

Litera

Правильная ссылка на статью:

Морозов Д.А. Опыт корпусного анализа компаративов в современном чешском письменном дискурсе // Litera. 2025. № 7. DOI: 10.25136/2409-8698.2025.7.75167 EDN: ATSHHJ URL: https://nbpublish.com/library_read_article.php?id=75167

Опыт корпусного анализа компаративов в современном чешском письменном дискурсе

Морозов Даниил Александрович

ORCID: 0009-0004-2654-8477

аспирант; Филологический факультет; Московский государственный университет им. М.В. Ломоносова
Технический писатель; ООО "Датадванс"

606015, Россия, Нижегородская обл., г. Дзержинск, ул. Самохвалова, д. 3, кв. 68

✉ mda1998@yandex.ru



[Статья из рубрики "Лингвистика"](#)

DOI:

10.25136/2409-8698.2025.7.75167

EDN:

ATSHHJ

Дата направления статьи в редакцию:

13-07-2025

Дата публикации:

20-07-2025

Аннотация: В статье представлен опыт корпусного анализа компаративов прилагательных в современном чешском письменном дискурсе. Предметом исследования являются формы сравнительной степени прилагательных, образованные аналитическим и синтетическим способами. Исследование сфокусировано на возможности функционирования аналитических компаративных конструкций прилагательных в современном чешском языке, где такие формы считаются допустимыми только в случае невозможности образования синтетического компаратива. Тем не менее, в русском языке и в других языках славянской группы наблюдается сосуществование сложных и простых форм сравнительной степени, вследствие чего была выдвинута гипотеза об аналогичном сосуществовании компаративных форм чешских прилагательных. Целью исследования являлось обнаружение контекстов, где аналитические формы

используются вместо синтетических, а также определение возможности конкуренции таких форм. Источником фактического материала послужил корпус Synek, являющийся пропорционально редуцированным корпусом современных письменных текстов SYN2000. Отобранные контексты обрабатывались системным и функциональным методами, разработанными советской школой богоемистики. Построение таблицы конкуренции осуществлялось с помощью статистического метода. Научная новизна исследования заключается в надежном документировании выявленной конкуренции благодаря средствам компьютерного анализа данных. На материале корпуса Synek показано, что в современном чешском письменном дискурсе аналитические формы типа *více populární* могут употребляться и встречаются даже в тех случаях, когда существуют их синтетические аналоги типа *populárnější*. То есть вопреки чешской лингвистической традиции аналитические и синтетические компаративы могут конкурировать между собой, и контексты таких употреблений обнаружены в чешском национальном корпусе. В статье описан алгоритм построения таблицы конкуренции, а также приведены 72 найденных соответствия между аналитическими и синтетическими компаративными конструкциями – от формы *více rovný* «более равный», встретившейся в 12 контекстах и имеющей 5 соотносимых синтетических компаративов, до формы *víc silný* «более сильный», употребленной в 1 контексте и имеющей 343 соотносимые простые формы. Полученные данные позволяют судить об абсолютной и относительной частотности употребления аналитических компаративов в современном чешском языке.

Ключевые слова:

корпусные исследования, славистика, чешский национальный корпус, чешский язык, чешский письменный дискурс, сравнительная степень, компаратив, степени сравнения прилагательных, аналитические формы компаративов, конкуренция форм компаративов

Введение

Актуальность настоящего исследования определена глобальным интересом современных лингвистов к корпусным исследованиям и использованию компьютерных технологий для анализа больших данных. Обращение к корпусу как основному источнику фактического материала позволяет находить случаи употребления грамматических моделей в большом объеме текстов, а статистический анализ выявленных включений – судить о вероятности их появления в тех или иных контекстах. В условиях современного развития информационных технологий это на порядок повышает надежность полученных данных [1]. Например, исследование функционирования суперлативов на материале корпуса Synek показало, что аналитические суперлативные конструкции прилагательных в современном чешском языке образуются не только в случаях, когда образование соответствующих синтетических форм недопустимо, вопреки мнению авторов пражской «Русской грамматики» 1979 г. [2, 3].

