

## БИБЛИОГРАФИЧЕСКИЕ ССЫЛКИ

1. Мечковская, Н.Б. История языка и история коммуникации: от клинописи до Интернета / Н.Б. Мечковская. – М.: Флинта: Наука, 2009. – 584 с.
2. Потапова, Н.В. Характеристики гипотекста электронного новостного дискурса, обусловленные возрастом адресата и не зависящие от него. Автореф. дис. ... канд. филол. наук: 10.02.19 / Н.В. Потапова; ФГБОУВО «Кемеровский государственный университет». – Кемерово, 2019. – 24 с.

### ВЫКАРЫСТАННЕ КАМП'ЮТАРНА-ЛІНГВІСТЫЧНЫХ СЭРВІСАЎ ДЛІЯ ЗАБЕСПЯЧЭННЯ ДАСТУПНАСЦІ ЗАКАНАДАЎЧЫХ ТЭКСТАЎ НА РОЗНЫХ МОВАХ У МАЎЛЕНЧАЙ ФОРМЕ

### USING COMPUTER-LINGUISTIC SERVICES TO ENSURE THE AVAILABILITY OF LEGISLATIVE TEXTS IN DIFFERENT LANGUAGES IN ORAL FORM

*Ю. С. Геўэвіч<sup>1)</sup>, В. В. Варановіч<sup>2)</sup>, А. У. Бабкоў<sup>3)</sup>, С. С. Маеўскі<sup>4)</sup>,  
Дз. А. Дзенісюк<sup>5)</sup>*

*Y. S. Hetsevich<sup>1)</sup>, V. V. Varanovich<sup>2)</sup>, A. U. Babkou<sup>3)</sup>, S. S. Mayeuskii<sup>4)</sup>,  
D. A. Dzenisiuk<sup>5)</sup>*

<sup>1), 3), 4), 5)</sup> Аб'яднаны інстытут праблем інфарматыкі НАН Беларусі  
Мінск, Беларусь

<sup>2)</sup> Беларускі дзяржаўны ўніверсітэт  
Мінск, Беларусь

<sup>1), 3), 4), 5)</sup> United Institute of Informatics Problems of the National Academy of Sciences  
of Belarus  
Minsk, Belarus

<sup>2)</sup> Belarusian State University  
Minsk, Belarus

*e-mail: <sup>1)</sup>yuras.hetsevich@gmail.com, <sup>2)</sup>gamrat@tut.by, <sup>3)</sup>andr1877@gmail.com,  
<sup>4)</sup>maevskiiss@gmail.com, <sup>5)</sup>d.denisjuk@gmail.com*

Падымаецца пытанне даступнасці юрыдычных тэкстаў на дзвюх дзяржаўных мовах на прыкладзе кодэксаў Рэспублікі Беларусь. Паказваюцца звязаныя з распаўсюджваннем інтэрнэту змяненні ў адносінах грамадства да людзей, якія слаба бачаць. Звяртаецца ўвага на комплекснасць паняцця «даступнасць заканадаўчых тэкстаў» у білінгвістычным грамадстве; разглядаюцца магчымасці машыннага перакладу на сайце etalonline.by і магчымасці выкарыстання партала cogrus.by у мэтах лінгвістычнай апрацоўкі тэкстаў.

*Ключавыя словы:* камп'ютарная апрацоўка тэкстаў; сінтэз маўлення; білінгвізм; даступнасць; заканадаўчы тэкст; інваліднасць па зроку.

The article is dedicated to accessibility of legal texts in two official languages through the example of Law Codes of the Republic of Belarus. Changes in the society's attitude towards visually impaired people associated with spreading of the internet are indicated. The article draws attention to the complexity of the concept of availability of legislative texts in bilingual society; the possibilities of machine translation on the etalonline.by site and the possibility of using the portal corpus.by for purposes of linguistic processing of texts are considered.

*Keywords:* computer text processing; speech synthesis; bilingualism; accessibility; legislative text; visual impairment.

