

М.В. Макарыч, Д.А. Міронава (Мінск, Беларусь)

КАНТРАСТЫЎНЫЯ ДАСЛЕДАВАННІ ТЭРМІНАЛАГІЧНЫХ СЛОВАЗЛУЧЭННЯЎ У АНГЛІЙСКОЙ І БЕЛАРУСКАЙ МОВАХ

Артыкул прысвечаны праблеме вывучэння паняткава-тэрміналагічнай сістэмы ў галіне інфармацыйных тэхналогій пры дапамозе кантрастыўнага аналізу. Разглядаюцца моўныя спосабы выразу паняткаў і логіка-лінгвістычная арганізацыя тэрміналагічных словазлучэнняў у англійскай і беларускай мовах. Вынікі даследавання падаюцца ў выглядзе дзвюх табліц: «сыходжанняў» і «фразыходжанняў» у пабудове вышэйазначаных словазлучэнняў. Зроблены колькасны і якасны аналіз па выніках даследавання і даюцца рэкамендацыі па іх эфектыўнаму выкарыстанню ў сістэмах штучнага інтэлекту.

Ключавыя словы: *кантрастыўны аналіз, логіка-лінгвістычная арга-нізацыя тэрміналагічных словазлучэнняў, паралельныя тэрміналагічныя словазлучэнні.*

Апошнім часам шмат лінгвістычных тэарэтычных і прыкладных даследаванняў выконваюцца на грунце дасягненняў у галіне кантрастыўнага аналізу моў. Кантрастыўная лінгвістыка ўзнікла на перакрываванні структурнай тыпалогіі моў і лінгвадыдактыкі. Яе асноўная мэта – выяўленне моўных кантрастаў і падабенстваў. Першыя дасягненні ў дадзенай галіне належаць В. Матэзіусу (1882–1945), заснавальніку Пражскай лінгвістычнай школы, які ўжо ў 1926 годзе усведаміў неабходнасць кантрастыўных даследаванняў падчас параўнання шматстатнасці выразных сродкаў існуючых моў.

Часам даследчыкі звужалі кола магчымасцяў кантрастыўнай лінгвістыкі і супрацьстаўлялі яе параўнальнай лінгвістыцы. Менавіта такі пункт гладжання дэманстраваў Г. Хельбіг. «Параўнальная лінгвістыка мае больш тэарэтычную арыентацыю і займаецца як падабенствамі, так і адрозненнямі паміж мовамі, тым часам як кантрастыўная лінгвістыка мае больш практычны накірунак і займаецца пераважна моўнымі адрозненнямі» [1].

У працэсе практычнага выкарыстання дадзеных кантрастыўнага аналізу, напрыклад, падчас выкладання замежных моў, неабходна больш засяродзіцца на адрозненнях, а не на падабенствах. Аднак падчас параўнання моў непазбежна наяўнасць як падабенстваў, так і адрозненняў пры адсутнасці прынцыповай розніцы. Такім чынам, праводзіць мяжу паміж параўнальным і кантрастыўным мовазнаўствам няма падстаў з прычыны таго, што абодва падыходы карыстаюцца аднолькавымі метадамі даследавання і асноўнае іх адрозненне – гэта дадзеныя, якія мусяць улічвацца ў першую чаргу падчас выкарыстання вынікаў даследавання.

Агульныя мэты кантрастыўнага вывучэння моў фармулююцца наступным чынам:

- ✓ выяўленне сыходжанняў і разыходжанняў падчас выкарыстання моўных сродкаў у мовах, якія супрацьстаўляюцца;
- ✓ азначэнне моўных асаблівасцяў, якія могуць застацца па-за ўвагай даследчыка, калі ён засяродзіцца толькі на «унутраным» вывучэнні мовы;
- ✓ назапашванне надзейнай лінгвістычнай базы для прыкладнога мовазнаўства;
- ✓ збор матэрыялаў тыпалогіі для выяўлення універсалій [2].

У якасці аб'ектаў супастаўляльнага аналізу ў кантрастыўных даследаваннях вылучаюцца адзінкі любога ўзроўню мовы: фанемы, словы, словазлучэнні, граматычныя з'явы і суцэльныя тэксты. Часам кантрастыўны метада датычыцца не толькі дзвюх, але трох, ці нават большай колькасці моў, бо ён спалучае якасныя і колькасныя метады аналізу і забяспечвае высокую эфектыўнасць і дзейнасць даследаванняў. Кантрастыўны аналіз непасрэдна звязаны з рознымі відамі прыкладнога мовазнаўства, якое займаецца фармаваннем шматфункцыянальных комплексных баз даных.

