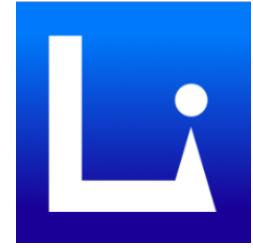


Investigating Language Preference of Multilingual RAG Systems

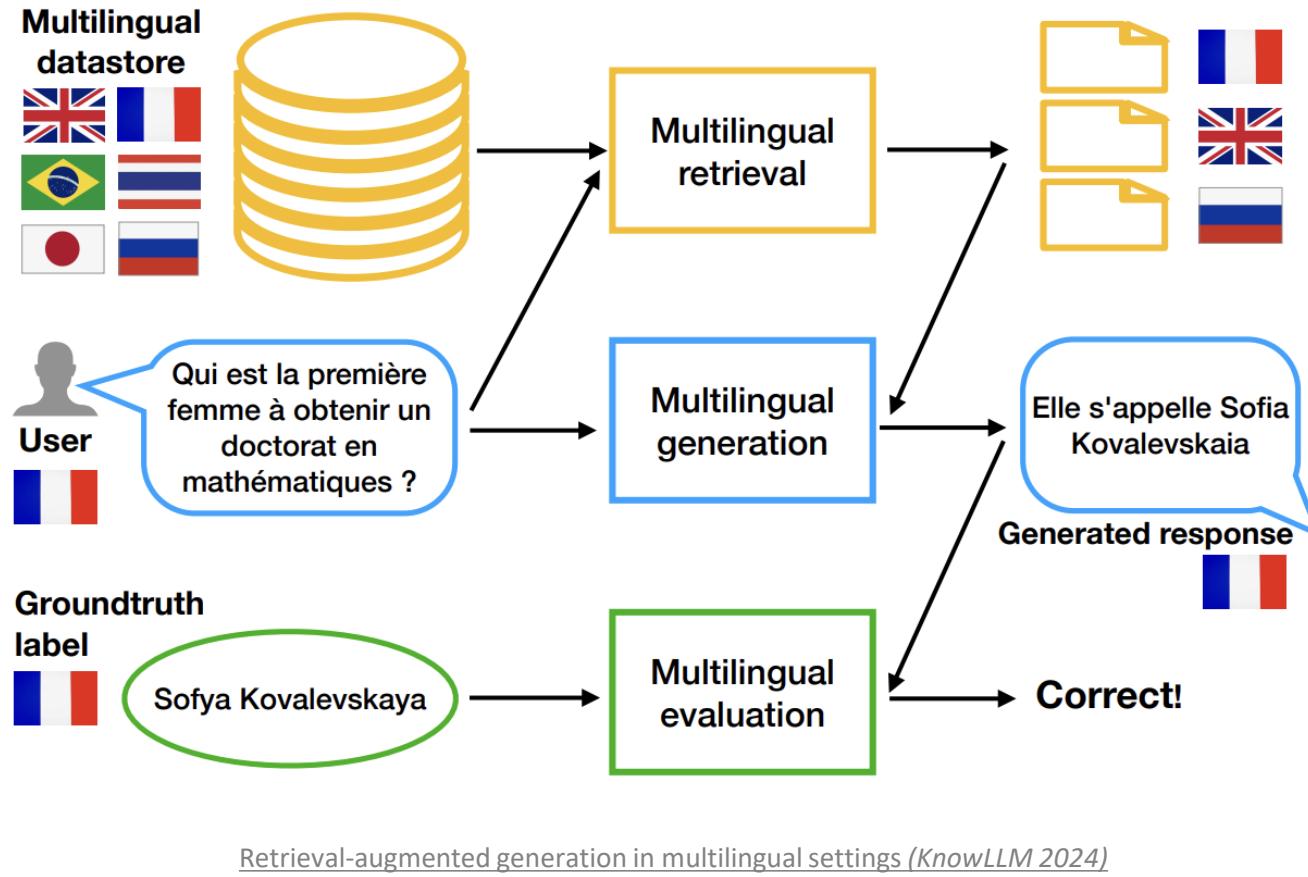
The 63rd Annual Meeting of the Association for Computational Linguistics
(ACL 2025)

Jeonghyun Park

LILAB
Department of
Artificial Intelligence
Chung-Ang University



Multilingual RAG (mRAG) System

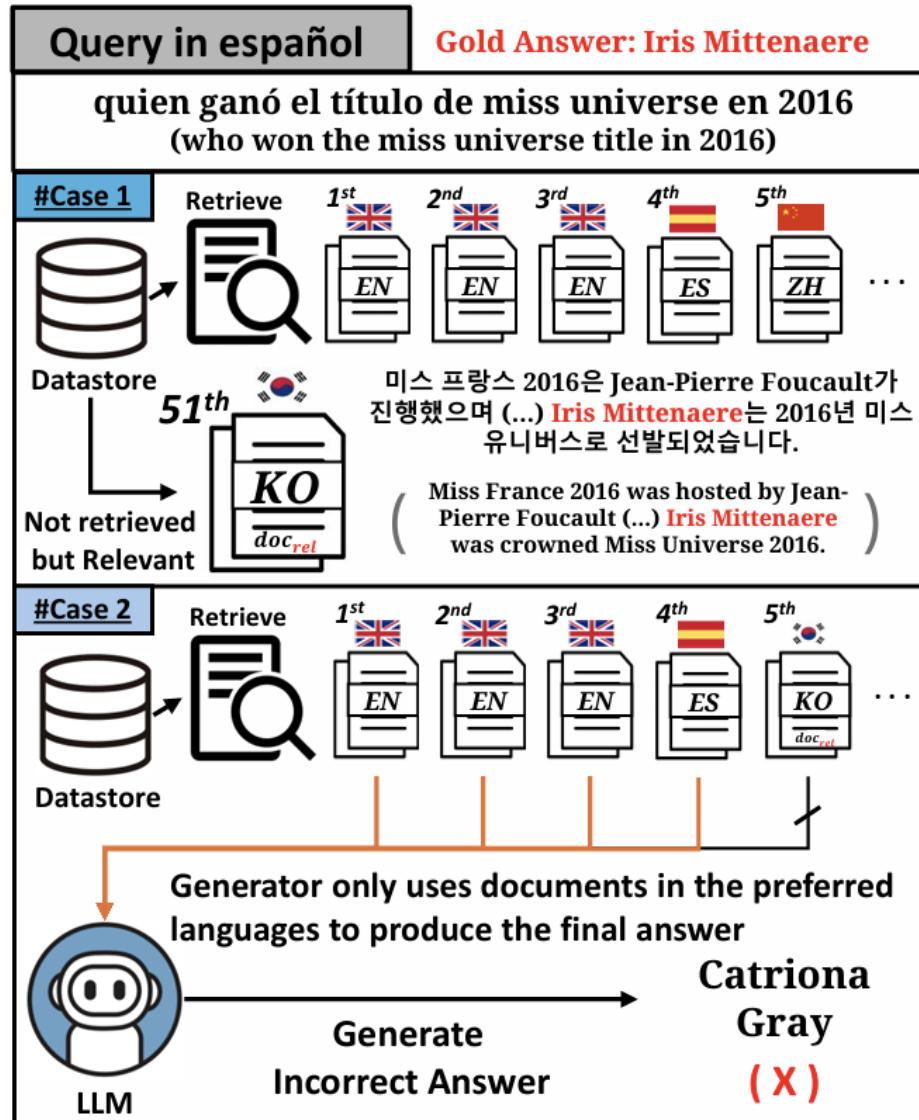


Retrieval-augmented Generation (RAG) pipeline that:

- For *queries and knowledge (data)* stores spanning multiple languages, retrieves *relevant context* across those languages
- *Multilingual Retrieval-Augmented Generation (mRAG) systems enhance language models by integrating external multilingual information to produce context-aware responses.*

- Generates a grounded answer based on retrieved contexts in the *user's language*

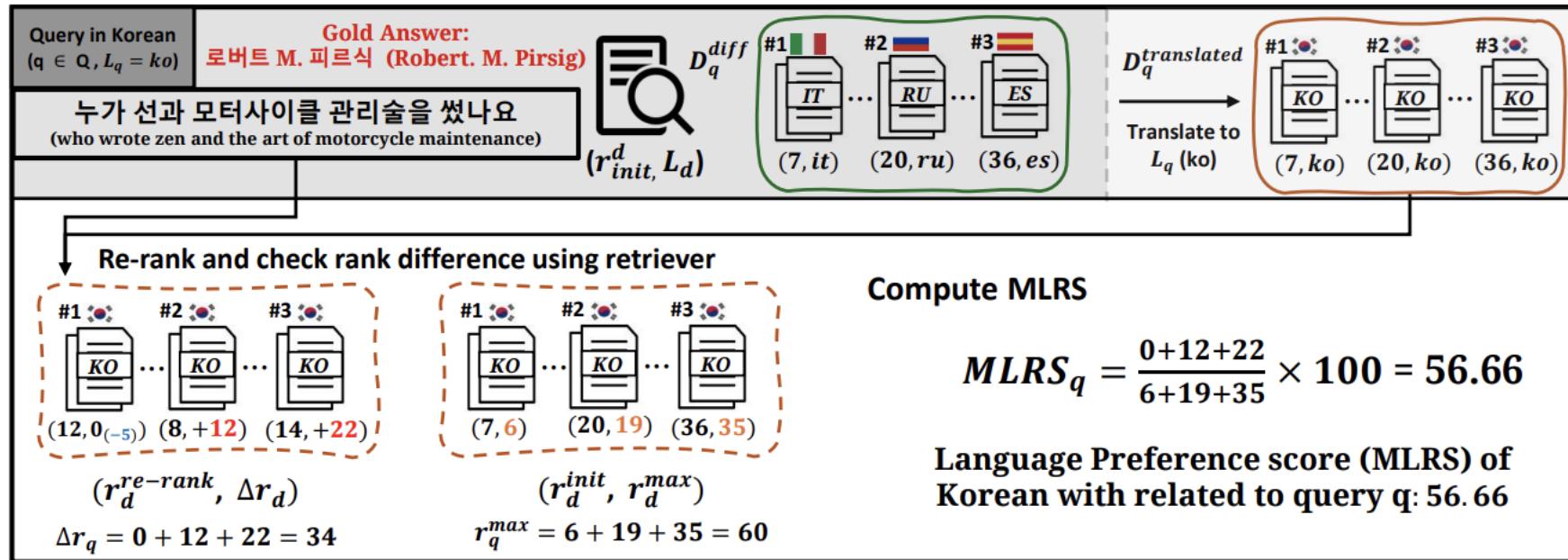
Multilingual RAG – Motivations (*Challenges*)



Failure cases of mRAG system

- However, because mRAG systems favor certain languages, the retriever often pulls in irrelevant contexts and this language preference present in both the retriever and the generator ultimately degrades the system's generation quality.
- (Case 1) Retriever may prioritize **high-resource or query-language** documents—at the expense of truly relevant information in low-resource language.
- (Case 2) Even if relevant documents are retrieved, the generator might favor passages in the **query language** or **Latin scripts**, ignoring essential evidence in lower-resource languages and resulting in inaccurate outputs

Language Preference – Retriever Level (Proposed Metric)



- Proposed metric, *MultiLingualRankShift (MLRS)* that measures **language preference at the retriever level**
 - Compute **ranking improvements** when documents in the *non-query language* are translated *into query language*
 - To normalize the score, MLRS also reflect **maximum (possible) rank improvement** for each document

Language Preference – Retriever Level (Results)