Па інфармацыі Сусветнай арганізацыі аховы здароўя ў свеце налічваецца больш за 40 млн людзей, якія не бачаць, і больш за 250 млн людзей з вельмі дрэнным зрокам. Ёсць яшчэ людзі з парушэннямі каляровых адчуванняў, колькасць якіх вар'іруецца ад 4 да 8% насельніцтва нашай планеты. У Беларусі больш за 30 000 слабабачачых, калі дадаць да гэтага 1,9 млн беларусаў, старэйшых за 60 гадоў, большасць з якіх слаба бачыць, то стане відавочна, што дадзеная праблема датычыцца практычна чвэрці жыхароў Беларусі.

Раней людзі, якія слаба бачаць, былі амаль цалкам ізаляваныя ад грамадства. Цяпер грамадства пачынае разумець: людзі з інваліднасцю па зроку існуюць не асобна, а з'яўляюцца яго часткай. Камп'ютар дазваляе ім не замыкацца ў сваім вузкім асяроддзі.

А з 1 студзеня 2019 г. уступіла ў сілу пастанова Савета міністраў Рэспублікі Беларусь № 797 ад 23 кастрычніка 2017 г., паводле якой унесены змяненне і дапаўненні ў Палажэнне аб парадку функцыянавання інтэрнэт-сайтаў дзяржаўных органаў і арганізацый [7]. Удакладнена, што інфармацыя на інтэрнэт-сайтах размяшчаецца на рускай і (або) беларускай мовах, а пры неабходнасці таксама на адной або некалькіх замежных мовах. Прымяненне і беларускай, і рускай моў неабходна пры размяшчэнні ўсёй інфармацыі, якая з'яўляецца абавязковай для галоўнай старонкі інтэрнэт-сайта. Устаноўлена, што на інтэрнэт-сайтах неабходна прадугледзець даступнасць інфармацыі ў адпаведнасці з патрабаваннямі, вызначанымі ў тэхнічных нарматыўна-прававых актах. Пры гэтым інтэрнэт-сайт павінен падтрымліваць версію для людзей з інваліднасцю па зроку.

Так, з мэтай забеспячэння агульнадаступнасці інфармацыі камітэтам па ахове здароўя Мінгарвыканкама была распрацавана версія інтэрнэт-сайта для людзей з інваліднасцю па зроку, якая дазваляе: уключаць ці выключаць галасавыя падказкі, якія прагаворваюць назвы раздзелаў інтэрнэт-сайта, кнопак і пунктаў меню; агучваць навіны па-беларуску і па-руску з дапамогай усталяванай сістэмы сінтэзу маўлення; а таксама

выбіраць камфортныя для ўспрымання памер шрыфту і каляровую палітру [6].

Для грамадзян Беларусі, маючых праблемы са зрокам, акрамя агульных, існуе дадатковая праблема, звязаная з тым, што ў нашай краіне дзве дзяржаўныя мовы: беларуская і руская. Але ж доўгі час аб даступнасці нарматыўных актаў толькі ўзгадвалася ў кантэксце патрабаванняў да якасці іх моўнага выкладу. Напрыклад, у Законе Рэспублікі Беларусь ад 10 студзеня 2000 г. «Аб нарматыўных прававых актах Рэспублікі Беларусь» у ліку патрабаванняў да мовы дакументаў вызначаны і такія, як яе яснасць, прастата і даступнасць [5]. Забеспячэнне даступнасці і зразумеласці для кожнага грамадзяніна – гэта канчатковая мэта працэсу ўдасканалення заканадаўства ў любой прававой дзяржаве. Аднак дадзеная мэта заўсёды будзе заставацца недасягальнай без уліку моўнай сітуацыі ў краіне, без арыентавання на кожнага адрасата прававой інфармацыі (у тым ліку на чалавека з інваліднасцю) як на моўную асобу. Інакш кажучы, якім бы ні быў заканадаўчы тэкст дасканалы апрацаваным з пункту гледжання дакладнасці, лагічнасці, яснасці і г. д., як бы добра не былі забяспечаны ўсе неабходныя ўмовы для азнаямлення з ім, даступнасць тэксту ўсё роўна не будзе поўнай, калі ён створаны і існуе на адной мове ў білінгвістычным грамадстве. А ў Рэспубліцы Беларусь, нягледзячы на дзяржаўнае двухмоўе, пераважная большасць заканадаўчых дакументаў рэалізавана толькі на адной мове. Так, з 26 кодэксаў Рэспублікі Беларусь, тэксты якіх прадстаўлены на Нацыянальным прававым інтэрнэт-партале [pravo.by](http://pravo.by), 25 апублікаваны толькі па-руску, 1 – толькі па-беларуску.

