ONLINE LANGUAGE ADAPTIVE SAMPLING FOR BETTER DISTRIBUTED CROSS-LINGUAL GAINS

Quang Phuoc Nguyen[1*] Felix Gaschi[2*] David Anugraha[3] Santiago Martínez Novoa[4] En-Shiun Annie Lee[1,5]

[1]Ontario Tech University [2]Doctrine [3]Stanford University [4]University of the Andes [5]University of Toronto [*]Equal contribution

1 Motivation: Realignment Gains Are Uneven

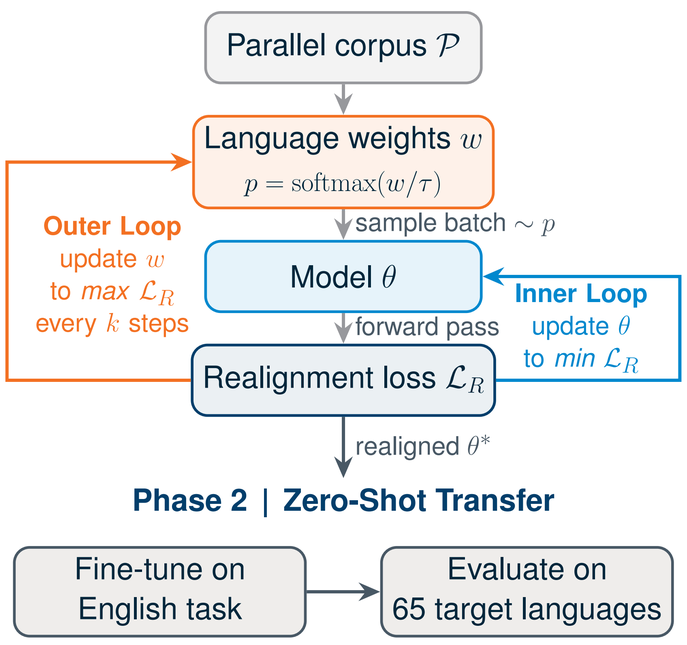
A Multilingual LMs still struggle on linguistically distant **low-resource languages (LRLs)**; prior work links much of this to weak cross-lingual **representation alignment**.

B Contrastive **realignment** on parallel data is a data-efficient fix, but gains are **inconsistent**: some languages improve, others regress.

C Existing realignment samples languages **uniformly**, or uses a **static** subset fixed before training.

Hypothesis: realignment benefits from more exposure to *hard-to-align* languages, and *which* languages are hard changes as training progresses.

2 Our Contribution: Online Adaptive Language Sampling for Realignment



Two Variants of the Outer-Loop Update

UCB bandit — exploit high loss + explore rarely seen

$$w_\ell \leftarrow \bar{r}_\ell + c\sqrt{\frac{\ln t}{n_\ell + 1}}$$

Gradient-based — smooth, no explicit exploration

$$w \leftarrow w + \eta \nabla_w \textstyle\sum_\ell w_\ell L_\ell$$

$\bar{r}_\ell$, L_ℓ: mean realignment loss of language ℓ over the last k inner steps.

Setting

Models	XLM-R Large, mBERT, mDeBERTa v3, mmBERT; Gemma 2 9B, Llama 3.1 8B (LoRA)
Tasks	POS (UDPOS, MasakhaPOS), NER (WikiANN, MasakhaNER), NLI (XNLI, IndoNLI, Myanmar-XNLI, AfriXNLI)
Data	OPUS-100 + NLLB parallel data; 16k steps, 3 seeds
Baselines	Fine-tuning only, Uniform realignment, Most-URIEL (40 most typologically diverse languages)

3 Results

	FT only	Uniform	M-URIEL	UCB	Grad.
Encoder-only					
mBERT	56.27	**61.18**	60.70	61.17	61.05
mDeBERTa v3	66.36	66.83	66.09	**66.96**	66.85
mmBERT	54.72	61.21	61.01	61.48	**61.91**
XLM-R Large	64.09	67.08	66.97	67.73	**67.75**
Decoder-only (LoRA)					
Llama 3.1 8B	42.30	45.10	45.48	45.20	**45.80**
Gemma 2 9B	49.49	51.57	**53.18**	52.16	51.33
Mean	55.54	58.83	58.91	**59.12**	59.11

Table 1. *Macro-average over NER / POS / NLI (3 seeds).* ***Bold****: best,* underline*: second-best per model.*

Key Points

A A dynamic method is **best or second-best on all 6 models**, and best on 4.

B The static **Most-URIEL** subset helps decoder-only models but falls *below uniform* on encoder-only ones.

C **Gains are spread out:** UCB wins most head-to-head comparisons (Fig. 1).

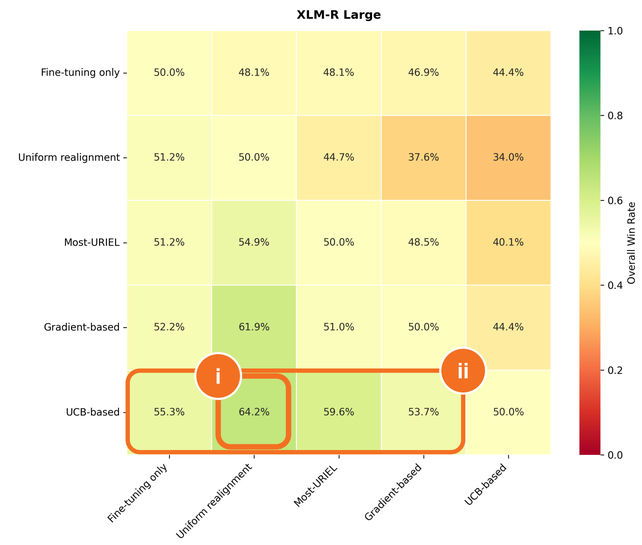


Fig. 1. *XLM-R win rate of row vs. column method (N=441 triplets).*

i UCB beats uniform realignment on **64.2%** of triplets.

ii UCB's row is **above 50%** against every other method.

