

УДК 167.6:80

ЛІНГВО-СТАТИСТИЧНІ ХАРАКТЕРИСТИКИ ВОКАБУЛЯРІВ ПЕРЕКЛАДІВ СВ. ПИСЬМА НА ПРИКЛАДІ СХІДНИХ МОВ

Богдан Анатолійович Рудий

*Київський національний університет ім. Тараса Шевченка, Музей лінгвістики,
м. Київ, вул. Володимирська 60,*

Вісник Львівського університету. Серія філологічна. 2008. Вип. 45. – С. 77-84.

У статті наводяться результати частотного аналізу однакового тексту (Євангеліє від Івана), перекладеного різними, в тому числі східними, мовами. З порівняльної таблиці випливає, що найбільш розвинену морфологію серед східних мов мають тюркські мови (турецька, азербайджанська, узбецька тощо), а найменш розвинену – чаморро, тагалог та цебуано. Китайська та корейська мови представлені тут умовно, адже для них не вдалося побудувати частотний список слів, а лише ієрогліфів/складів внаслідок відсутності комп'ютерного алгоритму групування ієрогліфів/складів у слова.

Ключові слова: лінгво-статистичні характеристики, частотний словник, переклади Св. Письма.

Постановка проблеми, аналіз публікацій та формулювання цілей статті

В мовознавстві існує проблема побудувати частотні словники (вокабуляри) для перекладів одного й того самого тексту, порівняти між собою лінгво-статистичні характеристики цих вокабулярів й з єдиної порівняльної таблиці зробити висновки про те, наскільки ці характеристики відрізняються одна від одної, як вони корелюють з типологізацією мов, які тенденції простежуються тощо. Ця проблема є науковою і стосується галузі мовної типологізації. Але одночасно вона є дидактичною проблемою в рамках розроблених нами нових методів вивчення та викладання іноземних мов – „методу частотного штурму” та методу „познайтеся з мовами!”.

Дослідження провадиться вперше. Подібне за масштабом та сутністю дослідження – порівняння лінгво-статистичних характеристик багатьох

перекладів *однакового* тексту – відсутнє у світовій літературі. В літературі присутні лише частотні словники різних текстів (чи текстових корпусів) для окремих мов, наприклад, загальномовний словник російської мови – Засорина Л. „Частотный словарь русского языка”. – М., 1985; фінської мови – Pauli Saukkonen, Marjatta Haipus, Antero Niemikorpi, Helena Sulkala. „Suomen kielen taajuussanasto. A Frequency Dictionary of Finnish”. – Helsinki, 1979; англійської мови – частотний словник на основі the Brown Corpus; та інших мов.

Ціллю статті є презентувати результати дослідження, яке значною мірою вирішує описану вище проблему. створення

Виклад основного матеріалу дослідження

Нами було створено частотні словники (вокабуляри) для перекладів одного й того самого тексту (Євангеліє від Івана) східними та іншими мовами. Частотні словники є нерафінованими, тобто словоформи у них незведені до вихідних форм. До кожного словника створено діаграму покриття тексту частотними словами. Для понад чверті мов виконано переклад верхівки списку (N_{50} чи більше) на українську, російську чи англійську мову. Результати роботи зведено у порівняльну таблицю (табл. 1). Результати роботи виставлено на сайті www.Bowfree.com/data/_Frequency%20dictionaries/Freq_John-u.htm.

Табл. 1. Частотні характеристики однакового тексту (Євангеліє від Івана), перекладеного 60-ма різними мовами. Список мов є відранжованим за кількістю найчастіше уживаних слів тексту, необхідних для покриття 50% тексту (N_{50}). Східні мови позначені виділені жирним.