Достижение упомянутой в аннотации цели исследования предполагало решение следующих задач:

- поиск в корпусе Synek контекстов, в которых употреблены аналитические формы сравнительной степени чешских прилагательных;
- исключение в процессе визуального анализа нерепрезентативных контекстов и лемматизация обнаруженных словоформ;

- построение таблицы конкуренции аналитических и синтетических компаративов прилагательных в современном чешском письменном дискурсе.

Выбор методов исследования обусловлен целью и поставленными задачами. Отбор фактического материала осуществлялся методом корпусного анализа: поиск в корпусе Synek производился с помощью корпусного менеджера Bonito (Manatee) v. 1.10.usck. Для обработки собранный в корпусе материал применялись системный и функциональный методы в том виде, в котором они были разработаны советской школой богоемистики, основанной заслуженным профессором МГУ им. М. В. Ломоносова А. Г. Широковой [4]. Построение таблицы и подсчет частотности осуществлялись с помощью статистического метода.

Материалом исследования послужил электронный корпус Synek (11 959 431 токен), являющийся десятикратно пропорционально редуцированным корпусом современных письменных текстов SYN2000 (120 908 724 токена). Данный корпус отличается от других корпусов в составе Чешского национального корпуса принципами отбора входящих в него текстов. Так как письменный текст не только отражает языковую ситуацию, но и формирует ее, влияя на идиолект читателя, для отбора текстов создатели корпуса SYN2000 опирались на результаты социологических исследований о читательских предпочтениях своих современников. Это обусловило степень представленности в SYN2000 текстов различного рода: основную часть языкового материала SYN2000 составляют публицистические тексты (60%), на втором месте – специальные тексты, справочники, энциклопедии и т.д. (25%), на третьем – беллетристика (15%). При этом корпус Synek достаточно велик (запрос на простые компаративы [tag="AA.....2.*"] дал 27207 контекстов употребления), чтобы адекватно представить функционирование таких единиц в современном чешском письменном дискурсе.

Теоретической базой исследования стали работы лингвистов Московской школы богоемистики [4], а также труды сотрудников отдела грамматики Института чешского языка Академии наук Чешской Республики, в частности, «Большая академическая грамматика литературного чешского языка» [5, 6] и выпускаемые ими журналы «Наша речь» [7], «Корпус-грамматика-аксиология» [8], «Слово и словесность» [9]. В качестве tertium comparationis полезными были также предшествующий вариант академической грамматики чешского языка [10] и «Грамматика современного чешского языка» для высших учебных заведений [11].

Практическая значимость исследования заключается в применении его результатов в прикладных лингвистических дисциплинах, таких как автоматическая обработка естественного языка (NLP) и машинное обучение (ML): при совершенствовании систем машинного перевода и обучении больших языковых моделей (LLM). Полученные данные могут быть полезны в академических курсах стилистики и грамматики чешского языка.

Экскурс в теорию проблемы

Среди языков славянской группы чешский в большей мере обладает чертами синтетизма, что проявляется в том числе и в вопросе градационности качественного признака прилагательных [12]. Это создает пресуппозицию о тенденции чешского языка использовать синтетический способ выражения компаративности везде, где это возможно. Синтетические модели образования степеней сравнения в славянских языках происходят из общих архетипов *j̊s, *ěj̊s [13, с. 59], но если русский язык продуктивно образует компаратив двумя способами: аналитическим сочетанием «более +

положительная степень» или синтетической формой, образованной суффиксом *-ejš/-ajš-*, за которым следуют флексии положительной степени – то в современном чешском языке «степени сравнения прилагательных представлены системой простых форм», а «сравнительная степень образуется посредством суффиксов *-ejš/-ajš-* и *-š-*: *povější*». Если «по формально-техническим причинам» чешское прилагательное не образует простых форм сравнительной степени, пражская «Русская грамматика» допускает возможность образования «описательных сочетаний *v i c e <...>* + положительная степень: *více překvapující*» [14, с. 349-350].

В академической грамматике чешского языка конца XX века, в разделе «Морфология», компаративу (чеш. komparativ) отводится один абзац. В нем указано, что эти формы прилагательных противопоставлены суперлативу (чеш. superlativ) и позитиву (чеш. pozitiv), поскольку выражают более высокую степень качества, обозначаемого прилагательным, по сравнению с позитивом, например: *Jan je vysoký, Jiří vyšší a Zdeněk nejvyšší* 'Ян высокий, Иржи выше, а Зденек самый высокий' [10, с. 80].

