

Metadata

Version	1
Created by	Translators Without Borders and University College London
Uploaded on	Jan-19
Notes and caveats	<p>Language usage in the Democratic Republic of the Congo is extremely complex and suitable data is limited. The issues noted here will be reviewed and resolved where possible in future version of this dataset. For now this dataset should be considered only indicative of the language context of DRC. All data confidence values are low.</p> <p>Percentage values for languages include all speakers in that area, regardless if they have previously been included in other languages. No primary language specific data is currently available.</p> <p>This data draws on that of CAID, a government department focused on collection and analysis of development sector data in DRC, and is subject to a number of issues and inconsistencies present in that source data.</p> <p>The methodology used to collect the source data is unclear and appears to differ across administrative areas.</p> <p>Different naming conventions have been used for languages and dialects, and so it is likely that some appear more than once in the data under different names. For example, Nyabwisha is a dialect of Kinyarwanda, but is listed here as a separate language. This has also made matching languages to ISO 639-3 codes challenging, and so many language fields do not currently have a code attributed to them.</p> <p>The prefix 'ki' which is used to denote a language has been removed for uniformity. This may create some confusion when referring to languages that typically retain the prefix, such as Kinyarwanda or Kiswahili.</p> <p>Admin 0 and 1 data has not been included due to limitations of source data.</p> <p>Language shares below 0.1% of an area population have been discarded. As a result, languages may not be present at higher administrative levels.</p> <p>All decimal values have been rounded to a maximum of 3 decimal places.</p>
Copyright and terms of use	<p>© Translators without Borders 2019.</p> <p>Data is available under a Attribution-NonCommercial-ShareAlike 4.0 International license (CC BY-NC-SA 4.0). You are free to share and adapt the data subject to requirements for attribution and non-commercial use. Any derivative work must be distributed under the same license as the original. Full terms at https://creativecommons.org/licenses/by-nc-sa/4.0/.</p>

Source name	Last accessed	Source URL
UN OCHA administrative boundaries via HDX (version at 13 September 2019)	10-Jan-20	https://data.humdata.org/dataset/drc-administrative-boundaries-levels-0-2
UN OCHA 2019 population estimates via HDX (version at 12 Sep 2019)	10-Jan-20	https://data.humdata.org/dataset/democratic-republic-of-the-congo-administrative-level-0-2-population-statistics
CAID territory language data	10-Jan-20	https://www.caid.cd/index.php/donnees-par-province-administrative/
Open Language Archives Community (OLAC) data on ISO 639-3 codes and alternative language names	10-Jan-20	http://www.language-archives.org/index.html

Field heading	HXL code	Description	Sources
admin2_name	#adm2+name	Name of administrative unit	UN OCHA administrative boundaries
admin2_pcode	#adm2+code	Unique geographic identification code	UN OCHA administrative boundaries
admin1_name	#adm1+name	Name of administrative unit	UN OCHA administrative boundaries
admin1_pcode	#adm1+code	Unique geographic identification code	UN OCHA administrative boundaries
admin0_name	#country+name	Name of administrative unit	UN OCHA administrative boundaries
admin0_pcode	#country+code	Unique geographic identification code	UN OCHA administrative boundaries
main_language	#indicator+lang+main	Name of reported language with highest percentage of speakers in area	Calculated field
main_language_share	#indicator+lang+main+pct	Percentage value of language with highest percentage of speakers in area	Calculated field
..
Language fields	#indicator+lang+pct+iso639-3_ABC	Percentage of population in area who can speak this language	CAID territory language data
Other languages	#indicator+lang+pct+iso639-3_unknown	Percentage of population in area who can speak a non-named (Other) language	CAID territory language data
..
pop_total	#population+total	Number of total population	UN OCHA 2019 population estimates
language_data_yn	#meta+data+bool	Binary field describing whether language data is available in this record	-
data_confidence	#meta+confidence	Assessment of data quality (high/medium/low)	-
notes	#meta+text	Notes and clarifications on data	-

Data confidence descriptions	Examples
High	Data from well regarded government census or other specialist organisation.
Medium	Data from quality source but without clear methodology, or inverse.
Low	Data known to be erroneous, or based on non-statistical or imprecise data.
Factors increasing confidence	Data is corroborated. Data can be directly applied to administrative areas without aggregation or manipulation.
Factors decreasing confidence	Data is disputed. Data requires transformation that may affect accuracy, for example combining admin 2 data to form admin 1 data.