Query Lang.	Encoder	$L_q = L_d$	$L_q \neq L_d$							
			en	ko	zh	fr	ja	it	pt	es
en	bge-m3	56.03	–	33.02 (-23.01)	33.10 (-22.93)	36.61 (-19.42)	33.36 (-22.67)	35.89 (-20.14)	35.86 (-20.17)	<u>36.62</u> (-19.41)
	p-mMiniLM	56.85	–	34.34 (-22.51)	34.61 (-22.24)	<u>38.17</u> (-18.68)	34.52 (-22.33)	37.15 (-19.70)	36.73 (-20.12)	37.96 (-18.89)
	p-mMpNet	57.49	–	34.45 (-23.04)	34.27 (-23.22)	<u>37.94</u> (-19.55)	34.67 (-22.82)	37.34 (-20.15)	37.02 (-20.47)	37.90 (-19.59)
ko	bge-m3	<u>41.15</u>	43.49 (+2.34)	–	34.42 (-6.73)	36.42 (-4.73)	37.18 (-3.97)	35.72 (-5.43)	35.30 (-5.85)	35.93 (-5.22)
	p-mMiniLM	<u>42.95</u>	44.62 (+1.67)	–	36.04 (-6.91)	37.08 (-5.87)	38.47 (-4.48)	36.07 (-6.88)	36.18 (-6.77)	36.45 (-6.50)
	p-mMpNet	<u>42.53</u>	44.98 (+2.45)	–	35.85 (-6.68)	37.20 (-5.33)	39.01 (-3.52)	36.21 (-6.32)	35.65 (-6.88)	36.34 (-6.19)
zh	bge-m3	<u>44.98</u>	45.26 (+0.28)	34.52 (-10.46)	–	36.34 (-8.64)	36.05 (-8.93)	35.86 (-9.12)	35.73 (-9.25)	36.45 (-8.53)
	p-mMiniLM	<u>46.18</u>	<u>45.39</u> (-0.79)	35.46 (-10.72)	–	36.98 (-9.20)	36.77 (-9.41)	36.38 (-9.80)	36.05 (-10.13)	36.85 (-9.33)
	p-mMpNet	<u>46.27</u>	<u>45.41</u> (-0.86)	35.21 (-11.06)	–	36.87 (-9.40)	36.71 (-9.56)	36.28 (-9.99)	35.94 (-10.33)	36.78 (-9.49)
fr	bge-m3	<u>43.18</u>	47.23 (+4.05)	33.29 (-9.89)	33.58 (-9.60)	–	34.07 (-9.11)	36.70 (-6.48)	36.30 (-6.88)	37.25 (-5.93)
	p-mMiniLM	<u>44.09</u>	48.15 (+4.06)	34.54 (-9.55)	34.52 (-9.57)	–	34.83 (-9.26)	37.65 (-6.44)	37.05 (-7.04)	38.03 (-6.06)
	p-mMpNet	<u>43.96</u>	48.14 (+4.18)	34.25 (-9.71)	34.37 (-9.59)	–	34.61 (-9.35)	37.59 (-6.37)	36.93 (-7.03)	38.01 (-5.95)
ja	bge-m3	<u>45.03</u>	45.18 (+0.15)	35.45 (-9.58)	34.86 (-10.17)	36.71 (-8.32)	–	36.11 (-8.92)	35.88 (-9.15)	36.56 (-8.47)
	p-mMiniLM	<u>45.80</u>	<u>45.54</u> (-0.26)	35.90 (-9.90)	35.57 (-10.23)	37.18 (-8.62)	–	36.53 (-9.27)	36.25 (-9.55)	36.91 (-8.89)
	p-mMpNet	<u>45.67</u>	<u>45.39</u> (-0.28)	35.73 (-9.94)	35.30 (-10.37)	36.94 (-8.73)	–	36.24 (-9.43)	35.98 (-9.69)	36.62 (-9.05)
it	bge-m3	<u>41.06</u>	46.63 (+5.57)	33.30 (-7.76)	33.47 (-7.59)	37.92 (-3.14)	33.86 (-7.20)	–	36.44 (-4.62)	37.68 (-3.38)
	p-mMiniLM	<u>42.11</u>	47.69 (+5.58)	34.57 (-7.54)	34.59 (-7.52)	39.07 (-3.04)	34.80 (-7.31)	–	37.55 (-4.56)	38.83 (-3.28)
	p-mMpNet	<u>41.98</u>	47.59 (+5.61)	34.48 (-7.50)	34.68 (-7.30)	38.94 (-3.04)	34.67 (-7.31)	–	37.27 (-4.71)	38.67 (-3.31)
pt	bge-m3	<u>39.19</u>	46.64 (+7.45)	33.37 (-5.82)	33.46 (-5.73)	37.83 (-1.36)	34.02 (-5.17)	37.13 (-2.06)	–	38.61 (-0.58)
	p-mMiniLM	<u>40.17</u>	47.75 (+7.58)	34.67 (-5.50)	34.91 (-5.26)	39.02 (-1.15)	35.03 (-5.14)	38.25 (-1.92)	–	39.68 (-0.49)
	p-mMpNet	<u>39.91</u>	47.30 (+7.39)	34.68 (-5.23)	34.50 (-5.41)	38.70 (-1.21)	34.72 (-5.19)	38.01 (-1.90)	–	39.35 (-0.56)
es	bge-m3	<u>40.76</u>	46.93 (+6.17)	33.36 (-7.40)	33.42 (-7.34)	37.73 (-3.03)	33.87 (-6.89)	37.22 (-3.54)	36.88 (-3.88)	–
	p-mMiniLM	<u>41.81</u>	47.90 (+6.09)	34.63 (-7.18)	34.52 (-7.29)	38.86 (-2.95)	34.76 (-7.05)	38.33 (-3.48)	37.84 (-3.97)	–
	p-mMpNet	<u>41.33</u>	47.34 (+6.01)	34.39 (-6.94)	34.19 (-7.14)	38.34 (-2.99)	34.39 (-6.94)	37.73 (-3.60)	37.25 (-4.08)	–

Language Preference – Retriever Level (Results)