	Мова	Сер. довжина слова	N_{25}	N_{50}	$N_{66,7}$	N_{80}	N_{90}	N_{100} – к-ть неповторюваних слів (порівняно з заг. к-тю слів у тексті)	Заг. к-ть слів у тексті	Наявність перекладу верхівки списку (N_{50} чи більше)
1.	<i>Maori</i>	6,2	5	16	37	85	191	944 (3,8%)	24961 (100%)	
2.	<i>Afrikaans</i>	6,2	7	23	58	138	317	1320 (6,5%)	20204 (100%)	+
3.	<i>Korean</i>	1,0	8	26	46	84	150	627 (1,8%)	35065 (100%)	
4.	<i>Chamorro</i>	7,7	6	28	86	251	681	2109 (12,2%)	17263 (100%)	

5.	Tagalog	8,2	6	28	84	228	565	2011 (10,0%)	20110 (100%)	
6.	<i>Esperanto</i>	6,6	7	28	91	238	602	1925 (11,0%)	17493 (100%)	+
7.	Cebuano	7,9	5	30	84	235	627	2193 (10,8%)	20380 (100%)	
8.	<i>Norwegian</i>	6,0	10	32	89	205	437	1540 (8,6%)	18008 (100%)	+
9.	<i>English</i>	5,9	9	33	80	182	391	1412 (7,6%)	18684 (100%)	+
10.	<i>English, literal transl</i>	6,0	9	34	77	164	337	1279 (6,3%)	20430 (100%)	
11.	Chinese	1,0	9	36	86	165	307	1041 (4,4%)	23640 (100%)	
12.	<i>Danish</i>	6,2	11	36	97	224	503	1676 (9,3%)	17942 (100%)	
13.	<i>Dutch</i>	6,7	11	37	100	231	497	1698 (9,0%)	18832 (100%)	+
14.	<i>Spanish</i>	6,7	8	39	116	288	689	2103 (11,8%)	17761 (100%)	+
15.	<i>Swedish</i>	6,2	11	39	113	256	557	1853 (9,8%)	18951 (100%)	+
16.	<i>Albanian</i>	6,2	9	40	143	389	929	2514 (13,6%)	18487 (100%)	
17.	<i>Gaelic manx</i>	6,4	9	41	104	238	511	1732 (8,6%)	20156 (100%)	
18.	<i>Breton</i>	6,3	9	44	127	294	654	2027 (10,4%)	19479 (100%)	
19.	Gypsy	6,7	10	45	120	268	561	1766 (9,7%)	18134 (100%)	
20.	Indonesian	8,1	11	45	117	281	613	1812 (10,6%)	17080 (100%)	
21.	Hebrew	4,3	6	46	173	484	1083	2503 (17,1%)	14669 (100%)	+
22.	<i>Hungarian</i>	7,6	7	47	213	649	1593	3191 (20,0%)	15986 (100%)	
23.	<i>French</i>	6,9	11	47	128	307	694	2115 (11,6%)	18311 (100%)	+
24.	<i>Portuguese</i>	6,9	8	48	142	339	769	2161 (12,9%)	16704 (100%)	+
25.	<i>Greek old</i>	8,0	10	48	216	682	1712	3159 (21,8%)	14475 (100%)	+
26.	<i>German</i>	6,8	13	50	129	296	661	2014 (11,0%)	18265 (100%)	+
27.	<i>Icelandic</i>	6,1	11	51	163	408	920	2376 (13,9%)	17133 (100%)	
28.	Vietnamic	6,4	15	54	117	251	581	2153 (10,8%)	19942 (100%)	
29.	<i>Italian</i>	6,9	13	55	144	338	771	2204 (12,5%)	17601 (100%)	+
