

УДК 167.6:80

ЛІНГВО-СТАТИСТИЧНІ ХАРАКТЕРИСТИКИ ВОКАБУЛЯРІВ ПЕРЕКЛАДІВ СВЯТОГО ПИСЬМА НА ПРИКЛАДІ СХІДНИХ МОВ

Богдан РУДИЙ

*Київський національний університет ім. Тараса Шевченка, Музей лінгвістики,
вул. Володимирська, 60, 01033, Київ, Україна,
тел.: (0038044) 2393497, моб. тел.: (38068)1239318,
e-mail: cri-evol@ukr.net*

Подано результати частотного аналізу однакового тексту (Євангеліє від Івана), перекладеного різними, у тому числі східними, мовами. З порівняльної таблиці випливає, що найрозвиненішу 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 укласти частотний список *слів*, а лише *ієрогліфів/складів* унаслідок відсутності комп'ютерного алгоритму групування ієрогліфів/складів у слова.

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

У мовознавстві існують проблеми побудови частотних словників (вокабулярів) для перекладів одного й того ж тексту, порівняння між собою лінгво-статистичних характеристик цих вокабулярів та складності 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 у межах розроблюваних нових методів вивчення та викладання іноземних мов – “методу частотного штурму” та методу “познайомтеся з мовами!”.

Дослідження проводиться вперше. Подібне за масштабом та суттю дослідження – порівняння лінгво-статистичних характеристик *багатьох* перекладів *однакового* тексту – відсутнє у світовій літературі. У літературі наявні лише частотні словники різних текстів (чи текстових корпусів) для *окремих* мов, наприклад, загальномовний словник російської мови – Засорина Л. Частотний 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 тощо.

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

Ми створили частотні словники (вокабуляри) для перекладів одного й того ж тексту (Євангеліє від Івана) східними та іншими мовами. Частотні словники є нерафінованими, тобто словоформи у них не зведені 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 1. Результати роботи виставлено на сайті: www.Bowfree.com/data/_Frequency%20dictionaries/Freq_John-u.htm

Таблиця 1

Частотні характеристики однакового тексту (Євангеліє від Івана), перекладеного 60-ма різними мовами*

