



Observatorio de la diversidad lingüística y cultural en la Internet
Observatoire de la diversité linguistique et culturelle dans l'Internet
Observatory of the Linguistic and Cultural Diversity in the Internet

<https://OBDILCI.ORG/>

WEB Multilingualism report #5, Version 2, 9/2025 (Version 2 correct errors in the computations per region.)

Web Multilingualism analyzed by ccTLD, languages, gTLDs and much more: winners and losers, Daniel Pimienta, OBDILCI, August 2025

This series of report has been made possible thanks to the courtesy access granted by DataProvider.com to OBDILCI. DataProvider.com maintains a database gathering information related to a series of websites close to the total WWW. Among the many information kept for each web site, there is the unique main language of the site and the presence of hreflang tags, which a proportion of sites uses to list the linguistic versions of the site.

PREVIOUS REPORTS

- Informes sobre el multilingüismo de la Web #4: [Una evaluación de la reciprocidad en el uso mutuo del español y del portugués en las Web lusófona e hispanófona.](#) 6/2025
- L'état de multilinguisme de la Toile, rapport technique #3 : [Une caractérisation du Web francophone à partir d'une série de paramètres, en comparaison avec d'autres langues dominantes sur la Toile.](#) 5/2025
- The state of Multilingualism of the WWW, Technical Report #2: [Exploring web presence and multilingualism of European minority languages with associated gTLDs.](#) 3/2025
- The state of Multilingualism of the WWW, Technical Report #1: [Approximation of the Rate of multilingualism of the WWW,](#) 2/2025

MAIN FINDINGS

- ✓ High web multilingualism is correlated with: ecommerce, high economic footprint, European countries, Arabic countries,
- ✓ Low web multilingualism is correlated with English-speaking countries and some Asian countries (China, North Korea, India, Japan, Indonesia)
- ✓ Low web multilingualism, 48% of countries but 76% of speakers
- ✓ Winners: Basque, Ukrainian, Latvian, Catalan, Estonian and Greek
- ✓ Losers: Chinese, Korean, Japanese
- ✓ Winners: Monaco, Moldova, Kuwait, Ukraine and Mauritania
- ✓ Losers: Zimbabwe, China, South Africa, South Korea
- ✓ Winners: Montreal, Kiev, Barcelona, Brussels, Hamburg, Lisbon
- ✓ Losers: Los Angeles, New York, Houston, Moscow, Tokyo
- ✓ Winners: Most linguistic and geographic gTLDs, ccTLDs, business-oriented sites
- ✓ Losers: The rest of gTLDs, non-business-oriented sites
- ✓ Singularity: Luxemburg very high but Luxemburgish very low
- ✓ Singularity: Portuguese low in spite Portugal high because Brazil very low
- ✓ Singularity: North America low in spite Canada fair because USA very low

CONTENT

PREVIOUS REPORTS	1
MAIN FINDINGS	1
FOCUS OF THIS REPORT	3
METHOD, SOURCES AND BIASES	3
ANALYSIS PER COUNTRY CODE TOP LEVEL DOMAIN.....	3
ANALYSIS BY LANGUAGE.....	9
ANALYSIS FROM OTHER CONTEXTS.....	12
gTLDs.....	12
CITIE'S WEBSITES	13
OTHER GROUPS.....	14
 Table 1: Results per country top level domain (ccTLD)	3
Table 2: Web multilingualism per continent	8
Table 3: Web multilingualism per region.....	9
Table 4: Web multilingualism per language	9
Table 5: Web multilingualism per gTLD	12
Table 6: Web multilingualism per city	13
Table 7: Web multilingualism other groups.....	14
Table 8: Web multilingualism per economic impact range	15
 Figure 1: Histogram number of countries per multilingualism score	7
Figure 2: Histogram %population per multilingualism range	7

FOCUS OF THIS REPORT

The two first reports have addressed the question of web multilingualism in particular contexts. This report is systematizing the analysis of this indicator in all possible contexts : all countries via ccTLD, all the languages processed by the source of data, generic TLDs and some additional ones, where applicable and feasible.

METHOD, SOURCES AND BIASES

The method stands on DataProvider.com database and the parameter Hreflang= which is collected by this database.

The rate of multilingualism of the group of websites analyzed is computed on the basis that 40%, in average, of multilingual websites use the hreflang= instruction to specify the list of linguistic versions of the website. A simple prorating operation is made (multiplying by 1/0.40, = x 2.5) the percentage of sites using href= in the sampling.

Obviously, this is a very rough manner to compute this indicator and the absolute values remain questionable, but some reasonable trust could be given for comparison purpose, on the hypothesis that the bias is evenly existing and does not affect therefore comparisons.

Biases which could distort the results in favor or some country or language cannot be however totally discarded if, for some reasons a particular language or country is prone to a use of this html instruction very different from the world average.

As for the language part of the study, it is subject to the limitations from the language detection algorithm used by Data Provider, which identify only some 100 languages. This part is also subject of the possible biases of the same algorithm; in the second study of the series, we did identify some situations with some of the languages.