Language	ISO code 639-3	Alternative names and spellings
Alur	alz	Lur, Aloro, Alua, Alulu, Luri, Dho Alur, Jo Alur, Aluru
Bakutshu	kvs	Kutsu
Balika	Unknown	
Balobo	lie	Likila
Bandaka	ndk	Ndaka, Indaaka, Ndaaka
Banunu	Unknown	
Bangala	bvg	Ngala, Mangala, Bamangala, Lingala
Bango	bbm	Babango, Mobango
Batabwa	tap	
Bati	Unknown	
Baloi	biz	Loi, Baato Baloi, Rebu, Bangi-Loi, Bangi-Loi, Ngiri
Bemba	bem	Chibemba, Ichibemba, Wemba, Chiwemba, Bemba
Bembe	bmb	Ibembe, Beembe, Ebembe, Kibembe
Benge	Unknown	
Benza	lgz	Gendza, Ligendza
Bila	bip	Bira, Kibila, Forest Bira, Ebila, Western Bila
Boa	bww	Bua, Boua, Libua, Libwali, Libenge, Kibua, Kibwa, Bua, li-, Balj, Bango
Boba	bws	
Bobayi	Unknown	
Boma	bws	
Bombo	bml	Bomboli, Bombongo
Budja	bja	Budza, Embudja, Limbudza
Budu	buu	Ebudu, Kibudu, Bodo
Bunda	mck	Mbunda
Dinga	diz	Ding, Di, Dzing
Djonga	Unknown	
Ekonda	Unknown	
Eleku	Unknown	
Fuliru	flr	Furiiru, Kifuliiru, Fulero
Geyna	gey	Enya, Genya
Gilima	Unknown	
Gobu	gox	Gubu, Ngobo, Ngobu, Gabou, Gabu
Havu	hav	Kihavu, Haavu
Hema	nix	Hema-Sud, Southern Hema, Kihema, Congo Nyoro, Nyoro, Runyoro
Hemba	hem	Kihemba, Emba, Kiemba, Luba-Hemba, Eastern Luba, Luba
Hesoo	soc	So, Eso, Gesogo, Soa
Humbu	Unknown	
Hunde	hke	Ihunde, Kobi, Rukobi
Kalanga	kck	Kalanga, Chikalanga, Tjikalanga, Kanana, Sekalaña, Kalana, Western Shona, Bakaa, Makalaka, Wakalanga, Kalaka, Sekalaka, Kalanja
Kango	kty and kzy	
Kanyoka	kny	Kanyok, Kanioka
Kaonde	kqn	Chikaonde, Chikahonde, Kawonde, Luba Kaonde, Kaonde, ci-, Kahonde
Kebu-tuu	ndp	Ndo, Ke'bu, Kebutu
Kere	mdj	Mangbetu, Nemangbetu, Mangbettu, Mambetto, Amangbetu, Kingbetu
Kilese	Unknown	

Kiongo	Unknown	
Kitwa	Unknown	
Koka	Unknown	
Kongo	kon	Kikongo
Kuba	kxx	Likuba
Kumu	kmw	Komo, Kuumu
Kusu	ksv	Kutsu
Kwala	kwc	Likwala, Likouala, Kwala, Ekwala
Kwange	Unknown	
Kwese	kws	Pindi, Kikwese, Ukwese
Lalia	lal	
Lande	Unknown	
Langa	Unknown	
Langbasi	lna	Langbashe, Langbase, Langbasi, Langwasi, Langbwasse
Leboale	Unknown	
Lega	Unknown	
Lemfu	Unknown	
Lendu	led	Bbadha, Bbaledha, Kilendu, Baletha, Batha, Balendru, Bale, Hema-Nord, Kihema-Nord, Baledha, Badha
Libinza	liz	
Liboko	bkp	Iboko, Boko
Likumbe	Unknown	
Lingala	lin	
Lobala	loq	
Logo	log	
Lokele	khy	Kele
Lokonda	Unknown	
Lonkundo	Unknown	
Londengese	dez	Dengese, Ileo
Longando	nxd	Ngando.
Lontomba	nto	Ntomba, Ntumba, Luntumba
Lothsua	Unknown	
Lotwa	Unknown	
Loyembe	sop	Songe, Sonye, Lusonge, Kalebwe
Luba	lua	CiLuba, Luba-Lulua, Tshiluba, Western Luba, Luva, Luba, Bashingu
Luba lubangule	lub	Luba-Katanga, Kiluba, Luba-Shaba, Luba
Lulua	lua	CiLuba, Luba-Lulua, Tshiluba, Western Luba, Luva, Luba, Bashingu
Lunda	lun	
Lungu	mgr	Mambwe-Lungu
Mabale	mmz	Mabaale, Lomabaale, Mbali
Makutu	Unknown	
Manianga	Unknown	
Manga	mba	Mba, Bamanga, Kimanga, Manga, Kimbanga, Mbayi
Manyanga	Unknown	
Mashi	shr	Shi
Mbala	mdp	Gimbala, Rumbala
Mbanza	Unknown	

