Breaking Language Preference in Multilingual RAG via Language-Controllable Retrieval and Language-Agnostic Reasoning

Wenshuai Huo[1,2], Xiaocheng Feng[1,2], Baohang Li[1], Chengpeng Fu[1,2], Yichang Huang[1], Hui Wang[2], Bing Qin[1,2]
[1]Harbin Institute of Technology, [2]Pengcheng Laboratory
{wshuo, xcfeng, baohangli, cpfu, ychuang, qinb}@ir.hit.edu.cn,
wangh06@pcl.ac.cn

Abstract

Retrieval-Augmented Generation (RAG) significantly improves the factual accuracy and generation quality of large language models by incorporating external knowledge. However, in multilingual settings, RAG systems suffer from severe language preference. On the one hand, the retrieval stage is sensitive to the query language: semantically equivalent queries expressed in different languages often lead to substantially different retrieval results. On the other hand, when retrieved documents contain knowledge written in multiple languages, large language models tend to be influenced by surface-level language forms, rather than reasoning solely based on semantic relevance to the query. To address these challenges, we propose a unified optimization framework that explicitly disentangles multilingual RAG into language-controllable retrieval and language-agnostic reasoning. Our framework allows LLM to adaptively select retrieval languages while enforcing cross-lingual consistency during reasoning, thereby mitigating language bias without modifying existing retrievers or translators. Experimental results demonstrate that our approach effectively reduces language bias in multilingual RAG and consistently outperforms baselines across multiple multilingual benchmarks .

1 Introduction

Large Language Models (LLMs) (Ouyang et al., 2022; Touvron et al., 2023; Liu et al., 2024; Yang et al., 2025) have achieved strong performance and have been widely applied across various domains. Retrieval-Augmented Generation (RAG) (Lewis et al., 2020; Ram et al., 2023; Chen et al., 2024b; Jin et al., 2025) further enhances generation quality by grounding model outputs in externally retrieved knowledge. However, in many practical settings, user queries may be expressed in low-resource languages, while external knowledge bases often contain documents written in diverse languages. This motivates Multilingual Retrieval-Augmented Generation (mRAG) (Chirkova et al., 2024; Ranaldi et al., 2025), which aims to support robust retrieval and generation in multilingual settings.

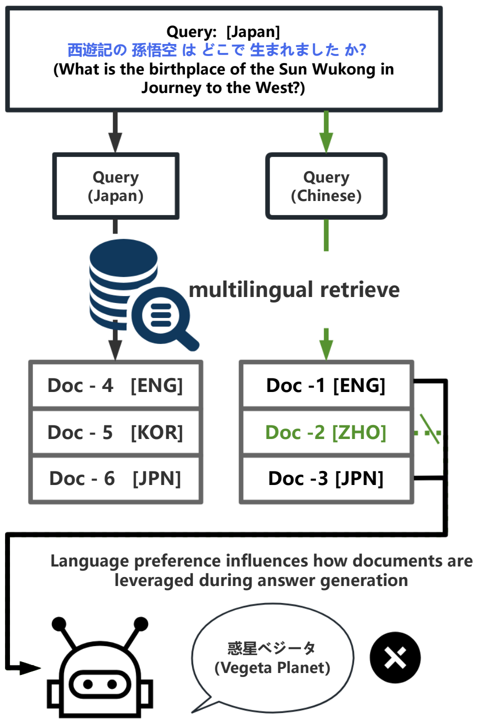

Figure 1: Language preference in mRAG. The example shows a Japanese query explicitly asking about the birthplace of Sun Wukong in Journey to the West. Using different retrieval languages leads to different retrieved documents from the same multilingual corpus. Doc-2 (in Chinese) contains directly relevant evidence from the original text, while other retrieved documents describe culturally related adaptations (e.g., Goku in Dragon Ball, inspired by Journey to the West). Language preferences during generation influence how these documents are weighted, affecting answer quality.

Language preference (Sharma et al., 2025; Park and Lee, 2025) poses a critical challenge for mRAG

Findings of the Association for Computational Linguistics: ACL 2026, pages 7579–7589

as illustrated in Figure 1. On the one hand, the retrieval stage is highly sensitive to the query language: semantically equivalent queries expressed in different languages often lead to substantially different retrieved document sets. On the other hand, when retrieved documents are multilingual, LLMs tend to be influenced by surface-level language forms during reasoning and generation, rather than strictly following semantic relevance. Such language bias can accumulate across retrieval and reasoning stages, significantly degrading the overall performance of multilingual RAG systems. To mitigate this issue, prior work primarily relies on heuristic translation strategies that unify multilingual content into a single language through translation. One line of work translates queries into a high-resource language before retrieval (Bi et al., 2020; Gao et al., 2025; Ranaldi et al., 2025), while another translates retrieved multilingual documents into a unified language before generation (Park and Lee, 2025; Moon et al., 2025; Ki et al., 2025). Although these methods alleviate language inconsistency to some extent, they rely on a static language unification strategy, which limits their flexibility under diverse language distributions.

In this work, we propose to explicitly disentangle multilingual RAG into language-controllable retrieval and language-agnostic reasoning. Before knowledge retrieval, the LLM is first prompted to select an appropriate retrieval language conditioned on the input query. The query is then issued in the selected language to the retriever, enabling more effective access to language-specific knowledge resources. For generation, the model is encouraged to perform reasoning based solely on the semantic relevance of the retrieved knowledge, rather than its surface language form. To achieve this objective, we design a reinforcement learning framework that provides stage-specific guidance and optimizes the model based on downstream task performance. Under this framework, the generator is the only learnable component and interacts with a fixed retriever and a fixed translation tool. This design avoids modifying existing retrieval systems, while enabling the model to adapt its behavior to diverse language distributions in multilingual RAG settings.

