

ParaRater: Enhancing Cross-Lingual Transfer in LLMs with Meta-Learning

Wenhan Han¹ Binbin Liu² Zhixun Chen³ Yifan Zhang²

Mykola Pechenizkiy¹ Meng Fang^{4,1*} Yin Zheng^{2*}

¹Eindhoven University of Technology ²ByteDance

³The Hong Kong University of Science and Technology ⁴University of Liverpool

w.han@tue.nl

Abstract

Multilingual LLMs increasingly claim broad language coverage, yet their performance still lags far behind English in many languages. Because low resource data is scarce and noisy, improving these languages directly is difficult, making cross lingual transfer from English a natural alternative. Parallel corpora are central to such transfer, but it remains unclear which kinds of parallel data matter most. We propose ParaRater, a corpus selection method that identifies high value English text to translate into target languages, building parallel corpora that most efficiently improve target language performance. ParaRater uses meta learning to tie selection directly to downstream accuracy on native target language data, and applies a two stage filter to find samples that only help when both language versions are used in training, capturing truly impactful parallel data. Experiments across Chinese, German, French, Japanese, Thai, Arabic, Tagalog, and Russian, together with qualitative analyses, support these findings.

1 Introduction

Multilingual large language models (LLMs) are rapidly proliferating and increasingly claim broad language coverage. Gemma 3 (Team et al., 2025) supports over 140 languages, and Qwen 3 (Yang et al., 2025) extends its predecessor by covering 119 languages and dialects. Closed-source models such as GPT-5 (OpenAI, 2025) and Claude-4 (Anthropic, 2025) also demonstrate strong multilingual capabilities. Despite these advances, substantial performance gaps remain between English and other languages, especially low-resource ones (Han et al., 2026). Prior methods for improving non-English performance mainly follow two directions: independently strengthening target-language ability by scaling and cleaning its training data (Zhao

et al., 2025; Dou et al., 2025), or transferring capabilities learned from abundant high-quality English data to other languages (Shen et al., 2025; Fujii et al., 2024; Lu et al., 2023; Qorib et al., 2025).

A key question remains unanswered: which parallel corpora drive target-language gains from English-originating training signals, and how can we identify them systematically? Parallel data is widely used for transfer, but most methods treat all pairs as equally helpful, relying on heuristics or post training adaptation, with gains observed mainly on translation tasks. Existing studies are largely empirical and focus on post training, leaving parallel corpus selection and its impact on general purpose language capabilities underexplored. Moreover, parallel corpora are not uniformly effective. Some English texts improve target language performance even without their translations, whereas others help only when both sides of the pair are seen during training. Distinguishing these cases is essential: only the latter constitutes the core parallel data that genuinely induces transfer. This motivates a principled framework to identify and exploit the most impactful subsets for efficient multilingual pretraining.

To address these challenges, we propose ParaRater, a meta-learning framework that directly aligns the value of parallel data with LLM performance on a target-language validation set.¹ ParaRater aims to identify the subset of parallel data that truly drives transfer, which we call the core of the parallel corpus. We define this core as the portion of source-language data that yields a positive impact on the target language only when paired with its translation. For example, for English-Chinese, we seek English documents that have little effect on Chinese when trained as English alone, but become beneficial once their Chi-

*Corresponding authors: Meng.Fang@liverpool.ac.uk, yzheng3xg@gmail.com.

¹We release our dataset and code at <https://pararater.github.io/>.

nese translations are included, such that the English side begins to contribute to Chinese capability.

We identify this core via a two-stage filtering strategy. Each stage trains a rater that assigns importance weights to documents based on their contribution to target-language performance. In Stage 1, the rater selects documents that receive high weights in both the source and target languages. In Stage 2, we remove the target-language versions from training and re-estimate weights, then select the source-language documents whose weights drop substantially. These are the core documents, whose benefits are genuinely parallel rather than arising from the monolingual side alone. This enables more efficient use of parallel data, either by extracting core subsets from existing parallel corpora or by selecting high-value English data to translate when building new corpora.

We conduct controlled pretraining experiments by training 1.2B-parameter LLMs from scratch on multiple bilingual pairs, showing that ParaRater consistently outperforms strong baselines. We further provide quantitative analyses and case studies of the selected data, offering new insights into cross-lingual transfer in LLMs.

To summarize, our contributions are as follows:

- We propose ParaRater, which leverages meta-learning and a two-stage strategy to separate genuinely effective parallel data from pseudo-parallel data, enabling more efficient exploitation of parallel corpora.
- We construct the ParaCore training set, consisting of core English-parallel data with several languages, which substantially improves non-English LLM performance.
- We develop ParaTool, a meta-learning framework for training parallel-data selectors, enabling practical identification of high-value data for cross-lingual transfer across languages and datasets.
- We conduct controlled experiments on multiple bilingual pairs with clean parallel data, and provide qualitative and quantitative analyses that clarify how parallel corpora contribute to multilingual pretraining.

2 Related Work

Cross-lingual Capability Transferring with Parallel Corpora Because high-quality non-English

data is much scarcer than English data, and aggressively oversampling low-resource languages can degrade English performance, many studies leverage parallel corpora to improve multilingual LLMs. Pretraining is a key stage in shaping general-purpose multilingual competence. Existing parallel-corpus methods address complementary questions. [Lu et al. \(2023\)](#) introduce TRIP, a trilingual grafting strategy for sequence-to-sequence pretraining; [Qorib et al. \(2025\)](#) study the format and timing for incorporating existing bilingual corpora; and [Shen et al. \(2025\)](#) study continued pretraining and instruction tuning with a multi-way TED corpus. These methods examine how available parallel corpora should be used, whereas ParaRater learns a document-level rule for selecting which source documents to translate under a fixed budget.

Data Curation High-quality data selection is central to LLM pretraining. Existing pipelines typically use heuristics or LLM-as-judge methods. Heuristic filtering removes duplicates, very short, or toxic text via rules and regular expressions ([Laurençon et al., 2023](#); [Weber et al., 2024](#); [Penedo et al., 2023](#); [Soldaini et al., 2024](#)), sometimes augmented with lightweight classifiers based on features such as perplexity or text type ([Chowdhery et al., 2022](#); [Touvron et al., 2023a](#); [Xie et al., 2023](#); [Muennighoff et al., 2023](#)). LLM-as-judge approaches score richer quality dimensions, e.g., QuRating ([Wettig et al., 2024](#)), FineWeb-Edu ([Penedo et al., 2024](#)), FIRE ([Xu et al., 2025](#)), and MuRating ([Chen et al., 2025](#)), but still depend on human-specified criteria. Learned selection policies have also been explored in active learning: [Fang et al. \(2017\)](#) train a reinforcement-learning policy that transfers across languages for named entity recognition. DataRater ([Calian et al., 2025](#)) instead learns a rater via meta-learning by directly optimizing for downstream pretraining effectiveness. Building on this idea, we adopt a meta-learning framework to study parallel corpora in LLM pretraining and train raters for principled data selection.

