

**МИНИСТЕРСТВО ОБРАЗОВАНИЯ
РЕСПУБЛИКИ БЕЛАРУСЬ
БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ФАКУЛЬТЕТ СОЦИОКУЛЬТУРНЫХ КОММУНИКАЦИЙ**

ЛИНГВИСТИКА, ЛИНГВОДИДАКТИКА, ЛИНГВОКУЛЬТУРОЛОГИЯ

АКТУАЛЬНЫЕ ВОПРОСЫ И ПЕРСПЕКТИВЫ РАЗВИТИЯ

**Материалы II Международной
научно-практической конференции**

Минск, 1–2 марта 2018 г.

Минск
Издательский центр БГУ
2018

УДК 81'1(06)

ББК 81я431

Л59

Редакционная коллегия:

О. Г. Прохоренко (ответственный редактор), *О. В. Дубровина*,
Т. В. Володина, *А. О. Долгова*, *О. И. Копач*, *З. И. Канапацкая*,
Л. М. Блинкова, *В. В. Воронович*

Рецензенты:

доцент кафедры делового английского языка
факультета международных бизнес-коммуникаций
УО «Белорусский государственный экономический университет»
кандидат филологических наук *Н. А. Куркович*;
доцент кафедры теории и практики перевода № 1
переводческого факультета УО «Минский государственный
лингвистический университет» *А. В. Вдовичев*

Лингвистика, лингводидактика, лингвокультурология: актуальные вопросы и перспективы развития : материалы II Международ. науч.-практ. конф., Минск, 1–2 марта 2018 г. / редкол. : О. Г. Прохоренко (отв. ред.) [и др.]. – Минск : Изд. центр БГУ, 2018. – 327 с.

ISBN 978-985-553-509-7.

Рассматриваются современные направления лингвистических исследований, традиции и инновации в обучении иностранным языкам, проблемы языковой картины мира и взаимодействия культур, освещаются вопросы литературоведения, теории перевода и интерпретации текста.

Адресуется преподавателям вузов, научным сотрудникам, аспирантам, магистрантам, студентам.

Авторы несут ответственность за достоверность и качество представленных материалов.

При полном или частичном использовании материалов ссылка на сайт Электронной библиотеки БГУ обязательна (www.elib.bsu.by).

УДК 81'1(06)

ББК 81я431

ISBN 978-985-553-509-7

© БГУ, 2018

© Оформление. РУП «Издательский центр БГУ», 2018

СОДЕРЖАНИЕ

РАЗДЕЛ 1. СОВРЕМЕННЫЕ НАПРАВЛЕНИЯ ЛИНГВИСТИЧЕСКИХ ИССЛЕДОВАНИЙ

Yang Yang. Классификация современных китайских междометий	3
Власюк О.И. Кластеры служебных слов во фразе: типичные морфолого-семантические модели соединений и их просодическая структура	8
Варановіч В.В. Трансфармацыйныя здольнасці сказаў з апісальнымі прэдыкатамі	12
Вострикова И.Ю. О специфике структурной и зонной организации лексико-семантического поля «трудовая деятельность» в русском языке	16
Еремеева М., Мацук А. Трудности разграничения паремиологии и фразеологии	20
Жевнерович Е.Э. Корпус текстов в научном исследовании	25
Зубовская Н.К., Смирнова Н.М., Ковган И.И. Изучение контекстной семантики слова как способ и средство познания его значения	32
Зяноўка Я.С. Сучасныя сродкі апрацоўкі інтанацыйных канструкцый	36
Канторович Т.М. Языковые предпочтения гендеров в ситуации конфликта (на уровне эмоционально-экспрессивных средств)	43
Климченко С.Г. Контрастивные исследования в контексте сопоставительного языкознания	49
Копач О.И. Вопрос об ономаσιологической структуре топонима: симметрия образования собственных и нарицательных имён	55
Марчык М.У., Лысы С.І., Гецэвіч Ю.С. Вычытка тэксту вялікага памеру на беларускай мове	58
Митюкова Е.А. Метафорические интернет-термины в английском языке	64
Прыгодзіч А.А. Намінацыйныя працэсы ў тапаніміі	68
Русак В.П. Перспектыўнасць міждысцыплінарных даследаванняў	71
Рыжкович А.Ч. Фазовые субстантивы как базовый компонент темпоральных предложных единиц	75