We evaluate our approach on widely used multilingual mRAG benchmarks. The results show that our method achieves consistent performance improvements across languages, with particularly notable gains for low-resource languages. Further analysis indicates that our approach leads to more effective retrieval and reasoning. These findings suggest that applying targeted optimization to both the retrieval and reasoning stages can mitigate the impact of language preference and enhance the capabilities of LLMs in multilingual RAG settings. Our main contributions are summarized as follows:

- We explicitly disentangle multilingual RAG into language-controllable retrieval and language-agnostic reasoning, based on the stage-specific characteristics of mRAG.;
- We propose a unified framework that applies targeted optimization to retrieval and reasoning, guiding the model to adaptively adjust its behavior across stages;
- Extensive experiments across multiple settings demonstrate that our approach consistently improves multilingual RAG performance, with particularly strong gains in low-resource language settings, indicating its effectiveness in mitigating language preference.

2 Preliminary

Retrieval-Augmented Generation (RAG) augments LLMs with external knowledge. Given a query q, a retriever R retrieves relevant documents D from a corpus, and generates an answer a based on both q and D:

$$D = R(q), \quad a = \text{LLM}(q, D). \tag{1}$$

In multilingual settings, both the query and the document corpus may involve multiple languages. Due to inherent language preferences in retrieval systems, semantically equivalent queries expressed in different languages (q_{l_1} and q_{l_2}) often result in retrieval outputs that differ substantially, leading to systematically poorer retrieval performance for low-resource languages. Naive multilingual solutions include translating the query into a fixed target language (e.g., English) before retrieval:

$$\tilde{q} = \mathcal{T}(q, l_{\text{en}}), \quad D = R(\tilde{q}), \tag{2}$$

or translating retrieved documents into a unified language prior to generation:

$$\tilde{D} = \{\mathcal{T}(d_i, l_{\text{en}})\}, \quad a = \text{LLM}(q, \tilde{D}). \tag{3}$$

Although these methods improve multilingual RAG performance to some extent, they rely on a

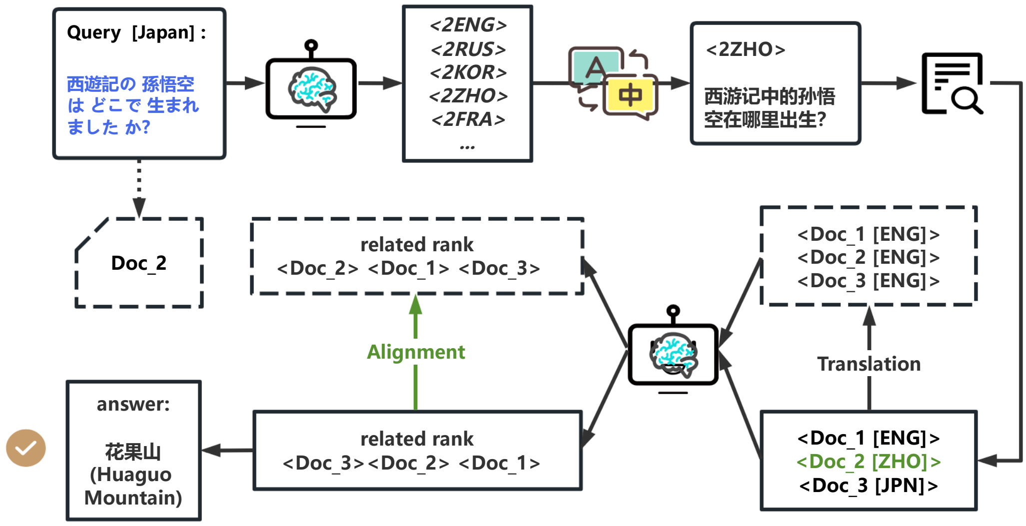

Figure 2: The illustration of the proposed methods. For each query, the model selects an appropriate retrieval language to facilitate effective document retrieval from a multilingual corpus. Language-agnostic reasoning is then achieved by comparing relevance rankings induced from original and translated document representations, thereby enforcing cross-lingual alignment in reasoning. Additional feedback from downstream task accuracy is used to further refine the model.

pre-defined fixed target language for translation, which may not adapt well to diverse language distributions and does not directly optimize for downstream task performance.

3 Method

3.1 Framework

This section presents our proposed method. To address language preference in mRAG, we propose a reinforcement-learning–based optimization framework, as illustrated in Figure 2. Considering the distinct characteristics of the retrieval and generation stages in mRAG, we guide the model's behavior from two complementary perspectives. At the retrieval stage, the model is guided to adaptively select an appropriate retrieval language in order to obtain more suitable document retrieval results. At the reasoning stage, the model is encouraged to rely primarily on semantic information when judging document relevance and generating answers, rather than being influenced by the surface language forms of the documents. The generator serves as the only learnable component and interacts with a fixed retriever and a fixed translation tool. The model is trained using reinforcement learning and optimized with downstream task-level feedback, and achieves improved retrieval and generation performance in multilingual settings.

3.2 Language-Controllable Retrieval

In our framework, the retrieval-stage behavior is guided to allow the model to control the language used for retrieval in multilingual settings. Given an input query q, the model is prompted to generate a retrieval language $\ell_r \in \mathcal{L}$, where $\mathcal{L}$ denotes the set of candidate retrieval languages:

Conditioned on the selected language ℓ_r, the query is either translated into ℓ_r or retained in its original form, and the retriever returns a set of top-k candidate documents D from the same multilingual corpus. Although the retrieval corpus itself remains unchanged, different choices of retrieval language can lead to substantially different document recall distributions. Therefore, selecting an appropriate language for retrieval is crucial. To guide this process, we incorporate a retrieval-level feedback signal that evaluates whether the retrieved top-k documents contain relevant evidence for the input query. Specifically, the reward r_{LcR} measures whether the ground-truth answer, or its translations into multiple languages, is present in the retrieved documents. By optimizing retrieval language choices with respect to these signals, the model progressively favors languages that are more likely to surface useful evidence and support subsequent reasoning and answer generation.