30.	<i>Bulgarian</i>	6,5	13	56	176	443	986	2422 (15,2%)	15933 (100%)	
31.	<i>Latvian</i>	6,7	13	56	192	528	1148	2602 (17,8%)	14615 (100%)	
32.	<i>Welsh</i>	6,2	10	57	145	335	738	2206 (11,3%)	19581 (100%)	
33.	<i>Wolof</i>	5,4	15	57	132	283	591	1816 (10,1%)	18018 (100%)	
34.	<i>Greek modern</i>	6,8	13	60	198	507	1118	2714 (15,5%)	17459 (100%)	
35.	<i>Gothic</i>	7,0	13	60	181	410	857	1771 (19,4%)	9145 (100%)	
36.	<i>Estonian</i>	6,6	14	62	191	468	1036	2487 (15,6%)	15906 (100%)	
37.	<i>Romanian</i>	6,3	14	63	170	379	782	2118 (12,0%)	17703 (100%)	
38.	<i>Ukrainian</i>	6,6	14	64	217	576	1254	2806 (18,1%)	15520 (100%)	+
39.	<i>Belorussian</i>	6,7	15	70	219	555	1165	2629 (17,9%)	14654 (100%)	
40.	<i>Latin</i>	7,2	13	71	226	545	1131	2520 (17,9%)	14057 (100%)	
41.	<i>Croatian</i>	6,2	13	72	207	489	1042	2413 (15,9%)	15174 (100%)	
42.	Georgian	7,6	10	75	256	646	1406	2790 (20,2%)	13842 (100%)	
43.	<i>Russian</i>	6,6	14	76	236	582	1195	2640 (18,2%)	14469 (100%)	+
44.	Uma	7,7	13	77	213	511	1142	2755 (15,8%)	17385 (100%)	
45.	<i>Finnish</i>	7,9	14	80	268	684	1497	3013 (19,9%)	15169 (100%)	
46.	<i>Slovak</i>	6,3	17	80	219	532	1141	2621 (16,7%)	15672 (100%)	
47.	Arabic	4,9	12	82	285	709	1508	2831 (21,4%)	13231 (100%)	+
48.	<i>Polish</i>	7,1	17	96	296	699	1423	2941 (19,4%)	15186 (100%)	
49.	Armenian	6,7	18	103	328	817	1751	3206 (22,0%)	14560 (100%)	+
50.	<i>Czech</i>	6,3	21	106	292	669	1370	2785 (19,7%)	14150 (100%)	
51.	<i>Swahili</i>	7,9	22	110	294	731	1670	3034 (22,2%)	13645 (100%)	
52.	<i>Lithuanian</i>	7,2	20	113	354	896	1902	3255 (24,1%)	13529 (100%)	
53.	<i>Basque of 1571</i>	7,9	22	125	350	762	1522	2980 (20,4%)	14588 (100%)	
54.	Kyrgyz	7,3	26	130	383	925	1891	3302 (23,4%)	14114 (100%)	
55.	Azerbaijan	7,4	28	150	436	1013	2003	3351 (24,8%)	13485 (100%)	
56.	<i>Basque</i>	7,6	24	153	443	957	1875	3325 (23,0%)	14487 (100%)	
57.	Uzbek	7,7	31	188	577	1300	2508	3881 (28,3%)	13733 (100%)	
58.	Turkish	7,9	36	205	581	1248	2354	3559 (29,5%)	12056 (100%)	+
59.	Xhosa	8,4	30	236	757	1770	2961	4152 (34,9%)	11911 (100%)	
60.	Thai	18,9	28	766	1812	2612	3228	3843 (62,4%)	6155 (100%)	