ВЫЧЫТКА ТЭКСТУ ВЯЛІКАГА ПАМЕРУ НА БЕЛАРУСКАЙ МОВЕ

PROOFREADING OF THE LARGE TEXT IN BELARUSIAN

М.У. Марчык, С.І. Лысы, Ю.С. Гецэвіч

M. Marchyk, S. Lysy, Y. Hetsevich

Аб'яднаны інстытут праблем інфарматыкі НАН Беларусі

Мінск, Беларусь

United Institute of Informatics Problems of National Academy of Sciences of Belarus

Minsk, Belarus

E-mail: marynamarczyk@gmail.com

Сёння правільная моўная практыка важная для тых, хто ў прафесійнай дзейнасці штодзённа карыстаецца вуснай і пісьмовай формамі літаратурнай мовы – настаўнікаў, выкладчыкаў, рэдакцыйна-выдавецкіх работнікаў, прадстаўнікоў інтэлігенцыі. Арфаграфічная правільнасць тэксту і маўлення – гэта таксама ўмова якаснай апрацоўкі інфармацыі і правільнага функцыянавання камп'ютарных сістэм пры чалавека-машынных камунікацыях.

Перакрыжаванне інтарэсаў вучоных філалагічнага і тэхнічнага профілю ў сферы камп'ютарнай лінгвістыкі выглядае ўзаемадапаўняльным і перспектыўным. У гэтым рэчышчы супрацоўнікамі АПП НАН Беларусі распрацавана метадыка вычиткі тэксту вялікага памеру на беларускай мове.

Сваім з'яўленнем метадыка абавязаная распрацоўцы сінтэзатара маўлення па тэксце. Адным з этапаў сінтэзу маўлення з'яўляецца праца тэкставага працэсара, напрацоўкі для якога магчыма было прымяніць і для вырашэння іншых задач сімвальнай і паслоўнай апрацоўкі. Такім чынам яны сталі самастойнымі сэрвісамі на платформе для апрацоўкі тэксту і маўлення corpus.by.

Метадыка працуе на сімвальным і паслоўным узроўні. Для атрымання арфаграфічна правільнага тэксту неабходна апрацаваць тэкст праз 5 анлайн-сэрвісаў, размешчаных на інтэрнэт-платформе corpus.by.

Парадак працы з кожным сэрвісам наступны: неабходна ўставіць тэкст у поле ўводу, абраць тыя ці іншыя наладкі, націснуць кнопку атрымання вынікаў. Па выніках апрацоўкі тэксту кожным сэрвісам карыстальнік павінен праглядаць атрыманыя вынікі і ўнесці ў тэкст адпаведныя праўкі праз функцыю «знайсці-замяніць» тэкставага рэдактара ці акна браўзера.

Першим етапом з'являється вычитка праз сэрвіс «Падлік частотнасці сімвалаў» [2]. Гэты сэрвіс прадэманструе спіс усіх сімвалаў, выкарыстаных у тэксце, і дазволіць выявіць і выправіць іх памылковае выкарыстанне (малюнак 1).

Напрыклад, сэрвіс паказвае колькасць выкарыстаных дужак, і калі колькасць левых і правых дужак не супадае, то, хутчэй за ўсё, у тэксце ёсць пунктуацыйныя памылкі. Таксама праз сэрвіс можна правярыць правільнасць выкарыстання дэфісаў, слэшаў, розных відаў двукоссяў і працяжнікаў.

Падлік частотнасці сімвалаў
BE EN RU ?

Калі ласка, увядзіце тэкст

Сірыус (лац. Sirius), таксама Альфа Вялікага Собаі (лац. Alpha Canis Majoris) – найярчэйшая зорка начнога неба. Бачная зорная велічыня -1,46. Сірыус аддалены ад Сонца на 8,60±0,04 св. гадоў і з'яўляецца адной з бліжэйшых да нас зорак. Нібесныя каардынаты: прамое ўзыходжанне 6 гадз. 45 хв. 9 сек., схіленне -16°42'58". Назва паходзіць з грэцкай мовы і азначае "бліскучы" альбо "пякучы".

