Fumbling in Babel:
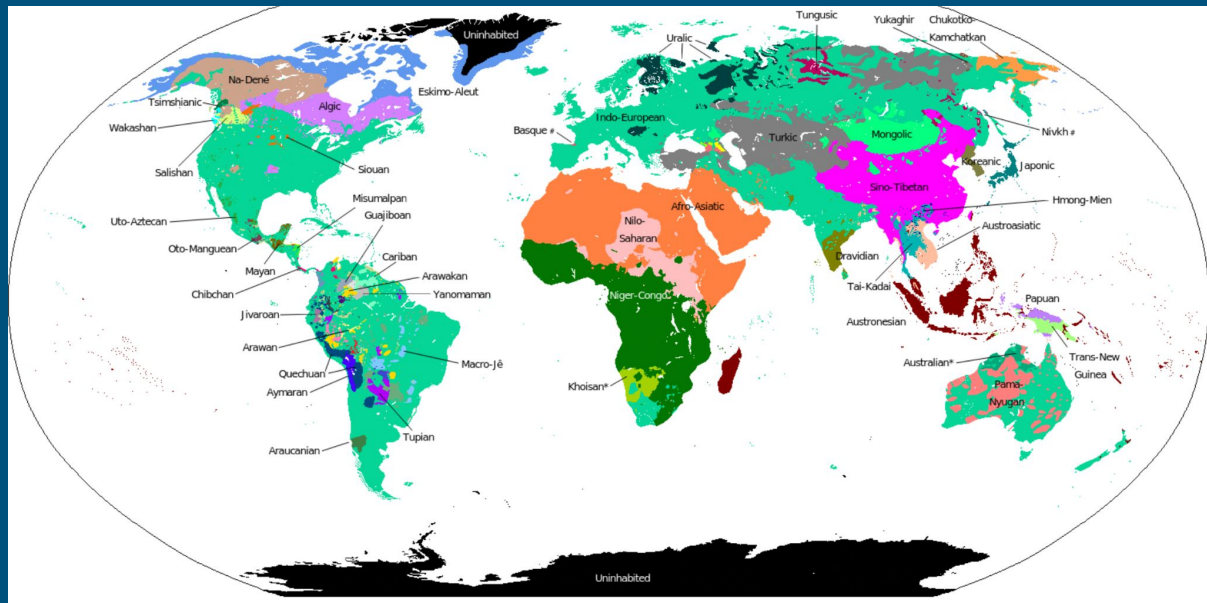# An Investigation into ChatGPT's Language Identification Ability

Wei-Rui Chen, Ife Adebara, Khai Duy Doan, Qisheng Liao, Muhammad Abdul-Mageed

Deep Learning & Natural Language Processing Group , UBC

Department of Natural Language Processing & Department of Machine Learning, MBZUAI

How Many Languages Does ChatGPT know?

Language Identification

A task to determine the language that a piece of text is written in (Jauhiainen et al. 2018)

Babel-670

670 languages

24 language families

Afro-Asiatic, Austro-Asiatic, Austronesian, Aymaran, Chibchan, Creole, constructed language, Dravidian, Indo-European, Japonic, Kartvelian, Khoe-Kwadi, Koreanic, Kra-Dai, Language Isolate, Mongolic, Niger Congo, Nilo-Saharan, Quechuan, Sino-Tibetan, Tucanoan, Tupian, Turkic, and Uralic