З самага пачатку кантрастыўныя тэрміналагічныя даследаванні былі накіраваныя на стварэнне банка тэхнічных ды электронных тэрмінаў і былі абмежаваныя патрэбамі перакладчыцкай і лексікаграфічнай практыкі. Дадзены падыход знайшоў адлюстраванне ў такіх навукоўцаў, як Е. Вюстэр (1931), Е. Андрус (1947), Р. Алан (1947), А. Рэй (1990). Аднак с цягам часу з'явіліся працы па вывучэнні ўзаемадзеяння тэрмінаў з агульналітаратурнай лексікай, у якіх даследавалася паняткава-тэрміналагічная сістэма розных галін ведаў з пункту гледжання моўных спосабаў выразу паняткаў, логіка-лінгвістычная арганізацыя сістэм тэрмінаў і выпрацоўваліся рэкамендацыі па іх эфектыўным выкарыстанні ў сістэмах штучнага інтэлекту [3]. Апошнія ўяўляецца найбольш важкім на сучасным этапе з прычыны таго, што праблема камп'ютарнай апрацоўкі натуральных моў грунтуецца на выкарыстанні лінгвістычных рэсурсаў – корпусы тэкстаў на розных мовах. Напрыклад, сістэма апрацоўкі англамоўных публіцыстычных тэкстаў будзеца на корпусе тэкстаў, якія выкарыстоўваюцца падчас практычных заняткаў па вывучэнні англійскай мовы [4].

Сістэмы для аўтаматычнай апрацоўкі тэкстаў выкарыстоўваюць шматлікія метады апрацоўкі мовы на розных узроўнях. Метады, якія арыентуюцца на ўзровень сказаў, сутыкаюцца з шэрагам складанасцяў, асноўная з якіх – іменная канструкцыя сказа. Больш эфектыўнымі з'яўляюцца метады, з дапамогай якіх мова апрацоўваецца на ўзроўні словазлучэнняў. Такія метады шырока распаўсюджаны ў сістэмах штучнага інтэлекту, дзе натуральна-моўны інтэрфейс выкарыстоўвае дадатковую інфармацыю з лінгвістычнай базы даных. Асноўная частка базы – гэта тэрміналагічныя словазлучэнні і іх перакладныя мадэлі на розных мовах. Прыкладам такой сістэмы можа быць сістэма аўтаматычнага выяўлення

англійскіх тэрміналагічных словазлучэнняў і іх перакладу на рускую мову [5].

На сучасным этапе ў Рэспубліцы Беларусь існуе патрэба ў правядзенні тэрміналагічных кантрастыўных даследаванняў англійскай і беларускай моў з мэтай стварэння адпаведных баз даных, на падмурку якіх можа быць распрацаваная не толькі сістэма аўтаматычнай апрацоўкі тэкста, але і электронны слоўнік ці перакладчык. Прыкладам ёсць англа-беларуска-рускі электронны тэрміналагічны слоўнік, які сёння ствараецца у Беларускім нацыянальным тэхнічным ўніверсітэце на кафедры англійскай мовы №2 сумесна з кафедрай праграмага забеспячэння вылічальнай тэхнікі і аўтаматычных сістэм [6]. З гэтай нагоды выкладчыкамі кафедра праводзіцца шэраг кантрастыўных даследаванняў.

Мэтай дадзенага артыкулу ёсць структурна-квантытатыўны аналіз паралельных тэрміналагічных словазлучэнняў (далей – *c/c*) у англійскай і беларускай мовах. Аб'ект даследавання – тэрміналогія з галіны інфармацыйных тэхналогій (ІТ), якая падаецца ў падручніках «JUnit in Action», «TestComplete Made Easy», «English for Computer Engineering» [7; 8; 9]. Вынікі аналізу прыведзены ў фрагментах табліц 1 і 2. Умоўныя скарачэнні: А – прыметнік; N – назоўнік; V – дзеяслоў; R – прыслоў'е; В – артыкль.