Скарына (3283 Skorina, 1979 QA10) – астероід № 3283. Належыць да Галоўнага пояса астероідаў Сонечнай сістэмы. Дыяметр – 12,65 км. Сярэдняя аддаленасць астероіда ад Сонца складае 2,396 астранамічнай адзінкі (~ 360 млн. км.). Перыяд поўнага абарачэння вакол Сонца – 3,71 зямнога года. Адкрыты астраномам Мікалаем Іванавічам Чарных у Крымскай астрафізічнай абсерваторыі 27 жніўня 1979

Атрымаць інфармацыю пра частотнасць сімвалаў!

ІНФАРМАЦЫЯ ПРА ўСЕ ЗНОЙДЗЕНЫЯ СІМВАЛЫ:

Агульная колькасць сімвалаў у тэксце: **836**
Колькасць унікальных сімвалаў у тэксце: **86**

№ С.	Код	Назва	Частата		Кантэкст
	U+000D	ВЯРТАННЕ КАРЭТКІ	2	0.24%	ліскучы" альбо "пякучы".
	U+000A	ПЕРАВОД РАДКА	2	0.24%	ліскучы" альбо "пякучы".
	U+0020	ПРАБЕЛ	118	14.11%	Сірыус (лац. Sirius), таксам
"	U+0022	ЗНАК ДВУКОСНЯ	4	0.48%	з грэцкай мовы і азначае "б
*	U+0027	АПСТРАФ	4	0.48%	на 8,60±0,04 св. гадоў і з'я
(	U+0028	ЛЕВАЯ КРУГЛАЯ ДУЖКА	4	0.48%	Сірыус (лац. Sirius), таксам
)	U+0029	ПРАВAYA КРУГЛАЯ ДУЖКА	4	0.48%	Сірыус (лац. Sirius), таксам
,	U+002C	КОСКА	9	1.08%	Сірыус (лац. Sirius), таксам
-	U+002D	ЗЛУЧОК-МІНУС	2	0.24%	Бачная зорная велічыня -1,4
.	U+002E	КРОПКА	20	2.39%	Сірыус (лац. Sirius), таксам
0	U+0030	ЛІЧБА НУЛЬ	5	0.6%	на 8,60±0,04 св. гадоў і з'я
1	U+0031	ЛІЧБА АДЗІН	7	0.84%	Бачная зорная велічыня -1,4

Малюнак 1 – Знешні інтэрфейс і вынікі працы сэрвіса «Падлік частотнасці сімвалаў»

Другім этапом з'яўляецца вычитка праз сэрвіс «Падлік частотнасці слоў» [3]. Сэрвіс дазваляе выявіць словы з канкрэтным наборам сімвалаў і ўбачыць кантэкст іх ужывання. Падчас вычиткі неабходна правярыць тэкст на прадмет ужывання лацінскіх сімвалаў, а таксама лічбаў.

Калі ў словах кірылічнага напісання выкарыстоўваюцца візуальна аднолькавыя лацінскія сімвалы (/a/, /i/, /c/ і г.д.), то тэкст будзе ў далейшым паўсюль апрацоўвацца недакладна (напрыклад, пры працы са шрыфтамі падчас вёрсткі, а таксама на наступных этапах вычыткі). Такія сімвалы неабходна выявіць з дапамогай дадзенага сэрвіса і замяніць кірылічнымі.

Напрыклад, у тэксце *Груша цвіла апошні год* у некаторых словах выкарыстаныя лацінскія сімвалы /i/, /o/, /a/. Каб іх выявіць, трэба ўстаіць дадзены тэкст у поле ўводу, а ў полі «Сімвалы, з якіх можа складацца слова» пакінуць толькі сімвалы лацінскага алфавіту. Адноструюцца вынікі, прадстаўленыя на малюнку 2.