4 Impacts toward Resource Level

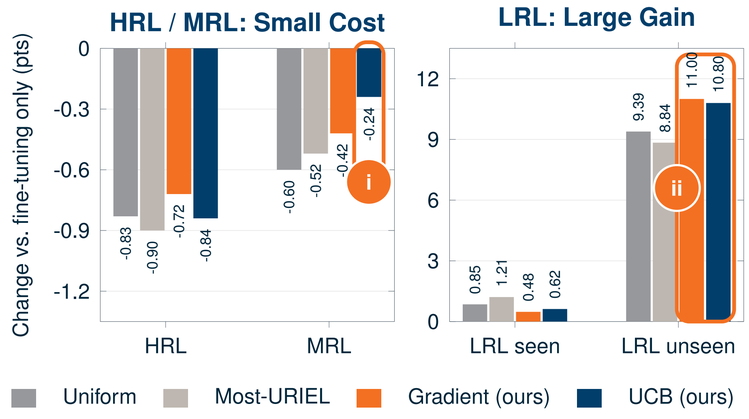


Fig. 2. *Change vs. fine-tuning only (XLM-R Large, 3 seeds); note the different y-scales.*

i UCB has the **smallest MRL drop** (-0.24 vs. -0.60 uniform).

ii Dynamic methods give the **largest unseen-LRL gains** (up to +11).

5 Analysis

Co-Evolution Matters

Question: do the gains come from finding a better language distribution, or from adapting it *during* training? **Ablation:** learn the weights with the model frozen, then realign with that *fixed* distribution ("decoupled").

Task scores are close i, but online adaptation wins far more task-wise comparisons ii. As realignment reshapes the embedding space, **which languages are hard keeps changing**, so any fixed distribution gradually goes stale.

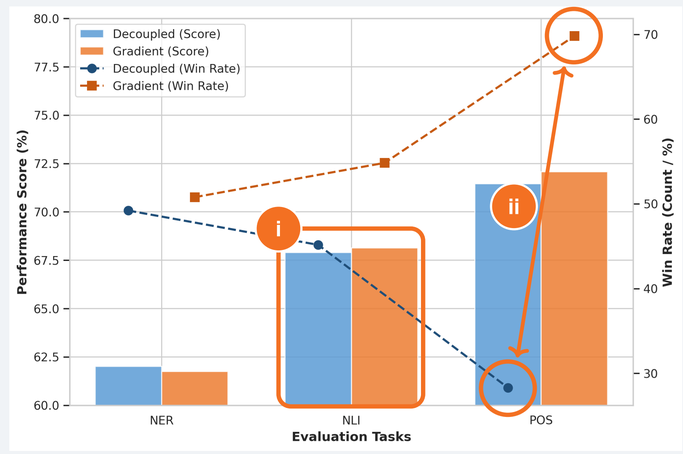
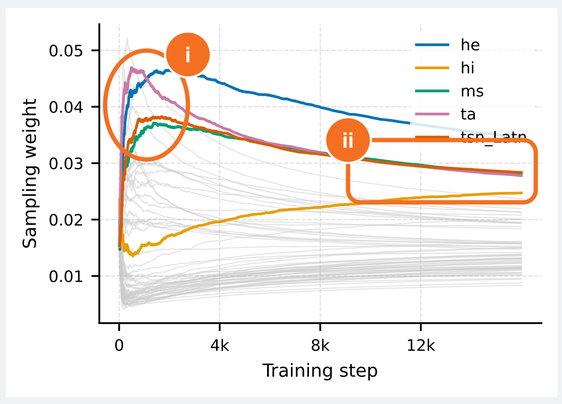


Fig. 3. *Online gradient-based vs. decoupled (fixed learned weights), XLM-R, 32k steps. Bars: task score; lines: win rate.*

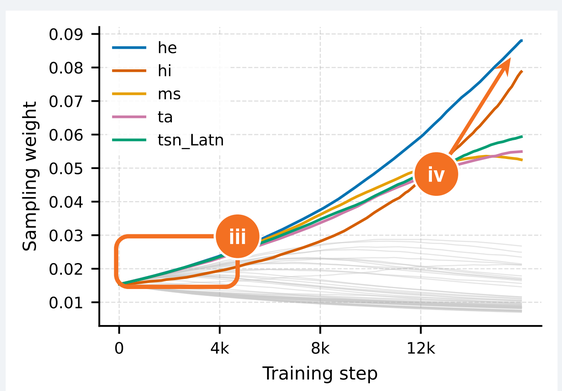
What Gets Oversampled?

Both variants converge on roughly the **same top-5 languages** across seeds: Hebrew, Hindi, Malay, Tamil and Tswana, a mix of mid-resource, low-resource and typologically distant languages.

They differ in *how* they get there: UCB concentrates on hard languages within ∼1k steps i and then settles ii; gradient-based starts ∼10× slower iii and keeps accelerating iv, which is why it degrades with longer training while UCB stays stable.



(a) UCB



(b) Gradient-based

Fig. 4. *Learned sampling weights over training (XLM-R, seed 31); top-5 languages coloured, all others grey.*

6 Conclusion:

Adapt Online

Loss-driven language sampling beats uniform and static selection: a dynamic method is in the top 2 on every model.

Spread the Gains

Improvements reach **more languages**: LRLs improve while the drop on mid-resource languages shrinks.

Pick UCB by Default

UCB is robust with no per-model tuning; gradient-based can gain more but needs tuning.