3 Preliminary: Data selection with meta-learning

Meta-learning for data selection trains a rater that scores data to minimize pretrained-model loss on a held-out validation dataset. Given a training corpus $\mathcal{D}_{\text{train}}$, a rater model R_η scores documents $d \in \mathcal{D}_{\text{train}}$. A language model M_θ is trained with the

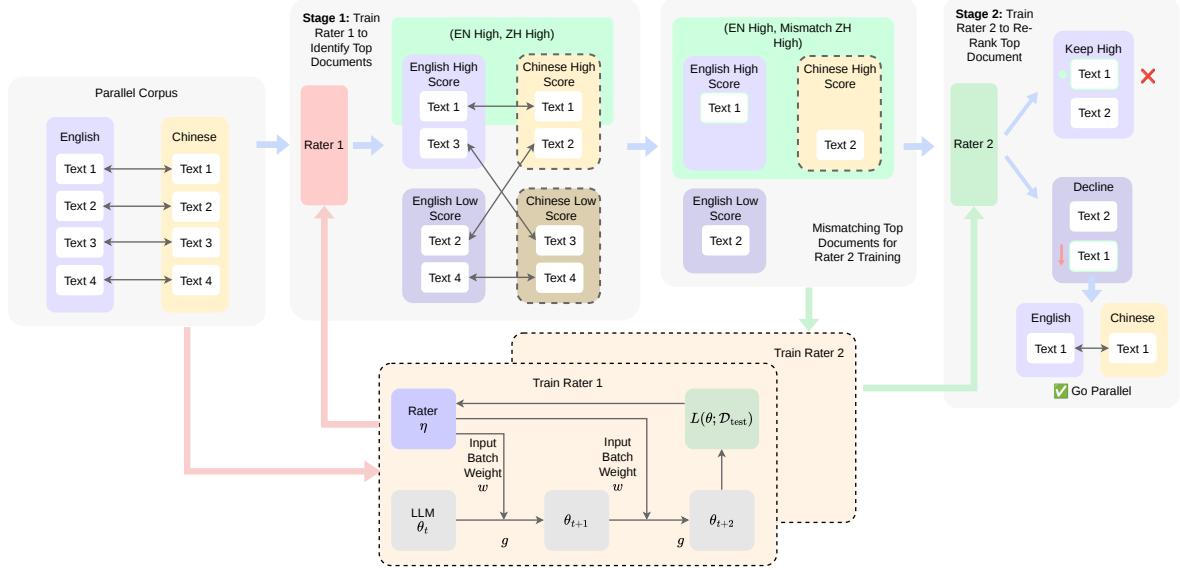


Figure 1: ParaRater uses two stages. Stage 1 trains Rater 1 to rank parallel documents that benefit both languages, then reshuffles the corpus by mismatching the top documents across languages. Stage 2 trains Rater 2 on this mismatched data to find source documents whose ranks drop without their true target counterparts. Those pairs are the highest value parallel data for pretraining.

curated data and a loss function l . The gradient of the language model in a step t is defined as:

$$g_t = \frac{1}{n} \sum_{i=1}^n \nabla_{\theta} l(d_i; \theta_t), \quad (1)$$

where the rater first scores documents and the selected or weighted batch is denoted as $d_i \in R_{\eta}(\mathcal{D}_{\text{train}})$ at step t . After T steps, the optimized parameters are denoted as $\theta^* = \theta(R_{\eta}(\mathcal{D}_{\text{train}}))$. The optimized language model is evaluated by a loss on held-out validation data $L(\theta_t; \mathcal{D}_{\text{val}})$. The optimization objective of the rater R is:

$$\eta^* = \arg \min_{\eta} L(\theta(R_{\eta}(\mathcal{D}_{\text{train}})); \mathcal{D}_{\text{val}}) \quad (2)$$

The task is to filter a predefined size of subset $\mathcal{D}'$, where the documents are highly effective for pre-training the language model. Appendix A provides the detailed meta-learning procedure.

4 ParaRater

Unlike standard data selection, ParaRater aims to extract a high value subset from source data to build target language corpora through parallel construction. This setting requires more than finding samples that improve the target language. We also need to separate the gains that come from the source side alone from the gains that rely on bilingual coupling in the paired data.

4.1 Parallel Corpus Filtering Problem

Given a training dataset in the source language $\mathcal{D}_s$ and a held-out validation dataset $\mathcal{D}_{\text{val}}$ in the target language, the rater selects a high-value subset of pre-determined size

$$\mathbf{d}_s = R(\mathcal{D}_s).$$

After the selection, $\mathbf{d}_s$ is used to build its target language version $\mathbf{d}_t$, forming the parallel corpus. Here, “parallel corpus” refers to dataset composition: source and target documents are both included in standard causal language modeling, but not concatenated into explicit bilingual prompts. We rewrite Eq. 1 as

$$g = \nabla_{\theta} l(\mathbf{d}_s \cup \mathbf{d}_t; \theta), \quad (3)$$

where we omit the step annotation for simplicity. Following the approach in section 3, we use the Eq. 2 to optimize the rater, with the $R(\mathcal{D}_s)$ replaced by $\mathbf{d}_s \cup \mathbf{d}_t$.

However, this objective does not guarantee that the constructed subset achieves the highest efficiency for language capability transfer. Parallel corpus that can drive effective cross-lingual transfer should manifest in the way that the two language versions of the same data mutually influence their effectiveness during model training. What we seek is to identify cases where the combination of

two languages produces high value when they appear together, rather than data that is already highly valuable in a monolingual setting. Therefore, we need to introduce additional constraints. The final optimization objective for the rater is then as follows:

$$\eta^* = \arg \min_{\eta} \left(L(\theta(\mathbf{d}_s \cup \mathbf{d}_t)) - L(\theta(\mathbf{d}_s \cup \mathbf{d}'_t)) \right), \quad (4)$$

where $\mathbf{d}'_t$ denotes the target-language counterpart that is mismatched with $\mathbf{d}_s$ and $\mathbf{d}'_t$ satisfies $L(\theta(\mathbf{d}_t)) = L(\theta(\mathbf{d}'_t))$, i.e., $\mathbf{d}'_t$ is of equal quality to $\mathbf{d}_t$. We omit the explicit dependence of L on $\mathcal{D}_{\text{val}}$. In this objective, the selected subset of source corpus only contributes to the language model performance when the model is jointly trained on the subset and its target-language counterpart.

4.2 Dual-rater cross filtering

Since the selection from the corpus is a discrete sampling process, directly optimizing Eq. 4 is a nondeterministic polynomial-time hard problem. Inspired by DataRater (Calian et al., 2025), we instead optimize the rater by assigning continuous weight values to batches of training samples. Moreover, directly optimizing Eq. 4 would require maintaining multiple language models and computing their gradients simultaneously, which is computationally sophisticated. Therefore, we adopt a two-stage strategy, optimizing two raters to cross-filter the data as an approximation. Figure 1 depicts the two stages.