☒ Адчувальны да рэгістру
 ☐ Выводзіць асобна левыя і правыя кантэксты

Усяго знойдзена слоў у тэксце: **6**
 Па Вашым запыце знойдзена ўнікальных слоў: **3**

↑ Частата	Слова	Кантэксты (макс. 2)
2	о	... а цвіла апошні год. ... а апошні год.
2	і	Груша цвіла апошні г іла апошні год.
2	а	Груша цвіла апошн ... Груша цвіла апошні год.

Малюнак 2 – Вынікі працы сэрвісу «Падлік частотнасці слоў»

Такім самым чынам неабходна здзейсніць пошук лічбаў. Так, у адным з мастацкіх тэкстаў аўтарам артыкула было знойдзена 14 ужыванняў лічбавага запісу лічэбнікаў замест належнага запісу пропісам. Напрыклад, *30 срэбнікаў* замест правільнага варыянту *трыццаць срэбнікаў*, з *запарожцамі ў 1651-м* замест правільнага з *запарожцамі ў 1651 годзе*, *каля 20-ці тысячаў жаўнераў* замест *каля дваццаці тысячаў жаўнераў*.

Па знойдзеных памылковых ужываннях лацінскіх сімвалаў і лічбаў неабходна ўнесці праўкі ў асноўны тэкст і перайсці да наступнага этапу вычыткі.

Трэцім этапам з’яўляецца вычытка праз сэрвіс «Праверка правапісу» [4]. Сэрвіс правярае тэкст паводле двух аспектаў:

1. Правільнасць напісання слоў паводле слоўнікаў.
2. Ужыванне лацінскіх і кірылічных сімвалаў у адным слове.

Слова, напісанае з арфаграфічнай памылкай ці памылкай друку, адсутнічае ў падключаных слоўніках і будзе пазначана сэрвісам як невядомае. Невядомыя словы выводзяцца ў выглядзе спісу з магчымасцю адзначыць правільнасць напісання таго ці іншага слова і пабачыць кантэкст яго выкарыстання (малюнак 3).

Калі ласка, увядзіце тэкст для праверкі

↶

✕

Ігнараваць словы

↶

✕

Груша цвіла апошні год. Усе галіны яе, усе валікія расо́кі, да апошняга пру́ціка, былі ўсыпаны буйным бела-ружовым цвeтам. Яна кіпeла, млeла і раскiшавалася ў пчaліным звoне, цягну́ла да сонца сталья лaпы і распу́скала ў яго зня́ні маленькія, кво́лыя пaльцы нoвых пaрaсткaў. І бы́ла яна тaкaя мaгўтнaя і свeжaя, тaк утpaпéна спpaчaлі́ся ў яe ружoвым paі пчoлы, штo, здавaлaсja, нe будзе ёй звoду і нe будзе ёй кaнцa.

сонца
пчолы

Максимум кантэкстаў:

10

Абраць слоўнікі

☒ SBM1987 (паводле публікацыі "Слоўнік беларускай мовы. Арфаграфік, Афэпты, Акцэнтуацыя, Словазмяненне / пад рэд. М.Б. Сірылы. - Мінск, 1987")

☒ SBM2008 (слоўнік беларускай мовы паводле Belarussian Grammar Database (<http://belarusian-grammar.org/>). Authors: Symon Kakeza, Alai Bukejka, Uladz Kozłanka. На ўмовах ліцэнзіі CC BY-SA 4.0)

☒ SBM2012initial (пачатковыя формы паводле публікацыі "Слоўнік беларускай мовы. / навуц. рэд. А.А. Лукашэвіч, В.П. Русак. - Мінск : Беларус. навука, 2012")

☒ NOUN2013 (паводле публікацыі "Граматычны слоўнік назоўніка / навуц. рэд. В.П. Русак. - Мінск : Беларус. навука, 2013")

☒ ZALIZNIAK (паводле публікацыі "Граматычны слоўнік рускаго языка: Словазмяненне / А.А. Зализняк. — Москва / Русский язык, 1980. — 880 с.")