В современном издании «Большой академической грамматики литературного чешского языка» компаратив описан более подробно. Подчеркивается возможность его образования преимущественно у качественных прилагательных, а редкие случаи формирования компаратива от относительных прилагательных трактуются как результат окказионального перехода последних в разряд качественных, например: *písek stříbrnější a květy mimoz zlatější* 'песок **более серебряный** и цветы мимоз **более золотые**' [5, с. 54]. Авторы грамматики отмечают, что основа сравнения в компаративе («более интенсивное качество по сравнению с кем-либо/чем-либо») каждый раз устанавливается контекстуально. Это позволяет считать вполне нормативными конструкции типа *Pan Tomášek byl čtyřikrát starší, než jeho nejstarší dcera* 'Господин Томашек был вчетверо старше своей старшей дочери' [6, с. 854]. В данном случае суперлатив *nejstarší* указывает на высшую степень признака лишь в пределах ограниченной группы (дочери господина Томашека), в то время как за ее границами (включая самого господина Томашека) могут существовать носители этого признака в еще большей степени. Ограничение по аспекту сравнения выделяется как один из критериев разграничения омографичных форм компаратива и суперлатива во всех славянских языках (ср. рус. форму *старший*, которая может выражать как сравнительную степень, так и превосходную) [13, с. 62].

В «Грамматике современного чешского языка» отмечается, что лишь около 6% прилагательных способны образовывать формы сравнительной степени. При этом полный парадигматический ряд, включающий все три степени сравнения, в современном чешском языке характерен только для 3% прилагательных, что составляет около двух тысяч лексем [11, с. 205].

Итак, обзор справочных материалов подчеркнул стремление чешского языка к синтетическим средствам выражения степени сравнения, но также указал на ограниченность образования компаративных форм качественными прилагательными. Логично предположить, что при таком ограничении в чешском языке могут использоваться аналитические компаративные конструкции, однако особенно любопытно проверить возможность функционирование сложных форм сравнительной степени в контекстах, где употребление простых форм является возможным.

Корpusное исследование частотности компаративов

Корпусный менеджер Bonito позволяет производить поиск по словоформе, по лемме, по грамматической матрице и их сочетаниям. Для поиска простых форм компаративов использовался запрос `[tag="AA.....2.*"]`, а поиск аналитических форм сравнительной степени чешских прилагательных в корпусе Synek осуществлялся с помощью запросов `([word="[Vv]íc"] [tag="AA.....1.*"])` и `([word="[Vv]íce"] [tag="AA.....1.*"])`. Об ограниченной употребительности таких аналитических компаративных конструкций свидетельствует тот факт, что на 27207 контекстах, полученных по запросу на простые компаративы, приходится 306 и 132 контекста, полученных по запросам на сочетания «více + положительная степень» и «víc + положительная степень» соответственно. Из этих 438 контекстов продуктивными оказались 174, другие же были исключены, так как в них квантификатор více/víc относился не к прилагательному, ср. *Ve 20. letech bylo v Německu <víc obrázkových> časopisů než kdekoli jinde ve světě* 'В 1920-е гг. в Германии было больше иллюстрированных журналов, чем где бы то ни было в мире'.

Полученные результаты необходимо было выверить посредством визуального контроля и лемматизировать, так как, например, запрос на простые компаративы `[tag="AA.....2.*"]` дал 27207 контекстов употребления 2346 прилагательных в форме сравнительной степени, среди которых встречались дублеты и контексты с опечатками. Составители корпуса исходили из того, что опечатки в письменном тексте заслуживают внимания как потенциальный объект корпусного анализа, и сознательно оставили во входящих в состав корпусов SYN2000 и Synek текстах опечатки. Такие опечатки могли скрывать нерепрезентативные сочетания букв типа *ml*, *věro*, ошибочно выделенные в отдельные прилагательные части слов типа *jší*, *nější* или же слитые с предыдущей словоформой компаративы типа *přesvědčen*, *že mohu zažít <i>ilepší</i> sezónu*.