Stage 1: High-Value Bilingual Subset Selection

At this stage, we train rater R_{η_1} to optimize the first term of Eq. 4, in order to select the documents that get high rank in language model pretraining in both languages. Here we have

$$\eta_1 = \arg \min_{\eta} \left(L(\theta(\tilde{\mathbf{d}}_s \cup \tilde{\mathbf{d}}_t)) \right),$$

where $\tilde{\mathbf{d}}_s$ and $\tilde{\mathbf{d}}_t$ denote the same coarse-grained high-value source and target subsets produced from $\mathbf{d}_s$ and $\mathbf{d}_t$. In practice, we first construct parallel data $\mathcal{D}_t$ in the target language from source data $\mathcal{D}_s$. The language model is trained on independently mixed source and target documents, and the rater meta-learns from the validation loss $L(\cdot; \mathcal{D}_{\text{val}})$. We then score $\mathcal{D}_t$ and $\mathcal{D}_s$ independently and select samples that rank high in both languages. Thus, Stage 1 is a standard meta-learning procedure over parallel corpus members, guided by target-language performance.

Stage 2: Low-Value Monolingual Source Language Subset Selection

The bilingual data found in Stage 1 may still improve each language independently rather than through transfer. Parallel data should enable interoperability: learning content in one language should interact with the model’s understanding of the same content in the other. For example, if a Chinese document improves Chinese performance, its English version should also help; in rater meta-learning, this appears as the English document’s score being influenced by the presence of its Chinese counterpart.

At this stage, our goal is to identify source-language data that improves model performance only when its target-language counterpart is also present. To do this, we start with the source-language documents that ranked highly in both languages in the previous stage. We then construct a new training set by combining: **1)** these high-ranking source documents, **2)** high-ranking target-language documents that are not parallel to them and **3)** a small subset of low-ranking source documents (as a comparison group).

We meta learn a second rater on this mixture and track how the top ranked source documents shift. A large rank drop implies the document is valuable mainly when paired with its target translation. Such pairs define the core parallel corpus, exposing the strongest cross lingual dependency.

Formally, a new rater R_{η_2} is meta-trained on the high-value source language subset $\tilde{\mathbf{d}}_s$, the high-value mismatched target language subset $\tilde{\mathbf{d}}'_t$, and the low-value source language subset $\tilde{\mathbf{d}}'_s$. Following Eq. 3 the language model optimization is denoted as

$$g = \nabla_{\theta} l(\tilde{\mathbf{d}}_s^* \cup \tilde{\mathbf{d}}'_t \cup \tilde{\mathbf{d}}'_s; \theta), \quad (5)$$

where $\tilde{\mathbf{d}}_s^* = R_{\eta_2}(\tilde{\mathbf{d}}_s)$, $\tilde{\mathbf{d}}'_t = R_{\eta_2}(\tilde{\mathbf{d}}'_t)$, and $\tilde{\mathbf{d}}'_s = R_{\eta_2}(\tilde{\mathbf{d}}'_s)$ are the subsets selected by Rater 2. Note again that this is an iterative process, and for simplicity we have omitted the step indices. The rater optimization is then defined as:

$$\eta_2 = \arg \min_{\eta} \left(L(\theta(\tilde{\mathbf{d}}_s^* \cup \tilde{\mathbf{d}}'_t \cup \tilde{\mathbf{d}}'_s)) \right).$$

$\tilde{\mathbf{d}}_s^*$ represents the source-language samples that remain highly valuable despite lacking a corresponding target-language counterpart. Since these samples already enhance the language model’s capability in the target language on their own, there is no need to use them as seeds for constructing

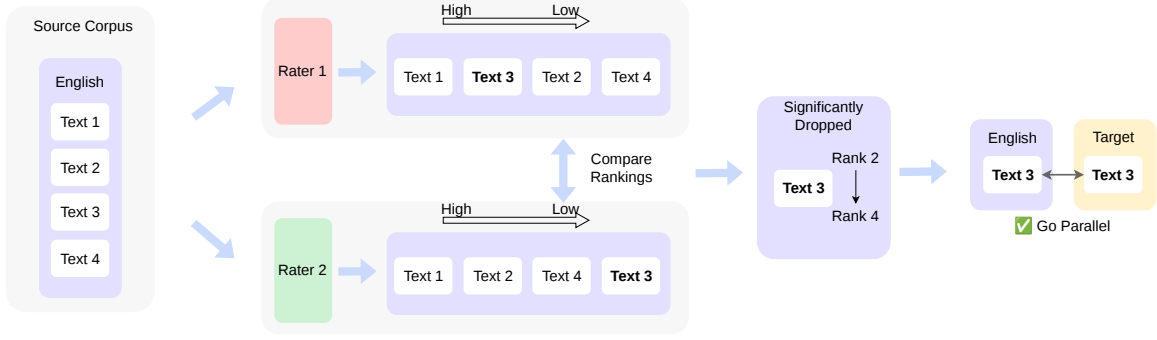


Figure 2: Source-only selection. Score the source-language data using Rater 1 and Rater 2, compare the differences in their scores, and identify the documents with significant drops to construct the core of parallel data.

additional parallel data. Our focus is on

$$\mathbf{d}_s = \tilde{\mathbf{d}}_s - \tilde{\mathbf{d}}_s^*,$$

which is the core of the parallel corpus: source-language data whose usefulness relies on the presence of their parallel target-language counterparts.

Source-only Selection With both raters, we can filter an existing parallel corpus to extract its core. In practice, parallel data are scarce and large-scale translation is costly, so the more useful setting is to first select a high-value subset from the source corpus and then translate it to form core parallel data. ParaRater supports this source-only workflow, as shown in Figure 2.

The only change is that we do not observe target-side rankings before translation. Each rater still scores individual documents rather than inferring whether a paired counterpart exists. Stage 1 therefore serves as a source-side shortlist, while final selection depends on the score drop between raters. It can admit source documents that would score poorly on the target side, mostly due to faulty translations; improving translation quality can reduce this issue.

5 Experiments

In this section, we conduct experiments under the Source-only Selection setting to evaluate the effectiveness of ParaRater in improving multilingual capabilities.

5.1 Setup

Dataset Since naturally occurring parallel data are relatively scarce, we use machine-translated parallel data when training the rater. Unlike some previous work, we do not use concatenated parallel pairs in order to eliminate potential bias introduced

by the data format; instead, all parallel documents are distributed independently within the training data. The experiment is conducted on multilingual data drawn from **Fineweb-2** (Penedo et al., 2025). For rater training, we sample 20B tokens from English corpora and translate them into 8 languages — Chinese, German, French, Japanese, Thai, Arabic, Tagalog and Russian — thereby ensuring coverage across multiple linguistic typology dimensions. Translations are generated using the Qwen3-8B² model, and documents with COMET (Rei et al., 2020) scores above 0.8 are retained to control and maintain high translation quality. We sample 2B tokens from the native corpus of each of the nine target languages to serve as the validation set in the meta-learning process.