ANALYSIS PER COUNTRY CODE TOP LEVEL DOMAIN

Table 1: Results per country top level domain (ccTLD)

			ALL	87 199 527	4 226 881	14,39%	
Country_Name	ITU	#Lang.	ccTLD	#WEB	HREFLANG	MULTIL	W.AVG
Monaco	99,15	5	MC	1 069	308	72,03%	
Moldova	80,21	10	MD	15 186	4278	70,43%	
Kuwait	99,75	18	KW	1 091	279	63,93%	
Ukraine	82,38	32	UA	211 516	52234	61,74%	
Mauritania	37,38	8	MR	561	136	60,61%	
Luxembourg	98,76	20	LU	22 859	5199	56,86%	
Latvia	92,19	15	LV	46 142	10142	54,95%	
Georgia	81,88	15	GE	17 345	3793	54,67%	
Saudi Arabia	100,00	28	SA	27 756	5994	53,99%	
Andorra	95,45	5	AD	1 052	225	53,47%	
Armenia	80,00	6	AM	11 182	2390	53,43%	Hacked radio am
Greece	85,01	26	GR	243 677	52077	53,43%	
Jordan	92,53	9	JO	1 552	329	53,00%	
Estonia	93,18	14	EE	65 173	13588	52,12%	
Iraq	81,73	10	IQ	3 185	652	51,18%	
Greenland	72,00	2	GL	1 062	217	51,08%	
Cuba	71,28	5	CU	358	72	50,28%	
Cyprus	91,22	21	CY	7 704	1508	48,94%	

Egypt	72,69	17	EG	2 591	507	48,92%	
Yemen	17,69	11	YE	209	40	47,85%	
Bahrain	100,00	15	BH	1 005	187	46,52%	
Puerto Rico	77,74	6	PR	360	65	45,14%	
Tajikistan	56,79	11	TJ	1 633	287	43,94%	
Myanmar	58,54	21	MM	1 052	184	43,73%	
Bulgaria	80,39	17	BG	40 863	7132	43,63%	
Albania	83,14	8	AL	11 114	1926	43,32%	
Oman	95,25	20	OM	1 796	307	42,73%	
North Macedonia	87,18	10	MK	13 737	2315	42,13%	
Azerbaijan	89,00	16	AZ	14 081	2273	40,36%	
Lithuania	88,50	10	LT	84 366	13478	39,94%	
San Marino	87,03	1	SM	580	91	39,22%	
Cambodia	60,72	10	KH	1 305	203	38,89%	
China–Hong Kong	95,98	12	HK	47 269	7296	38,59%	
Portugal	85,79	19	PT	152 372	22253	36,51%	
Panama	77,95	10	PA	2 654	380	35,80%	
Slovenia	90,38	12	SI	68 650	9606	34,98%	Some level of hacking
Croatia	83,24	16	HR	68 411	9439	34,49%	
Kyrgyzstan	88,47	16	KG	5 893	813	34,49%	
Belgium	94,63	52	BE	477 415	63182	33,09%	
China–Macao	89,18	4	MO	883	115	32,56%	
Aruba	97,17	4	AW	143	18	31,47%	
Bahamas	94,82	3	BS	185	23	31,08%	
Uzbekistan	89,01	28	UZ	23 545	2908	30,88%	
Viet Nam	78,08	15	VN	196 929	23691	30,08%	
French Polynesia	75,00	2	PF	549	66	30,05%	
Mauritius	79,50	14	MU	3 499	417	29,79%	
Finland	93,51	48	FI	206 111	24542	29,77%	
Qatar	99,65	15	QA	4 573	544	29,74%	
Nauru	82,67	1	NR	17	2	29,41%	Too few domains
Bosnia and Herzegovina	83,38	10	BA	16 623	1915	28,80%	
Palestine	86,64	3	PS	2 897	324	27,96%	
Nicaragua	58,23	3	NI	834	93	27,88%	
Turkmenistan	25,00	27	TM	190	21	27,63%	
Spain	95,45	49	ES	538 377	58925	27,36%	
Thailand	89,54	27	TH	37 027	4003	27,03%	
Serbia	85,36	20	RS	70 043	7556	26,97%	
Costa Rica	85,40	5	CR	4 268	460	26,94%	
Italy	87,03	49	IT	1 204 635	129005	26,77%	
Switzerland	97,34	35	CH	696 369	73800	26,49%	
Haiti	38,95	4	HT	362	37	25,55%	
Algeria	76,91	12	DZ	3 746	380	25,36%	
Iceland	99,83	5	IS	22 470	2234	24,86%	
Lebanon	83,49	11	LB	1 048	102	24,33%	
Antigua and Barbuda	77,58	1	AG	3 514	341	24,26%	English speaking high
Sudan	28,70	14	SD	689	66	23,95%	
Honduras	58,28	4	HN	1 671	158	23,64%	
New Caledonia	85,00	2	NC	2 345	221	23,56%	
Tunisia	72,35	5	TN	14 149	1326	23,43%	
United Arab Emirates	100,00	26	AE	69 386	6321	22,77%	
Mongolia	83,02	7	MN	6 905	620	22,45%	
Turkey	87,31	32	TR	283 043	25319	22,36%	
Faroe Islands	98,00	1	FO	2 176	194	22,29%	
Mozambique	19,84	24	MZ	4 243	374	22,04%	
Madagascar	20,37	9	MG	1 759	155	22,03%	
Kazakhstan	92,88	32	KZ	61 603	5353	21,72%	
Netherlands	97,01	57	NL	1 430 664	124129	21,69%	
Chad	13,18	6	TD	151	13	21,52%	
Romania	89,20	21	RO	280 905	24006	21,36%	
Suriname	78,40	6	SR	1 100	94	21,36%	Some level of hacking
Dominican Rep.	84,62	8	DO	9 140	780	21,33%	
Malta	92,07	8	MT	4 494	370	20,58%	
Vanuatu	45,73	3	VU	317	26	20,50%	
Guinea-Bissau	32,47	13	GW	110	9	20,45%	
Belize	72,36	5	BZ	4 235	346	20,43%	
Libya	88,50	19	LY	8 287	676	20,39%	
Niger	23,19	12	NE	149	12	20,13%	
Morocco	91,00	9	MA	36 525	2939	20,12%	
Slovakia	89,83	17	SK	191 889	15226	19,84%	
Syria	36,00	11	SY	530	42	19,81%	