1	Мова	Середня довжина слова	N_{25}	N_{50}	$N_{66,7}$	N_{80}	N_{90}	N_{100} – кількість неповторюваних слів (порівняно з загальною кількістю слів у тексті)	Загальна кількість слів у тексті	Наявність перекладу верхівки списку (N_{50} чи більше)
2	<i>Maori</i>	6,2	5	16	37	85	191	944 (3,8%)	24961 (100%)	
3	<i>Afrikaans</i>	6,2	7	23	58	138	317	1320 (6,5%)	20204 (100%)	+
4	<i>Korean</i>	1,0	8	26	46	84	150	627 (1,8%)	35065 (100%)	
5	<i>Chamorro</i>	7,7	6	28	86	251	681	2109 (12,2%)	17263 (100%)	
6	<i>Tagalog</i>	8,2	6	28	84	228	565	2011 (10,0%)	20110 (100%)	
7	<i>Esperanto</i>	6,6	7	28	91	238	602	1925 (11,0%)	17493 (100%)	+
8	<i>Cebuano</i>	7,9	5	30	84	235	627	2193 (10,8%)	20380 (100%)	
9	<i>Norwegian</i>	6,0	10	32	89	205	437	1540 (8,6%)	18008 (100%)	+
10	<i>English</i>	5,9	9	33	80	182	391	1412 (7,6%)	18684 (100%)	+
11	<i>English, literal transl</i>	6,0	9	34	77	164	337	1279 (6,3%)	20430 (100%)	
12	<i>Chinese</i>	1,0	9	36	86	165	307	1041 (4,4%)	23640 (100%)	
13	<i>Danish</i>	6,2	11	36	97	224	503	1676 (9,3%)	17942 (100%)	
14	<i>Dutch</i>	6,7	11	37	100	231	497	1698 (9,0%)	18832 (100%)	+
15	<i>Spanish</i>	6,7	8	39	116	288	689	2103 (11,8%)	17761 (100%)	+
16	<i>Swedish</i>	6,2	11	39	113	256	557	1853 (9,8%)	18951 (100%)	+
17	<i>Albanian</i>	6,2	9	40	143	389	929	2514 (13,6%)	18487 (100%)	
18	<i>Gaelic manx</i>	6,4	9	41	104	238	511	1732 (8,6%)	20156 (100%)	
19	<i>Breton</i>	6,3	9	44	127	294	654	2027 (10,4%)	19479 (100%)	
20	<i>Gypsy</i>	6,7	10	45	120	268	561	1766 (9,7%)	18134 (100%)	
21	<i>Indonesian</i>	8,1	11	45	117	281	613	1812 (10,6%)	17080 (100%)	
22	<i>Hebrew</i>	4,3	6	46	173	484	1083	2503 (17,1%)	14669 (100%)	+
23	<i>Hungarian</i>	7,6	7	47	213	649	1593	3191 (20,0%)	15986 (100%)	
24	<i>French</i>	6,9	11	47	128	307	694	2115 (11,6%)	18311 (100%)	+
25	<i>Portuguese</i>	6,9	8	48	142	339	769	2161 (12,9%)	16704 (100%)	+
26	<i>Greek old</i>	8,0	10	48	216	682	1712	3159 (21,8%)	14475 (100%)	+
27	<i>German</i>	6,8	13	50	129	296	661	2014 (11,0%)	18265 (100%)	+
28	<i>Icelandic</i>	6,1	11	51	163	408	920	2376 (13,9%)	17133 (100%)	
29	<i>Vietnamic</i>	6,4	15	54	117	251	581	2153 (10,8%)	19942 (100%)	
30	<i>Italian</i>	6,9	13	55	144	338	771	2204 (12,5%)	17601 (100%)	+
31	<i>Bulgarian</i>	6,5	13	56	176	443	986	2422 (15,2%)	15933 (100%)	
32	<i>Latvian</i>	6,7	13	56	192	528	1148	2602 (17,8%)	14615 (100%)	
33	<i>Welsh</i>	6,2	10	57	145	335	738	2206 (11,3%)	19581 (100%)	
34	<i>Wolof</i>	5,4	15	57	132	283	591	1816 (10,1%)	18018 (100%)	
35	<i>Greek modern</i>	6,8	13	60	198	507	1118	2714 (15,5%)	17459 (100%)	
36	<i>Gothic</i>	7,0	13	60	181	410	857	1771 (19,4%)	9145 (100%)	

Закінчення табл. 1

37	<i>Estonian</i>	6,6	14	62	191	468	1036	2487 (15,6%)	15906 (100%)	
38	<i>Romanian</i>	6,3	14	63	170	379	782	2118 (12,0%)	17703 (100%)	
39	<i>Ukrainian</i>	6,6	14	64	217	576	1254	2806 (18,1%)	15520 (100%)	+
40	<i>Belorussian</i>	6,7	15	70	219	555	1165	2629 (17,9%)	14654 (100%)	
41	<i>Latin</i>	7,2	13	71	226	545	1131	2520 (17,9%)	14057 (100%)	
42	<i>Croatian</i>	6,2	13	72	207	489	1042	2413 (15,9%)	15174 (100%)	
43	Georgian	7,6	10	75	256	646	1406	2790 (20,2%)	13842 (100%)	
44	<i>Russian</i>	6,6	14	76	236	582	1195	2640 (18,2%)	14469 (100%)	+
45	Uma	7,7	13	77	213	511	1142	2755 (15,8%)	17385 (100%)	
46	<i>Finnish</i>	7,9	14	80	268	684	1497	3013 (19,9%)	15169 (100%)	
47	<i>Slovak</i>	6,3	17	80	219	532	1141	2621 (16,7%)	15672 (100%)	
48	Arabic	4,9	12	82	285	709	1508	2831 (21,4%)	13231 (100%)	+
49	<i>Polish</i>	7,1	17	96	296	699	1423	2941 (19,4%)	15186 (100%)	
50	Armenian	6,7	18	103	328	817	1751	3206 (22,0%)	14560 (100%)	+
51	<i>Czech</i>	6,3	21	106	292	669	1370	2785 (19,7%)	14150 (100%)	
52	<i>Swahili</i>	7,9	22	110	294	731	1670	3034 (22,2%)	13645 (100%)	
53	<i>Lithuanian</i>	7,2	20	113	354	896	1902	3255 (24,1%)	13529 (100%)	
54	<i>Basque of 1571</i>	7,9	22	125	350	762	1522	2980 (20,4%)	14588 (100%)	
55	Kyrgyz	7,3	26	130	383	925	1891	3302 (23,4%)	14114 (100%)	
56	Azerbaijan	7,4	28	150	436	1013	2003	3351 (24,8%)	13485 (100%)	
57	<i>Basque</i>	7,6	24	153	443	957	1875	3325 (23,0%)	14487 (100%)	
58	Uzbek	7,7	31	188	577	1300	2508	3881 (28,3%)	13733 (100%)	
59	Turkish	7,9	36	205	581	1248	2354	3559 (29,5%)	12056 (100%)	+
60	Xhosa	8,4	30	236	757	1770	2961	4152 (34,9%)	11911 (100%)	
61	Thai	18,9	28	766	1812	2612	3228	3843 (62,4%)	6155 (100%)	