Query Lang.	Encoder	$L_q = L_d$	$L_q \neq L_d$							
		en	ko	zh	fr	ja	it	pt	es	
en	bge-m3	56.03	–	33.02 (-23.01)	33.10 (-22.93)	36.61 (-19.42)	33.36 (-22.67)	35.89 (-20.14)	35.86 (-20.17)	<u>36.62</u> (-19.41)
	p-mMiniLM	56.85	–	34.34 (-22.51)	34.61 (-22.24)	<u>38.17</u> (-18.68)	34.52 (-22.33)	37.15 (-19.70)	36.73 (-20.12)	37.96 (-18.89)
	p-mMpNet	57.49	–	34.45 (-23.04)	34.27 (-23.22)	<u>37.94</u> (-19.55)	34.67 (-22.82)	37.34 (-20.15)	37.02 (-20.47)	37.90 (-19.59)
ko	bge-m3	<u>41.15</u>	43.49 (+2.34)	–	34.42 (-6.73)	36.42 (-4.73)	37.18 (-3.97)	35.72 (-5.43)	35.30 (-5.85)	35.93 (-5.22)
	p-mMiniLM	<u>42.95</u>	44.62 (+1.67)	–	36.04 (-6.91)	37.08 (-5.87)	38.47 (-4.48)	36.07 (-6.88)	36.18 (-6.77)	36.45 (-6.50)
	p-mMpNet	<u>42.53</u>	44.98 (+2.45)	–	35.85 (-6.68)	37.20 (-5.33)	39.01 (-3.52)	36.21 (-6.32)	35.65 (-6.88)	36.34 (-6.19)
zh	bge-m3	<u>44.98</u>	45.26 (+0.28)	34.52 (-10.46)	–	36.34 (-8.64)	36.05 (-8.93)	35.86 (-9.12)	35.73 (-9.25)	36.45 (-8.53)
	p-mMiniLM	46.18	<u>45.39</u> (-0.79)	35.46 (-10.72)	–	36.98 (-9.20)	36.77 (-9.41)	36.38 (-9.80)	36.05 (-10.13)	36.85 (-9.33)
	p-mMpNet	46.27	<u>45.41</u> (-0.86)	35.21 (-11.06)	–	36.87 (-9.40)	36.71 (-9.56)	36.28 (-9.99)	35.94 (-10.33)	36.78 (-9.49)
fr	bge-m3	<u>43.18</u>	47.23 (+4.05)	33.29 (-9.89)	33.58 (-9.60)	–	34.07 (-9.11)	36.70 (-6.48)	36.30 (-6.88)	37.25 (-5.93)
	p-mMiniLM	<u>44.09</u>	48.15 (+4.06)	34.54 (-9.55)	34.52 (-9.57)	–	34.83 (-9.26)	37.65 (-6.44)	37.05 (-7.04)	38.03 (-6.06)
	p-mMpNet	<u>43.96</u>	48.14 (+4.18)	34.25 (-9.71)	34.37 (-9.59)	–	34.61 (-9.35)	37.59 (-6.37)	36.93 (-7.03)	38.01 (-5.95)
ja	bge-m3	<u>45.03</u>	45.18 (+0.15)	35.45 (-9.58)	34.86 (-10.17)	36.71 (-8.32)	–	36.11 (-8.92)	35.88 (-9.15)	36.56 (-8.47)
	p-mMiniLM	45.80	<u>45.54</u> (-0.26)	35.90 (-9.90)	35.57 (-10.23)	37.18 (-8.62)	–	36.53 (-9.27)	36.25 (-9.55)	36.91 (-8.89)
	p-mMpNet	45.67	<u>45.39</u> (-0.28)	35.73 (-9.94)	35.30 (-10.37)	36.94 (-8.73)	–	36.24 (-9.43)	35.98 (-9.69)	36.62 (-9.05)
it	bge-m3	<u>41.06</u>	46.63 (+5.57)	33.30 (-7.76)	33.47 (-7.59)	37.92 (-3.14)	33.86 (-7.20)	–	36.44 (-4.62)	37.68 (-3.38)
	p-mMiniLM	<u>42.11</u>	47.69 (+5.58)	34.57 (-7.54)	34.59 (-7.52)	39.07 (-3.04)	34.80 (-7.31)	–	37.55 (-4.56)	38.83 (-3.28)
	p-mMpNet	<u>41.98</u>	47.59 (+5.61)	34.48 (-7.50)	34.68 (-7.30)	38.94 (-3.04)	34.67 (-7.31)	–	37.27 (-4.71)	38.67 (-3.31)
pt	bge-m3	<u>39.19</u>	46.64 (+7.45)	33.37 (-5.82)	33.46 (-5.73)	37.83 (-1.36)	34.02 (-5.17)	37.13 (-2.06)	–	38.61 (-0.58)
	p-mMiniLM	<u>40.17</u>	47.75 (+7.58)	34.67 (-5.50)	34.91 (-5.26)	39.02 (-1.15)	35.03 (-5.14)	38.25 (-1.92)	–	39.68 (-0.49)
	p-mMpNet	<u>39.91</u>	47.30 (+7.39)	34.68 (-5.23)	34.50 (-5.41)	38.70 (-1.21)	34.72 (-5.19)	38.01 (-1.90)	–	39.35 (-0.56)
es	bge-m3	<u>40.76</u>	46.93 (+6.17)	33.36 (-7.40)	33.42 (-7.34)	37.73 (-3.03)	33.87 (-6.89)	37.22 (-3.54)	36.88 (-3.88)	–
	p-mMiniLM	<u>41.81</u>	47.90 (+6.09)	34.63 (-7.18)	34.52 (-7.29)	38.86 (-2.95)	34.76 (-7.05)	38.33 (-3.48)	37.84 (-3.97)	–
	p-mMpNet	<u>41.33</u>	47.34 (+6.01)	34.39 (-6.94)	34.19 (-7.14)	38.34 (-2.99)	34.39 (-6.94)	37.73 (-3.60)	37.25 (-4.08)	–

Language Preference – Retriever Level (Results)

Query Lang.	Encoder	$L_q = L_d$	$L_q \neq L_d$							
			en	ko	zh	fr	ja	it	pt	es
en	bge-m3	56.03	–	33.02 (-23.01)	33.10 (-22.93)	36.61 (-19.42)	33.36 (-22.67)	35.89 (-20.14)	35.86 (-20.17)	<u>36.62</u> (-19.41)
	p-mMiniLM	56.85	–	34.34 (-22.51)	34.61 (-22.24)	<u>38.17</u> (-18.68)	34.52 (-22.33)	37.15 (-19.70)	36.73 (-20.12)	37.96 (-18.89)
	p-mMpNet	57.49	–	34.45 (-23.04)	34.27 (-23.22)	<u>37.94</u> (-19.55)	34.67 (-22.82)	37.34 (-20.15)	37.02 (-20.47)	37.90 (-19.59)
ko	bge-m3	<u>41.15</u>	43.49 (+2.34)	–	34.42 (-6.73)	36.42 (-4.73)	37.18 (-3.97)	35.72 (-5.43)	35.30 (-5.85)	35.93 (-5.22)
	p-mMiniLM	<u>42.95</u>	44.62 (+1.67)	–	36.04 (-6.91)	37.08 (-5.87)	38.47 (-4.48)	36.07 (-6.88)	36.18 (-6.77)	36.45 (-6.50)
	p-mMpNet	<u>42.53</u>	44.98 (+2.45)	–	35.85 (-6.68)	37.20 (-5.33)	39.01 (-3.52)	36.21 (-6.32)	35.65 (-6.88)	36.34 (-6.19)
zh	bge-m3	<u>44.98</u>	45.26 (+0.28)	34.52 (-10.46)	–	36.34 (-8.64)	36.05 (-8.93)	35.86 (-9.12)	35.73 (-9.25)	36.45 (-8.53)
	p-mMiniLM	46.18	<u>45.39</u> (-0.79)	35.46 (-10.72)	–	36.98 (-9.20)	36.77 (-9.41)	36.38 (-9.80)	36.05 (-10.13)	36.85 (-9.33)
	p-mMpNet	46.27	<u>45.41</u> (-0.86)	35.21 (-11.06)	–	36.87 (-9.40)	36.71 (-9.56)	36.28 (-9.99)	35.94 (-10.33)	36.78 (-9.49)
fr	bge-m3	<u>43.18</u>	47.23 (+4.05)	33.29 (-9.89)	33.58 (-9.60)	–	34.07 (-9.11)	36.70 (-6.48)	36.30 (-6.88)	37.25 (-5.93)
	p-mMiniLM	<u>44.09</u>	48.15 (+4.06)	34.54 (-9.55)	34.52 (-9.57)	–	34.83 (-9.26)	37.65 (-6.44)	37.05 (-7.04)	38.03 (-6.06)
	p-mMpNet	<u>43.96</u>	48.14 (+4.18)	34.25 (-9.71)	34.37 (-9.59)	–	34.61 (-9.35)	37.59 (-6.37)	36.93 (-7.03)	38.01 (-5.95)
ja	bge-m3	<u>45.03</u>	45.18 (+0.15)	35.45 (-9.58)	34.86 (-10.17)	36.71 (-8.32)	–	36.11 (-8.92)	35.88 (-9.15)	36.56 (-8.47)
	p-mMiniLM	45.80	<u>45.54</u> (-0.26)	35.90 (-9.90)	35.57 (-10.23)	37.18 (-8.62)	–	36.53 (-9.27)	36.25 (-9.55)	36.91 (-8.89)
	p-mMpNet	45.67	<u>45.39</u> (-0.28)	35.73 (-9.94)	35.30 (-10.37)	36.94 (-8.73)	–	36.24 (-9.43)	35.98 (-9.69)	36.62 (-9.05)
it	bge-m3	<u>41.06</u>	46.63 (+5.57)	33.30 (-7.76)	33.47 (-7.59)	37.92 (-3.14)	33.86 (-7.20)	–	36.44 (-4.62)	37.68 (-3.38)
	p-mMiniLM	<u>42.11</u>	47.69 (+5.58)	34.57 (-7.54)	34.59 (-7.52)	39.07 (-3.04)	34.80 (-7.31)	–	37.55 (-4.56)	38.83 (-3.28)
	p-mMpNet	<u>41.98</u>	47.59 (+5.61)	34.48 (-7.50)	34.68 (-7.30)	38.94 (-3.04)	34.67 (-7.31)	–	37.27 (-4.71)	38.67 (-3.31)
pt	bge-m3	<u>39.19</u>	46.64 (+7.45)	33.37 (-5.82)	33.46 (-5.73)	37.83 (-1.36)	34.02 (-5.17)	37.13 (-2.06)	–	38.61 (-0.58)
	p-mMiniLM	<u>40.17</u>	47.75 (+7.58)	34.67 (-5.50)	34.91 (-5.26)	39.02 (-1.15)	35.03 (-5.14)	38.25 (-1.92)	–	39.68 (-0.49)
	p-mMpNet	<u>39.91</u>	47.30 (+7.39)	34.68 (-5.23)	34.50 (-5.41)	38.70 (-1.21)	34.72 (-5.19)	38.01 (-1.90)	–	39.35 (-0.56)
es	bge-m3	<u>40.76</u>	46.93 (+6.17)	33.36 (-7.40)	33.42 (-7.34)	37.73 (-3.03)	33.87 (-6.89)	37.22 (-3.54)	36.88 (-3.88)	–
	p-mMiniLM	<u>41.81</u>	47.90 (+6.09)	34.63 (-7.18)	34.52 (-7.29)	38.86 (-2.95)	34.76 (-7.05)	38.33 (-3.48)	37.84 (-3.97)	–
	p-mMpNet	<u>41.33</u>	47.34 (+6.01)	34.39 (-6.94)	34.19 (-7.14)	38.34 (-2.99)	34.39 (-6.94)	37.73 (-3.60)	37.25 (-4.08)	–