30 scripts

Arabic, Armenian, Bengali, Burmese, Coptic, Cyrillic, Devanagari, Ethiopic, Georgian, Greek, Gujarati, Gurmukhi, Hangul, Hans , Hant, Hebrew, Japanese, Kannada, Khmer, Laoo, Latin, Malayalam, Odia, Ol Chiki, Sinhala, Tamil, Telugu, Thai, Tibetan, Vai
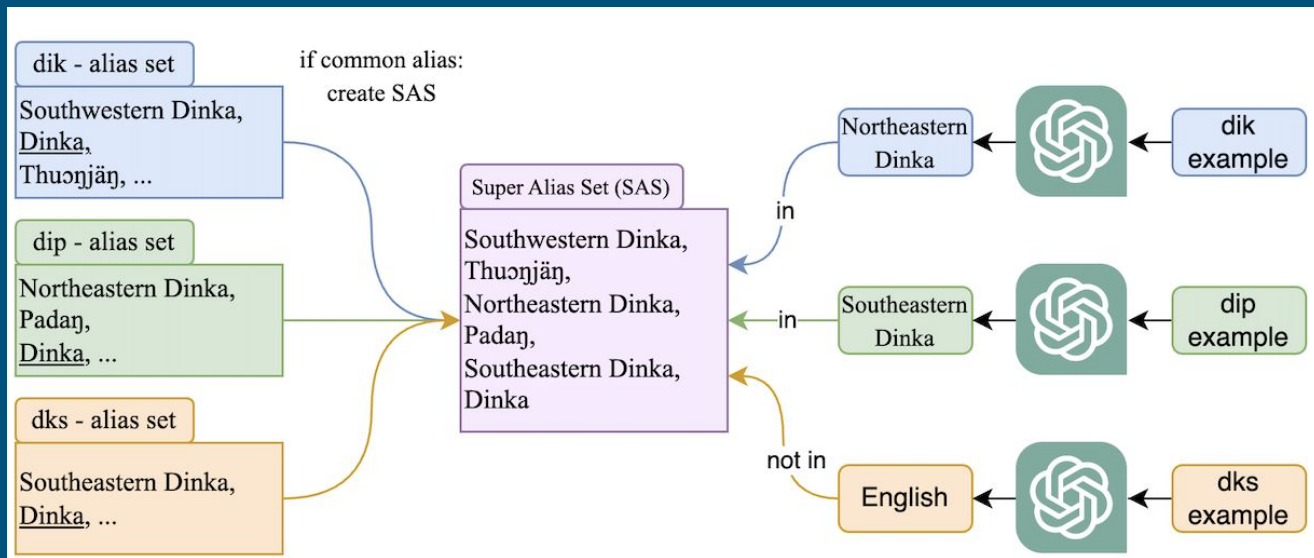
Methodology
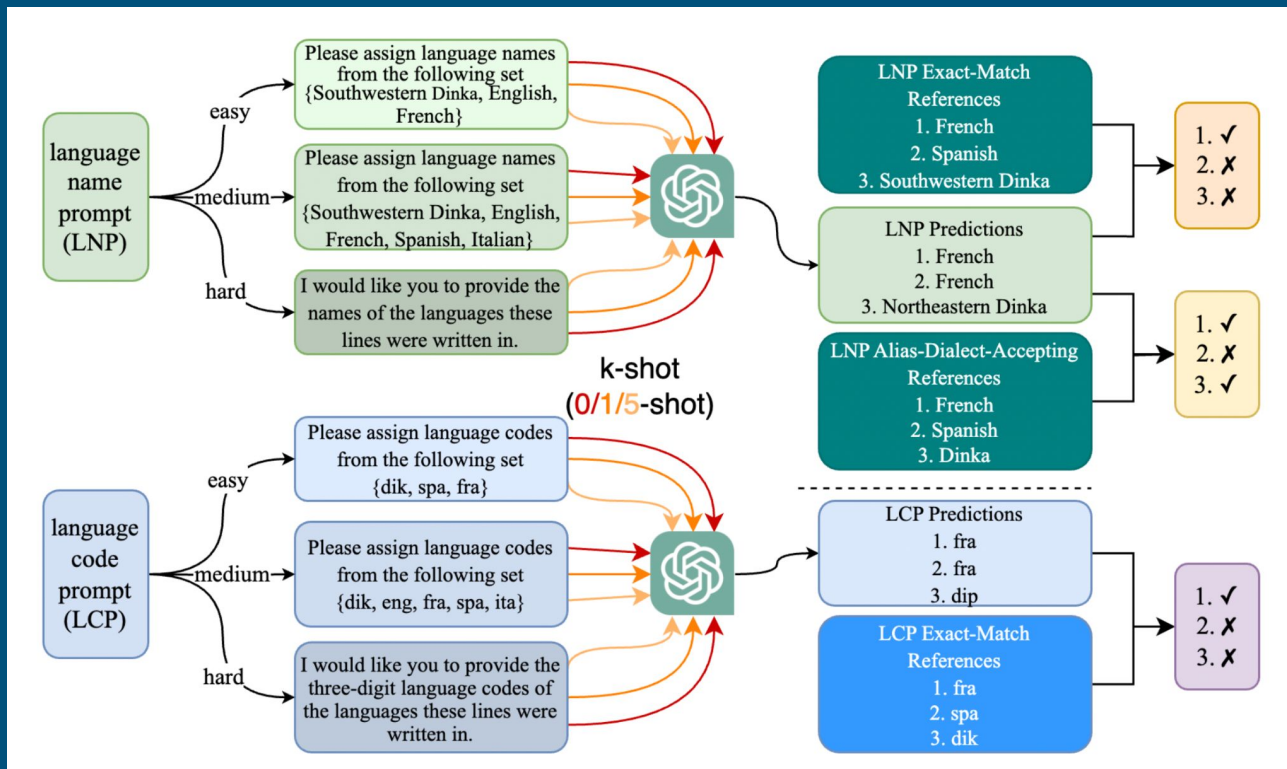
Alias-Dialect-Accepting Evaluation Method