Адзін з асноўных фактараў, якія стрымліваюць практычнае забеспячэнне двухмоўя ў прававой сферы Рэспублікі Беларусь, – нявырашанасць праблемы якаснай і хуткай лінгвістычнай апрацоўкі тэкстаў вялікага памеру, што сведчыць аб актуальнасці якаснага машыннага перакладу.

Атрымаць неафіцыйную беларускамоўную версію рускамоўнага нарматыўнага акта або рускамоўную версію беларускамоўнага нарматыўнага акта сёння магчыма, калі звярнуцца на сайт Нацыянальнага цэнтра прававой інфармацыі Рэспублікі Беларусь [etalonline.by](http://etalonline.by) і націснуць кнопку «Машынны пераклад». Пераклад ажыццяўляецца за некалькі хвілін, а яго якасць дазваляе ў цэлым зразумець дакумент, але звяртае на сябе ўвагу сэнсавая нераўнацэннасць лексічных адзінак або сінтаксічных канструкцый зыходнай рускай мовы і беларускай мовы, напрыклад (руская лексема – прапанаваная машынным перакладчыкам лексема – правільны аналаг): *настоящий – сапраўдны – дадзены; управ-*

леннямі по труду – кіраваннямі па працы – упраўленнямі па працы; за-  
конодацельствам, дзействующим на яе тэрыторыі, – заканадаўствам,  
якое дзее на яе тэрыторыі, – заканадаўствам, якое дзейнічае на яе  
тэрыторыі.

Менавіта таму пострэдагаванне тэкстаў лінгвістамі – гэта неабход-  
ная праца, якая можа быць аптымізавана пры дапамозе праграмнага за-  
беспячэння Coprus.by [1], распрацаванага супрацоўнікамі лабараторыі  
распазнавання і сінтэзу маўлення Аб'яднанага інстытута праблем  
інфарматыкі (АІПІ) НАН Беларусі. У прыватнасці, пры перакладзе  
тэкстаў кодэксаў Рэспублікі Беларусь карыснымі з'яўляюцца алгарытмы  
сістэмы праверкі правапісу, статыстычнай і даведчнай інфармацыі аб  
сімвалах, статыстыкі ўжывання адвольных сімвальных паслядоўнасцей  
у электронным тэксце, пошуку і выпраўлення памылак у напісанні літар  
«у» і «ў», распазнавання і вылучэння ў тэксце амогафаў.

Машынны перакладчык можа быць удасканалены, калі ён будзе  
мець магчымасць пастаянна напрацоўваць свае навыкі ў адной і той жа  
прадметнай галіне, грунтуючыся на правілах, якія выбіраюцца з задад-  
зеных крыніц інфармацыі, а з удасканаленнем машыннага перакладчыка  
пачне вырашацца і праблема забеспячэння даступнасці заканадаўства  
для носьбітаў розных моў (у тым ліку сярод людзей з інваліднасцю па  
зroku) – адна з галоўных задач, што вырашаюцца ў кожнай дзяржаве ў  
мэтах інфармацыйнай бяспекі [3].

У лабараторыі распазнавання і сінтэзу маўлення АІПІ НАН Беларусі  
існуе шэраг распрацаваных сінтэзатараў беларускага і рускага  
маўлення, сярод якіх стацыянарны [2], мабільны і анлайн-сінтэзатар [4],  
што дазваляе людзям з інваліднасцю па зроку праз аўтаматычнае агуч-  
ванне тэксту сінтэзатарам праслухоўваць патрэбны ім кантэнт на роз-  
ных прыстасаваннях.