To evaluate ParaRater, we use the trained ParaRater for each language to source-filter FineWeb English data, selecting 1B tokens per language. These samples are translated with the same pipeline used for all baselines to form the core parallel corpus. We then add 100B English tokens and pre-train a 1.2B-parameter LLaMA-style model. This unbalanced design simulates target-language data scarcity and evaluates the effect of transferring knowledge from English through translated parallel data. Appendix C reports implementation details.

Baselines We compare ParaRater to other data selection methods, including **FineWeb-Edu Classifier** (Penedo et al., 2024), which is trained on LLM labeled data to identify educational content, and **DCLM** (Li et al., 2024), a FastText classifier trained to distinguish high-quality data from raw web data. We also include random sampling from the same pool as an uncured baseline. All methods select 1B English tokens and share the same

²<https://huggingface.co/Qwen/Qwen3-8B>

Table 1: Model performance trained on parallel data selected by different methods.

	DCLM					FineWeb-Edu					ParaRater				
ARC-E	EN	ZH	JA	FR	TH	EN	ZH	JA	FR	TH	EN	ZH	JA	FR	TH
	58.41	48.28	45.74	44.89	39.85	59.52	54.85	49.43	48.28	45.36	62.23	49.81	45.82	45.82	41.42
	TL	DE	RU	AR		TL	DE	RU	AR		TL	DE	RU	AR	
	41.08	44.47	42.43	38.83		42.31	47.99	45.70	42.94		39.93	46.20	43.92	38.58	
MMMLU	EN	ZH	JA	FR	TH	EN	ZH	JA	FR	TH	EN	ZH	JA	FR	TH
	28.93	27.34	27.10	27.40	26.59	29.66	27.88	27.74	27.98	27.28	29.80	27.63	26.97	27.41	26.78
	TL	DE	RU	AR		TL	DE	RU	AR		TL	DE	RU	AR	
	27.58	27.21	26.55	25.89		28.11	27.33	26.65	26.12		28.38	27.33	26.90	26.33	
BMLAMA	EN	ZH	JA	FR	TH	EN	ZH	JA	FR	TH	EN	ZH	JA	FR	TH
	69.35	43.42	36.15	55.22	32.80	69.28	46.43	40.39	57.23	34.96	70.04	44.95	41.06	57.85	39.30
	TL	DE	RU	AR		TL	DE	RU	AR		TL	DE	RU	AR	
	63.88	60.11	35.24	38.55		64.03	63.07	39.48	41.29		63.91	62.81	39.78	43.22	
HellaSwag	EN	ZH	JA	FR	TH	EN	ZH	JA	FR	TH	EN	ZH	JA	FR	TH
	53.51	36.28	32.09	38.13	32.30	55.15	38.32	32.80	40.14	33.71	55.13	40.85	35.65	42.29	34.11
	TL	DE	RU	AR		TL	DE	RU	AR		TL	DE	RU	AR	
	36.47	34.14	36.17	33.18		38.62	36.26	37.85	35.50		39.22	37.35	40.37	37.01	
StoryCloze	EN	ZH	JA	FR	TH	EN	ZH	JA	FR	TH	EN	ZH	JA	FR	TH
	71.98	60.29	59.06	62.69	57.35	73.45	64.16	59.67	64.78	58.20	73.30	65.25	62.00	66.25	60.14
	TL	DE	RU	AR		TL	DE	RU	AR		TL	DE	RU	AR	
	56.04	62.38	61.53	54.41		56.58	64.09	64.40	56.50		57.04	65.94	65.40	58.36	
XWinograd	EN	ZH	JA	FR	RU	EN	ZH	JA	FR	RU	EN	ZH	JA	FR	RU
	73.12	54.96	54.01	55.42	54.29	78.37	58.93	54.01	57.83	53.65	78.97	59.92	55.16	59.03	58.41
Flores		ZH	JA	FR	TH		ZH	JA	FR	TH		ZH	JA	FR	TH
		8.08	8.40	18.54	10.56		10.30	9.54	24.81	12.53		11.05	9.93	29.89	14.41
	TL	DE	RU	AR		TL	DE	RU	AR		TL	DE	RU	AR	
	31.53	19.06	15.35	14.55		35.15	23.12	17.24	17.17		36.75	30.15	23.76	19.11	

Table 2: Comparison between ParaRater Stage 1 and the full two-stage ParaRater pipeline. Results are averaged across all evaluated languages.

	ARC-E	BMLAMA	MMLU	HellaSwag	StoryCloze	XWinoGrad	Flores
Stage1	46.44	52.15	27.44	40.01	62.70	62.30	20.13
Stage2	45.97	51.44	27.50	40.22	63.74	62.30	21.88

translation, pretraining pipeline, and 100B-token English mixture.

Evaluation We evaluate the pre-trained language models on both English and the target languages. For English evaluation, we adopt **ARC-E** (Clark et al., 2018), **Hellaswag** (Zellers et al., 2019), **StoryCloze** (Mostafazadeh et al., 2016), **BMLAMA** (Qi et al., 2023), **MMLU** (Hendrycks et al., 2021), covering multiple capabilities ranging from natural language understanding, commonsense reasoning, and factoid knowledge. For evaluation on the target languages, the widely used multilingual version of **Hellaswag** (Lai et al., 2023), **StoryCloze** (Lin et al., 2022) and **MMMLU**³ are adopted. For **BMLAMA**, we observed that the original multilingual version contained a substantial number of errors, so we used the MuBench (Han et al., 2026) version instead. For the languages not covered by the above

multilingual version, we also supplement them using MuBench. These target-language benchmarks measure multilingual generalization, while **FLORES** (Goyal et al., 2022) ChrF++ scores explicitly evaluate EN-to-target translation capabilities.

5.2 Results

Table 1 shows the result. ParaRater demonstrates clear advantages on most benchmarks, covering NLU tasks (e.g., HellaSwag, StoryCloze, xWinoGrad), factual knowledge tasks, and translation, aligning with prior observations that parallel corpora boost translation performance. Compared with other data selection approaches for constructing parallel corpora, ParaRater achieves higher data efficiency, delivering notable gains using only 1B tokens per language.

All runs share the same 100B-token English component and differ only in which 1B tokens populate the selected-and-translated component,

³<https://huggingface.co/datasets/openai/MMMLU>

Benchmark	Random	ParaRater
ARC-E	44.39	45.97
MMMLU	27.14	27.50
BMLAMA	50.19	51.44
HellaSwag	38.64	40.22
StoryCloze	61.96	63.74
XWinograd	51.45	62.30
Flores	15.42	21.88

Table 3: Average performance of random sampling and ParaRater across target languages under the same data and training budget.

so fewer than 1% of training tokens change. Under this stringent control, ParaRater improves over random sampling on all seven aggregate metrics in Table 3. Consistency across heterogeneous tasks is therefore more informative here than a large gain on every individual benchmark.