A language can have multiple names

Spanish (spa) -> Spanish, Español

Results: LNP vs. LCP

Level	#	GPT-3.5-turbo-0613						GPT-4-0613					
		LNP (exact)		LNP (ADA)		LCP (exact)		LNP (exact)		LNP (ADA)		LCP (exact)	
		ACC	F_1	ACC	F_1	ACC	F_1	ACC	F_1	ACC	F_1	ACC	F_1
easy	0	24.86	28.28	24.90	28.14	14.63	17.63	65.82	65.36	65.97	65.92	47.16	46.29
	1	28.13	31.18	28.23	31.12	27.32	27.51	65.19	65.70	65.29	66.16	45.88	45.96
	5	33.82	34.46	33.96	34.97	27.37	27.17	68.29	68.05	68.42	68.69	47.95	47.31
medium	0	19.81	21.09	20.26	21.62	10.27	12.30	48.66	45.39	49.08	46.28	37.79	34.24
	1	25.30	22.63	26.32	23.58	22.88	19.60	48.02	45.57	48.40	46.22	38.19	34.64
	5	26.22	23.02	26.94	24.08	22.76	19.64	50.49	47.63	50.89	48.32	39.39	35.62
hard	0	12.70	12.17	17.39	16.36	2.47	3.82	20.02	17.80	28.32	24.16	21.47	18.93
	1	16.05	13.81	23.82	19.36	16.66	14.34	20.40	18.26	28.58	24.76	22.10	19.72
	5	17.12	14.65	25.36	20.25	17.71	15.09	20.94	18.97	28.79	25.31	22.52	20.26

Results: Difficulty Levels

Level	#	GPT-3.5-turbo-0613						GPT-4-0613					
		LNP (exact)		LNP (ADA)		LCP (exact)		LNP (exact)		LNP (ADA)		LCP (exact)	
		ACC	F_1	ACC	F_1	ACC	F_1	ACC	F_1	ACC	F_1	ACC	F_1
easy	0	24.86	28.28	24.90	28.14	14.63	17.63	65.82	65.36	65.97	65.92	47.16	46.29
	1	28.13	31.18	28.23	31.12	27.32	27.51	65.19	65.70	65.29	66.16	45.88	45.96
	5	33.82	34.46	33.96	34.97	27.37	27.17	68.29	68.05	68.42	68.69	47.95	47.31
medium	0	19.81	21.09	20.26	21.62	10.27	12.30	48.66	45.39	49.08	46.28	37.79	34.24
	1	25.30	22.63	26.32	23.58	22.88	19.60	48.02	45.57	48.40	46.22	38.19	34.64
	5	26.22	23.02	26.94	24.08	22.76	19.64	50.49	47.63	50.89	48.32	39.39	35.62
hard	0	12.70	12.17	17.39	16.36	2.47	3.82	20.02	17.80	28.32	24.16	21.47	18.93
	1	16.05	13.81	23.82	19.36	16.66	14.34	20.40	18.26	28.58	24.76	22.10	19.72
	5	17.12	14.65	25.36	20.25	17.71	15.09	20.94	18.97	28.79	25.31	22.52	20.26