Austria	94,92	37	AT	435 276	33976	19,51%	
Togo	37,02	13	TG	678	50	18,44%	
Israel	87,04	20	IL	130 731	9639	18,43%	
Seychelles	87,40	2	SC	771	56	18,16%	
Cameroon	41,91	14	CM	3 273	237	18,10%	
Burundi	11,08	7	BI	768	55	17,90%	
Gabon	71,93	6	GA	522	37	17,72%	
Congo	38,38	10	CG	438	31	17,69%	
China-Taiwan	93,14	9	TW	127 416	8900	17,46%	
Afghanistan	17,71	16	AF	1 350	94	17,41%	
El Salvador	67,66	5	SV	2 426	167	17,21%	
Ethiopia	19,38	17	ET	1 589	109	17,15%	
Poland	86,41	25	PL	839 242	56984	16,97%	
British Virgin Islands	83,00	2	VG	390	26	16,67%	Hacked
France	86,84	57	FR	1 409 670	91475	16,22%	
Fiji	79,35	2	FJ	682	44	16,13%	
Liechtenstein	97,34	2	LI	8 736	561	16,05%	
Hungary	91,45	28	HU	322 873	20408	15,80%	
Mali	35,09	11	ML	905	56	15,47%	
Canada	93,96	105	CA	892 374	55005	15,41%	
Guatemala	56,05	4	GT	8 134	480	14,75%	
Dominica	83,83	3	DM	119	7	14,71%	
Czechia	85,99	23	CZ	575 432	33693	14,64%	
AVERAGE ccTLD						14,39%	
Burkina Faso	17,02	11	BF	648	37	14,27%	
Guinea	26,50	8	GN	158	9	14,24%	
Guernsey	87,00	1	GG	8 837	503	14,23%	Hacked Good Game
Tonga	58,53	1	TO	4 541	258	14,20%	Hacked
Côte d'Ivoire	40,65	19	CI	3 842	218	14,19%	
Bolivia	70,24	7	BO	4 656	263	14,12%	
Brunei	99,03	9	BN	374	21	14,04%	
Bhutan	88,43	5	BT	557	31	13,91%	
Guyana	81,72	3	GY	602	33	13,70%	
Denmark	99,77	45	DK	351 942	19199	13,64%	
São Tomé e Príncipe	61,46	5	ST	1 761	96	13,63%	
Norway	99,00	44	NO	212 544	11571	13,61%	
Venezuela	67,00	9	VE	5 943	323	13,59%	
Ecuador	77,17	7	EC	18 016	978	13,57%	
Germany	93,50	67	DE	4 156 621	224514	13,50%	
Singapore	94,29	25	SG	57 731	3107	13,45%	
Angola	44,76	12	AO	4 836	258	13,34%	
Benin	32,21	10	BJ	844	45	13,33%	
Malaysia	97,69	28	MY	84 826	4440	13,09%	
Uruguay	89,90	16	UY	26 208	1366	13,03%	
Ghana	69,94	19	GH	1 672	87	13,01%	
Eswatini	57,62	8	SZ	404	21	13,00%	
Djibouti	65,02	7	DJ	366	19	12,98%	
Saint Lucia	70,08	3	LC	352	18	12,78%	
Senegal	60,61	14	SE	464 959	23012	12,37%	
Sweden	95,70	56	SE	464 959	23012	12,37%	
AVERAGE GLOBAL						12,12%	
Laos	63,63	10	LA	6 852	331	12,08%	Hacked Los Angeles, Louisiana...
Peru	79,48	7	PE	51 233	2450	11,96%	
Mexico	81,18	16	MX	306 474	14599	11,91%	
Gibraltar	94,44	3	GI	696	32	11,49%	
Saint Kitts and Nevis	76,42	2	KN	109	5	11,47%	
Paraguay	78,09	11	PY	11 692	514	10,99%	
Solomon Islands	42,46	2	SB	210	9	10,71%	
Democratic Republic of the Congo	30,55	19	CD	1 353	57	10,53%	Hacked cd
Montserrat	0,00	1	MS	1 167	49	10,50%	
Wallis and Futuna	20,00	1	WF	239	10	10,46%	
Malawi	17,99	15	MW	599	25	10,43%	
Anguilla	85,00	1	AI	136 384	5665	10,38%	hacked Artificial intelligence
Tuvalu	74,31	1	TV	51 937	2156	10,38%	Hacked TV
Chile	94,46	12	CL	217 534	8882	10,21%	
Cape Verde Islands	73,54	2	CV	2 551	102	10,00%	
Iran	79,63	28	IR	208 872	8178	9,79%	
Colombia	77,34	5	CO	578 846	22454	9,70%	Domain hacked com
Somalia	27,60	9	SO	3 586	137	9,55%	
Kiribati	88,00	1	KI	105	4	9,52%	