Следует упомянуть и особенности поиска по корпусу, которые повлияли на абсолютные и относительные результаты корпусного исследования:

1. Корпусный менеджер Bonito не смог соотнести с формой позитива = формы с лишними или пропущенными буквами (*nižších* < *nižší* < *nízký* 'низкий'; *vyší* < *vyšší* < *vysoký* 'высокий'), с употребленным дефисом вместо знака переноса (*lep-ší* < *lepší* < *dobrý* 'хороший'), с пропущенной или неверной диакритикой (*menších* < *menší* < *malý* 'маленький'), а также с фиксацией нелитературной речи (*lepčím* < лит. *lepší* < *dobrý* 'хороший').
2. В корпусе Synek был обнаружен один контекст с огласовкой *exkluzívnejší* и один контекст с огласовкой *exkluzivnější*, из которых программа восстановила формы позитивов *exkluzívni* и *exkluzivní* 'эксклюзивный'. При этом имеется в виду одна и та же лексема, приведенная в соответствии со старой (*exkluzivní*, *exkluzivnější*) и новой (*exkluzívni*, *exkluzivnější*) орфографическими нормами. Интересно, что чешский модуль орфографии текстового процессора Microsoft Word 1997 выделяет формы с долгим í (*exkluzívni*, *exkluzivnější*) как ошибочные, одобряя при этом формы с кратким í (*exkluzivní*, *exkluzivnější*).
3. С позитивом *šťastný* 'счастливый' менеджер Bonito соотнес не только формы компаратива *šťastnější* (38 контекстов обнаружено в результате запроса `[(tag="AA.....2.*") & (lemma="šťastný")]`), но и формы компаратива *nešťastnější* (3 контекста в корпусе Synek), безусловно связанные с другим позитивом, а именно – *nešťastný* 'несчастный'. Сходная ситуация наблюдалась с позитивом *bezpečný* 'безопасный': программа соотнесла с ним не только формы компаратива *bezpečnější* (55 контекстов в корпусе Synek), но и формы компаратива *nebezpečnější* (41 контекст в корпусе Synek), связанного с позитивом *nebezpečný* 'опасный'.

4. В корпусе Synek обнаружены два контекста с формой *významější* и один контекст с формой *významějšího*, которые должны выглядеть как *významnější* и *významnějšího* соответственно и включаться в статистику употребления компаратива *významnější* и позитива *významný* 'значительный'. Причиной ошибочных написаний в данном случае является одинаковое чтение сочетаний *tě* и *tpě* в чешском языке.

Результаты экспериментального исследования

Исключив неподходящие контексты и используя леммы, то есть приводя словоформы прилагательных к основе, удалось сократить дублеты и выделить 1330 прилагательных, которые возвращались на запрос [tag="AA.....2.*"]. Аналогичным образом было выявлено 240 прилагательных в коллокации с *více* и 113 – в коллокации с *víc*. 72 из них были сопоставлены на основе общего позитива, к которому восходят простые и сложные формы.

В результате была составлена *Таблица 1*, свидетельствующая о сосуществовании и конкуренции соотносимых аналитических и синтетических форм сравнительной степени в современном чешском языке по данным корпуса Synek. В третьей колонке указано число обнаруженных программой аналитических конструкций, в пятой – число соотносимых с ними синтетических компаративов, а в шестой – доля аналитических конструкций. Во втором столбце представленной таблицы приводится русский лексический эквивалент, зафиксированный в классическом чешско-русском словаре, изданным пражским «Государственным педагогическим издательством» совместно с московским издательством «Просвещение» [15]. Из предлагаемых в данном словаре русских лексических соответствий выбирались те, которые максимально соотносились с обнаруженными контекстами [16].

více rovný	более равный	12	rovnější	5	0,706
víc závislý (1)	более зависимый	3	závislejší	2	0,600
více závislý (2)					
víc soustředěný	более сосредоточенный	2	soustředěnější	2	0,500
více vyhledávaný	более популярный	2	vyhledávanější	2	0,500
více energický	более энергичный	1	energičejší	1	0,500
více hravý	более игривый	1	hravější	1	0,500
více nepostradatelný	более незаменимый	1	neposradatelnější	1	0,500
víc nepravděpodobný	более неправдоподобный	1	nepravděpodobnější	1	0,500
více nepříznivý	более неблагоприятный	1	nepříznivější	1	0,500
více postižený	более пострадавший	1	postiženější	1	0,500
více povedený	более удачный	1	povedenější	1	0,500
víc povědomý	более знакомый	1	povědomější	1	0,500
více průměrný	более средний	1	průměrnější	1	0,500
více roztríštěný	более раздробленный	1	roztríštěnější	1	0,500
více smyslný	более чувственный	1	smyslnější	1	0,500
víc špinavý	более грязный	1	špinavější	1	0,500
více vyvinutý	более развитый	1	vyvinutější	1	0,500
více harmonický	более гармоничный	1	harmoničtější	2	0,333