Табліца 1

Сыходжанні ў пабудове англійскіх і беларускіх тэрміналагічных *c/c*²⁶

№	Яксны характар сыходжанняў і прыклады, якія яго адлюстроўваюць	Сінтаксічная мадэль
1	En: join Бел: злучэнне	N
2	En: iteration Бел: ітэрацыя	N
3	En: interrogation Бел: апытанне	N

21	En: foolproof Бел: памылкаўстойлівы	A
22	En: implicit Бел: няяўны	A
23	En: invalid Бел: памылковы	A

²⁶ Выбарка складае 430 лексічных адзінак (ЛА), з іх сыходжанняў – 163 ЛА (38%).

65	En: install Бел: усталёўваць	V
66	En: kill Бел: знішчаць	V
67	En: fill Бел: запаўняць	V

102	En: job priority Бел: прыярытэт задання	N+N
103	En: interrupt vector Бел: вектар перапыну	N+N
104	En: inference chain Бел: ланцужок высновы	N+N
105	En: kernel operation Бел: аперацыя ядра	N+N
106	En: learning program Бел: саманавучальная праграма	N+N

128	En: imperative language Бел: імператыўная мова	A+N
129	En: interactive processing Бел: дыялогавая апрацоўка	A+N
130	En: internal timer Бел: ўбудаваны таймер	A+N

Табліца 2

Разыходжанні ў пабудове англійскіх і беларускіх тэрміналагічных с/с²⁷

№	Якасны характар сыходжанняў і прыклады, якія яго адлюстроўваюць	Сінтаксічная мадэль
1	En: kludge Бел: варыянтны запіс	N↔A+N
2	En: keyword Бел: ключавое слова	N↔A+N
3	En: inverse Бел: інвертаваная выява	N↔A+N
4	En: internals	N↔A+N

²⁷ Выбарка складае 430 ЛА, разыходжанняў – 267 ЛА (62%).

	Бел: унутраная арганізацыя	

44	En: hard copy Бел: друкаванка	A+N↔ N
45	En: graphic pallet Бел: палітра	A+N↔ N

87	En: justification Бел: выраўноўванне радкоў	N↔N+N
88	En: hibernation Бел: стан чаканні	N↔N+N
89	En: holding Бел: захоўванне дадзеных	N↔N+N

12 3	En: radix Бел: падстава сістэмы злічэння	N ↔ N+N+N
12 4	En: program origin Бел: пачатковы адрас праграмы	N+N ↔ A+N+N

19 1	En: job mix Бел: загрузка	N+N↔N
19 2	En: hanging indent Бел: выступ	N+N↔N

23 4	En: iterate Бел: выконваць ітэрацыю	V↔V+N
23 5	En: listen Бел: чакаць сігналу	V↔V+N
23 6	En: item size Бел: памер элемента дадзеных	N+N ↔ N+N+N
23 7	En: least significant digit Бел: малодшы разрад	N+N+N ↔ A+N
23 8	En: gray scale image Бел: паўтонавая выява	A+A+N ↔ A+N

Назіранні за колькаснымі сыходжаньнямі і разыходжаньнямі даследуемых тэрміналагічных словазлучэнняў дазваляюць зрабіць наступныя высновы:

✓ большая частка англійскіх і беларускіх тэрмінаў мае разыходжанні ў пабудове адпаведных сінтаксічных мадэляў – 267 ЛА (62%);

✓ 163 ЛА (38%) ад агульнай колькасці даследуемых ЛА выяўляюць паралельнасць у пабудове сінтаксічных мадэляў, 20 % з якіх – аднакампанентныя тэрміны, а 18 % – шматкампанентныя тэрміны;

✓ асноўную частку паралельных аднакампанентных тэрмінаў складаюць назоўнікі (89% ад агульнай колькасці);

✓ інтэрнацыянальныя тэрміны, наўпрост запазычаныя з англійскай мовы складаюць 64% (*timer – таймер, iteration – ітэрацыя, invariant – інварыянт, indicator – індыкатар, quantifier – квантар, program – праграма, iterative – ітэрацыйны, interface – інтэрфейс, file – файл, operator – аператар*)

✓ найчасцей паралельныя двухкампанентныя тэрміны маюць у сваім складзе прыметнікі (45% ад агульнай колькасці двухкампанентных тэрмінаў), напрыклад:

$A+N$

En: immediate addressing

Бел: наўпростae адрасаванне

✓ назіраецца ўскладненне сінтаксічных структур, якія маюць у складзе спецыяльныя назвы адмысловых інфармацыйных або тэхналагічных паняццяў, напрыклад:

$N \leftrightarrow A+N$

En: tock

Бел: падстаўны аб'ект

$A+N \leftrightarrow A+A+N$

En: petal printer

Бел: п'ялёсткавая друкавальная прылада

✓ неабходна таксама зазначыць нешматлікія выпадкі спрашчэння сінтаксічных структур (8%) з прычыны наяўнасці ў беларускай мове адпаведных лексічных адзінак:

$A+N \leftrightarrow N$

En: pick device

Бел: указка

$N+N \leftrightarrow N$

En: program scheduler

Бел: планавальнік

Колькасныя дадзеныя, якія былі атрыманы падчас кантрастыўных даследаванняў, дазваляюць прааналізаваць ужыванне рэлятараў (да іх можна аднесці падставы, артыклі, парадак словаў) у вышэйзгаданых мовах. Такім чынам, прапанаваны эксперымент дае магчымасць для эфектыўнага паўнаваргаснага даследавання тыпалогіі стварэння тэрміналагічных словазлучэнняў у галіне інфармацыйных тэхналогій у англійскай і беларускай мовах.

Бібліяграфічныя спасылкі

1. Helbig G. Kurze deutsche Grammatik für Ausländer / G. Helbig, G.J. Buscha. Leipzig, 1976.

2. Гак В.Г. О контрастивной лингвистике / В.Г.Гак // Новое в зарубежной лингвистике. Вып. 25. Контрастивная лингвистика. М., 1989. С. 5–17.

3. *Teresa Cabre, M. Terminology: Theory, Methods and Applications / M. Teresa Cabre, J. Sager. // Language Arts & Dicipines, 1999. 247 p.*
4. *Макарыч М.В.* Рэалізацыя навучальнай камп’ютарнай праграмы на падмурку лінгвістычнай базы ведаў аўтаматычнай сістэмы апрацоўкі тэкста // “Сістэмны аналіз і прыкладная інфарматыка”. Мінск, 2016, №1, с. 78–83.
5. *Степанова Д.В.* Специфика организации англо-русского корпуса параллельных текстов по вычислительной технике для перевода терминологических словосочетаний / Д.В.Степанова // Вестн. Минск. гос. лингв. ун-та. Сер.1, Филология, 2009. № 6 (43). С.185–193.
6. *Макарыч М.В.* Лингвистическое и программное обеспечение англо-беларуско-русского словаря технических терминов / М.В. Макарыч, Ю.Б. Попова, М.А. Швед // Системный анализ и прикладная информатика, 2018. № 4. С.74–82.
7. *Tahchiev P. JUnit in Action / P. Tahchiev, F. Leme and others. Stamford : Manning Publications Co, 2011. 492 p.*
8. *Tadros L. TestComplete Made Easy / L. Tadros, E. Holton and others, AutomatedQA Corp. - 1.0 Release, 2008. 284 p.*
9. *Макарыч М.В.* English for Computer Engineering: дапаможнік з электронным дадаткам для студэнтаў спецыяльнасці 1-08 01 01-07 Прафесійнае навучанне (інфарматыка) / М.В. Макарыч, В.В. Пытко. Мінск : РІВШ, 2019.

Y.B. Popova, M.V. Makarych, M.A. Shved (Minsk, Belarus)

MOBILE APPLICATION FOR THE ENGLISH-BELARUSIAN-RUSSIAN TECHNICAL DICTIONARY

The article is devoted to the problem of creating multilingual electronic dictionaries. The authors offer an overview of current approaches in the field of computer lexicography. A mobile application to the electronic English-Belarusian-Russian dictionary of technical terms TechLex is presented. The developed application covers the following subject areas: architecture and construction, water supply, information technology, pedagogy, transport communications, economics and energy construction. It is created in Belarusian National Technical University by the 2nd English Department together with the Department of Software for Information Systems and Technologies.

Key words: *a multilingual electronic dictionary, computer lexicography, an electronic technical dictionary, a client-server application.*

Introduction. Modern computer lexicography has almost completely replaced traditional manual and handwritten lexicographical methods that existed for many centuries. Today there are a lot of new paperless information technologies. Modern applied linguistics offers a set of methods and software for creating dictionaries. The Central object of computer lexicography is an electronic dictionary, which can be defined as «any lexicographic work on computer media and equipped with automatic processing and filling programs» [1, c. 36].

The idea of creating electronic dictionaries arose from scientific investigations