Saint Helena	41,00	1	SH	3 968	140	8,82%	used in Switzerland
Russian Federation	92,25	48	RU	1 709 307	59892	8,76%	
Gambia	45,91	13	GM	400	14	8,75%	
Montenegro	89,81	12	ME	187 346	6544	8,73%	Hacked me
Uganda	15,25	22	UG	3 411	118	8,65%	
Saint Vincent and the Grenadines	75,99	2	VC	6 074	206	8,48%	Hacked
Belarus	91,51	19	BY	67 067	2259	8,42%	
Ireland	96,50	31	IE	108 291	3647	8,42%	
Rwanda	34,20	9	RW	2 273	76	8,36%	
Zambia	33,04	19	ZM	1 527	51	8,35%	
Argentina	89,23	28	AR	251 744	8328	8,27%	
Grenada	74,12	1	GD	488	16	8,20%	
Trinidad and Tobago	84,70	7	TT	665	21	7,89%	
Maldives	84,69	1	MV	1 765	55	7,79%	
Niue	85,00	1	NU	44 566	1358	7,62%	Hacked nu=naked
East Timor	34,02	3	TL	1 598	48	7,51%	Hacked
Jamaica	83,39	7	JM	514	15	7,30%	
Namibia	64,41	6	NA	1 883	53	7,04%	
Brazil	84,15	29	BR	1 767 293	49222	6,96%	
Micronesia	40,54	2	FM	4 648	129	6,94%	Hacked radio fm
Philippines	83,77	18	PH	25 248	694	6,87%	
Sri Lanka	51,20	9	LK	21 103	570	6,75%	
Indonesia	69,21	30	ID	265 085	7026	6,63%	
Barbados	79,99	1	BB	272	7	6,43%	
Samoa	58,14	1	WS	8 441	216	6,40%	Hacked
United States	93,14	102	US	270 045	6248	5,78%	
United States (country)	93,14	102	ALL	26 955 911	518791	4,81%	
New Zealand	96,16	49	NZ	223 600	4968	5,55%	
Pakistan	27,38	25	PK	55 940	1215	5,43%	
Cayman Islands	85,00	3	KY	1 888	41	5,43%	
Equatorial Guinea	60,37	7	GQ	1 286	27	5,25%	
Tanzania	29,06	26	TZ	13 773	287	5,21%	
Japan	86,98	17	JP	900 409	18411	5,11%	
Jersey	90,00	2	JE	1 594	32	5,02%	Hacked je
United Kingdom	96,30	74	UK	2 178 399	43330	4,97%	
South Sudan	12,14	7	SS	151	3	4,97%	
Central African Rep.	10,58	9	CF	2 517	46	4,57%	Free domain
India	48,08	72	IN	903 802	16233	4,49%	
Falkland Islands	99,02	1	FK	56	1	4,46%	
Kenya	35,00	30	KE	42 705	715	4,19%	
U.S. Virgin Islands	70,00	3	VI	122	2	4,10%	
Sierra Leone	20,64	11	SL	316	5	3,96%	
Australia	97,06	79	AU	1 121 532	17619	3,93%	
Bermuda	98,37	2	BM	1 081	16	3,70%	
Papua New Guinea	24,07	3	PG	952	14	3,68%	
Nigeria	39,21	28	NG	59 913	865	3,61%	
Bangladesh	44,50	18	BD	15 848	223	3,52%	
Nepal	55,77	19	NP	32 255	444	3,44%	
South Korea	97,42	15	KR	249 401	3426	3,43%	
South Africa	75,66	31	ZA	406 452	5265	3,24%	
Botswana	81,36	9	BW	3 038	39	3,21%	
Lesotho	47,99	7	LS	803	10	3,11%	
China	77,48	23	CN	861 667	9407	2,73%	
Zimbabwe	38,43	15	ZW	13 422	118	2,20%	
Liberia	23,50	10	LR	131	1	1,91%	
Comoros	35,67	4	KM	36		0,00%	Too few domains
Eritrea	20,01	6	ER	1		0,00%	Too few domains
Marshall Islands	65,72	1	MH	3		0,00%	Too few domains
						20,42%	Average
						0,76	Variance coefficient