☒ CMU (паводле "Cambridge English University Pronunciation Dictionary")

☐ NEW (новыя словы для CUD)

☐ S2016_01 (беларуская мова, слоўнік даступны па [ссылка](#))

☐ S2016_02 (руская мова, слоўнік даступны па [ссылка](#))

☒ S2016_03 (беларуская мова, слоўнік даступны па [ссылка](#))

☐ S2017_04 (слоўнік даступны па [ссылка](#))

☐ S2017_05 (слоўнік даступны па [ссылка](#))

☐ UWP_BE (беларускія словы са слоўніка, які даступны па [ссылка](#))

☐ UWP_RU (рускія словы са слоўніка, які даступны па [ссылка](#))

Праверыць!

СЛОВЫ І З КІРЫЛІЧНЫМІ, І З ЛАЦІНСКІМІ СІМВАЛАМІ (1)

апошні	1	Груша цвіла апошні год. Усе галіны ...	апошні
--------	---	--	--------

ПАЗНАЧЭ СЛОВЫ БЕЗ ПАМЫЛАК І КЛІКНІЦЕ "ПЕРАПРАВЕРЫЦЫ" (2)

☐ буйным	1	... пру́ціка, былі ўсыпаны буйным бела-ружовым цвeтам. Яна ...	магчыма, буйным
☐ парасткаў	1	... кво́лыя пaльцы нoвых пaрaсткaў. І бы́ла яна ...	магчыма, парасткаў

Малюнак 3 – Знешні інтэрфейс і вынікі працы сэрвіса «Праверка правапісу»

Сэрвіс дубліруе працу на папярэднім этапе, робячы пошук слоў з кірылічнымі і лацінскімі сімваламі. На цяперашні час гэта магчымасць пераправерыць тэкст, а ў будучым функцыянал «Праверкі правапісу» можа быць сумешчаны з «Падлікам частотнасці слоў».

Таксама гэты сэрвіс дае магчымасць ігнараваць пэўныя словы. Гэтая магчымасць спатрэбіцца пры вычытцы вузкасפעцыяльнага тэксту, каб выключыць пападанне загадзя невядомага сэрвісу слова ў спіс невядомых і паскоруць прагляд гэтага спісу.

61

Чацвёртым этапам з’яўляецца вычытка праз сэрвіс «Праверка правапісу “Ў”» [5]. Алгарытм праверкі правільнасці ўжывання ў тэксце літар «у» і «ў» дзейнічае наступным чынам:

1. Пошук сімвалаў /y/, /ÿ/, /У/, /Ў/.
2. Счытванне папярэдняга сімвала (літары ці знака прыпынку).
3. Праверка правільнасці ўжывання /y/, /ÿ/, зважаючы на папярэдні сімвал.

Па выніках апрацоўкі неабходна прагледзець спіс месцаў, дзе сэрвіс прапанаваў карыстальніку спраўдзіць ужыванне «у» і «ў» (малюнак 4).

Сустрэлася	Каментар
«А у»: ...А у іх ёсць пчылы...»	(«у» пасля галоснай «А» без знакаў прыпынку)
«а ў»: ...«Рама» у краме...»	(«у» пасля галоснай «а» без знакаў прыпынку)
«а-у»: ...На Украіне паўднёва-усходні вечер...»	(«у» пасля галоснай «а» і злучка)
«аў»: ...Сястра ёсць аўсянку...»	(«у» пасля галоснай «а»)

Малюнак 4 – Знешні інтэрфейс і вынікі працы сэрвіса «Праверка правапісу “Ў”»

Сэрвіс таксама мае поле «Выключэнні», дзе змяшчаюцца актуальныя выключэнні з правілаў правапісу літар «у» і «ў». Поле можна рэдагаваць: выдаляць уведзеныя па змаўчанні і дадаваць неабходныя карыстальніку выключэнні.

Пятым этапам з’яўляецца вычытка праз сэрвіс «Ідэнтыфікатар амографаў» [1]. Сэрвіс паказвае карыстальніку амографы – словы, якія маюць аднолькавае напісанне і рознае вымаўленне, напрыклад, *музыка* і *музыка* (малюнак 5).