Дадзеныя распрацоўкі таксама выкарыстоўваюцца для працы  
з юрыдычным даменам беларускай лексікі пры перакладзе з рускай мо-  
вы на беларускую і далейшым агучванні кодэксаў Рэспублікі Беларусь,  
у тым ліку Жыллёвага, Банкаўскага, Ляснога, Кодэкса аб зямлі, Кодэкса  
аб шлобе і сям'і, Кодэкса аб адміністрацыйных правапарушэннях.

У прыватнасці, падчас працы над Банкаўскім кодэксам (колькасць  
слоў у рускамоўным тэксце – 43 507) найбольшую цяжкасць выклікаў  
пераклад спецыяльных тэрмінаў, прыйшоўшых з замежных моў: *інсай-  
дер – інсайдар, пруденциальные – прудэнцыяльныя, бенефициар –  
бенефіцыяр, факторинг – факторынг, отзыв (лицензии) – адкліканне  
(ліцэнзіі)*. Таксама цяжка паддаваліся алгарытмаванню ў сістэмах ма-  
шыннага перакладу наступныя асаблівасці беларускай мовы  
(у параўнанні з рускай):

– канчаткі назоўнікаў мужчынскага роду ў родным склоне: *фактару* (у значэнні ‘чыннік’) і *фактара* (у значэнні ‘фізічная ці юрыдычная асоба’);

– адрозненні ў граматычным афармленні канструкцый з ступенямі параўнання прыметнікаў і прыслоўяў: *цена не ніжже минимальной – цена, не ніжэйшая за мінімальную*;

– дзеепрыметнікі, дзеепрыметныя звароты: *нормы банковского законодательства, содержащиеся в международных договорах – нормы банкаўскага заканадаўства, якія змяшчаюцца ў міжнародных дагаворах*;

– пераважнае выкарыстанне меснага склону пасля прыназоўніка «па» (у рускай мове – давальны склон): *правовые акты по бухгалтерскому учету – прававыя акты па бухгалтарскім уліку*.

Пры перакладзе іншых кодэксаў сустрэліся выпадкі, калі ў беларускай мове адсутнічаюць пэўныя тэрміны або розныя рэаліі абазначаюцца адным словам (лексічныя лакуны), у той час як у дакументах гэтыя рэаліі маюць адрозненне і неабходна назваць іх рознымі словамі:

– *супруги* ў рускай мове ў слоўніках перакладаюцца як *муж і жонка*, у той жа час у кодэксах сустракаецца паняцце *супруг* у адзіночным ліку; у такіх выпадках магчымы пераклад як *муж або жонка* ці выкарыстанне неалагізма *сужэнец (сужэнцы)*;

– у рускай мове існуе розніца паміж паняццямі *хозяйственное общество* і *хозяйственное товарищество*, у беларускай жа выкарыстоўваецца толькі тэрмін *таварыства*; у гэтым выпадку магчыма ўжыванне тэрміна *суполка* або *супольнасць*;

– маюць адрозненні тэрміны *бюджетная ссуда* і *бюджетный заём*, якія на беларускую мову перакладаюцца адным словам *пазыка*.

Пасля перакладу ўсе тэксты кодэксаў агучваюцца мадэрнізаваным анлайн-сінтэзатарам *BarysBelHigh*, які дае магчымасць атрымліваць больш чыстае гучанне за кошт выдалення часткі басовага спектра.