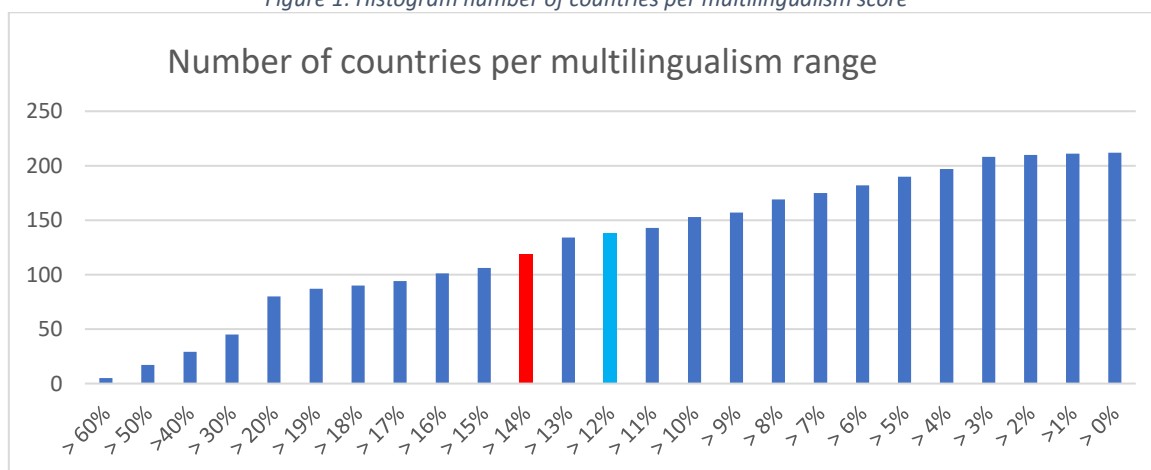
Notes:

- The ccTLD .us having a limited use compared to the size of the web in the US, the measurement of the country is noted below and used in some of the stats.
- It is an existing practice that some ccTLD are “hacked”, e.g. sold to foreign interests, they are marked in the table and excluded from some of the statistics.

The winners are : Monaco, Moldova, Kuwait, Ukraine and Mauritania.
The losers are: Liberia, Zimbabwe and China.

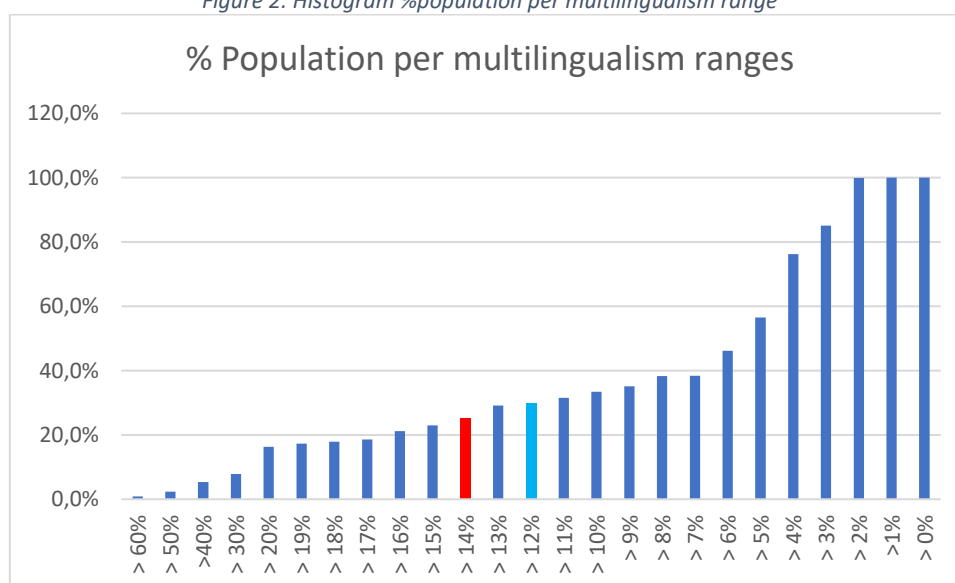
The histogram of number of countries per multilingualism range does not show anything special; but the histogram per percentage of population reveals that the low level of web multilingualism is more a question of populated countries than of number of countries. The countries above average represent 52% of the total number of countries but only 24% of the total L1+L2 speaker's population. This is the reflect of the fact that many most populated countries are in the bottom of the table (below 7%): China, South Africa, Bangladesh, Nigeria, Kenya, India, UK, Japan, Pakistan, USA, Indonesia, Philippines and Brazil¹.

Figure 1: Histogram number of countries per multilingualism score



While the weighted average for ccTLDs is higher than the global web one (14,39% versus 12,12%), the dispersion is huge with values from around 2% to more than 70%. At first glance, the large English-speaking countries, most African English-speaking countries and some large Asian countries show low figures, while Arabic countries and many European ones show high figures.

Figure 2: Histogram %population per multilingualism range



¹ This group of countries represents together 58% of the L1+L2 speakers in the world.

The analysis by region confirms the diagnostic of the strength of Europe in terms of multilingualism and the poor performance in terms of multilingualism of North America, to be imputed to the US (4.8%) in spite of Canada above average performance (15.4%).

Table 2: Web multilingualism per continent

REGION	#WEB	#Hreflang	MULTIL%
North America	32 046 995	811655	6.33%
Europe	29 749 927	2479658	20.84%
Asia	17 568 560	670162	9.54%
South America	3 263 826	126465	9.69%
Africa	1 059 767	35323	8.33%
Oceania	753	19	6.31%

Notes:

- This table is made directly from the split offered by DataProvider.com which is not ccTLD based but country based.
- North America includes Mexico.

In Europe, the country with below average performance are, first, alone in the bottom of the ranking the UK (5%), then Ireland (8,4%); Gibraltar (11,4%), Germany (13,5%), Norway (13,6%), Denmark (13,6%), all other countries are above the ccTLD average of 14,4%. The highest multilingual performance, behind the mentioned winners (Monaco, Moldova and Ukraine) are Luxembourg, Latvia, Georgia, Andorra, Greece and Estonia.