* Список мов є відранжованим за кількістю найчастіше уживаних слів тексту, необхідних для покриття 50% тексту (N_{50}). Східні мови виділено жирним шрифтом.

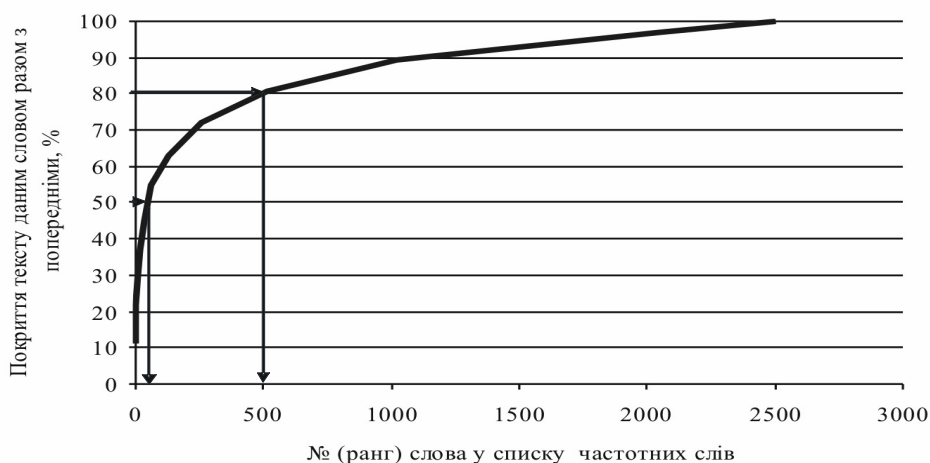


Рис. 1. Діаграма покриття тексту частотними словами
(для Євангелія від Івана на івриті)

Таблиця 2

Найчастіше уживані слова тексту, необхідні для покриття 50% тексту (N_{50})

Ранг	Слово	Транскрипція	Переклад	Частота	Скільки % покриває 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 падежа	378	22
5	לא	[ло]	не, нет	319	24,2
6	אשר	[ашёр]	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	[йамар]	он сказал	176	32,1
11	אני	[ану]	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	[атём]	вы (мужской род)	91	37,8
18	גם	[гам]	также	89	38,4
19	לכם	[лэхём]	вам (мужской род)	81	39
20	מן	[мин]	из, от	75	39,5
21	יען	[йаан]	отвечая	71	40
22	אמר	[амёр]	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	[элогум]	Бог	68	43,2
29	אבי	[аву]	мой отец	68	43,7
30	יאמרו ויאמרו	[йамеру]	они сказали	67	44,2
31	אליהם	[элейгём]	им	67	44,6
32	היה	[гайá]	был, была	64	45,1
33	עולם	[олám]	мир, век	63	45,5
34	למען	[лемаан]	для	62	45,9
35	יהודים	[йагудум]	иудеи	61	46,3

Закінчення табл. 2

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.