ParaRater underperforms on science-oriented tasks (e.g., ARC-E), likely because we use multilingual FineWeb-2 as validation data, giving ParaRater a broader optimization target, whereas FineWeb-Edu explicitly selects educational content and is advantaged on science-focused benchmarks. Nevertheless, ParaRater still delivers measurable English improvement, suggesting that its selection process filters noisy data from low-quality translations and prevents negative impacts on English capabilities.

5.3 Ablation Study: Stage 1

We conduct an ablation study by removing Stage 2 to investigate the effect of using only Rater 1. Following the same setup, Rater 1 selects 1B tokens per language, which are combined with 100B English tokens for training. Table 2 shows that Stage 2 changes the performance profile rather than uniformly improving every task: it improves HellaSwag, StoryCloze, and Flores, while Stage 1 is slightly stronger on ARC-E and BMLAMA. Stage 1 mainly filters low-quality texts and reduces distributional noise; Stage 2 maintains marginal distributions while perturbing the joint distribution between languages to identify strongly interacting parallel corpora. Since both raters share the same validation set, Stage 2 tends to bias selection toward language understanding rather than knowledge acquisition, making validation set design a promising direction.

6 Case Study and Discussion

In stage 1, the rater assigns scores to the bilingual corpus, which are then ranked separately within each language. We designate the top 30% as high-ranked and the bottom 30% as low-ranked, yielding four cross-lingual ranking combinations. The upper section of Table 4 illustrates representative cases from these four categories for Chinese–English data. We find that the four categories display clear patterns. In particular, documents ranked highly in both Chinese and English tend to be fluent, knowledge-bearing explanatory texts with relatively greater length. For documents that receive high rankings in English but low rankings in Chinese ($\text{EN} \uparrow \text{ZH} \downarrow$), we observe that many instances, consistent with our earlier discussion, are attributable to poor or incorrect translations. In the case, “are king” is a figurative expression meaning “to dominate” or “to be the most important or influential thing.” However, in Chinese this was translated literally as “king,” which makes the sentence sound very confusing. Therefore, improving translation quality would reduce the number of samples falling into this category and potentially shift them toward the first category ($\text{EN} \uparrow \text{ZH} \uparrow$). In the category where English ranks low but Chinese ranks high ($\text{EN} \downarrow \text{ZH} \uparrow$), we observe that most samples pertain to topics such as Western geography, institutions, and culture. However, the domains of these texts may be relatively distant from the distribution of the Chinese corpus, making the original English texts less effective in improving the model’s Chinese capability. Nevertheless, these texts are generally fluent, and thus, once translated into Chinese, they may be recognized as high-quality Chinese data. For data that receive low scores in both languages ($\text{EN} \downarrow \text{ZH} \downarrow$), we generally observe issues with fluency and completeness; the corresponding English texts are often disorganized and difficult to understand.

Stage 2 performs a decomposition of the subset that ranked highly in both languages ($\text{EN} \uparrow \text{ZH} \uparrow$). In Table 4, we categorize samples whose rankings drop by more than 20% as having a significant decline ($\text{EN} \downarrow$); otherwise, they are considered to have maintained their ranking ($\text{EN} \uparrow$). In Stage 2, the samples that maintain high rankings ($\text{EN} \uparrow$) tend to be more well-structured and formal, with overall very fluent translations. In contrast, the samples whose rankings drop ($\text{EN} \downarrow$), although still high-quality in English, often have translations

Table 4: Samples classified according to ParaRater rankings in Stage 1 and Stage 2. In stage 1, $\uparrow$ and $\downarrow$ denotes high and low rank respectively. In Stage 2, $\uparrow$ denotes samples that keep the rank, whereas $\downarrow$ denotes samples whose ranks have declined.

Category	English	Chinese
<i>Stage 1</i>		
EN $\uparrow$ ZH $\uparrow$	Amplifier circuits can be described in terms of the layout or topology of components which plays a crucial part in shaping the sound ... using a lead made from this between guitar and amp, while not dangerous, will usually be very noisy.	放大器电路可以用元件的布局或拓扑来描述，这在塑造声音方面起着至关重要的作用。大多数经典的电吉他放大器采用A类设计 ... 使用这种线材连接吉他和音箱，虽然不会有危险，但通常会很嘈杂。
EN $\uparrow$ ZH $\downarrow$	It's 1928 in Harlem, New York. Jazz and gangsters are king in an interesting lifestyle. Richard Gere plays Dixie Dwyer, a musician who witnesses a murder. ...	1928年，纽约哈莱姆。爵士乐和黑帮是一个有趣的生活方式的国王。理查德·吉尔饰演迪克西·德怀尔，目击了一起谋杀案。 ...
EN $\downarrow$ ZH $\uparrow$	The "Foothill Village" of Sierra Madre is a small, quaint and friendly town with storefronts reminiscent of the 1920s and 1930s, and is often thought of as an "artist" colony" ...	"塞拉玛德拉的"山麓村"是一个小，古朴而友好的小镇，店面令人联想到20世纪20年代和30年代，并经常被认为是一个"艺术家的殖民地"。超过40%的城镇的房屋超过50岁 ...
EN $\downarrow$ ZH $\downarrow$	Materials: Chart paper with the riddle: "Should Sheri share her shoes with her sister Shelly?" on it. Class set of the book by Dr. Seuss, One Fish, Two Fish, Red Fish, Blue Fish.	材料：一张纸上写着谜语："Sheri应该和她的妹妹 Shelly 分享她的鞋吗？" Dr. Seuss 的书，《一只鱼，两只鱼，红鱼，蓝鱼》的班级套书。
<i>Stage 2</i>		
EN $\downarrow$	People often message me on my blog with questions about why their cakes have gone wrong and how they can prevent it happening again. ...	人们经常在我的博客上给我留言，问为什么他们的蛋糕出了问题，以及如何防止再次发生。 ...
EN $\uparrow$	Over the past century, the neighborhood of Little Italy, Manhattan, has evolved from a cultural center for Italian immigrants to a tourist destination with very little Italian culture at all. ...	在过去的一个世纪里，曼哈顿的小意大利区从意大利移民的文化中心演变为一个几乎没有意大利文化的旅游目的地 ...

that appear somewhat stiff or unnatural. While more regular and polished text generally signals higher quality, it also reduces diversity to some extent. We take a reverse perspective to analyze the factors contributing to these translation quality outcomes. Fluently translated text is easier for the model to learn — after all, a translation model is itself an LLM. Consequently, the samples that drop in ranking in Stage 2 are harder to learn in the target language, and thus require the presence of parallel data to enable effective transfer.

ParaRater can either extract the core from existing parallel corpora or pre-select source documents for targeted translation, maximizing corpus construction efficiency and cross-lingual transfer. Through meta-learning, it learns language-specific

selection strategies, providing an adaptive, fine-grained alternative to fixed heuristics and rigid filtering rules.

7 Conclusion

This paper introduces ParaRater for selecting and constructing parallel corpora. By comparing score differences between two raters, ParaRater isolates parallel pairs with strong cross-lingual coupling and improves transfer efficiency in LLM training. Experiments from English to eight target languages show consistent gains on NLU and translation while preserving English performance. Future work includes better choices and configurations of ParaRater training objectives.

