Table 36: Language families and language codes supported by Qwen3 in Belebele Benchmark

Language family	# Langs	Language code (ISO 639-3_ISO 15924)
Indo-European	40	por_Latn, deu_Latn, tgk_Cyrl, ces_Latn, nob_Latn, dan_Latn, snd_Arab, spa_Latn, isl_Latn, slv_Latn, eng_Latn, ory_Orya, hrv_Latn, ell_Grek, ukr_Cyrl, pan_Guru, srp_Cyrl, npi_Deva, mkd_Cyrl, guj_Gujr, nld_Latn, swe_Latn, hin_Deva, rus_Cyrl, asm_Beng, cat_Latn, als_Latn, sin_Sinh, urd_Arab, mar_Deva, lit_Latn, slk_Latn, ita_Latn, pol_Latn, bul_Cyrl, afr_Latn, ron_Latn, fra_Latn, ben_Beng, hye_Armn
Sino-Tibetan	3	zho_Hans, mya_Mymr, zho_Hant
Afro-Asiatic	8	heb_Hebr, apc_Arab, acm_Arab, ary_Arab, ars_Arab, arb_Arab, mlt_Latn, erz_Arab
Austronesian	7	ilo_Latn, ceb_Latn, tgl_Latn, sun_Latn, jav_Latn, war_Latn, ind_Latn
Dravidian	4	mal_Mlym, kan_Knda, tel_Telu, tam_Taml
Turkic	4	kaz_Cyrl, azj_Latn, tur_Latn, uzn_Latn
Tai-Kadai	2	tha_Thai, lao_Laoo
Uralic	3	fin_Latn, hun_Latn, est_Latn
Austroasiatic	2	vie_Latn, khm_Khmr
Other	7	eus_Latn, kor_Hang, hat_Latn, swh_Latn, kea_Latn, jpn_Jpan, kat_Geor

Table 37: **Comparison of Belebele Benchmark performance between Qwen3 and other baseline models.** Scores are highlighted with the highest in **bold** and the second-best underlined.

Model	Indo-European	Sino-Tibetan	Afro-Asiatic	Austronesian	Dravidian	Turkic	Tai-Kadai	Uralic	Austroasiatic	Other
Gemma-3-27B-IT	89.2	86.3	**85.9**	84.1	83.5	86.8	81.0	91.0	86.5	**87.0**
Qwen2.5-32B-Instruct	85.5	82.3	80.4	70.6	67.8	80.8	74.5	87.0	79.0	72.6
QwQ-32B	86.1	83.7	81.9	71.3	69.3	80.3	77.0	88.0	83.0	74.0
Qwen3-32B (Thinking)	**90.7**	**89.7**	84.8	**86.7**	84.5	**89.3**	83.5	**91.3**	**88.0**	83.1
Qwen3-32B (Non-thinking)	89.1	88.0	82.3	83.7	84.0	85.0	**85.0**	88.7	**88.0**	81.3
Gemma-3-12B-IT	85.8	83.3	**83.4**	79.3	79.0	82.8	77.5	89.0	**83.0**	81.6
Qwen2.5-14B-Instruct	82.7	78.9	80.4	69.1	66.2	74.2	72.2	83.9	77.9	70.4
Qwen3-14B (Thinking)	**88.6**	**87.3**	82.4	**82.4**	**81.0**	**83.8**	83.5	**91.0**	82.5	**81.7**
Qwen3-14B (Non-thinking)	87.4	82.7	80.1	80.7	78.0	81.8	80.5	87.7	81.5	77.0
Gemma-3-4B-IT	71.8	72.0	63.5	61.7	64.8	64.0	61.5	70.7	71.0	62.6
Qwen2.5-3B-Instruct	58.0	62.3	57.2	47.9	36.9	45.1	49.8	50.6	56.8	48.4
Qwen3-4B (Thinking)	**82.2**	**77.7**	**74.1**	**73.0**	**74.3**	**76.3**	68.5	**83.0**	**74.5**	**67.9**
Qwen3-4B (Non-thinking)	76.0	77.0	65.6	65.6	65.5	64.0	60.5	74.0	74.0	61.0
Gemma-3-1B-IT	36.5	36.0	30.0	29.1	28.8	27.3	28.0	32.7	33.0	30.9
Qwen2.5-1.5B-Instruct	41.5	43.0	39.6	34.8	28.6	29.7	39.4	33.8	42.0	36.0
Qwen3-1.7B (Thinking)	**69.7**	**66.0**	**59.4**	**58.6**	**52.8**	**57.8**	**53.5**	**70.3**	**63.5**	**53.4**
Qwen3-1.7B (Non-thinking)	58.8	62.7	50.8	53.0	43.3	48.0	46.0	54.3	54.0	43.9