Далей лабараторыяй распазнавання і сінтэзу маўлення плануецца прапаноўваць для тэсціравання і наступнай практычнай эксплуатацыі сістэмы сінтэзу і распазнавання маўлення па тэксце, сістэмы машыннага перакладу тэкстаў юрыдычнага і медыцынскага даменаў, удзельнічаць у развіцці і нападзенні дзяржаўных інфармацыйна-прававых рэсурсаў. Для гэтага яшчэ 7 жніўня 2018 г. АІПІ НАН Беларусі і Нацыянальным цэнтрам прававой інфармацыі Рэспублікі Беларусь быў заключаны дагавор аб супрацоўніцтве ў галіне навуковых распрацовак па развіцці і ўдасканаленні інфармацыйных тэхналогій і сістэм, нацыянальных інфармацыйных рэсурсаў і іх рацыянальным выкарыстанні, у тым ліку

па выпрацоўцы карпусоў беларускай, рускай і англійскай лексікі з дапамогай партала corpus.by.

Такім чынам, ва ўмовах новага інфармацыйнага асяроддзя чалавечай цывілізацыі – інфасферы, якая цягне за сабой радыкальныя сацыяльныя змены і істотным чынам датычыцца жыцця практычна кожнага грамадзяніна, – у Беларусі робяцца вельмі сур’ёзныя і патрэбныя крокі ў бок дапамогі людзям з інваліднасцю па зроку. Значную ролю пры гэтым адыгрывае забеспячэнне даступнасці інфармацыі для безбар’ернай навігацыі па заканадаўстве Рэспублікі Беларусь усіх катэгорый моўных носьбітаў.

### БІБЛІАГРАФІЧНЫЯ СПАСЫЛКІ

1. Corpus.by. Платформа для апрацоўкі тэкставай і гукавой інфармацыі для розных тэматычных даменаў [Электронны рэсурс]. – 2020. Рэжым доступу: <https://corpus.by>. – Дата доступу: 14.01.2020.
2. Гецевіч, Ю.С. Система синтеза беларускай рэчы па тэксце / Ю.С. Гецевіч, Б.М. Лобанов // Речевые технологии. – 2010. – № 1. – С. 91-100.
3. Гецэвіч, Ю.С. Выкарыстанне сістэм машыннага перакладу і сінтэзу маўлення для забеспячэння даступнасці заканадаўчых тэкстаў на розных мовах у Рэспубліцы Беларусь / Ю.С. Гецэвіч, А.А. Кірдуноў // Информационные технологии и право (Правовая информатизация-2018): материалы VI Международ. науч.-практ. конф., Минск, 17 мая 2018 г. / под общ. ред. Е.И. Коваленко; Нац. центр правовой информ. Респ. Беларусь. – Минск, 2018. – С. 123-128.
4. Гецэвіч, Ю.С. Распрацоўка сінтэзатара беларускага і рускага маўленняў па тэксце для мабільных і інтэрнэт-платформаў / Ю.С. Гецэвіч, Д.А. Пакладок, Д.В. Брэж // Развитие информатизации и государственной системы научно-технической информации (РИНТИ–2012): докл. XI Международ. конф., Минск, 15 нояб. 2012 г. – Минск: ОИПИ НАН Беларусі, 2012. – С. 254-259.
5. Закон Рэспублікі Беларусь от 10 января 2000 г. № 361-З «О нормативно-правовых актах Республики Беларусь» // Информационно-поисковая система «ЭТАЛОН-ONLINE» [Электронный ресурс]. – 2020. Режим доступа: <http://etalonline.by/document/?regnum=h10000361>. – Дата доступа: 14.01.2020.
6. Камітэт па ахове здароўя Мінгарвыканкама [Электронны рэсурс]. – 2020. Рэжым доступу: <https://komzdrav-minsk.gov.by>. – Дата доступу: 14.01.2020.
7. Постановление Совета министров Республики Беларусь от 23 октября 2017 г. № 797 «О внесении изменения и дополнений в Положение о порядке функционирования интернет-сайтов государственных органов и организаций» // Информационно-поисковая система «ЭТАЛОН-ONLINE» [Электронный ресурс]. – 2020. Режим доступа: [http://etalonline.by/document/?regnum=c21700797&q\\_id=643794](http://etalonline.by/document/?regnum=c21700797&q_id=643794). – Дата доступа: 14.01.2020.