Таблиця 3

Уривок з Євангелія від Івана (перші 5 віршів) 18-ма східними мовами*

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

Продовження табл. 3

Иврит	<p>1 בראשית היה הדבר והדבר היה את האלהים ואלהים היה הדבר: 2 הוא היה בראשית אצל האלהים: 3 הכל נהיה על ידו ומבלעדיו לא נהיה כל אשר נהיה: 4 בו היו חיים והחיים היו אור בני האדם: 5 והאור בחשך זרה והחשך לא השיגו:</p>
Турецька	<p>1 Başlangıçta Söz vardı. Söz Tanrı'yla birlikteydi ve Söz Tanrı'ydı. 2 Başlangıçta O, Tanrı'yla birlikteydi. 3 Her şey O'nun aracılığıyla var oldu, var olan hiçbir şey O'nusuz olmadı. 4 Yaşam O'ndaydı ve yaşam insanların ışığıydı. 5 Işık karanlıkta parlar ve karanlık onu alt edememiştir.</p>
Азербайджанська	<p>1 Əzəldən Kəlam mövcuddur; Kəlam Allahda olub və Kəlam Allah olub. 2 O əzəldən Allahda olub. 3 Hər şey Onun vasitəsilə yarandı və yaranan şeylərin heç biri Onsuz yaranmadı. 4 Həyat onda olub, həyat isə insanların nuru olub. 5 Nur qaranlıqda parlayır və qaranlıq onu dərk etməmişdir.</p>
Узбецька	<p>1 Азалда Калом бор эди. Калом Худо хузурида эди, Калом -- Худо эди. 2 Азалданок У Худо хузурида эди. 3 Борлик У оркали бор бўлган, Усиз ҳеч нарса бор бўлган эмас. Неки бор бўлган бўлса, У оркали бор бўлган. 4 Унда ҳаёт бор эди, Ҳаёт инсонларнинг нури эди. 5 Нур зулматда порлайди, Зулмат эса нурни қоплай олмади.</p>
Киргизька	<p>1 Эң башында Сөз бар болгон, Сөз Кудайда болгон жана Сөз Кудай болгон. 2 Сөз эң башында Кудайда болгон. 3 Бардыгы Ал аркылуу башталган, эмне башталса да, Ансыз башталган эмес. 4 Анда өмүр болгон, ал өмүр адамдар үчүн жарык болгон. 5 Жарык караңгыда жаркырап турат, караңгылык жарыкты каптаган жок.</p>
Вірменська	<p>1 Սկիզբից էր Խօսքը, ու Խօսքը Աստուծոյ քով էր, եւ Խօսքը Աստուծօ էր. 2 աւ սկիզբից Աստուծոյ քով էր: 3 Ասէն ինչ եղաւ աւն վ. եղածներէն ո՛չ սէկը եղաւ առանց աւնը: 4 Կէանքը աւով էր, ու կէանքը մարդոց լոյսն էր. 5 լոյսը կը փայլէր խաւարի մէջ, բայց խաւարը չճանչցաւ զայն:</p>
В'єтнамська	<p>1 Ban ñaàu còu Ngoàì Lôøì, Ngoàì Lôøì òu cùøng Nõuc Chuùa Trðøì, vaø Ngoàì Lôøì laø Nõuc Chuùa Trðøì. 2 Ban ñaàu Ngaøì òu cùøng Nõuc Chuùa Trðøì. 3 Muoàn vaät bõuì Ngaøì laøm neân, chaúng vaät chi ñaø laøm neân maø khaøng bõuì Ngaøì. 4 Trong Ngaøì còu sõi soáng, sõi soáng laø sõi saúng cuùa loaøì ngõðøì. 5 Sõi saúng soi trong toái taêm, toái taêm chaúng heà nhaän laáy sõi saúng.</p>
Індонезійська	<p>1 Pada mulanya adalah Firman; Firman itu bersama-sama dengan Allah dan Firman itu adalah Allah. 2 Ia pada mulanya bersama-sama dengan Allah. 3 Segala sesuatu dijadikan oleh Dia dan tanpa Dia tidak ada suatupun yang telah jadi dari segala yang telah dijadikan. 4 Dalam Dia ada hidup dan hidup itu adalah terang manusia. 5 Terang itu bercahaya di dalam kegelapan dan kegelapan itu tidak menguasainya.</p>

Закінчення табл. 3

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

* На початку подано український варіант як переклад уривка.

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

З таблиць 1, 2 та рисунка 1 видно велику дидактичну користь створення частотних словників, бо студент, скажімо, івриту, завчивши лише перші 46 слів, здатний упізнавати 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.

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

Нам вдалося в загальних рисах розробити напівавтоматичний алгоритм створення частотних словників китайської та корейської мов. Тому одним із 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

*Taras Shevchenko Kyiv National University, Linguistics' museum,
Volodymyrska str., 60, 01033, Kyiv, Ukraine,
tel.: (0038044)2393497; (38068)1239318; e-mail: cri-evol@ukr.net*

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.

Стаття надійшла до редколегії 27.09.2007

Прийнята до друку 03.03.2008