3.3 Language-Agnostic Reasoning

To guide language-agnostic reasoning, we introduce a dual ranking mechanism that evaluates document relevance under different language representations and enforces consistency between the induced rankings, thereby ensuring that the LLM relies on semantic relevance rather than surface language forms when reasoning over multilingual external knowledge.

3.3.1 Document Ranking

Given an input query q and a retrieved document set $D = \{d_1, \dots, d_N\}$, the generator first assesses relevance based on the original documents and assigns a relevance score to each document, thereby producing a relevance ranking:

$$\pi^{\text{orig}} = Rank(q, D), \tag{4}$$

where $\text{Rank}(\cdot)$ denotes the ranking induced by sorting documents according to their relevance scores predicted by the model.

In parallel, each document d_i is translated into a unified language to isolate semantic relevance from linguistic variation, yielding a translated document set $\tilde{\mathcal{D}} = \{\tilde{d}_1, \dots, \tilde{d}_N\}$. Using the same query q, the generator forms a second relevance ranking based on the translated documents:

$$\pi^{\text{trans}} = Rank(q, \tilde{D}), \tag{5}$$

3.3.2 Ranking Consistency

To encourage the model to perform reasoning that is insensitive to language-specific surface forms, we align its relevance judgments across the original and unified-language document representations. Specifically, we measure the consistency between the two induced rankings using Rank-Biased Overlap (RBO). Given a persistence parameter $p \in (0, 1)$, RBO is defined as:

$$\text{RBO} = (1 - p) \sum_{k=1}^{\infty} p^{k-1} \cdot \frac{|\pi_{1:k}^{\text{orig}} \cap \pi_{1:k}^{\text{trans}}|}{k}, \tag{6}$$

where $\pi_{1:k}$ denotes the top-k prefix of a ranking. Higher RBO values indicate stronger agreement between relevance judgments across different language representations. We use the resulting RBO score as the language-agnostic ranking reward, denoted as r_{LaR}.

This ranking consistency signal is used to guide the model toward making relevance judgments that are less sensitive to surface language forms.

3.4 Accuracy-Based Reward

In addition to the feedback signals used for retrieval and reasoning, we further optimize the model with respect to downstream task accuracy. Given the input query q, the retrieved document set D, and the final generated answer a, we compute a task-level accuracy reward that reflects whether the generated answer matches the ground-truth answer. Formally, the accuracy reward is defined as:

$$r_{\text{acc}} = \mathbb{I}(a = a^*), \tag{7}$$

where a^* denotes the reference answer and $\mathbb{I}(\cdot)$ is the indicator function that returns 1 if the condition is satisfied and 0 otherwise. In our implementation, this reward is computed using Exact Match between the generated answer and the reference.

This accuracy-based reward provides direct supervision for the end task and serves as an optimization signal for both retrieval and reasoning behaviors.

3.5 Optimization Strategy

We optimize the language model using reinforcement learning with Proximal Policy Optimization (PPO). During training, only the language model parameters are updated, while the retriever and translation components remain fixed. All reward terms are assigned equal weights throughout training.

4 Experiments

4.1 Experimental Setup

We evaluate the proposed method on two representative large language models: LLaMA-3.2[1] and Qwen-2.5[2]. We follow the setup described in Chirkova et al. (2024). We adopt BGE-M3[3] as the multilingual dense retriever and use bge-reranker-v2-m3[4] as the reranking model. Document embeddings are indexed using FAISS with an IVF8192,PQ32 configuration for efficient retrieval. For translation, we use NLLB-3.3B[5] as the translation tool to translate queries or documents when required. To mitigate the impact of

[1] https://huggingface.co/collections/meta-llama/llama-32

[2] https://huggingface.co/collections/Qwen/qwen25

[3] https://huggingface.co/BAAI/bge-m3

[4] https://huggingface.co/BAAI/bge-reranker-v2-m3

[5] https://huggingface.co/facebook/nllb-200-3.3B

Method	ar	de	en	es	fi	fr	it	ja	ko	pt	ru	th	zh	avg
Performance on Qwen2.5-7B-Instruct														
MRAG	0.09	0.21	0.31	0.22	0.12	0.21	0.21	0.09	0.07	0.19	0.10	0.07	0.14	0.16
TRANS_RAG	0.09	0.22	0.31	0.24	0.15	0.24	0.23	0.10	0.08	0.19	0.09	0.08	0.16	0.17
DOC_TRANS	0.09	0.19	0.33	0.25	0.12	0.22	0.22	0.09	0.08	0.22	0.11	0.09	0.16	0.17
DKM_RAG	0.08	0.23	0.33	0.25	0.12	0.23	0.21	0.10	0.07	0.21	0.13	0.06	0.15	0.17
OURS	**0.15**	**0.31**	**0.37**	**0.33**	**0.26**	**0.32**	**0.29**	**0.20**	**0.15**	**0.32**	**0.19**	**0.14**	**0.23**	**0.25**
Performance on LLaMA3.2-8B-Instruct														
MRAG	0.06	0.20	0.35	0.25	0.14	0.25	0.23	0.10	0.07	0.20	0.11	0.06	0.09	0.16
TRANS_RAG	0.09	0.25	0.35	0.27	0.16	0.26	0.26	0.13	0.09	0.22	0.12	0.09	0.11	0.19
DOC_TRANS	0.07	0.27	0.37	0.26	0.13	0.27	0.24	0.11	0.08	0.24	0.12	0.05	0.09	0.18
DKM_RAG	0.07	0.29	**0.46**	0.32	0.17	0.33	0.32	0.13	0.09	0.27	0.14	0.07	0.12	0.21
OURS	**0.17**	**0.37**	0.42	**0.37**	**0.34**	**0.37**	**0.38**	**0.24**	**0.18**	**0.36**	**0.27**	**0.14**	**0.18**	**0.29**