Language Preference – Retriever Level (Results)

Query Lang.	Encoder	$L_q = L_d$	$L_q \neq L_d$							
		en	ko	zh	fr	ja	it	pt	es	
en	bge-m3	56.03	–	33.02 (-23.01)	33.10 (-22.93)	36.61 (-19.42)	33.36 (-22.67)	35.89 (-20.14)	35.86 (-20.17)	<u>36.62</u> (-19.41)
	p-mMiniLM	56.85	–	34.34 (-22.51)	34.61 (-22.24)	38.17 (-18.68)	34.52 (-22.33)	37.15 (-19.70)	36.73 (-20.12)	37.96 (-18.89)
	p-mMpNet	57.49	–	34.45 (-23.04)	34.27 (-23.22)	37.94 (-19.55)	34.67 (-22.82)	37.34 (-20.15)	37.02 (-20.47)	37.90 (-19.59)
ko	bge-m3	<u>41.15</u>	43.49 (+2.34)	–	34.42 (-6.73)	36.42 (-4.73)	37.18 (-3.97)	35.72 (-5.43)	35.30 (-5.85)	35.93 (-5.22)
	p-mMiniLM	<u>42.95</u>	44.62 (+1.67)	–	36.04 (-6.91)	37.08 (-5.87)	38.47 (-4.48)	36.07 (-6.88)	36.18 (-6.77)	36.45 (-6.50)
	p-mMpNet	<u>42.53</u>	44.98 (+2.45)	–	35.85 (-6.68)	37.20 (-5.33)	39.01 (-3.52)	36.21 (-6.32)	35.65 (-6.88)	36.34 (-6.19)
zh	bge-m3	<u>44.98</u>	45.26 (+0.28)	34.52 (-10.46)	–	36.34 (-8.64)	36.05 (-8.93)	35.86 (-9.12)	35.73 (-9.25)	36.45 (-8.53)
	p-mMiniLM	46.18	<u>45.39</u> (-0.79)	35.46 (-10.72)	–	36.98 (-9.20)	36.77 (-9.41)	36.38 (-9.80)	36.05 (-10.13)	36.85 (-9.33)
	p-mMpNet	46.27	<u>45.41</u> (-0.86)	35.21 (-11.06)	–	36.87 (-9.40)	36.71 (-9.56)	36.28 (-9.99)	35.94 (-10.33)	36.78 (-9.49)
fr	bge-m3	<u>43.18</u>	47.23 (+4.05)	33.29 (-9.89)	33.58 (-9.60)	–	34.07 (-9.11)	36.70 (-6.48)	36.30 (-6.88)	37.25 (-5.93)
	p-mMiniLM	<u>44.09</u>	48.15 (+4.06)	34.54 (-9.55)	34.52 (-9.57)	–	34.83 (-9.26)	37.65 (-6.44)	37.05 (-7.04)	38.03 (-6.06)
	p-mMpNet	<u>43.96</u>	48.14 (+4.18)	34.25 (-9.71)	34.37 (-9.59)	–	34.61 (-9.35)	37.59 (-6.37)	36.93 (-7.03)	38.01 (-5.95)
ja	bge-m3	<u>45.03</u>	45.18 (+0.15)	35.45 (-9.58)	34.86 (-10.17)	36.71 (-8.32)	–	36.11 (-8.92)	35.88 (-9.15)	36.56 (-8.47)
	p-mMiniLM	45.80	<u>45.54</u> (-0.26)	35.90 (-9.90)	35.57 (-10.23)	37.18 (-8.62)	–	36.53 (-9.27)	36.25 (-9.55)	36.91 (-8.89)
	p-mMpNet	45.67	<u>45.39</u> (-0.28)	35.73 (-9.94)	35.30 (-10.37)	36.94 (-8.73)	–	36.24 (-9.43)	35.98 (-9.69)	36.62 (-9.05)
it	bge-m3	<u>41.06</u>	46.63 (+5.57)	33.30 (-7.76)	33.47 (-7.59)	37.92 (-3.14)	33.86 (-7.20)	–	36.44 (-4.62)	37.68 (-3.38)
	p-mMiniLM	<u>42.11</u>	47.69 (+5.58)	34.57 (-7.54)	34.59 (-7.52)	39.07 (-3.04)	34.80 (-7.31)	–	37.55 (-4.56)	38.83 (-3.28)
	p-mMpNet	<u>41.98</u>	47.59 (+5.61)	34.48 (-7.50)	34.68 (-7.30)	38.94 (-3.04)	34.67 (-7.31)	–	37.27 (-4.71)	38.67 (-3.31)
pt	bge-m3	<u>39.19</u>	46.64 (+7.45)	33.37 (-5.82)	33.46 (-5.73)	37.83 (-1.36)	34.02 (-5.17)	37.13 (-2.06)	–	38.61 (-0.58)
	p-mMiniLM	<u>40.17</u>	47.75 (+7.58)	34.67 (-5.50)	34.91 (-5.26)	39.02 (-1.15)	35.03 (-5.14)	38.25 (-1.92)	–	39.68 (-0.49)
	p-mMpNet	<u>39.91</u>	47.30 (+7.39)	34.68 (-5.23)	34.50 (-5.41)	38.70 (-1.21)	34.72 (-5.19)	38.01 (-1.90)	–	39.35 (-0.56)
es	bge-m3	<u>40.76</u>	46.93 (+6.17)	33.36 (-7.40)	33.42 (-7.34)	37.73 (-3.03)	33.87 (-6.89)	37.22 (-3.54)	36.88 (-3.88)	–
	p-mMiniLM	<u>41.81</u>	47.90 (+6.09)	34.63 (-7.18)	34.52 (-7.29)	38.86 (-2.95)	34.76 (-7.05)	38.33 (-3.48)	37.84 (-3.97)	–
	p-mMpNet	<u>41.33</u>	47.34 (+6.01)	34.39 (-6.94)	34.19 (-7.14)	38.34 (-2.99)	34.39 (-6.94)	37.73 (-3.60)	37.25 (-4.08)	–