In Latin America, Brazil is singled out with a very low performance, close to that of the USA (7%) followed by Argentina (8,3%), Chile (10,2%), Paraguay (11%), Mexico (11.9%)) and Peru (12%). In the other side of the ranking, Cuba is the leader, followed by Puerto Rico, Panama, Aruba and Bahamas, which make an exception as most English-speaking Caribbean islands, except British Virgin island are below average.

In Asia, the variance is high, with Tajikistan, Myanmar, Azerbaijan, Cambodia, Hong Kong, Macao, Uzbekistan and Viet Nam in the high part of the ranking, above 30%, and China (2,7%) a huge loser, followed by South Korea, Bangladesh, India, Japan and Pakistan, all below 6%.

The Arabic countries display a stable high ranking with all countries between 20% and 64%.

Finally, African countries show a large dispersion with most English-speaking ones in the very low part of the ranking while Sudan, Mozambique, Madagascar are in the top part.

The split made directly from the collected data and with different categories is worth analyzing and confirm a clear diagnostic: English-speaking country are, from far, the less multilingual, with few exceptions (Puerto Rico, Bahamas, Mauritius, Malta, Belize, Seychelles, Cameroon, Burundi, British Virgin island and Canada) most marked by multilingualism.

Table 3: Web multilingualism per region

REGION (*)	#WEB	MULTIL%
Europe	18 987 454	17.8%
North America	28 119 411	5,2%
Latin American & Caribbean	2 703 407	8,4%
Africa	1 055 493	7,9%
Oceania	1 350 548	4,3%
Arabic countries	417 094	43,3%
Asia	4 424 336	8,9%
English Official Language	33 291 720	5,1%

Notes:

- This table is built from ccTLD figures and hacked domains are filtered out. Some errors in regional assignment to countries were detected and corrected in this new version.
- Oceania includes Australia and New Zealand

ANALYSIS BY LANGUAGE

Table 4: Web multilingualism per language

LANGUAGE	#WEB	HREFLANG	MULTIL
ALL	95 612 368	4 186 176	10,95%
Basque	8 072	2051	63,52%
Ukrainian	221 043	51875	58,67%
Latvian Macro	43 428	9348	53,81%
Catalan	82 966	17272	52,05%
Estonian Macro	57 245	11653	50,89%
Greek	231 074	46771	50,60%
Kazakh	9 285	1693	45,58%
Lao	1 109	201	45,31%
Macedonian	10 607	1661	39,15%
Lithuanian	91 380	14143	38,69%
Sindhi	9 350	1315	35,16%
Slovene	86 936	12069	34,71%
Bulgarian	130 100	17816	34,24%
Azerbaijani Macro	21 617	2834	32,78%
Tajik	1 613	207	32,08%
Arabic Macro	332 071	41924	31,56%
Welsh	12 642	1579	31,23%
Croatian	86 665	10035	28,95%
Yiddish	400	46	28,75%
Irish	2 510	267	26,59%
Bosnian	119 615	12673	26,49%
Amharic	1 210	125	25,83%
Finnish	258 164	25358	24,56%
Dutch	1 976 285	186686	23,62%

Romanian	278 962	26071	23,36%
Italian	2 092 445	191189	22,84%
Mongolian	9 289	828	22,28%
Burmese	3 016	265	21,97%
Serbo-Croatian Macro	11 821	1037	21,93%
Belarusian	1 477	129	21,83%
Hungarian	268 370	23263	21,67%
French	3 616 611	306840	21,21%
Pashto	584	46	19,69%
Vietnamese	561 515	42307	18,84%
Spanish	4 012 087	299332	18,65%
Galician	27 669	2034	18,38%
Hebrew	181 342	13313	18,35%
Khmer	6 641	475	17,88%
Slovak	243 400	17339	17,81%
Urdu	5 110	353	17,27%
Turkish	1 118 202	76972	17,21%
Swahili Macro	2 428	163	16,78%
Persian Macro	341 046	21488	15,75%
Polish	1 037 982	64913	15,63%
Uzbek Macro	13 334	824	15,45%
Norwegian	236 437	14052	14,86%
Thai	391 757	23210	14,81%
German, Standard	7 193 295	422076	14,67%
Punjabi, Eastern	1 237	72	14,55%
Yoruba	269	15	13,94%
Kurdish Macro	856	47	13,73%
Swedish	567 574	30331	13,36%
Czech	747 855	39936	13,35%
Danish	406 386	20832	12,82%
Kannada	2 459	125	12,71%
Telugu	3 682	179	12,15%
Hindi	93 415	4535	12,14%
Frisian	10 453	486	11,62%
Gujarati	2 668	120	11,24%
Portuguese	2 858 885	126227	11,04%
Russian	2 670 930	117623	11,01%
AVERAGE			10,95%
Tamil	7 033	270	9,60%
Tamir	7 033	270	9,60%
Maltese	6 593	247	9,37%
Malay Macro	43 569	1611	9,24%
Malayalam	3 308	122	9,22%
English	52 456 229	1677816	8,00%
Indonesian	1 220 050	38737	7,94%
Xhosa	979	27	6,89%

Hausa	3 237	87	6,72%
Afrikaans	20 039	537	6,70%
Esperanto	1 757	44	6,26%
Somali	2 622	65	6,20%
Haitian	1 856	46	6,20%
Nepali Macro	19 233	463	6,02%
Scots	1 408	33	5,86%
Hm	1 365	28	5,13%
Luxemburgish	31 987	655	5,12%
Sunda	6 099	106	4,34%
Marathi	33 902	585	4,31%
Japanese	3 341 482	55568	4,16%
Shona	2 648	42	3,97%
Korean	596 650	8582	3,60%
Javanese	4 853	68	3,50%
Corsican	3 353	46	3,43%
Zulu	2 935	31	2,64%
Samoan	1 061	11	2,59%
Chinese Macro	4 969 986	41400	2,08%
Malagasy	4 225	30	1,78%

The winners are Basque, Ukrainian, Latvian, Catalan, Estonian and Greek.