Table 1: Performance on MKQA across different mRAG methods.

low-quality translations, we further apply an unsupervised quality estimation (QE) model[6] to filter out translations with low confidence. For the ranking consistency objective, we compute RBO with the persistence parameter set to $p = 0.9$. During training, only the parameters of the generator are updated, while the retriever and the translation system remain frozen throughout all experiments. Exact Match (EM) is used as the evaluation metric. The model is trained using Proximal Policy Optimization (PPO), with the reinforcement learning setup following Jin et al. (2025). All experiments are performed on 8 NVIDIA A100 GPUs with a total batch size of 512. The maximum sequence length is set to 4,096 tokens. For each query, the top-3 passages are retrieved.

4.2 Datasets

Our experiments are based on the MKQA dataset (Longpre et al., 2021), which contains approximately 10,000 samples per language. We filter out instances without annotated answers and, for each language, sample 500 examples as the test set, with the remaining data used for training. In addition, we evaluate our models on the XOR-TyDi dataset (Asai et al., 2021). XOR-TyDi is a multilingual question answering dataset that contains questions written natively in multiple languages along with their corresponding answers. For retrieval, we operate over a multilingual knowledge base composed of multilingual Wikipedia[7]. In our experiments, we use a total of 13 language directions for both the training data and the retrieval corpus, covering a diverse set of languages: Arabic (ar), German (de), English (en), Spanish (es), Finnish (fi), French (fr), Italian (it), Japanese (ja), Korean (ko), Portuguese (pt), Russian (ru), Thai (th), and Chinese (zh).

[6] https://huggingface.co/Unbabel/wmt22-cometkiwi-da

[7] https://huggingface.co/datasets/wikimedia/wikipedia

4.3 Baselines

To comprehensively evaluate the effectiveness of our proposed framework, we selected several commonly used multilingual RAG strategies as comparative baselines.

- **MRAG** is the vanilla baseline that performs retrieval and generation directly in a multilingual setting.
- **TRANS_RAG** translates questions into English before retrieval, aiming to leverage the stronger retrieval performance
- **DOC_TRANS** (Ranaldi et al., 2025) conducts retrieval using the original query language, while converting the retrieved documents into English before generation to reduce language inconsistency at the generation stage.
- **DKM-RAG** (Park and Lee, 2025) converts retrieved documents into the query language and further refines them using an LLM-based rewriting module, with the goal of mitigating language preference effects in multilingual RAG systems.

All baselines use the same backbone language models, retriever, and translation tool as our method to ensure fair and controlled comparison.

4.4 Main Results

We evaluate the overall performance of our method across different LLMs in the mRAG setting, as shown in Table 1. Our approach yields consistent improvements over the baselines in almost all evaluated languages. The gains are particularly substantial for low-resource languages, indicating that language preference has a stronger impact in these settings. In addition, both TRANS_RAG and DOC_TRANS yield improvements over the vanilla RAG baseline, suggesting that the language form used during retrieval and generation substantially influences mRAG performance. However, our method surpasses these translation-based approaches, suggesting that guiding the model's behavior, rather than statically unifying queries or documents into a fixed language, provides a broader exploration space and ultimately achieves higher performance. We further observe that query translation tends to be more beneficial for low-resource languages, while document translation yields more pronounced improvements for high-resource languages. A possible explanation is that translation errors at the document level are more severe for low-resource languages, which limits the effectiveness of document translation–based methods. In contrast, our approach achieves stable improvements across nearly all languages.

4.5 Ablation Experiments

To assess the contribution of each key component in our proposed framework, we conduct a series of ablation experiments by selectively removing individual optimization terms and evaluating model performance on the test set. Results on LLaMA3.2-1B-Instruct are summarized in Figure 3. We consider the following three ablation settings:

- **+Acc** Only the accuracy reward is used, measuring task-level correctness of the generated answers.
- **+LaR (Language-Agnostic Reasoning)** Further adds the ranking consistency reward to encourage cross-language alignment of relevance judgments.
- **+LcR (Language-Controllable Retrieval)** Further incorporates optimization for adaptive retrieval language selection to enable language-aware retrieval behavior.

From Figure 3, we observe that: (1) Using the accuracy reward alone already yields substantial performance improvements, particularly for high-resource languages. (2) Adding the cross-alignment reward further enhances performance by promoting language-agnostic reasoning, which is consistent with Liu et al.'s finding that the primary bottleneck of cross-lingual RAG lies in reasoning over cross-lingual documents. (3) Incorporating adaptive retrieval leads to stable improvements across nearly all languages and achieves the best overall performance. These results indicate that our method effectively improves multilingual RAG performance while balancing language preference in both retrieval and reasoning.