Limitations

While ParaRater offers an initial step toward analyzing which parts of parallel data matter most for cross lingual transfer, several aspects remain to be improved. Our experiments cover English paired with eight target languages under specific model sizes, data mixtures, and training budgets. It remains to be validated in larger scale experiments or across a broader set of languages. In addition, we do not apply strict curation to the validation set, and how to select the validation set is an important direction for optimization. Finally, currently ParaRater needs meta learning and training raters per language pair. Since ParaRater is trained per English-target pair, the same English source can be selected for multiple target languages when it provides broadly useful transfer; joint multilingual selection is a natural next step. We hope future work will explore more efficient training, broader language coverage, and better objective design for ParaRater.

References

- Anthropic. 2025. [Claude 4](#).
- Dan A. Calian, Gregory Farquhar, Iurii Kemaev, Luisa M. Zintgraf, Matteo Hessel, Jeremy Shar, Junhyuk Oh, András György, Tom Schaul, Jeffrey Dean, Hado van Hasselt, and David Silver. 2025. [DataRater: Meta-Learned Dataset Curation](#). *Preprint*, arXiv:2505.17895.
- Zhixun Chen, Ping Guo, Wenhan Han, Yifan Zhang, Binbin Liu, Haobin Lin, Fengze Liu, Yan Zhao, Bingni Zhang, Taifeng Wang, Yin Zheng, and Meng Fang. 2025. [MuRating: A High Quality Data Selecting Approach to Multilingual Large Language Model Pretraining](#). *Preprint*, arXiv:2507.01785.
- Aakanksha Chowdhery, Sharan Narang, Jacob Devlin, Maarten Bosma, Gaurav Mishra, Adam Roberts, Paul Barham, Hyung Won Chung, Charles Sutton, Sebastian Gehrmann, Parker Schuh, Kensen Shi, Sasha Tsvyashchenko, Joshua Maynez, Abhishek Rao, Parker Barnes, Yi Tay, Noam Shazeer, Vinodkumar Prabhakaran, and 48 others. 2022. [PaLM: Scaling Language Modeling with Pathways](#). *Preprint*, arXiv:2204.02311.
- Peter Clark, Isaac Cowhey, Oren Etzioni, Tushar Khot, Ashish Sabharwal, Carissa Schoenick, and Oyvind Tafjord. 2018. [Think you have Solved Question Answering? Try ARC, the AI2 Reasoning Challenge](#). *Preprint*, arXiv:1803.05457.
- Longxu Dou, Qian Liu, Fan Zhou, Changyu Chen, Zili Wang, Ziqi Jin, Zichen Liu, Tongyao Zhu, Cunxiao Du, Penghui Yang, Haonan Wang, Jiaheng Liu, Yongchi Zhao, Xiachong Feng, Xin Mao, Man Tsung Yeung, Kunat Pipatanakul, Fajri Koto, Min Si Thu, and 22 others. 2025. [Sailor2: Sailing in South-East Asia with Inclusive Multilingual LLMs](#). *Preprint*, arXiv:2502.12982.
- Meng Fang, Yuan Li, and Trevor Cohn. 2017. [Learning how to active learn: A deep reinforcement learning approach](#). In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, pages 595–605, Copenhagen, Denmark. Association for Computational Linguistics.
- Kazuki Fujii, Taishi Nakamura, Mengsay Loem, Hiroki Iida, Masanari Ohi, Kakeru Hattori, Hirai Shota, Sakae Mizuki, Rio Yokota, and Naoaki Okazaki. 2024. [Continual Pre-Training for Cross-Lingual LLM Adaptation: Enhancing Japanese Language Capabilities](#). *Preprint*, arXiv:2404.17790.
- Naman Goyal, Cynthia Gao, Vishrav Chaudhary, Peng-Jen Chen, Guillaume Wenzek, Da Ju, Sanjana Krishnan, Marc’Aurelio Ranzato, Francisco Guzmán, and Angela Fan. 2022. [The Flores-101 Evaluation Benchmark for Low-Resource and Multilingual Machine Translation](#). *Transactions of the Association for Computational Linguistics*, 10:522–538.
- Wenhan Han, Yifan Zhang, Zhixun Chen, Binbin Liu, Mykola Pechenizkiy, Meng Fang, and Yin Zheng. 2026. [MuBench: Assessment of Multilingual Capabilities of Large Language Models Across 61 Languages](#). In *Findings of the Association for Computational Linguistics: ACL 2026*, pages 16163–16192, San Diego, California, United States. Association for Computational Linguistics.
- Dan Hendrycks, Collin Burns, Steven Basart, Andy Zou, Mantas Mazeika, Dawn Song, and Jacob Steinhardt. 2021. [Measuring Massive Multitask Language Understanding](#). *Preprint*, arXiv:2009.03300.
- Viet Lai, Chien Nguyen, Nghia Ngo, Thuat Nguyen, Franck Dernoncourt, Ryan Rossi, and Thien Nguyen. 2023. [Okapi: Instruction-tuned Large Language Models in Multiple Languages with Reinforcement Learning from Human Feedback](#). In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*, pages 318–327, Singapore. Association for Computational Linguistics.
- Hugo Laurençon, Lucile Saulnier, Thomas Wang, Christopher Akiki, Albert Villanova del Moral, Teven Le Scao, Leandro Von Werra, Chenghao Mou, Eduardo González Ponferrada, Huu Nguyen, Jörg Froberg, Mario Šaško, Quentin Lhoest, Angelina McMillan-Major, Gerard Dupont, Stella Biderman, Anna Rogers, Loubna Ben allal, Francesco De Toni, and 35 others. 2023. [The BigScience ROOTS Corpus: A 1.6TB Composite Multilingual Dataset](#). *Preprint*, arXiv:2303.03915.
- Jeffrey Li, Alex Fang, Georgios Smyrnis, Maor Ivgi, Matt Jordan, Samir Gadre, Hritik Bansal, Etash Guha,