Language Preference – Retriever Level (Results)

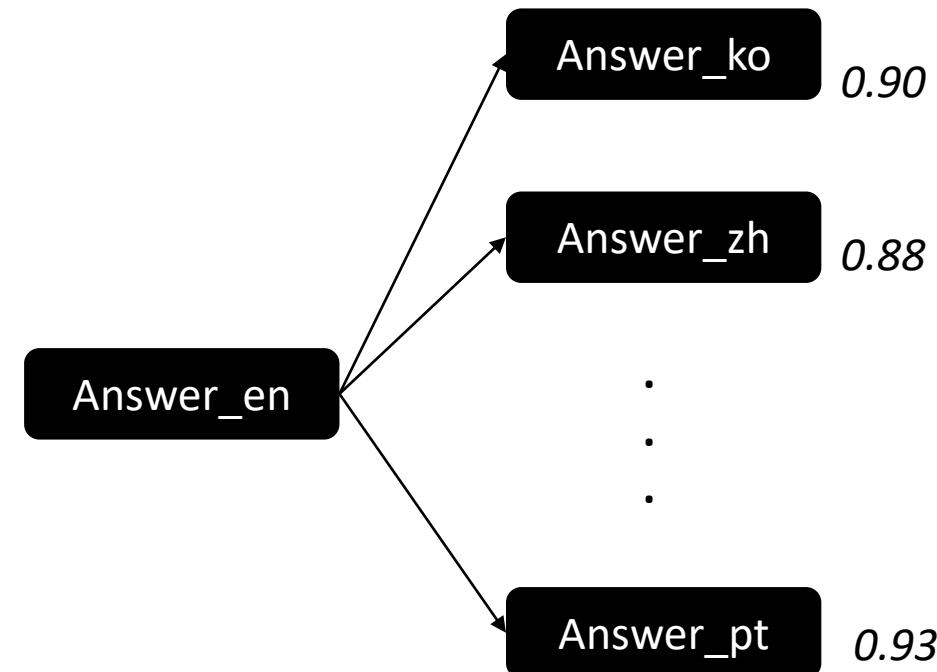
Query Lang.	Encoder	$L_q = L_d$		$L_q \neq L_d$						
		en	ko	zh	fr	ja	it	pt	es	
en	bge-m3	56.03	–	33.02 (-23.01)	33.10 (-22.93)	36.61 (-19.42)	33.36 (-22.67)	35.89 (-20.14)	35.86 (-20.17)	36.62 (-19.41)
	p-mMiniLM	56.85	–	34.34 (-22.51)	34.61 (-22.24)	38.17 (-18.68)	34.52 (-22.33)	37.15 (-19.70)	36.73 (-20.12)	37.96 (-18.89)
	p-mMpNet	57.49	–	34.45 (-23.04)	34.27 (-23.22)	37.94 (-19.55)	34.67 (-22.82)	37.34 (-20.15)	37.02 (-20.47)	37.90 (-19.59)
ko	bge-m3	<u>41.15</u>	43.49 (+2.34)	–	34.42 (-6.73)	36.42 (-4.73)	37.18 (-3.97)	35.72 (-5.43)	35.30 (-5.85)	35.93 (-5.22)
	p-mMiniLM	<u>42.95</u>	44.62 (+1.67)	–	36.04 (-6.91)	37.08 (-5.87)	38.47 (-4.48)	36.07 (-6.88)	36.18 (-6.77)	36.45 (-6.50)
	p-mMpNet	<u>42.53</u>	44.98 (+2.45)	–	35.85 (-6.68)	37.20 (-5.33)	39.01 (-3.52)	36.21 (-6.32)	35.65 (-6.88)	36.34 (-6.19)
zh	bge-m3	<u>44.98</u>	45.26 (+0.28)	34.52 (-10.46)	–	36.34 (-8.64)	36.05 (-8.93)	35.86 (-9.12)	35.73 (-9.25)	36.45 (-8.53)
	p-mMiniLM	46.18	<u>45.39 (-0.79)</u>	35.46 (-10.72)	–	36.98 (-9.20)	36.77 (-9.41)	36.38 (-9.80)	36.05 (-10.13)	36.85 (-9.33)
	p-mMpNet	46.27	<u>45.41 (-0.86)</u>	35.21 (-11.06)	–	36.87 (-9.40)	36.71 (-9.56)	36.28 (-9.99)	35.94 (-10.33)	36.78 (-9.49)
fr	bge-m3	<u>43.18</u>	47.23 (+4.05)	33.29 (-9.89)	33.58 (-9.60)	–	34.07 (-9.11)	36.70 (-6.48)	36.30 (-6.88)	37.25 (-5.93)
	p-mMiniLM	<u>44.09</u>	48.15 (+4.06)	34.54 (-9.55)	34.52 (-9.57)	–	34.83 (-9.26)	37.65 (-6.44)	37.05 (-7.04)	38.03 (-6.06)
	p-mMpNet	<u>43.96</u>	48.14 (+4.18)	34.25 (-9.71)	34.37 (-9.59)	–	34.61 (-9.35)	37.59 (-6.37)	36.93 (-7.03)	38.01 (-5.95)
ja	bge-m3	<u>45.03</u>	45.18 (+0.15)	35.45 (-9.58)	34.86 (-10.17)	36.71 (-8.32)	–	36.11 (-8.92)	35.88 (-9.15)	36.56 (-8.47)
	p-mMiniLM	45.80	<u>45.54 (-0.26)</u>	35.90 (-9.90)	35.57 (-10.23)	37.18 (-8.62)	–	36.53 (-9.27)	36.25 (-9.55)	36.91 (-8.89)
	p-mMpNet	45.67	<u>45.39 (-0.28)</u>	35.73 (-9.94)	35.30 (-10.37)	36.94 (-8.73)	–	36.24 (-9.43)	35.98 (-9.69)	36.62 (-9.05)
it	bge-m3	<u>41.06</u>	46.63 (+5.57)	33.30 (-7.76)	33.47 (-7.59)	37.92 (-3.14)	33.86 (-7.20)	–	36.44 (-4.62)	37.68 (-3.38)
	p-mMiniLM	<u>42.11</u>	47.69 (+5.58)	34.57 (-7.54)	34.59 (-7.52)	39.07 (-3.04)	34.80 (-7.31)	–	37.55 (-4.56)	38.83 (-3.28)
	p-mMpNet	<u>41.98</u>	47.59 (+5.61)	34.48 (-7.50)	34.68 (-7.30)	38.94 (-3.04)	34.67 (-7.31)	–	37.27 (-4.71)	38.67 (-3.31)
pt	bge-m3	<u>39.19</u>	46.64 (+7.45)	33.37 (-5.82)	33.46 (-5.73)	37.83 (-1.36)	34.02 (-5.17)	37.13 (-2.06)	–	38.61 (-0.58)
	p-mMiniLM	<u>40.17</u>	47.75 (+7.58)	34.67 (-5.50)	34.91 (-5.26)	39.02 (-1.15)	35.03 (-5.14)	38.25 (-1.92)	–	39.68 (-0.49)
	p-mMpNet	<u>39.91</u>	47.30 (+7.39)	34.68 (-5.23)	34.50 (-5.41)	38.70 (-1.21)	34.72 (-5.19)	38.01 (-1.90)	–	39.35 (-0.56)
es	bge-m3	<u>40.76</u>	46.93 (+6.17)	33.36 (-7.40)	33.42 (-7.34)	37.73 (-3.03)	33.87 (-6.89)	37.22 (-3.54)	36.88 (-3.88)	–
	p-mMiniLM	<u>41.81</u>	47.90 (+6.09)	34.63 (-7.18)	34.52 (-7.29)	38.86 (-2.95)	34.76 (-7.05)	38.33 (-3.48)	37.84 (-3.97)	–
	p-mMpNet	<u>41.33</u>	47.34 (+6.01)	34.39 (-6.94)	34.19 (-7.14)	38.34 (-2.99)	34.39 (-6.94)	37.73 (-3.60)	37.25 (-4.08)	–