Неабходна ўважліва прагледзець спіс знойдзеных амографаў і прыняць рашэнне, у якіх словах-амографах у тэксце трэба пазначыць націск.

Калі ласка, увядзіце тэкст (максімум 5000 сімвалаў)

Маленькаму прынцу, які ўжо сабраўся ў дарогу, зусім не хацелася засмучаць старога манарха.
 — Калі ваша виліасць хоча, каб яас безагаворачна слухалі, вы маглі б аддаць мне разумны загад. Вы, да прыкладу, маглі б загадаць мне адправіцца не пазней як праз хаіліну. Мне здаецца, умовы спрыяльныя.
 Кароль маўчаў. Маленькі прынц трошкі павягаўся, уздыжнуў і пайшоў.
 — Прызначаю цябе паслом! — паспешліва крыкнуў яму ўслед кароль.
 І выглад у яго быў пры гэтым такі, быццам ён не далучаў ніякіх парочанняў.
 «Дзіўныя людзі гэтыя дарослыя», — падумаў дарогаю Маленькі прынц.

Шукаць па наступных слоўніках:

- SBM1987 (спасылка на публікацыю "Слоўнік беларускай мовы. Аграфічны, алфавітны, лічэбны і гл. інш. / Мінск, 1997")
- SBM2012initial (аналітычныя формы паслужыць публікацыя "Слоўнік беларускай мовы. Г. навука, рэд. А.А. Лукашэвіч, 8.П. Рукі. - Мінск: Беларуская навука, 2012")
- ZALIZNIAK (спасылка на публікацыю "Граматычны слоўнік рускай мовы: слоўнік / А.А. Залізняк. - Мінск: Беларуская навука, 1988. - 880 с.")
- CMU (спасылка на "Corpus Mellon University WordFrequency Dictionary")
- NEW (новыя словы для СМД)

Шукаць амографію

Знойдзена амографіяў: 3 (паглядзець спіс амографіяў можна тут)

Амограф	Варыянты націску	Тып амографы	Кольк.	Кантэксты	Слоўнік
прыкладу	прыкладу прыкладу прыкладу	адна частына мовы	1	... загад. Вы, да прыкладу, маглі б загадаць ...	SBM1987
яму	яму яму	адна частына мовы	1	... паслом! — паспешліва крыкнуў яму ўслед кароль. І ...	SBM1987

Малюнак 5 – Знешні інтэрфейс сэрвіса «Ідэнтыфікатар амографіяў»

Выкананне ўсіх этапаў гэтай метадыкі дазваляе атрымаць арфаграфічна правільны тэкст на беларускай мове (методыка не ахоплівае пунктуацыйны і сінтаксічны раздзелы граматыкі). Таксама дадзены парадак вычыткі апрабаваны ў шматлікіх праектах лабараторыі распазнавання і сінтэзу маўлення: метадыкай вычытаны «Новы Запавет і Прыповесці» (2016), «Універсальная дзесятковая класіфікацыя» (2016).

ЛІТАРАТУРА

1. Ідэнтыфікатар амографіяў [Электронны рэсурс]. – 2018. Рэжым доступу: <http://corpus.by/HomographIdentifier/?lang=be>. – Дата доступу: 10.01.2018.
2. Падлік частотнасці сімвалаў [Электронны рэсурс]. – 2018. Рэжым доступу: <http://corpus.by/CharacterFrequencyCounter/?lang=be>. – Дата доступу: 10.01.2018.
3. Падлік частотнасці слоў [Электронны рэсурс]. – 2018. Рэжым доступу: <http://corpus.by/WordFrequencyCounter/?lang=be>. – Дата доступу: 10.01.2018.
4. Праверка правапісу [Электронны рэсурс]. – 2018. Рэжым доступу: <http://corpus.by/SpellChecker/?lang=be>. – Дата доступу: 10.01.2018.
5. Праверка правапісу «Ў» [Электронны рэсурс]. – 2018. Рэжым доступу: <http://corpus.by/ShortUSpellChecker/?lang=be>. – Дата доступу: 10.01.2018.