- Sedrick Keh, Kushal Arora, Saurabh Garg, Rui Xin, Niklas Muennighoff, Reinhard Heckel, Jean Mercat, Mayee Chen, Suchin Gururangan, Mitchell Wortsman, Alon Albalak, and 40 others. 2024. [DataComp-LM: In search of the next generation of training sets for language models](#). *Preprint*, arXiv:2406.11794.
- Xi Victoria Lin, Todor Mihaylov, Mikel Artetxe, Tianlu Wang, Shuohui Chen, Daniel Simig, Myle Ott, Naman Goyal, Shruti Bhosale, Jingfei Du, Ramakanth Pasunuru, Sam Shleifer, Punit Singh Koura, Vishrav Chaudhary, Brian O’Horo, Jeff Wang, Luke Zettlemoyer, Zornitsa Kozareva, Mona Diab, and 2 others. 2022. [Few-shot Learning with Multilingual Language Models](#). *Preprint*, arXiv:2112.10668.
- Hongyuan Lu, Haoyang Huang, Shuming Ma, Dongdong Zhang, Wai Lam, and Furu Wei. 2023. [Advancing Multilingual Pre-training: TRIP Triangular Document-level Pre-training for Multilingual Language Models](#). *Preprint*, arXiv:2212.07752.
- Nasrin Mostafazadeh, Nathanael Chambers, Xiaodong He, Devi Parikh, Dhruv Batra, Lucy Vanderwende, Pushmeet Kohli, and James Allen. 2016. [A Corpus and Evaluation Framework for Deeper Understanding of Commonsense Stories](#). *Preprint*, arXiv:1604.01696.
- Niklas Muennighoff, Alexander M. Rush, Boaz Barak, Teven Le Scao, Aleksandra Piktus, Nouamane Tazi, Sampo Pyysalo, Thomas Wolf, and Colin Raffel. 2023. [Scaling Data-Constrained Language Models](#). *Preprint*, arXiv:2305.16264.
- OpenAI. 2025. [Introducing GPT-5](#).
- Guilherme Penedo, Hynek Kydlíček, Loubna Ben al-lal, Anton Lozhkov, Margaret Mitchell, Colin Raffel, Leandro Von Werra, and Thomas Wolf. 2024. [The FineWeb Datasets: Decanting the Web for the Finest Text Data at Scale](#). *Preprint*, arXiv:2406.17557.
- Guilherme Penedo, Hynek Kydlíček, Vinko Sabolčec, Bettina Messmer, Negar Foroutan, Amir Hossein Kargaran, Colin Raffel, Martin Jaggi, Leandro Von Werra, and Thomas Wolf. 2025. [FineWeb2: One Pipeline to Scale Them All – Adapting Pre-Training Data Processing to Every Language](#). *Preprint*, arXiv:2506.20920.
- Guilherme Penedo, Quentin Malartic, Daniel Hesslow, Ruxandra Cojocaru, Hamza Alobeidli, Alessandro Cappelli, Baptiste Pannier, Ebtesam Almazrouei, and Julien Launay. 2023. The RefinedWeb Dataset for Falcon LLM: Outperforming Curated Corpora with Web Data Only. *Advances in Neural Information Processing Systems*, 36:79155–79172.
- Jirui Qi, Raquel Fernández, and Arianna Bisazza. 2023. [Cross-Lingual Consistency of Factual Knowledge in Multilingual Language Models](#). In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, pages 10650–10666, Singapore. Association for Computational Linguistics.
- Muhammad Reza Qorib, Junyi Li, and Hwee Tou Ng. 2025. [Just Go Parallel: Improving the Multilingual Capabilities of Large Language Models](#). In *Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 33411–33424, Vienna, Austria. Association for Computational Linguistics.
- Ricardo Rei, Craig Stewart, Ana C. Farinha, and Alon Lavie. 2020. [COMET: A Neural Framework for MT Evaluation](#). *Preprint*, arXiv:2009.09025.
- Yingli Shen, Wen Lai, Shuo Wang, Kangyang Luo, Alexander Fraser, and Maosong Sun. 2025. [From Unaligned to Aligned: Scaling Multilingual LLMs with Multi-Way Parallel Corpora](#). *Preprint*, arXiv:2505.14045.
- Luca Soldaini, Rodney Kinney, Akshita Bhagia, Dustin Schwenk, David Atkinson, Russell Authur, Ben Bogin, Khyathi Chandu, Jennifer Dumas, Yanai Elazar, Valentin Hofmann, Ananya Jha, Sachin Kumar, Li Lucy, Xinxu Lyu, Nathan Lambert, Ian Magnusson, Jacob Morrison, Niklas Muennighoff, and 17 others. 2024. [Dolma: An Open Corpus of Three Trillion Tokens for Language Model Pretraining Research](#). In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 15725–15788, Bangkok, Thailand. Association for Computational Linguistics.
- Gemma Team, Aishwarya Kamath, Johan Ferret, Shreya Pathak, Nino Vieillard, Ramona Merhej, Sarah Perrin, Tatiana Matejovicova, Alexandre Ramé, Morgane Rivière, Louis Rouillard, Thomas Mesnard, Geoffrey Cideron, Jean-bastien Grill, Sabela Ramos, Edouard Yvinec, Michelle Casbon, Etienne Pot, Ivo Penchev, and 197 others. 2025. [Gemma 3 Technical Report](#). *Preprint*, arXiv:2503.19786.
- Hugo Touvron, Thibaut Lavril, Gautier Izacard, Xavier Martinet, Marie-Anne Lachaux, Timothée Lacroix, Baptiste Rozière, Naman Goyal, Eric Hambro, Faisal Azhar, Aurelien Rodriguez, Armand Joulin, Edouard Grave, and Guillaume Lample. 2023a. [LLaMA: Open and Efficient Foundation Language Models](#). *Preprint*, arXiv:2302.13971.
- Hugo Touvron, Thibaut Lavril, Gautier Izacard, Xavier Martinet, Marie-Anne Lachaux, Timothée Lacroix, Baptiste Rozière, Naman Goyal, Eric Hambro, Faisal Azhar, and 1 others. 2023b. Llama: Open and efficient foundation language models. *arXiv preprint arXiv:2302.13971*.
- Maurice Weber, Daniel Fu, Quentin Anthony, Yonatan Oren, Shane Adams, Anton Alexandrov, Xiaozhong Lyu, Huu Nguyen, Xiaozhe Yao, Virginia Adams, Ben Athiwaratkun, Rahul Chalamala, Kezhen Chen, Max Ryabinin, Tri Dao, Percy Liang, Christopher Ré, Irina Rish, and Ce Zhang. 2024. [RedPajama: An Open Dataset for Training Large Language Models](#). *Preprint*, arXiv:2411.12372.
- Alexander Wettig, Aatmik Gupta, Saumya Malik, and Danqi Chen. 2024. [QuRating: Selecting](#)

High-Quality Data for Training Language Models. *Preprint*, arXiv:2402.09739.

Sang Michael Xie, Shibani Santurkar, Tengyu Ma, and Percy Liang. 2023. [Data Selection for Language Models via Importance Resampling](#). *Preprint*, arXiv:2302.03169.

Liangyu Xu, Xuemiao Zhang, Feiyu Duan, Sirui Wang, Rongxiang Weng, Jingang Wang, and Xunliang Cai. 2025. [FIRE: Flexible Integration of Data Quality Ratings for Effective Pre-Training](#). *Preprint*, arXiv:2502.00761.

An Yang, Anfeng Li, Baosong Yang, Beichen Zhang, Binyuan Hui, Bo Zheng, Bowen Yu, Chang Gao, Chengen Huang, Chenxu Lv, Chujie Zheng, Dayiheng Liu, Fan Zhou, Fei Huang, Feng Hu, Hao Ge, Haoran Wei, Huan Lin, Jialong Tang, and 41 others. 2025. [Qwen3 Technical Report](#). *Preprint*, arXiv:2505.09388.