Табл. 2. Діаграма покриття тексту частотними словами (для Єв. від Івана на івриті).

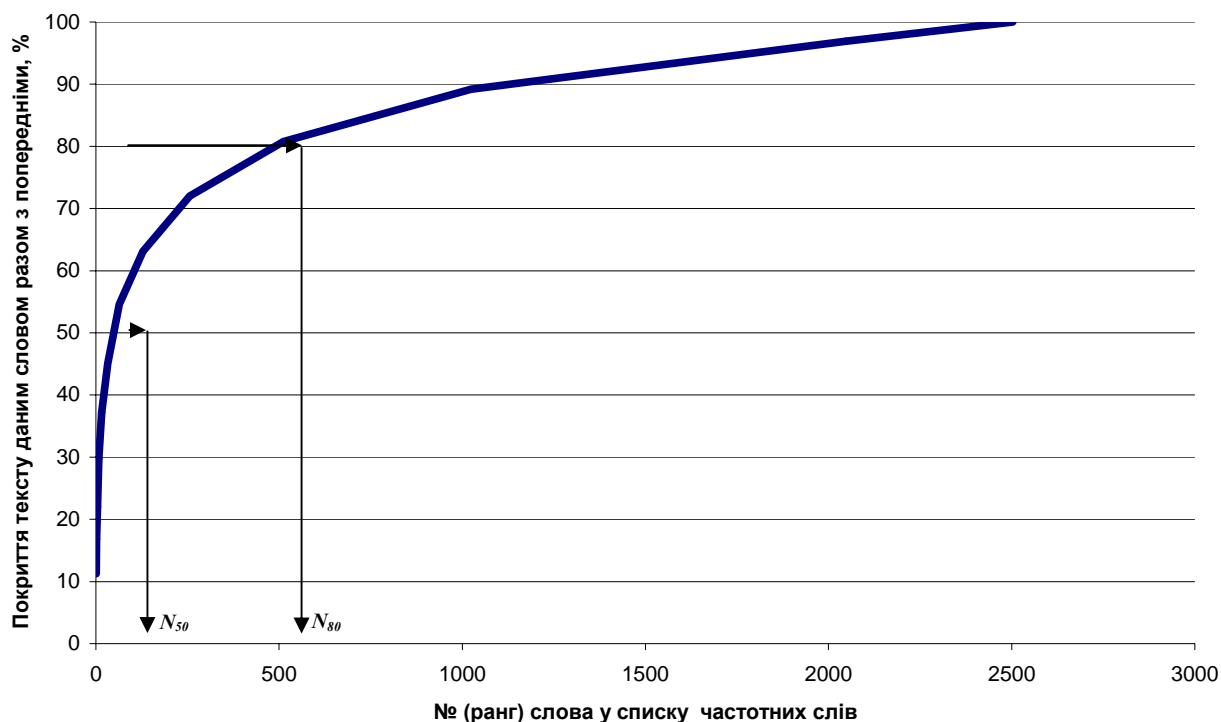


Табл. 3. Список найчастіше уживаних слів тексту, необхідних для покриття 50% тексту (N_{50}).

Ранг	Слово	Транскрипція	Переклад	Частота	Скільки відсотків покриває дане слово разом з попередніми
1	ו	[ве]	и	1651	11,3
2	ה	[га]	артикуль	820	16,8
3	כי	[ки]	ибо, потому что	379	19,4
4	את	[ат]	префикс винительного п	378	22
5	לא	[ло]	не, нет	319	24,2
6	אשר	[ашер]	который, что	314	26,3
7	ישוע	[йешуа]	Иисус	251	28
8	הוא	[гу]	он; он/это есть	246	29,7

9	אל	[аль]	не	182	30,9
10	יאמר	[йамáр]	он сказал	176	32,1
11	אני	[анú]	я	155	33,2
12	אם	[им]	если	127	34,1
13	אליו	[елив]	ему	122	34,9
14	איש	[иш]	мужчина	120	35,7
15	על	[аль]	на	108	36,5
16	זה	[зэ]	это, этот	107	37,2
17	אתם	[атém]	вы (м. р.)	91	37,8
18	גם	[гам]	также	89	38,4
19	לכם	[лэхém]	вам (м. р.)	81	39
20	מן	[мин]	из, от	75	39,5
21	יען	[йаан]	отвечая	71	40
22	אמר	[амér]	он сказал	71	40,4
23	יהי	[йегé]	есть, является	70	40,9
24	כל	[коль]	все, всё	69	41,4
25	דבר	[дувар]	слово, вещь	68	41,9
26	אתה	[атá]	ты	68	42,3
27	אתו	[итó; отó]	с ним; (н)его	68	42,8
28	אלהים	[элогúм]	Бог	68	43,2
29	אבי	[авú]	мой отец	68	43,7
30	יאמרו ויאמרו	[йамеру]	они сказали	67	44,2
31	אליהם	[элейгém]	им	67	44,6
32	היה	[гайá]	был, была	64	45,1
33	עולם	[ола́м]	мир, век	63	45,5
34	למען	[лемаан]	для	62	45,9
35	יהודים	[йагудúм]	иудеи	61	46,3