The losers are Malagasy², Chinese, Samoan and Zulu.

The same phenomenon occurs that in the previous analysis: the number of languages above the average is much higher than the language below, however, amongst the languages with low level of web multilingualism, there is those with highest population of connected speakers : Chinese, English, Japanese, Indonesian and Korean.

Luxemburgish is an interesting matter. Luxemburg, as a country, is in the side of the winners, in position 6, with a very high multilingualism rate in the web; however, the sites using the national language, Luxemburgish, seems to be hardly multilingual. Could be a paradox, but there is no contradiction here. It is a kind of unusual situation probably reflecting some cultural reality in this country where the population of resident foreigners is close to national's³.

Portuguese and Russia are very close to the average, the lowest European languages. The case of Portuguese is also an interesting one. Portugal have a high level of web multilingualism (36,5%), as well as Mozambique (22%); what makes the figure so low for Portuguese language is the combination of the fact that Brazil hold more than 90% of Portuguese speakers connected to the Internet while showing a very low level of multilingualism (7%).

² However, Madagascar is on the high multilingualism side ! Web sites in the island are fairly multilingual except those in Malagasy which remain heavily monolingual.

³ According to <https://luxembourg.public.lu/fr/societe-et-culture/population/demographie.html>, 47,6% of the resident population is composed of foreigners, without counting the 33% of boundary workers who spend their day working in Luxemburg and return home for the night in one of neighbor country (France, Belgium or Germany).

ANALYSIS FROM OTHER CONTEXTS

gTLDs

Table 5: Web multilingualism per gTLD

gTLD	#WEB	HREFLANG	MULTIL
ALL	54 606 539	2 374 632	10,87%
Eus	5 870	2463	104,90%
Cymru	1 854	486	65,53%
Gal	2 645	636	60,11%
Cat	34 400	7097	51,58%
Krd	341	42	30,79%
Eu	551 298	67261	30,50%
Wales	2 743	309	28,16%
Corsica	830	84	25,30%
Alsace	616	62	25,16%
Bzh	4 870	348	17,86%
Shop	550 109	38629	17,56%
Frl	1 288	86	16,69%
Irish	428	25	14,60%
Lat	13 071	742	14,19%
Asia	31 928	1685	13,19%
Art	84 054	4172	12,41%
Studio	41 766	1951	11,68%
App	124 475	5607	11,26%
Com	43 485 229	1925106	11,07%
World	37 798	1670	11,05%
Lol	25 231	1100	10,90%
AVERAGE			10,87%
One	31 585	1343	10,63%
Store	299 481	12656	10,56%
Pro	155 367	6533	10,51%
Tech	90 129	3588	9,95%
Org	2 939 856	115087	9,79%
Net	2 312 718	87163	9,42%
Click	55 041	2053	9,32%
Biz	165 922	6046	9,11%
Cloud	36 497	1269	8,69%
Online	427 165	14466	8,47%
Edu	19 060	644	8,45%
Africa	12 467	402	8,06%
Info	503 361	16173	8,03%
Today	14 223	432	7,59%
Live	82 478	2479	7,51%
Club	86 602	2596	7,49%
Cyou	20 159	601	7,45%
Space	69 864	1943	6,95%
Dev	98 565	2694	6,83%
Website	45 073	1198	6,64%

Site	363 696	9169	6,30%
Fun	62 012	1380	5,56%
Life	75 180	1543	5,13%
Vip	90 916	1856	5,10%
Сайт	150	3	5,00%
Scot	3 273	63	4,81%
Icu	36 269	694	4,78%
Хyz	393 086	7514	4,78%
Buzz	42 313	631	3,73%
Top	596 174	7925	3,32%
Рф	179 060	2275	3,18%
Blog	118 938	1219	2,56%
Link	59 296	544	2,29%
Shiksha	177	1	1,41%
Mobi	70 033	346	1,24%
Онлайн	427	2	1,17%
Mil	195		0,00%
みんな	106	0	0,00%

The winners are some of the linguistic domains, first of all, the absolute winner, with almost 100% (all websites are at least bilingual⁴), the linguistic domain of the Basque language (.eus), followed by the linguistic domains of Galician, Scottish, Catalan and Kurdish.

Most of the linguistic gTLDs are in the top part of the ranking, yet a subset is in the bottom part : みんな (for Japanese), Онлайн, for Cyrillic alphabets, Shiksha, for Hindi, Рф, for Russian, Scot and сайт, also for Cyrillic alphabets.

As for the geographic domains, excellent performance of the .eu, good performance of .lat, .asia and .world, and weak performance of .africa.

For the rest, .shop is the winner, logically as ecommerce sites are propense to multilingualism. Fair performance for .art and .studio, the rest is below average.