Mbanza Fula	Unknown	
Mbati	mdn	
Mbole	mdq	Lombole
Mboma	bws	
Mbuba	mxh	
Mituku	zmq	Kinya-Mituku, Metoko, Mjtũkũ
Moko	Unknown	
Mongo	lol	Lomongo, Lonkundu, Nkundo, Mongo-Nkundu
Mongando	Unknown	
Mongbandi	Unknown	
Mono	Unknown	
Mpama	Unknown	
Mpee	Unknown	
Nalengwe	Unknown	
Nande	nmb	
Nanyembo	Unknown	
Ndembo	Unknown	
Ndibu	Unknown	
Ndo-tuki	Unknown	
Ngando	nxd	Ngando.
Ngbaka	nga	Mbaka, Ngbaka-Minangende, Ngbaka-Manza, Ali, 'Bofi, Ngbaka Minangende, Ngbaka Gbaya, Ngbaka-Minagende, Klor/Ngeh, Kriang
Ngbaka Mabo	nbm	Ngbaka-Ma'bo, Bwaka, Ngbaka Limba, Mbaka, Mbacca, Bouaka, Nbwaka, Gbaka, Ma'bo, Gbala, Mbwaka, Ngbaka-ma'bo, Ngbaka-Mabo, Mbaca
Ngbandi	Unknown	
Ngbundu	Unknown	
Ngbungbu	Unknown	
Ngelema	Unknown	
Ngengele	Unknown	
Ngoli	Unknown	
Ngombe	ngc	Ngombe, Lingombe
Ngwaka Minagende	Unknown	
Ngwandi	Unknown	
Nkundo	Unknown	
Nkutshu	Unknown	
Ntandu	Unknown	
Nunu	Unknown	
Nyali	nlj	
Nyabwisha	kin	Kinyarwanda
Nyanga	nyj	Kinyanga, Inyanga, Nyanga
Nyarwanda	kin	Kinyarwanda
Nyindu	Unknown	
Nzakara	Unknown	
Ohendo	Unknown	
Okela	Unknown	
Okutsu	Unknown	
Pakombe	Unknown	

Pazande	Unknown	
Pende	Unknown	
Piri	bhy	Ebhele, Kipere, Ipere, Pere, Peri, Piri, Pili, Bili, Kipili, Perj
Popoyi	Unknown	
Portuguese	Unknown	
Probe	Unknown	
Pygmy	Unknown	
Sakata	Unknown	
Sanga	Unknown	
Sango	Unknown	
Songe	sop	Songe, Sonye, Lusonge, Kalebwe
Songola	Unknown	
Suku	Unknown	
Swahili	swc	
Tabwa	Unknown	
Teke	Unknown	
Tembo	tbt	Kitembo
Tete (South)	Unknown	
Tetela	tll	Otetela, Sungu, Tetela, Sungu
Togba	Unknown	
Topoke	Unknown	
Tshibindi	Unknown	
Tshikete	Unknown	
Tshikwamputu	Unknown	
Tsihilualua	Unknown	
Tshiluba	Unknown	
Tshokwe	Unknown	
Tua	Unknown	
Tungwa	Unknown	
Urund	Unknown	
Vungu	Unknown	
Waria	Unknown	
Yaelima	Unknown	
Yaka	yaf	
Yazi	Unknown	
Yanzi	Unknown	
Yogo	mdm	Mayogo, Kiyogo
Yombe	yom	
Zande	zne	
Zela	Unknown	
Zimba	zmb	
Zola	Unknown	