Rowan Zellers, Ari Holtzman, Yonatan Bisk, Ali Farhadi, and Yejin Choi. 2019. [HellaSwag: Can a Machine Really Finish Your Sentence?](#) In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 4791–4800, Florence, Italy. Association for Computational Linguistics.

Yiran Zhao, Chaoqun Liu, Yue Deng, Jiahao Ying, Mahani Aljunied, Zhaodonghui Li, Lidong Bing, Hou Pong Chan, Yu Rong, Deli Zhao, and Wenxuan Zhang. 2025. [Babel: Open Multilingual Large Language Models Serving Over 90% of Global Speakers](#). *Preprint*, arXiv:2503.00865.

A Details of Meta Learning ParaRater

Algorithm 1 outlines the rater training process, which is cast as a bi-level optimization problem. In the inner loop, the language model f_θ is updated on batches of English data from $\mathcal{D}_{\text{train}}$. For each batch, the rater ϕ_η assigns a score to every example, which is then converted into a non-negative weight. These weights determine how strongly each example influences the inner training loss $\mathcal{L}_{\text{inner}}$, encouraging the model to focus on English data that may be most valuable for transfer. In detail, we use the softmax function to get the normalized weights:

$$w(x) = \frac{e^{\phi_\eta(x)}}{\sum_{x' \in B_t} e^{\phi_\eta(x')}}.$$

In the outer loop, after T such inner updates, the adapted model $f_{\theta^{(T)}}$ is evaluated on held-out target-language validation data $\mathcal{D}_{\text{val}}$. The validation loss $\mathcal{L}_{\text{val}}$ captures how well the model transfers knowledge across languages. By differentiating through the inner steps, we obtain meta-gradients with respect to η , which are used to update the rater. Over time, ϕ_η learns to assign higher weights to English examples whose translations improve target-language performance, while down-weighting less useful ones.

Algorithm 1 Meta-Learning a ParaRater

Require: Train data $\mathcal{D}_{\text{train}}$, validation data $\mathcal{D}_{\text{val}}$; inner model f_θ ; rater ϕ_η producing per-sample scores/weights; inner steps T , outer steps K .

- 1: Initialize θ, η
 - 2: **for** $k = 1$ to K **do** $\triangleright$ Outer loop (meta step)
 - 3: $\theta^{(0)} \leftarrow \text{CLONE}(\theta)$
 - 4: **for** $t = 1$ to T **do** $\triangleright$ Inner loop (train LM with rater weights)
 - 5: Sample batch $B \subset \mathcal{D}_{\text{train}}$
 - 6: $s \leftarrow \phi_\eta(B)$ $\triangleright$ Rater scores
 - 7: $w \leftarrow \text{SCORESTOWEIGHTS}(s)$ $\triangleright$ Any monotone map/normalization
 - 8: $\mathcal{L}_{\text{inner}} \leftarrow \sum_{(x,y) \in B} w(x) \cdot \text{CE}(f_{\theta^{(t-1)}}(x), y)$
 - 9: $\theta^{(t)} \leftarrow \theta^{(t-1)} - \alpha \nabla_{\theta^{(t-1)}} \mathcal{L}_{\text{inner}}$
 - 10: **end for**
 - 11: Sample validation batch $\tilde{B} \subset \mathcal{D}_{\text{val}}$
 - 12: $\mathcal{L}_{\text{val}} \leftarrow \frac{1}{|\tilde{B}|} \sum_{(\tilde{x}, \tilde{y}) \in \tilde{B}} \text{CE}(f_{\theta^{(T)}}(\tilde{x}), \tilde{y})$
 - 13: $\triangleright$ Meta-update
 - 14: $\eta \leftarrow \eta - \beta \nabla_\eta \mathcal{L}_{\text{val}}(\theta^{(T)}(\eta))$
 - 15: $\theta \leftarrow \text{COPY}(\theta^{(T)})$
 - 16: **end for**
 - 17: **return** ϕ_η $\triangleright$ Trained rater
-

To further ensure the quality of selected data, ParaRater employs a two-stage filtering strategy. First, it selects examples that are helpful on their own when weighted by the rater. Second, it retains only those that prove effective when both the English and translated versions are included in training, thus capturing truly impactful parallel corpora. The final selected set can then be translated to build high-quality parallel corpora, enabling more efficient and effective cross-lingual transfer for multilingual LLMs.

B Additional Experimental Results

Table 5 provides an additional 3B-scale comparison with FineWeb-Edu. The trends remain broadly consistent with the 1.2B setting: ParaRater improves most NLU and translation metrics, while ARC-E favors FineWeb-Edu due to its educational-domain bias.

Benchmark	FineWeb-Edu	ParaRater
ARC-E	48.65	47.20
BMLAMA	51.97	52.69
HellaSwag	39.87	41.70
MMMLU	27.68	27.64
StoryCloze	63.45	64.30
XWinograd	63.55	64.13
Flores	19.84	21.12

Table 5: Average 3B model performance comparison with FineWeb-Edu.

C Experimental Details

We conducted both rater training and language model training on a cluster of 64 H100 GPUs. To avoid interference from cross-lingual factors and ensure fairness, we train a separate ParaRater for each En–Target language pair. For each language, ParaRater training consumed roughly 24×64 H100 GPU-hours, while training each final language model required 12×64 GPU-hours. As the initial rater, we used Qwen3-Embedding-0.6B⁴ with a linear head. The language model co-trained with the rater during meta-learning, as well as the final language model, share the same architecture with 1.2B parameters. Rater training is a one-time selection cost for each language pair; the released

⁴<https://huggingface.co/Qwen/Qwen3-Embedding-0.6B>

checkpoint can be reused to score new source documents without retraining. Inspired by DataRater, we developed **ParaTool**, a meta-learning toolkit for rater training, and applied several optimizations to make large-scale training practical. Both **ParaTool** and **ParaCore**—the parallel corpus we constructed through ParaRater—are open-sourced to foster further research on multilingual large language models.

The hyper-parameter settings employed for the meta-learning of ParaRater are listed below.

For the pretraining experiments, we employ a transformer architecture derived from the LLaMA-2 model (Touvron et al., 2023b), configured with roughly 1.2 billion parameters. All models are initialized randomly prior to pretraining. Table 6 provides the full specifications of the model architecture and training hyperparameters.

For rater training, we initialize with Qwen3-Embedding-0.6B and using a inner model with the same architecture of the language model as in the pretraining. During the meta-learning process, we set the number of inner steps to 2, and the maximum token length for both the rater and the inner model is fixed at 512.

Model configuration	Values
Attention head	16
Layers	24
Hiddent size	2048
Intermediate layer dimension	5504
maximum position embedding	4096
layer normalization epsilon	1×10^{-5}
Training Hyperparameters	Values
Batch size	3072
Sequence length	4096
Optimizer	AdamW
Learning rate	4.3×10^{-4}
Learning rate schedule	Cosine decay to 10% of inital value
Traning steps	Varied based on the total token budget
Precision	bf16(mxied-precision training)

Table 6: Model configuration and Training Hyperparameters for pretraining LLms