Results: GPT-3.5 vs. GPT-4

Level	#	GPT-3.5-turbo-0613						GPT-4-0613					
		LNP (exact)		LNP (ADA)		LCP (exact)		LNP (exact)		LNP (ADA)		LCP (exact)	
		ACC	F_1	ACC	F_1	ACC	F_1	ACC	F_1	ACC	F_1	ACC	F_1
easy	0	24.86	28.28	24.90	28.14	14.63	17.63	65.82	65.36	65.97	65.92	47.16	46.29
	1	28.13	31.18	28.23	31.12	27.32	27.51	65.19	65.70	65.29	66.16	45.88	45.96
	5	33.82	34.46	33.96	34.97	27.37	27.17	68.29	68.05	68.42	68.69	47.95	47.31
medium	0	19.81	21.09	20.26	21.62	10.27	12.30	48.66	45.39	49.08	46.28	37.79	34.24
	1	25.30	22.63	26.32	23.58	22.88	19.60	48.02	45.57	48.40	46.22	38.19	34.64
	5	26.22	23.02	26.94	24.08	22.76	19.64	50.49	47.63	50.89	48.32	39.39	35.62
hard	0	12.70	12.17	17.39	16.36	2.47	3.82	20.02	17.80	28.32	24.16	21.47	18.93
	1	16.05	13.81	23.82	19.36	16.66	14.34	20.40	18.26	28.58	24.76	22.10	19.72
	5	17.12	14.65	25.36	20.25	17.71	15.09	20.94	18.97	28.79	25.31	22.52	20.26