více ...nější,	...více ...nější...	-	...nější...	-	...
více katolický	более католический	1	katoličtější	2	0,333
více dominantní	более доминантный	1	dominantnější	3	0,250
víc nešťastný	более несчастный	1	nešťastnější	3	0,250
více přizpůsobivý	более легко приспособляющийся	1	přizpůsobivější	3	0,250
víc pyšný	богее гордый	1	pyšnější	3	0,250
víc ambiciozní	более амбициозный	1	ambicióznější	4	0,200
víc dráždivý	более раздражительный	1	dráždivější	4	0,200
více duchovní	более духовный	1	duchovnější	4	0,200
více kvalifikovaný	более квалифицированный	1	kvalifikovanější	4	0,200
víc vyvážený	более уравновешенный	1	vyváženější	4	0,200
více aktuální	более актуальный	1	aktuálnější	5	0,167
více patrný	более явный	1	patrnější	5	0,167
více podobný	более похожий	1	podobnější	5	0,167
více přiměřený	более соразмерный	1	přiměřenější	5	0,167
více sebevědomý	более самоуверенный	1	sebevědomější	5	0,167
více důvěryhodný	более достоверный	1	důvěryhodnější	6	0,143
více rozšířený	более распространенный	1	rozšířenější	6	0,143
víc zranitelný (1)	более ранимый	2	zranitelnější	13	0,133
více zranitelný (1)					
víc problematický	более проблемный	1	problematictější	7	0,125
víc smutný	более грустный	2	smutnější	16	0,111
více náchylný	более склонный	1	náchylnější	8	0,111
více nervózní	более нервный	1	nervóznější	8	0,111
více propracovaný	более проработанный	1	propracovanější	8	0,111
více racionální	более рациональный	1	racionálnější	8	0,111
více realistický	более реалистический	1	realističtější	8	0,111
víc přirozený	более естественный	1	přirozenější	9	0,100
více potřebný	более необходимый	1	potřebnější	10	0,091
více zřejmý	более очевидный	1	zřejmější	10	0,091
více živý	более живой	1	živější	11	0,083
více rafinovaný	более рафинированный	1	rafinovanější	12	0,077
víc agresivní	более агрессивный	1	agresivnější	14	0,067
více komplexní	более комплексный	1	komplexnější	14	0,067
více luxusní	более роскошный	1	luxusnější	14	0,067
více sympathetický	более симпатичный	1	sympatičtější	17	0,056
více spokojený	более довольный	1	spokojenější	21	0,045
více vyrovnaný	более сбалансированный	1	vyrovnanější	21	0,045
více důrazný	более настоятельный	1	důraznější	22	0,043
více otevřený	более открытый	1	otevřenější	25	0,038
více přitažlivý	более привлекательный	1	přitažlivější	26	0,037

více jistý	более несомненный	1	jistější	27	0,036
víc populární	более популярный	1	populárnější	31	0,031
více opatrný	более осторожный	1	opatrňejší	32	0,030
více citlivý	более чувствительный	1	citlivější	34	0,029
více pravděpodobný	более правдоподобный	1	pravděpodobnější	35	0,028
více efektivní	более эффективный	1	efektivnější	41	0,024
více nebezpečný	более опасный	1	nebezpečnější	42	0,023
více pohodlný	более удобный	1	pohodlnější	44	0,022
více aktivní	более активный	1	aktivnější	46	0,021
víc šťastný	более счастливый	1	šťastnější	52	0,019
více jasný	более ясный	1	jasnější	57	0,017
více vzdálený	более далекий	1	vzdálenější	63	0,016
více bohatý	более богатый	1	bohatější	101	0,010
více složitý	более сложный	1	složitější	255	0,004
víc silný	более сильный	1	silnější	343	0,003

Таблица 1. Конкуренция аналитических и синтетических компаративов чешских прилагательных

