that
	ternyata	diinginkan	[...]				
	apparently	wish-PASS					

Thank you aunt @- though invited to the party but @- is still opening the gifts she wished for after the party [...]

### **Conclusions**

This contribution has aimed at giving an organic, though partial, overview of the language used online by Indonesian young adults, specifically on Instagram. Although the data represented here belong to a specific group of people coming from a specific area, the results can be applied to other contact situations taking place online. The investigation of the data showed some peculiar characteristics that transcend the typical borders between formality/informality and writing/speaking.

As displayed in the excerpts, different grammar structures as well as the overall organisation of a sentence can be affected by switching and mixing phenomena, meaning that online users utilise all the linguistic resources at their disposal in various ways, at

times inserting alien material in specific construction from the base language, other times in a more chaotic way.

Broadening the point of view from these limited examples to a wider perspective that includes all the linguistic practices online, I can say that multilingual speakers make use of the same strategies used in real vis-à-vis conversations also online, even in a less structured way, without following particular patterns, but behaving accordingly to that moment and that context where the caption is being written.

Studies on virtual language still leave some open questions. How do the speakers behave in informal conversations? Do they use the same strategies they use online? Do they switch in instant messaging? How do they perceive their behaviour? How will these language practices evolve?

# The New Computational-Linguistics Author-Attribution Method Used to Reattribute the British Renaissance to Six Ghostwriters

Anna Faktorovich  
Anaphora Literary Press  
Quanah, TX, USA  
director@anaphoraliterary.com

**Abstract**—The *British Renaissance Re-Attribution and Modernization* (BRRAM) series solves the previously critically discussed mysteries concerning the authorship of British Renaissance texts (including the “William Shakespeare” and 122 other bylines) by applying to 303 texts a newly invented for this study computational-linguistics author-attribution method that uses a combination of 27 different linguistic tests. This testing has derived that six ghostwriters wrote all of the tested works. By in part comparing biographical dates against the publication dates, these ghostwriters were proven to be: Richard Verstegan, Josuah Sylvester, Gabriel Harvey, Benjamin Jonson, William Byrd and William Percy. This presentation will describe and show how to apply the computational steps anybody can follow to solve any attribution mystery. It will explain study findings and the supporting internal and external, biographical and historical, and various other types of evidence that supports the quantitative conclusions.

**Key words**—computational linguistics; attribution; authorship; British literature

[BRRAM](#) is the first accurate quantitative re-attribution of all central texts of the British Renaissance. It describes and applies the first unbiased and accurate method of computational-linguistics authorial-attribution. This method includes 27 tests for punctuation, lexical density, parts of speech, passive voice, characters and syllables per word, psychological word-choice, and patterns of the top-6 words and letters. It covers 303 texts with 8,106,059 words, 123 authorial bylines, a range of genres, and a timespan between 1510 and 1662.

The presentation will include helpful diagrams that visually show the quantitative-matches and the identical most-frequent phrases between the texts in each linguistic-signature-group. The GitHub files include

detailed chronologies for each of the six ghostwriters and the bylines they wrote under, including their dates of birth, death, publications, and other biographical markers that explain why each of them was the only logical attribution. All of the raw and processed data is fully publicly-accessible on [GitHub](#).

To explain why these are the first truly accurate re-attributions, numerous reasons for discrediting previous attribution claims are provided throughout BRRAM; in this presentation visuals of matching handwriting samples across multiple bylines will be shown. The accuracy of this attribution method will also be shown through the 300+ text corpuses of 18<sup>th</sup> and 19<sup>th</sup> century British texts. You can [read a sample](#) out of Volumes 1-2 of BRAM.