36	מה	[ма]	что?	56	46,7
37	זאת	[зот]	эта	56	47,1
38	לו	[ло]	ему	52	47,4
39	בו	[бо]	в нём; им	52	47,8
40	אמן	[амэн]	аминь, поправде	51	48,1
41	לי	[ли]	мне; мой	50	48,5
42	עתה	[атá]	сейчас	50	48,8
43	בן	[бэн]	сын	48	49,2
44	כאשר	[каáшер]	когда, поскольку	48	49,5
45	הנה	[гинé]	вот	44	49,8
46	עשה	[асá]	сделал	44	50,1

Наступна таблиця дає уявлення про мови, залучені до дослідження.

Табл. 4. Уривок з Євангелії від Івана (перші 5 віршів) 18-ма східними мовами. На початку даний український варіант в якості перекладу уривка.

Українська	<p>1 Споконвіку було Слово, а Слово в Бога було, і Бог було Слово.</p> <p>2 Воно в Бога було споконвіку.</p> <p>3 Усе через Нього повстало, і ніщо, що повстало, не повстало без Нього.</p> <p>4 І життя було в Нім, а життя було Світлом людей.</p> <p>5 А Світло у темряві світить, і темрява не обгорнула його.</p>
1. Китайська	<p>1 太初有道，道与神同在，道就是神。</p> <p>2 这道太初与神同在。</p> <p>3 万物是藉着他造的。凡被造的，没有一样不是藉着他造的。</p> <p>4 生命在他里头。这生命就是人的光。</p> <p>5 光照在黑暗里，黑暗却不接受光。</p>
2. Корейська	<p>1 태초에 말씀이 계시니라 이 말씀이 하나님과 함께 계셨으니 이 말씀은 곧 하나님이시니라</p> <p>2 그가 태초에 하나님과 함께 계셨고</p> <p>3 만물이 그로 말미암아 지은 바 되었으니 지은 것이 하나도 그가 없이는 된 것이 없느니라</p> <p>4 그 안에 생명이 있었으니 이 생명은 사람들의 빛이라</p> <p>5 빛이 어두움에 비취되 어두움이 깨달지 못하더라</p>
3. Арабська	<p>1 بولس وسلوانس وتيموثاوس الى كنيسة التسالونيكين في الله ابينا والرب يسوع المسيح</p> <p>2 نعمة لكم وسلام من الله ابينا والرب يسوع المسيح</p> <p>3 ينبغي لنا ان نشكر الله كل حين من جهتكم ايها الاخوة كما يحق لان ايمانكم ينمو كثيرا ومحبة كل واحد منكم جميعا بعضكم لبعض تزداد</p> <p>4 حتى اننا نحن انفسنا نفتخر بكم في كل كنائس الله من اجل صبركم وايمانكم في جميع اضطهاداتكم والضيقات التي تحتملونها</p> <p>5 بيينة على قضاء الله العادل انكم تؤهلون لملكوت الله الذي لاجله تتألّمون ايضا</p>