Анализируя данные построенной таблицы, можно заметить, что из 72 задокументированных в корпусе случаев сосуществования простых и сложных форм сравнительной степени у 15 прилагательных вероятность конкуренции составляет 50%, у 11 – от 20% до 33%, у 16 – от 10% до 20%; также в двух случаях аналитические компаративы оказались более частотными, чем синтетические. Безусловно, чешский язык остается преимущественно синтетическим, так как на 10 самых частотных качественных прилагательных (*velký, nový, dobrý, malý, vysoký, starý, mladý, dlouhý, důležitý, špatný*) корпус не дает примеров аналитических компаративных конструкций. Однако справедливо утверждать, что для ряда прилагательных (а возможно, и грамматических моделей) вероятность употребления аналитического компаратива соразмерна вероятности использования синтетической формы, что ранее считалось в принципе невозможным.

Заключение

Анализ данных электронного корпуса современных чешских письменных текстов *Synek*, представляющего собой пропорционально сокращенную версию SYN2000 в рамках Чешского национального корпуса, дает возможность изучить частотность синтетических и аналитических форм компаративов прилагательных в современном чешском письменном дискурсе, а также определить их функциональные особенности.

В отличие от традиционного взгляда чешской лингвистической школы, согласно которому аналитические формы степеней сравнения прилагательных в чешском языке возможны лишь при невозможности образования синтетических форм, анализ корпуса *Synek* выявил случаи параллельного существования аналитических и синтетических форм компаративов у 72 чешских прилагательных. Вкупе с данными об аналогичной конкуренции суперлативных форм можно выдвинуть предположения об абсолютной и относительной частотности их употребления в современном чешском языке.

Дальнейшие перспективы исследования предполагают обращение к другим корпусным ресурсам Чешского национального корпуса, включая синхронные и диахронные, письменные и устные, региональные, специализированные, двуязычные и многоязычные

корпусы. Это позволит углубить изучение различных аспектов рассматриваемой проблематики, а полученные результаты могут найти применение как в теоретических исследованиях способов выражения градации признака, так и в практической области преподавания чешского языка и совершенствовании систем машинного перевода.

Библиография

1. Изотов А.И. Корпусная революция: от искусства к науке // Филологические науки. Вопросы теории и практики. Тамбов: Грамота, 2013. № 4 (22): в 2-х ч. Ч. I. С. 68-71. EDN: PVXXUR.
2. Изотов А.И., Морозов Д.А. Суперлатив в современном чешском письменном дискурсе: опыт корпусного анализа // Славянский альманах. 2025. № 1-2. [В печати].
3. Изотов А.И. Опыт корпусного анализа степеней сравнения прилагательных (на материале современного чешского языка) // Филологические науки. Вопросы теории и практики. Тамбов: Грамота, 2025. № 7.
4. Широкова А.Г. Методы, принципы и условия сопоставительного изучения грамматического строя генетически родственных славянских языков // Широкова А.Г., Васильева В.Ф., Изотов А.И., Ананьева Н.Е. Сопоставительные исследования грамматики и лексики русского и западнославянских языков. М.: Изд-во Моск. ун-та, 1998. С. 10-99.
5. Velká akademická gramatika spisovné češtiny. I., Morfologie. Druhy slov, tvoření slov / Fr. Štícha a kol. Praha: Academia, 2018. [dva svazky] 1148 s.
6. Velká akademická gramatika spisovné češtiny. II., Morfologické kategorie, flexe / Fr. Štícha a kol. Praha: Academia, 2021. [dva svazky] 977 s.
7. Naše řeč // <https://asjournals.lib.cas.cz/naserec/home> (Последнее обращение 29.06.2025).
8. Korpus-Gramatika-Axiologie // <https://asjournals.lib.cas.cz/korpus-gramatika-axiologie/home> (Последнее обращение 29.06.2025).
9. Slovo a slovesnost // <https://asjournals.lib.cas.cz/slovoaslovesnost/home> (Последнее обращение 29.06.2025).
10. Mluvnice češtiny / M. Komárek, J. Kořenský, J. Petr, J. Veselková et al. Díl 2. Tvarosloví. Praha: Academia, 1986. 536 s.
11. Mluvnice současná čeština 1. Jak se píše a jak se mluví / V. Cvrček et al. Praha: Nakladatelství Karolinum, 2010. 353 s.
12. Рылов, С. А. Проблема типологии современных славянских языков: социально-функциональный аспект // Лингвистические традиции и современность: сборник статей, посвящённый 90-летию проф. В. Н. Немченко / Отв. ред. Л. В. Рацбурская. Нижний Новгород: ННГУ им. Н. И. Лобачевского, 2018. С. 113-120. EDN: XZXCDZ.
13. Морозов, Д. А., Рылов, С. А. Функционально-грамматическая градационность качественного признака в современных славянских языках: универсальное и локальное // Актуальные проблемы славянской филологии, культуры и журналистики: материалы I Международной научной конференции. Нижний Новгород: ННГУ им. Н. И. Лобачевского, 2021. С. 57-63. EDN: QXMDHF.
14. Русская грамматика / V. Barnetová, H. Běličová-Křížková, O. Leška, Z. Zkoumalová, V. Straková. Díl 1. Praha: Academia, 1979. 664 s.
15. Чешско-русский словарь / Под ред. Л.В. Копецкого и Й. Филипца. В 2-х томах. Прага: Государственное педагогическое издательство, 1976. Т. 1 580 с.; Т. 2 864 с.
16. Изотов А.И. Новый чешско-русский словарь: около 100 000 слов и выражений. М.: Издательство "Просвещение", 2021. 1023 с.

