Language	ISO code	Line Count	fastText	family
hindi	hi	11235	t	indo-european
albanian	sq	4532	t	indo-european
portuguese	pt	49244	t	indo-european
latin	la	23682	t	indo-european
spanish	es	80784	t	indo-european
french	fr	63340	t	indo-european
gujarati	gu	2828	t	indo-european
russian	ru	29896	t	indo-european
hebrew	he	2951	t	afro-asiatic
maltese	mt	5801	t	afro-asiatic
romanian	ro	78964	t	indo-european
serbo-croatian	sh	30793	t	indo-european
swedish	sv	5967	t	indo-european
slovene	sl	3263	t	indo-european
ukrainian	uk	10861	t	indo-european
sanskrit	sa	5357	t	indo-european
catalan	ca	21865	t	indo-european
latvian	lv	6829	t	indo-european
italian	it	113851	t	indo-european
slovak	sk	5056	t	indo-european
telugu	te	2228	t	dravidian
faroese	fo	4709	t	indo-european
irish	ga	14720	t	indo-european
bulgarian	bg	5780	t	indo-european
lithuanian	lt	3901	t	indo-european
macedonian	mk	22315	t	indo-european
sicilian	scn	1548	t	indo-european
dutch	nl	44072	t	indo-european
arabic	ar	12752	t	afro-asiatic

danish	da	8068	t	indo-european
belarusian	be	2695	t	indo-european
polish	pl	60076	t	indo-european
classical syriac	syc	2197	f	afro-asiatic
greek	el	16760	t	indo-european
occitan	oc	2681	t	indo-european
pashto	ps	844	t	indo-european
asturian	ast	3635	t	indo-european
czech	cs	24421	t	indo-european
zulu	zu	1634	t	niger-congo
welsh	cy	5928	t	indo-european
german	de	58545	t	indo-european
icelandic	is	12702	t	indo-european
galician	gl	10367	t	indo-european
xhosa	xh	1646	t	niger-congo
scottish	gd	6677	t	indo-european