CITIE'S WEBSITES

DataProvider.com is able, with an algorithm, to associate websites with particular cities; this has nothing to do with domains for cities which do exist and in all probability are included in the list of websites corresponding. When applicable, in the right column are the results for the corresponding city domain.

Table 6: Web multilingualism per city

CITY	#WEB	#HrefLang	MULTIL%	#WEB	#HrefLang	MULTIL%
Montreal	26 396	9135	86,52%			
Kiev	22 947	7901	86,08%			
Barcelona	50 243	14847	73,88%	711	170	59,77%
Brussels	15 749	4389	69,67%	1 519	480	79,00%

⁴ It is important to record that the absolute values are very approximative and obviously should not be higher than 100%. It makes sense that almost all websites are bilingual, with Spanish, at difference to the .cat which is higher than 50% but not reaching 100%. Almost of Catalan people speaks Catalan; this is not the case for Basque people and that could explain why the bilingualism is almost systematic.

Hamburg	177 825	45811	64,40%	2 155	108	12,53%
Lisbon	20 031	4510	56,29%			
Athens	23 742	4931	51,92%			
Milano	64 452	11472	44,50%			
Zurich	44 051	7729	43,86%			
Amsterdam	54 167	8916	41,15%	1 810	228	31,49%
Paris	142 970	21767	38,06%	3 950	502	31,77%
Madrid	95 843	14561	37,98%	502	45	22,41%
Roma	78 228	10952	35,00%			
Budapest	63 333	8327	32,87%			
Stockholm	32 310	4050	31,34%	66	0	0
Wien	101 019	12571	31,11%	2 791	164	14,69%
Melbourne	21 845	2709	31,00%	697	10	3,59%
Warszawa	95 641	11124	29,08%			
Roubaix	97 459	10613	27,22%			
München	131 874	14176	26,87%			
Oslo	22 068	2232	25,29%			
Berlin	251 270	23715	23,60%	8 107	755	23,28%
Marseille	16 568	1475	22,26%			
Koln	72 319	5834	20,17%	2 337	75	8,02%
Miami	58 550	4643	19,82%	1 436	52	9,05%
Kuala Lumpur	25 321	1957	19,32%			
Mexico	50 493	3494	17,30%			
Singapore	70 818	4533	16,00%			
Buenos Aires	15 619	984	15,75%			
Rio de Janeiro	16 062	943	14,68%	179	15	20,95%
Sao Paulo	73 374	4019	13,69%			
Toronto	36 152	1950	13,48%			
London	234 785	10455	11,13%	5 893	114	4,84%
San Francisco	73 795	2534	8,58%			
東京都 (Tokyo)	256 891	8309	8,09%	24 662	623	6,32%
Moscow	144 052	4659	8,09%	4 251	114	6,70%
Houston	75 575	2443	8,08%			
New York	164 834	4846	7,35%			
Los Angeles	68 524	1944	7,09%			

The winners are Montreal, Kiev, Barcelona, Brussels, Hamburg and Lisbon. The losers are US cities, Tokyo and Moscow. The rate of multilingualism is higher than in other category, probably, as a consequence of the touristic vocation of cities. The correspondence of results between both results is the rule with some exceptions.

OTHER GROUPS

Table 7: Web multilingualism other groups

GROUP	#WEB	HREFLANG	MULTIL
ccTLD	31 498 261	1855957	14,73%

gTLD	49 571 411	2156431	10,88%
New gTLD	5 521 510	197001	8,92%
sTLD	200 022	12191	15,24%
IDN	87 199 527	4226881	12,12%
Business	60 839 086	3 802 629	15,63%
No Business	21 532 154	314 490	3,65%
B2C	26 589 834	1 800 759	16,93%
B2B	8 971 766	653 977	18,22%
Online store	9 761 547	985 760	25,25%
Not online store	77 437 980	3 241 321	10,46%

Those figures tell some evidences which are sustained by data:

- ccTLD are more propense to multilingualism than gTLDs (except, as we have seen before, for the linguistic gTLDs).
- IDN (internationalized domain names) are in the average.
- Business oriented websites have 4 times more level of multilingualism than websites not conceived for business activity. This is still more true for B2C type of websites and still more for B2B.
- Online stores are at the very high level of multilingualism; all studies converge in that evidence, consumers want to use their mother tongue to buy online.

The last figure is a gift for people who are advocating in favor of multilingualism in the digital world and it serves for a strong conclusion of that study: **the more multilingual a site, the more economic impact**. Economic impact (EI) is one of the parameters stored in the DataProvider.com database, in association with each website; it is a value from 0 to 100 determined by an algorithm after careful analysis of the website.

Table 8: Web multilingualism per economic impact range

	#WEB	HREFLANG	MULTIL
EI > 90	110	20	45,45%
80 < EI < 90	1 111	340	76,51%
70 < EI < 80	7 270	1 883	64,75%
60 < EI < 70	40 895	9 268	56,66%
50 < EI < 60	165 685	26 362	39,78%
40 < EI < 50	903 613	135 309	37,44%
30 < EI < 40	4 026 546	425 074	26,39%
20 < EI < 30	17 424 223	1 200 260	17,22%
10 < EI < 20	51 084 470	1 876 194	9,18%
EI < 10	4 927 618	43 585	2,21%
TOTAL/AVG	78 581 541	3 718 295	11,83%