4. Іврит	<p>1 בראשית היה הדבר והדבר היה את האלהים ואלהים היה הדבר: 1</p> <p>2 הוא היה בראשית אצל האלהים: 2</p> <p>3 הכל נהיה על ידו ומבלעדיו לא נהיה כל אשר נהיה: 3</p> <p>4 בו היו חיים והחיים היו אור בני האדם: 4</p> <p>5 והאור בחשך זרח והחשך לא השיגו: 5</p>
5. Турецька	<p>1 Başlangıçta Söz vardı. Söz Tanrı'yla birlikteydi ve Söz Tanrı'ydı.</p> <p>2 Başlangıçta O, Tanrı'yla birlikteydi.</p> <p>3 Her şey O'nun aracılığıyla var oldu, var olan hiçbir şey O'nsuz olmadı.</p> <p>4 Yaşam O'ndaydı ve yaşam insanların ışığıydı.</p> <p>5 Işık karanlıkta parlar ve karanlık onu alt edememiştir.</p>
6. Азербайджанська	<p>1 Əzəldən Kəlam mövcuddur; Kəlam Allahda olub və Kəlam Allah olub.</p> <p>2 O əzəldən Allahda olub.</p> <p>3 Hər şey Onun vasitəsilə yarandı və yaranan şeylərin heç biri Onsuz yaranmadı.</p> <p>4 Həyat onda olub, həyat isə insanların nuru olub.</p> <p>5 Nur qaranlıqda parlayır və qaranlıq onu dərk etməmişdir.</p>
7. Узбецька	<p>1 Азалда Калом бор эди. Калом Худо хузурида эди, Калом -- Худо эди.</p> <p>2 Азалданок У Худо хузурида эди.</p> <p>3 Борлик У орқали бор бўлган, Усиз ҳеч нарса бор бўлган эмас. Неки бор бўлган бўлса, У орқали бор бўлган.</p> <p>4 Унда ҳаёт бор эди, Ҳаёт инсонларнинг нури эди.</p> <p>5 Нур зулматда порлайди, Зулмат эса нурни қоплай олмади.</p>
8. Киргизька	<p>1 Эң башында Сөз бар болгон, Сөз Кудайда болгон жана Сөз Кудай болгон.</p> <p>2 Сөз эң башында Кудайда болгон.</p> <p>3 Бардыгы Ал аркылуу башталган, эмне башталса да, Ансыз башталган эмес.</p> <p>4 Анда өмүр болгон, ал өмүр адамдар үчүн жарык болгон.</p> <p>5 Жарык караңгыда жаркырап турат, караңгылык жарыкты каптаган жок.</p>
9. Грузинська	<p>1 პირველითგან იყო სიტყუაჲ, და სიტყუაჲ იგი იყო ღმრთისა თანა, და ღმერთი იყო სიტყუაჲ იგი.</p> <p>2 ესე იყო პირველითგან ღმრთისა თანა.</p> <p>3 ყოველივე მის მიერ შეიქმნა, და თვნიერ მისა არცა ერთი რაჲ იქმნა, რაოდენი-რაჲ იქმნა.</p> <p>4 მის თანა ცხორებაჲ იყო და ცხორებაჲ იგი იყო ნათელ კაცთა.</p> <p>5 და ნათელი იგი ბნელსა შინა ჩანს, და ბნელი იგი მას ვერ ეწია.</p>
10. Вірменська	<p>1 Սկիզբէն էր Խօսքը, ու Խօսքը Աստուծոյ քով էր, եւ Խօսքը Աստուած էր.</p> <p>2 ան սկիզբէն Աստուծոյ քով էր:</p> <p>3 Ամէն ինչ եղաւ անո՛վ. եղածներէն ո՛չ մէկը եղաւ առանց անոր:</p> <p>4 Կեանքը անով էր, ու կեանքը մարդոց լոյսն էր.</p> <p>5 լոյսը կը փայլէր խաւարի մէջ, բայց խաւարը չճանչցաւ զայն:</p>
11. В'єтнамська	<p>1 Ban ñaàu cou Ngoài Lôøi, Ngoài Lôøi ôu cuøng Nõuc Chuùa Trôøi, vaø Ngoài Lôøi laø Nõuc Chuùa Trôøi.</p> <p>2 Ban ñaàu Ngaøi ôu cuøng Nõuc Chuùa Trôøi.</p> <p>3 Muoân vaät bôui Ngaøi laøm neân, chaúng vaät chi ñaø laøm neân maø khoâng bôui Ngaøi.</p> <p>4 Trong Ngaøi cou søi soáng, søi soáng laø søi saùng cuùa loaøi ngöðøi.</p> <p>5 Sõi saùng soi trong toái taêm, toái taêm chaúng heà nhaän laáy søi saùng.</p>
12. Індонезійська	<p>1 Pada mulanya adalah Firman; Firman itu bersama-sama dengan Allah dan Firman itu adalah Allah.</p> <p>2 Ia pada mulanya bersama-sama dengan Allah.</p> <p>3 Segala sesuatu dijadikan oleh Dia dan tanpa Dia tidak ada suatupun yang telah jadi dari segala yang telah dijadikan.</p> <p>4 Dalam Dia ada hidup dan hidup itu adalah terang manusia.</p> <p>5 Terang itu bercahaya di dalam kegelapan dan kegelapan itu tidak menguasainya.</p>