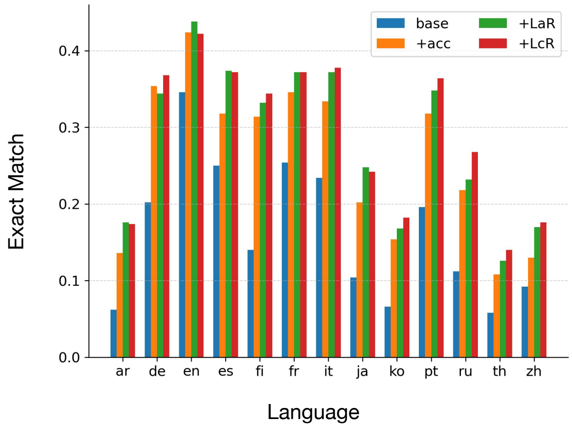

Figure 3: Ablation study results of the proposed framework for each key reward component.

5 Analysis

5.1 Analysis of Retrieval Language Selection

To verify the effectiveness of our retrieval language selection strategy, we compare the top-3 recall on the test set between the base method and our approach for each language, as illustrated in Figure 5.

For the base method, retrieval is performed by directly issuing the original input query to the multilingual corpus. Consequently, the retrieval behavior is entirely determined by the original query language, which may not align well with the language distribution of relevant documents in the corpus.In contrast, for our approach, given an input query, the optimized model first selects an appropriate language form for retrieval, and the query is then issued in the selected language to retrieve documents from the multilingual corpus. Both methods use the same retriever and retrieval corpus; the only difference lies in whether retrieval language selection is applied. Under this controlled setting, the results show that our approach achieves consistent improvements across all languages after optimization. This indicates that the model learns to select more suitable language forms for queries, leading to more effective multilingual retrieval.

a.	query [zh]	2015年的世界足球先生是谁	
	retrieved_documents	doc1 [ko]:	國際足協世界足球先生…教練、男子國家隊隊長、足球記者以及在國際足協官方網站上投票的球迷進行投票。選票上有第一、第二和第三選擇，…
		doc2 [zh]:	世界足球胜利十一人2015……里体育会、毕尔包、拜仁慕尼黑、祖云达斯及拿玻里、\n收录球队\n球会\n……
		doc3 [pt]:	Чемпионат мира по футболу 7 × 7 2015. Награды Лучший игрок турнира: Эдуард Рамонов Лучший вратарь турнира: ..
	translation [es]	quien ha ganado el premio al mejor jugador de la fifa en 2015	
	retrieved_documents	doc1 [es]:	Premios Individuales de Fútbol 2015 (Europa\"\n. ...Los Premios Individuales de Fútbol 2015 son Bota de Oro, Mejor Jugador del Mundo, Jugador europeo del Año y Trofeo Zamora..
		doc2 [en]:	2015 FIFA Ballon d'Or...The awards were given out in Zürich, Switzerland on 11 January 2016.Lionel Messi won the award as the World Player of the Year for the fifth time, extending his own record of five wins. ..
		doc3 [ru]:	Золотой мяч ФИФА. .. Обладателями «Золотого мяча ФИФА» стали только два футболиста — в 2010, 2011, 2012 и 2015 годах им оказался аргентинец Лионель Месси, а в..
b.	query [ru]	кто написал гитарное соло для beat it	
	retrieved_documents	doc1 [ko]:	〈Beat It〉...은 미국의 가수 마이클 잭슨의 여섯 번째 스튜디오 음반인 《Thriller》에 수록된 곡이다. 이 곡은 잭슨이 쓰고 잭슨과 퀸시 존스가 프로듀싱했다. . ...이 곡은 에디 반 헤일런의 기타 독주를 포함한다...
		doc2 [ru]:	\"Beat (альбом\"\n... — Э. Белью, Бруфорд, Фрип и Левин) — 3:23 «The Howler» — 4:13 «Requiem» — 6:48 Участники записи Роберт Фрипп — гитара, орган; Эдриан Белью — гитара, вокал...
		doc3 [pt]:	She Said She Said... cantar um refrão que diz \"I know what is like to be dead\" (Eu sei como é estar morto\"\n. Ficha Técnica John Lennon: Vocal, guitarra Base e órgão George Harrison...
	output	base	Джордж Харрисон
		ours	Ван Хален, Эдди

Figure 4: Case study comparing retrieval and reasoning behaviors of the base method and our approach.

5.2 Case Study

To better illustrate how our approach produces more effective retrieval and reasoning behaviors, we present several examples from the evaluation set in Figure 4.

In Case (a), when applying the base system, the query is issued in its original language to the multilingual corpus, which fails to retrieve relevant documents. In contrast, our approach assigns Spanish as the retrieval language at the initial stage and successfully retrieves documents containing relevant knowledge for the query. In Case (b), relevant documents are already retrieved and provided to the model. However, in the base system, the evidence is overlooked during reasoning due to its language form, leading to an incorrect answer. By comparison, our approach encourages language-agnostic reasoning, allowing the model to focus on semantic relevance rather than surface language cues, and consequently produces the correct answer.

5.3 Analysis of Out-of-Domain Generalization

We further analyze the out-of-domain generalization of our approach on the XOR-TyDi dataset. As shown in Table 2, our method consistently outperforms the vanilla mRAG baseline across languages, with particularly strong improvements in low-resource settings. For languages not seen during training, such as bn (Bengali) and te (Telugu), the baseline performance is relatively low, since the retrieval corpus used in our experiments does not include documents in these languages. After applying our optimization framework, both languages achieve substantial performance gains, indicating improved generalization to unseen languages.

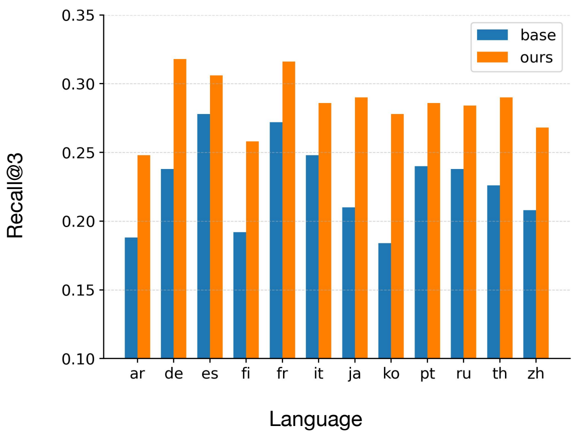

Figure 5: Comparison of retrieval recall on a multilingual corpus before and after optimization.