F1 score for each language
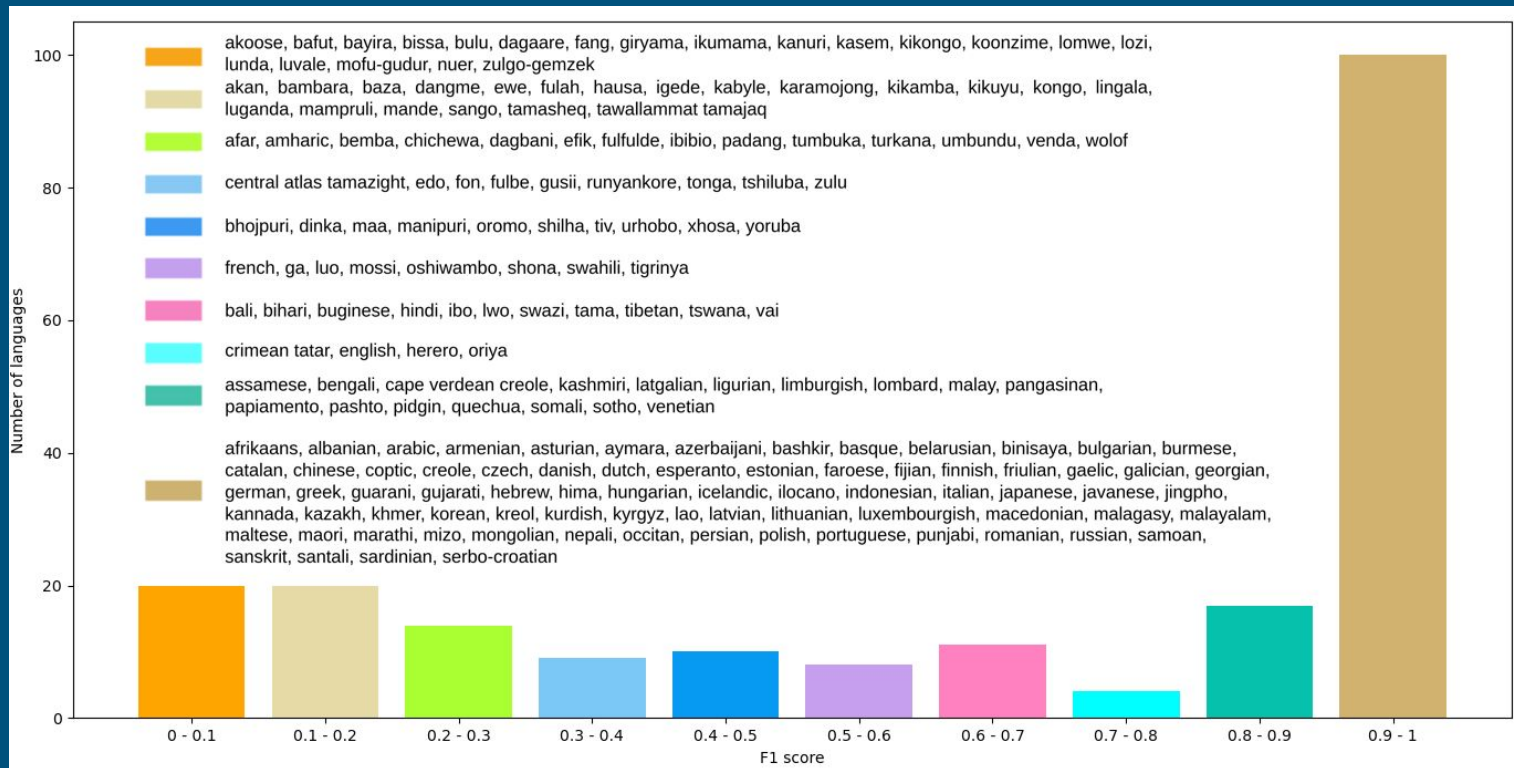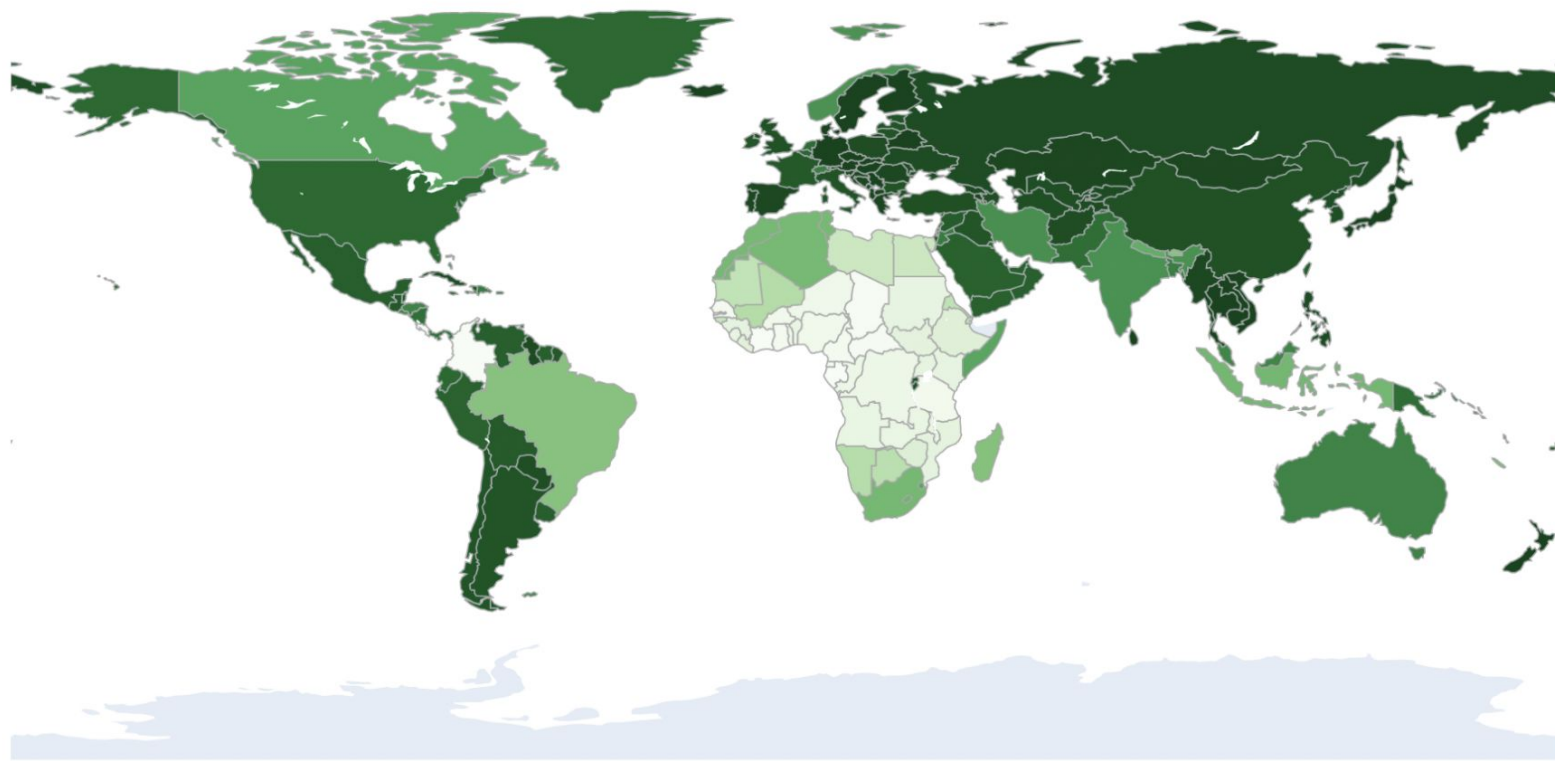
Thank you!
weirui.chen@ubc.ca