13. Цебуано	<p>1 Sa sinugdan mao na ang Pulong, ug ang Pulong uban sa Dios, ug Dios ang Pulong. 2 Kini siya sa sinugdan uban sa Dios. 3 Ang tanang mga butang nangahimo pinaagi kaniya; ug niadtong mga nangahimo na, walay bisan usa nga nahimo nga dili pinaagi kaniya. 4 Diha kaniya ang kinabuhi, ug ang maong kinabuhi mao ang kahayag alang sa mga tawo. 5 Ug ang kahayag nagadan-ag sa taliwala sa kangitngitan, ug ang kangitngitan wala makabuntog kaniya.</p>
14. Tagalog (Філіпіні)	<p>1 Nang pasimula siya ang Verbo, at ang Verbo ay sumasa Dios, at ang Verbo ay Dios. 2 Ito rin nang pasimula'y sumasa Dios. 3 Ang lahat ng mga bagay ay ginawa sa pamamagitan niya; at alin man sa lahat ng ginawa ay hindi ginawa kung wala siya. 4 Nasa kaniya ang buhay; at ang buhay ay siyang ilaw ng mga tao. 5 At ang ilaw ay lumiliwanag sa kadiliman; at ito'y hindi napagunawa ng kadiliman.</p>
15. Тайська	<p>1 ในเริ่มแรกนั้นพระวาหะทรงเป็นอยู่แล้ว และพระวาหะทรงอยู่กับพระเจ้า และพระวาหะทรงเป็นพระเจ้า</p> <p>2 ในเริ่มแรกนั้นพระองค์นั้นทรงอยู่กับพระเจ้า</p> <p>3 พระองค์ทรงสร้างสิ่งทั้งปวงขึ้นมา และในบรรดาสิ่งที่เป็นมานั้น ไม่มีสักสิ่งเดียวที่ได้เป็นมานอกเหนือพระองค์</p> <p>4 ในพระองค์มีชีวิต และชีวิตนั้นเป็นความสว่างของมนุษย์ทั้งปวง</p> <p>5 ความสว่างนั้นส่องเข้ามาในความมืด และความมืดหาได้เข้าใจความสว่างไม่</p>
16. Чаморро	<p>1 Y tutujonña gaegue y Finijo, ya y Finijo güiya yan si Yuus; ya y Finijo güiya si Yuus. 2 Güiya gaegue gui tutujonña yan si Yuus. 3 Todo y güinaja sija manmafatinas pot güiya; yaguin ti pot güiya, taya ni esta mafatinas, nu y gaegue gui finatinas sija. 4 Y linâlâ gaegue guiya güiya, ya y linâlâ, güiya y candet y taotao sija. 5 Ya y candet gui jalom jomjom maniina, ya y jemjom ti matungo.</p>
17. Ума	<p>1 Lomo' -na, kako'ia-na dunia' rapajadi', mpolia' ria-imi-hana Lolita. Lolita toei, dohe Alata'ala-i, pai' Lolita toei, Hi'a moto-mi Alata'ala. 2 Ngkai lomo' -na-imi dohe Alata'ala-e. 3 Hi'a-mi to napopale Alata'ala mpopajadi' hawe'ea to rapajadi'. Uma ria hanyalaa to jadi' ane uma Hi'a. 4 Hi'a-mi pehupaa' katuwua', pai' Hi'a-mi baja to mpobajahi nono manusia'. 5 Baja toe mehini ncuu hi rala kabengia-na, pai' kabengia uma mpakulei' mpopatehi-i.</p>
18. Хоса	<p>1 Ekuqalekeni ube ekho uLizwi, waye uLizwi ekuye uThixo, waye uLizwi enguThixo. 2 Yena lowo ebekho ekuqalekeni kuThixo. 3 Izinto zonke zabakho ngaye; engekho yena, akubangakho nanye into ethe yabakho. 4 Kuye kwakukho ubomi, baye ubomi obo bulukhanyiselo lwabantu. 5 Ukhanyiso ke olo lwabonakala ebumnyameni, basuka ubumnyama abaluyiqqa.</p>
19. Циганська	<p>1 Anda gor sas E Vorba, ai E Vorba sas le Devlesa, ai E Vorba sas O Del. 2 Wo anda gor sas le Devlesa. 3 Wo kerdia swako fielo, ai kanchi so kerdilo nas kerdo bi leste. 4 Ande leste sas o traio, ai o traio sas e vediara le manushenge. 5 E vediara strefial ando tuniariko, ai o tuniariko shoxar chi mudardia e vediara.</p>