6 Related Work

6.1 Multilingual Retrieval-Augmented Generation

Multilingual Retrieval-Augmented Generatio (mRAG) extends RAG (Lewis et al., 2020; Jin et al., 2025) to multilingual scenarios, enabling retrieval and generation across multiple languages. To facilitate systematic evaluation in multilingual settings, several works have proposed benchmarks (Thakur et al., 2024; Lim et al., 2025; Thakur et al., 2025). Beyond this, many studies explic-

Method	ar	bn	fi	ja	ko	ru	te	avg
				Performance on Qwen2.5-7B-Instruct				
MRAG	0.16	0.07	0.11	0.13	0.17	0.09	0.05	0.11
TRANS_RAG	0.16	0.12	0.13	0.11	0.18	0.08	0.07	0.12
DOC_TRANS	0.15	0.10	0.13	0.08	0.18	0.10	0.13	0.12
OURS	0.20	0.13	0.18	0.19	0.19	0.12	0.08	0.16
				Performance on LLaMA3.2-8B-Instruct				
MRAG	0.14	0.10	0.14	0.10	0.16	0.09	0.12	0.12
TRANS_RAG	0.15	0.12	0.15	0.11	0.18	0.10	0.16	0.14
DOC_TRANS	0.14	0.10	0.14	0.10	0.16	0.10	0.13	0.12
OURS	0.20	0.17	0.22	0.22	0.20	0.18	0.19	0.20

Table 2: Performance on XOR-TyDi across different mRAG methods.

itly consider real-world multilingual data and culture-specific multilingual scenarios (Li et al., 2025; Cruz Blandón et al., 2025). Meanwhile, some studies aim to support effective cross-lingual retrieval and reranking, with approaches that utilize multilingual representation learning to build shared semantic spaces across languages (Litschko et al., 2023; Chen et al., 2024a; Goworek et al., 2025).

While these efforts improve usability in multilingual scenarios, they typically assume that knowledge distributions across languages can be sufficiently aligned, which does not always hold in practice and may lead to unstable performance.

6.2 Language Preference in Multilingual RAG

Studies indicate that during retrieval and ranking, queries or documents in different languages may not be treated equally (Yang et al., 2024; Wu et al., 2024). On the generation side, models may favor linguistically familiar but less relevant information over more relevant evidence in other languages, even when valid multilingual evidence is available (Ki et al., 2025). These findings suggest that language systematically influences how information is retrieved and utilized in multilingual retrieval-augmented systems. To mitigate language preference, prior work has proposed various strategies. Some works assign language-specific weights or adjustments to balance inherent model biases(Telemala and Suleman, 2022; Yang et al., 2024), while others translate queries or documents into high-resource languages (Park and Lee, 2025; Moon et al., 2025).

However, most of them rely on heuristic design. In contrast, our approach is driven directly by downstream task performance.

6.3 Reinforcement Learning for Generation Models

Reinforcement learning has been widely applied to natural language generation to shape model behavior beyond supervised objectives, enabling optimization toward task-level utility and preference signals. In LLMs, Proximal Policy Optimization (PPO) has become a standard tool for aligning model outputs with desired behaviors (Schulman et al., 2017; Ouyang et al., 2022; Guo et al., 2025). Prior work has shown that reinforcement learning can improve factuality and reasoning, support controllable generation, and enhance safety and preference consistency (Xue et al., 2023; Sun et al., 2024; Rafailov et al., 2023; Chakraborty et al., 2024; Dai et al., 2023; Mu et al., 2024). More recently, reward designs that account for language disparities have been explored to reduce performance gaps between high- and low-resource languages (Yang et al., 2024; She et al., 2024).

In this work, we leverage Reinforcement learning to guide model behavior across stages in multilingual RAG, using downstream performance and multi-dimensional reward signals to inform decisions relevant to retrieval and reasoning.

7 Conclusion

In this work, we propose a framework that disentangles mRAG into language-controllable retrieval and language-agnostic reasoning, and applies targeted optimization to both stages. Experiments on widely used benchmarks show consistent improvements across languages, with particularly strong gains for low-resource languages. Further analyses indicate that the proposed framework leads to more effective retrieval and more stable reasoning behavior. Together, these results demonstrate the effectiveness of our approach for mitigating language preference in mRAG.

8 Acknowledgments

Xiaocheng Feng, and Bing Qin are the co-corresponding authors of this work. We thank the anonymous reviewers for their insightful comments. This work was supported by the National Natural Science Foundation of China (NSFC) (grant 62522603, 62276078), the Key R&D Program of Heilongjiang via grant 2022ZX01A32, the Fundamental Research Funds for the Central Universities (XNJKKGYDJ2024013), the Major Key Project of PCL (Grant No. PCL2025A12, PCL2025A03), the National Science and Technology Major Program (Grant No. 2024ZD01NL00101).

Limitations

Our method introduces additional steps for determining the retrieval language during inference, which incurs extra latency and computational overhead compared to standard RAG pipelines. Moreover, the approach relies on external translation tools, and translation errors may degrade performance.

References

Akari Asai, Jungo Kasai, Jonathan Clark, Kenton Lee, Eunsol Choi, and Hannaneh Hajishirzi. 2021. XOR QA: Cross-lingual open-retrieval question answering. In *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 547–564, Online. Association for Computational Linguistics.

