

# Early-Layer LoRA vs. Full Fine-Tuning for Swahili Lexical Alignment in XTREME-R

Assignee Research

June 8, 2026

## Abstract

This report synthesises findings from 12 peer-reviewed papers addressing the following research question: How does Early-Layer LoRA fine-tuning for lexical alignment in Swahili compare to full model fine-tuning on the XTREME-R natural language inference task in terms of accuracy and training efficiency. 9 claims were extracted from source literature; 1 was independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 4.2/10. This report is a machine-generated literature synthesis and does not constitute original research.

## 1 Introduction

This paper examines: Targeted Lexical Injection: Unlocking Latent Cross-Lingual Alignment in Lugha-Llama via Early-Layer LoRA Fine-Tuning. Research question: How does Early-Layer LoRA fine-tuning for lexical alignment in Swahili compare to full model fine-tuning on the XTREME-R natural language inference task in terms of accuracy and training efficiency?.

## 2 Methodology

Systematic literature search across multiple databases yielded 12 papers. Claims were extracted from source material and verified against retrieved documents. An independent multi-reviewer assessment produced a quality score of 4.2/10.

## 3 Results

12 papers retrieved. 9 claims extracted; 1 independently verified. Quality review score: 4.2/10.

## 4 Limitations

This report is a machine-generated literature synthesis and does not constitute original research. Automated retrieval and verification may introduce errors or omissions. Review scores reflect automated assessment, not human peer review. Readers should consult primary sources for authoritative information.

## 5 Extracted Claims

Claim	Verified	Confidence
Layer 0 (input embeddings) showed a modest average cosine similarity of approximately 0.3153.	×	0.07
Layer 1 showed an average cosine similarity of 0.9808.	×	0.09
Layer 2 exhibited the peak average cosine similarity, reaching 0.99998.	×	0.08
Layer 31 showed an average similarity of 0.9876.	×	0.04
The baseline output similarity observed on the full evaluation set was approximately 0.32.	×	0.09
The average cosine similarity at the final output layer (Layer 31) of the base model was approximately 0.3211 for the tr	×	0.15
The base model used is Lugha-Llama-8B-wura, an open-source LLM adapted for several African languages, including Swahili,	×	0.10
The model is loaded in 4-bit precision using bitsandbytes with NF4 quantization and torch.bfloat16 as the compute data t	×	0.02
The pilot study revealed that Lugha-Llama-8B-wura inherently achieves very high lexical alignment in its early layers, p	✓	0.20

## References

- <http://arxiv.org/abs/2510.27364v1>
- <http://arxiv.org/abs/2506.15415v1>
- <http://arxiv.org/abs/2412.09827v1>