Результаты процедуры рецензирования статьи

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

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

Статья "Опыт корпусного анализа компаративов в современном чешском письменном дискурсе" представляет собой исследование в области богемистики и грамматики чешского языка.

Актуальность исследования определена глобальным интересом современных лингвистов к корпусным исследованиям и использованию компьютерных технологий для анализа больших данных.

В исследовании реализовались следующие задачи:

- поиск в корпусе Synek контекстов, в которых употреблены аналитические формы сравнительной степени чешских прилагательных;
- исключение в процессе визуального анализа нерепрезентативных контекстов и лемматизация обнаруженных словоформ;
- построение таблицы конкуренции аналитических и синтетических компаративов прилагательных в современном чешском письменном дискурсе.

Теоретической базой исследования стали работы лингвистов Московской школы богемистики, а также труды сотрудников отдела грамматики Института чешского языка Академии наук Чешской республики, в частности, «Большая академическая грамматика литературного чешского языка» и выпускаемые ими журналы «Наша речь», «Корпус-грамматика-аксиология», «Слово и словесность».

В работе использованы методы корпусного исследования, а также статистические методы.

Практическая значимость исследования заключается в применении его результатов в прикладных лингвистических дисциплинах, таких как автоматическая обработка естественного языка (NLP) и машинное обучение (ML).

В основной части статьи автор описывает историю вопроса и особенности поиска по корпусу.

В результате автором была составлена таблица, свидетельствующая о существовании и конкуренции соотносимых аналитических и синтетических форм сравнительной степени в современном чешском языке по данным корпуса Synek.

Библиография содержит необходимое количество актуальных отечественных и зарубежных источников по теме исследования.

Основываясь на представленных проанализированных данных, в заключении автор статьи автор приходит к следующему выводу: "анализ данных электронного корпуса современных чешских письменных текстов Synek, представляющего собой пропорционально сокращенную версию SYN2000 в рамках Чешского национального корпуса, дает возможность изучить частотность синтетических и аналитических форм компаративов прилагательных в современном чешском письменном дискурсе, а также определить их функциональные особенности.

В отличие от традиционного взгляда чешской лингвистической школы, согласно которому аналитические формы степеней сравнения прилагательных в чешском языке возможны лишь при невозможности образования синтетических форм, анализ корпуса Synek выявил случаи параллельного существования аналитических и синтетических форм компаративов у 72 чешских прилагательных. Вкупе с данными об аналогичной конкуренции суперлативных форм можно выдвинуть предположения об абсолютной и относительной частотности их употребления в современном чешском языке".

На основании приведённого объёма и качества материала представленный вывод, опровергающий традиционный взгляд на современное состояние вопроса

использования синтетических и аналитических форм компаративов прилагательных можно считать достоверным.

Таким образом, исследование вносит вклад в современную богоемистику и может быть рекомендовано к публикации в журнале "Litera".