Tianchi Bi, Liang Yao, Baosong Yang, Haibo Zhang, Weihua Luo, and Boxing Chen. 2020. Constraint translation candidates: A bridge between neural query translation and cross-lingual information retrieval. *arXiv preprint arXiv:2010.13658*.

Souradip Chakraborty, Jiahao Qiu, Hui Yuan, Alec Koppel, Furong Huang, Dinesh Manocha, Amrit Singh Bedi, and Mengdi Wang. 2024. Maxmin-rlhf: Alignment with diverse human preferences. *arXiv preprint arXiv:2402.08925*.

Jianlyu Chen, Shitao Xiao, Peitian Zhang, Kun Luo, Defu Lian, and Zheng Liu. 2024a. M3-embedding: Multi-linguality, multi-functionality, multi-granularity text embeddings through self-knowledge distillation. In *Findings of the Association for Computational Linguistics: ACL 2024*, pages 2318–2335, Bangkok, Thailand. Association for Computational Linguistics.

Jiawei Chen, Hongyu Lin, Xianpei Han, and Le Sun. 2024b. Benchmarking large language models in retrieval-augmented generation. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 38, pages 17754–17762.

Nadezhda Chirkova, David Rau, Hervé Déjean, Thibault Formal, Stéphane Clinchant, and Vassilina Nikoulina. 2024. Retrieval-augmented generation in multilingual settings. In *Proceedings of the 1st Workshop on Towards Knowledgeable Language Models (KnowLLM 2024)*, pages 177–188.

María Andrea Cruz Blandón, Jayasimha Talur, Bruno Charron, Dong Liu, Saab Mansour, and Marcello Federico. 2025. MEMERAG: A multilingual end-to-end meta-evaluation benchmark for retrieval augmented generation. In *Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 22577–22595, Vienna, Austria. Association for Computational Linguistics.

Josef Dai, Xuehai Pan, Ruiyang Sun, Jiaming Ji, Xinbo Xu, Mickel Liu, Yizhou Wang, and Yaodong Yang. 2023. Safe rlhf: Safe reinforcement learning from human feedback. *arXiv preprint arXiv:2310.12773*.

Ziyang Gao, Yuanliang Qu, and Yi Han. 2025. Cross-lingual sponsored search via dual-encoder and graph neural networks for context-aware query translation in advertising platforms. *arXiv preprint arXiv:2510.22957*.

Roksana Goworek, Olivia Macmillan-Scott, and Eda B Özyiğit. 2025. What drives cross-lingual ranking? retrieval approaches with multilingual language models. *arXiv preprint arXiv:2511.19324*.

Daya Guo, Dejian Yang, Haowei Zhang, Junxiao Song, Ruoyu Zhang, Runxin Xu, Qihao Zhu, Shirong Ma, Peiyi Wang, Xiao Bi, and 1 others. 2025. Deepseek-r1: Incentivizing reasoning capability in llms via reinforcement learning. *arXiv preprint arXiv:2501.12948*.

Bowen Jin, Hansi Zeng, Zhenrui Yue, Jinsung Yoon, Sercan Arik, Dong Wang, Hamed Zamani, and Jiawei Han. 2025. Search-r1: Training llms to reason and leverage search engines with reinforcement learning. *arXiv preprint arXiv:2503.09516*.

Dayeon Ki, Marine Carpuat, Paul McNamee, Daniel Khashabi, Eugene Yang, Dawn Lawrie, and Kevin Duh. 2025. Linguistic nepotism: Trading-off quality for language preference in multilingual rag. *arXiv preprint arXiv:2509.13930*.

Patrick Lewis, Ethan Perez, Aleksandra Piktus, Fabio Petroni, Vladimir Karpukhin, Naman Goyal, Heinrich Küttler, Mike Lewis, Wen-tau Yih, Tim Rocktäschel, and 1 others. 2020. Retrieval-augmented generation for knowledge-intensive nlp tasks. *Advances in neural information processing systems*, 33:9459–9474.

Bryan Li, Fiona Luo, Samar Haider, Adwait Agashe, Siyu Li, Runqi Liu, Miranda Muqing Miao, Shriya

Ramakrishnan, Yuan Yuan, and Chris Callison-Burch. 2025. Multilingual retrieval augmented generation for culturally-sensitive tasks: A benchmark for cross-lingual robustness. In *Findings of the Association for Computational Linguistics: ACL 2025*, pages 4215–4241, Vienna, Austria. Association for Computational Linguistics.

Hyeonseok Lim, Dongjae Shin, Seohyun Song, Inho Won, Minjun Kim, Junghun Yuk, Haneol Jang, and KyungTae Lim. 2025. VLR-bench: Multilingual benchmark dataset for vision-language retrieval augmented generation. In *Proceedings of the 31st International Conference on Computational Linguistics*, pages 6150–6168, Abu Dhabi, UAE. Association for Computational Linguistics.

Robert Litschko, Ekaterina Artemova, and Barbara Plank. 2023. Boosting zero-shot cross-lingual retrieval by training on artificially code-switched data. *arXiv preprint arXiv:2305.05295*.

Aixin Liu, Bei Feng, Bing Xue, Bingxuan Wang, Bochao Wu, Chengda Lu, Chenggang Zhao, Chengqi Deng, Chenyu Zhang, Chong Ruan, and 1 others. 2024. Deepseek-v3 technical report. *arXiv preprint arXiv:2412.19437*.

Wei Liu, Sony Trenous, Leonardo FR Ribeiro, Bill Byrne, and Felix Hieber. Xrag: Cross-lingual retrieval-augmented generation.