Висновки

З табл. 1 випливає, що найбільш розвинену морфологію серед східних мов мають тюркські мови (турецька, азербайджанська, узбецька тощо), а найменш розвинену – чаморро, тагалог та цебуано.

З табл. 1-3 видно велику дидактичну користь створення частотних словників, бо студент, скажімо, івриту, завчивши лише перші 46 слів, стає здатним впізнавати кожне друге слово тексту.

Для китайської та корейської мов нагально стоїть проблема створення загальномовного частотного словника, а для цього необхідно винайти комп'ютерний алгоритм групування ієрогліфів/складів у слова.

Перспективи подальших розвідок

Наступним етапом планується опублікувати сутність нових методів вивчення та викладання іноземних мов – „метод частотного штурму” та метод „познайтеся з мовами!”, – та результати спроби побудови експериментального підручника на їхній основі.

Нам вдалося в загальних рисах розробити напів-автоматичний алгоритм створення частотних словників китайської та корейської мов. Тому одним з наступних завдань мусить бути опис і спроба реалізації такого алгоритму.

Lingvo-statistical characteristics of the vocabularies of the Holy Scriptures translations on the example of oriental languages

Bogdan Rudyj

The paper presents the results of comparative analyses of lingvo-statistical characteristics of the translations of the Gospel of John in 60 languages, including 19 oriental ones. For this reason frequency dictionaries for all this translations have been built. It is seen that of all oriental languages the Turkic languages have the most developed (complicated) morphology system, whereas the Chamorro, Tagalog and Cebuano – the most simple morphology system. It can be seen also that Chinese and Korean frequency dictionaries are not the word-ones, but

hieroglyph/syllable ones, – meaning that an algorithm for creating frequency dictionaries is still unrealized.

Key words: lingvo-statistical characteristics, frequency dictionary, translations of the Holy Scriptures.