Language Preference – Generator Level (way to measure)

*Generate answer in
different languages*



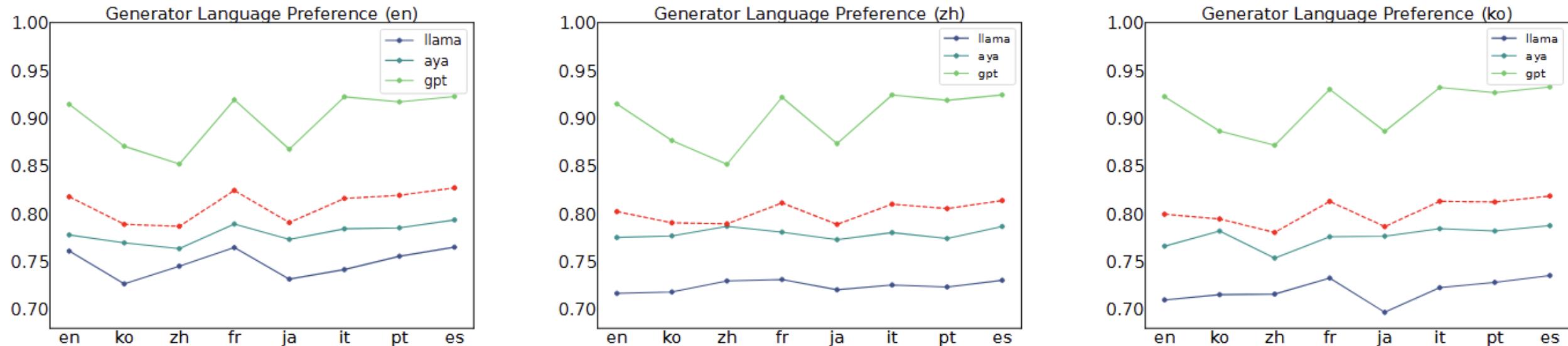
**English Preference
of Generator**



*Compute average multilingual
embedding similarity*

English Preference of Generator: 0.903

Language Preference – Generator Level (Results)