Shayne Longpre, Yi Lu, and Joachim Daiber. 2021. MKQA: A linguistically diverse benchmark for multilingual open domain question answering. *Transactions of the Association for Computational Linguistics*, 9:1389–1406.

Hoyeon Moon, Byeolhee Kim, and Nikhil Verma. 2025. Quality-aware translation tagging in multilingual rag system. In *Proceedings of the 5th Workshop on Multilingual Representation Learning (MRL 2025)*, pages 161–177.

Tong Mu, Alec Helyar, Johannes Heidecke, Joshua Achiam, Andrea Vallone, Ian Kivlichan, Molly Lin, Alex Beutel, John Schulman, and Lilian Weng. 2024. Rule based rewards for language model safety. *Advances in Neural Information Processing Systems*, 37:108877–108901.

Long Ouyang, Jeffrey Wu, Xu Jiang, Diogo Almeida, Carroll Wainwright, Pamela Mishkin, Chong Zhang, Sandhini Agarwal, Katarina Slama, Alex Ray, and 1 others. 2022. Training language models to follow instructions with human feedback. *Advances in neural information processing systems*, 35:27730–27744.

Jeonghyun Park and Hwanhee Lee. 2025. Investigating language preference of multilingual rag systems. *arXiv preprint arXiv:2502.11175*.

Rafael Rafailov, Archit Sharma, Eric Mitchell, Christopher D Manning, Stefano Ermon, and Chelsea Finn. 2023. Direct preference optimization: Your language model is secretly a reward model. *Advances in neural information processing systems*, 36:53728–53741.

Ori Ram, Yoav Levine, Itay Dalmedigos, Dor Muhlgay, Amnon Shashua, Kevin Leyton-Brown, and Yoav Shoham. 2023. In-context retrieval-augmented language models. *Transactions of the Association for Computational Linguistics*, 11:1316–1331.

Leonardo Ranaldi, Barry Haddow, and Alexandra Birch. 2025. Multilingual retrieval-augmented generation for knowledge-intensive task. *arXiv preprint arXiv:2504.03616*.

John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. 2017. Proximal policy optimization algorithms. *arXiv preprint arXiv:1707.06347*.

Nikhil Sharma, Kenton Murray, and Ziang Xiao. 2025. Faux polyglot: A study on information disparity in multilingual large language models. In *Proceedings of the 2025 Conference of the Nations of the Americas Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 1: Long Papers)*, pages 8090–8107, Albuquerque, New Mexico. Association for Computational Linguistics.

Shuaijie She, Wei Zou, Shujian Huang, Wenhao Zhu, Xiang Liu, Xiang Geng, and Jiajun Chen. 2024. Mapo: Advancing multilingual reasoning through multilingual alignment-as-preference optimization. *arXiv preprint arXiv:2401.06838*.

Zhiqing Sun, Sheng Shen, Shengcao Cao, Haotian Liu, Chunyuan Li, Yikang Shen, Chuang Gan, Liangyan Gui, Yu-Xiong Wang, Yiming Yang, and 1 others. 2024. Aligning large multimodal models with factually augmented rlhf. In *Findings of the Association for Computational Linguistics: ACL 2024*, pages 13088–13110.

Joseph P Telemala and Hussein Suleman. 2022. Language-preference-based re-ranking for multilingual swahili information retrieval. In *Proceedings of the 2022 ACM SIGIR International Conference on Theory of Information Retrieval*, pages 144–152.

Nandan Thakur, Luiz Bonifacio, Crystina Zhang, Odunayo Ogundepo, Ehsan Kamalloo, David Alfonso-Hermelo, Xiaoguang Li, Qun Liu, Boxing Chen, Mehdi Rezagholizadeh, and Jimmy Lin. 2024. "knowing when you don't know": A multilingual relevance assessment dataset for robust retrieval-augmented generation. In *Findings of the Association for Computational Linguistics: EMNLP 2024*, pages 12508–12526, Miami, Florida, USA. Association for Computational Linguistics.

Nandan Thakur, Suleman Kazi, Ge Luo, Jimmy Lin, and Amin Ahmad. 2025. Mirage-bench: Automatic multilingual benchmark arena for retrieval-augmented generation systems. In *Proceedings of the 2025 Conference of the Nations of the Americas Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 1: Long Papers)*, pages 274–298.

Hugo Touvron, Thibaut Lavril, Gautier Izacard, Xavier Martinet, Marie-Anne Lachaux, Timothée Lacroix, Baptiste Rozière, Naman Goyal, Eric Hambro, Faisal Azhar, and 1 others. 2023. Llama: Open and efficient foundation language models. *arXiv preprint arXiv:2302.13971*.

Jie Wu, Zhaochun Ren, and Suzan Verberne. 2024. What are the limits of cross-lingual dense passage retrieval for low-resource languages? *arXiv preprint arXiv:2408.11942*.

Boyang Xue, Weichao Wang, Hongru Wang, Fei Mi, Rui Wang, Yasheng Wang, Lifeng Shang, Xin Jiang, Qun Liu, and Kam-Fai Wong. 2023. Improving factual consistency for knowledge-grounded dialogue systems via knowledge enhancement and alignment. In *Findings of the Association for Computational Linguistics: EMNLP 2023*, pages 7829–7844.

An Yang, Anfeng Li, Baosong Yang, Beichen Zhang, Binyuan Hui, Bo Zheng, Bowen Yu, Chang Gao, Chengen Huang, Chenxu Lv, and 1 others. 2025. Qwen3 technical report. *arXiv preprint arXiv:2505.09388*.

Eugene Yang, Thomas Jänich, James Mayfield, and Dawn Lawrie. 2024. Language fairness in multilingual information retrieval. In *Proceedings of the 47th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 2487–2491.