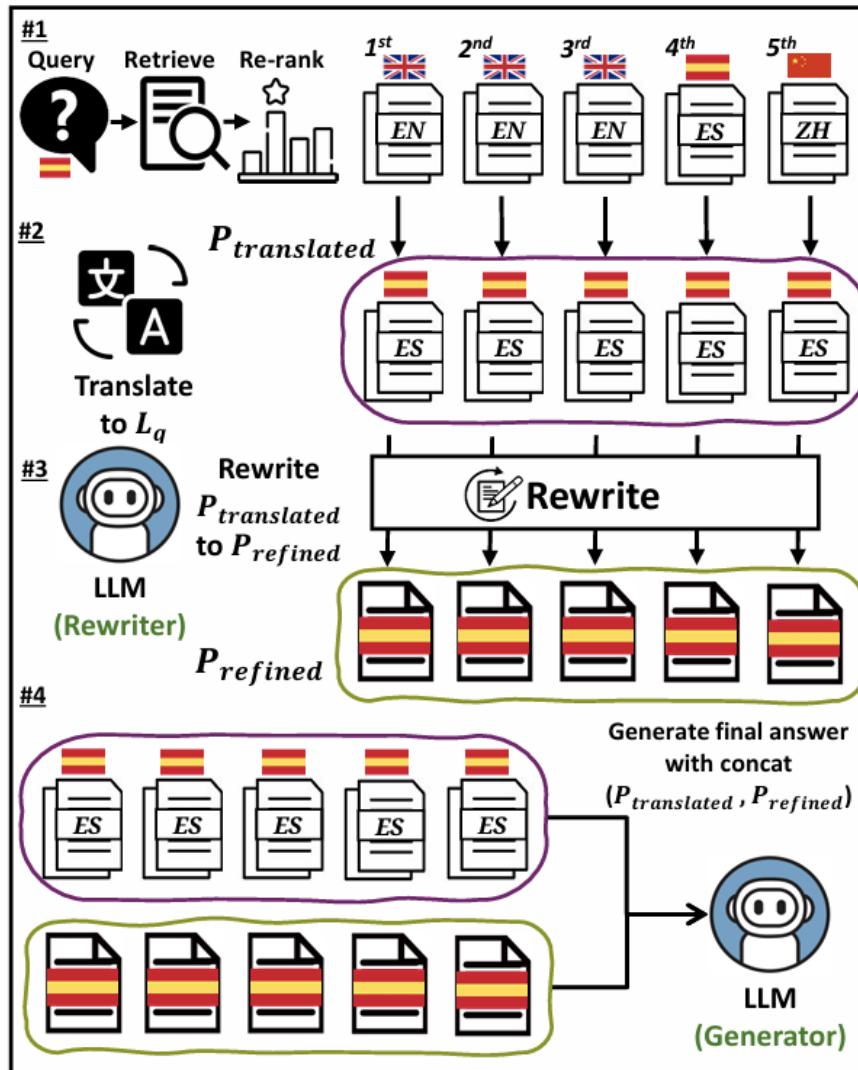
- Generators prefer:
 - Latin script languages (en, fr, it, pt, es)
 - Query language but not big effect

Language Preference – mitigating strategy (Proposed Method)

	all	en	zh	ko	fr	ja	it	pt	es
$L_q = \text{en}$									
aya-expanse-8b	<u>80.09</u>	79.34	63.08	64.46	76.13	61.20	75.47	75.65	76.32
Phi-4	<u>79.69</u>	78.89	63.06	52.30	74.43	48.86	74.02	74.39	75.32
Qwen2.5-7B-Instruct	<u>80.15</u>	79.11	50.31	64.90	76.28	62.62	75.47	75.97	76.54
Llama3.1-8B-Instruct	<u>80.25</u>	79.28	61.99	65.81	76.40	62.58	75.89	76.09	76.47
$L_q = \text{zh}$									
aya-expanse-8b	32.55	25.62	<u>38.31</u>	26.64	24.00	25.27	23.63	23.63	23.79
Phi-4	16.75	17.57	<u>36.76</u>	17.50	18.15	17.56	18.19	17.89	18.44
Qwen2.5-7B-Instruct	34.28	27.33	<u>38.31</u>	27.91	25.15	27.78	25.90	25.37	25.30
Llama3.1-8B-Instruct	28.50	24.36	<u>38.48</u>	23.84	22.48	23.78	23.18	23.32	23.02
$L_q = \text{ko}$									
aya-expanse-8b	40.60	38.08	26.01	<u>49.66</u>	25.37	26.82	24.98	25.26	25.51
Phi-4	26.80	20.24	17.54	<u>49.25</u>	19.03	17.91	18.93	19.19	19.19
Qwen2.5-7B-Instruct	36.50	22.87	20.08	<u>49.44</u>	21.79	20.94	21.65	21.44	21.52
Llama3.1-8B-Instruct	37.18	26.48	22.88	<u>49.87</u>	24.46	24.86	25.23	24.87	25.22

- For query in English:
retrieving from datastores
in various languages (all
strategy) is effective
- For query in non-English:
retrieving from datastores
in various languages and
translate to query language
is effective

Language Preference – mitigating strategy (Proposed Method)



Proposed Dual-Knowledge Multilingual RAG (DKM-RAG):

- For a given query, **retrieves** documents from datastores in various languages and **re-rank** documents
- **Translate** all non-query language documents into **query language**
- **Rewrite** translated passages to filter inaccuracies and enrich the retrieved information with more reliable content
- Finally, **concatenate** translated passage and refined passage, generating final answer with them

Language Preference – mitigating strategy (Results)

	all	en	zh	ko	fr	ja	it	pt	es	DKM-RAG
<i>L_q = en</i>										
aya-expanse-8b	80.09	79.34	63.08	64.46	76.13	61.20	75.47	75.65	76.32	82.60
Phi-4	79.69	78.89	63.06	52.30	74.43	48.86	74.02	74.39	75.32	82.59
Qwen2.5-7B-Instruct	80.15	79.11	50.31	64.90	76.28	62.62	75.47	75.97	76.54	82.60
Llama3.1-8B-Instruct	80.25	79.28	61.99	65.81	76.40	62.58	75.89	76.09	76.47	82.57
<i>L_q = zh</i>										
aya-expanse-8b	32.55	25.62	38.31	26.64	24.00	25.27	23.63	23.63	23.79	44.57
Phi-4	16.75	17.57	36.76	17.50	18.15	17.56	18.19	17.89	18.44	44.56
Qwen2.5-7B-Instruct	34.28	27.33	38.31	27.91	25.15	27.78	25.90	25.37	25.30	44.70
Llama3.1-8B-Instruct	28.50	24.36	38.48	23.84	22.48	23.78	23.18	23.32	23.02	44.51
<i>L_q = ko</i>										
aya-expanse-8b	40.60	38.08	26.01	49.66	25.37	26.82	24.98	25.26	25.51	55.01
Phi-4	26.80	20.24	17.54	49.25	19.03	17.91	18.93	19.19	19.19	54.82
Qwen2.5-7B-Instruct	36.50	22.87	20.08	49.44	21.79	20.94	21.65	21.44	21.52	54.85
Llama3.1-8B-Instruct	37.18	26.48	22.88	49.87	24.46	24.86	25.23	24.87	25.22	54.99
MLR (Preference)	-	47.70	35.90	35.47	37.94	37.59	37.66	37.15	<u>37.97</u>	-

Conclusion

- Investigated **language preferences** in mRAG systems
- Proposed a **metric that measures the language preference of retrievers** by checking the rank difference between the translated passage and the original one
- Experiments reveal that retrievers prefer **high-resource and query language** but do not always yield better generation performance
 - Generators often favor the **query language or Latin scripts**, resulting in inconsistent outputs
- Proposed DKM-RAG which integrates translated passages with internal knowledge
 - Empirical results show that DKM-RAG consistently enhances mRAG performance across diverse languages