

# Performance comparison of cross-lingual transfer learning methods for Named Entity Recognition in low-resource African languages

Assignee Research

July 5, 2026

## Abstract

Cross-lingual transfer learning enables NLP for low-resource languages by leveraging labeled data from higher-resource sources, yet existing comparisons of source language selection strategies do not control for total training data, confounding language selection effects with data quantity effects. We introduce Budget-Xfer, a framework that formulates multi-source cross-lingual transfer as a budget-constrained resource allocation problem. Given a fixed annotation budget  $B$ , our framework jointly optimizes which source languages to include and how much data to allocate from each. We evaluate fou

## 1 Introduction

This paper examines: Budget-Xfer: Budget-Constrained Source Language Selection for Cross-Lingual Transfer to African Languages. Research question: How does the performance of cross-lingual transfer learning for Named Entity Recognition in low-resource African languages compare when using Budget-Xfer versus traditional single-source or multi-source transfer methods, as measured by F1 scores on the MasakhaNER dataset?.

## 2 Methodology

Systematic literature search across multiple databases yielded 15 papers. Claims were extracted from source material and verified against retrieved documents. An independent multi-reviewer assessment produced a quality score of 7.7/10.

## 3 Results

15 papers retrieved. 6 claims extracted; 5 independently verified. Quality review score: 7.7/10.

## 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
cosine_gap predicts cross-lingual transfer with a correlation coefficient ( $\rho$ ) of 0.4–0.6.	×	0.10
The evaluation metric for NER is entity-level F1 (micro-averaged) computed with seqeval, requiring exact span and type m	✓	0.24
The evaluation metric for sentiment analysis is weighted F1 to account for class imbalance.	✓	0.18
Multi-source transfer significantly outperforms single-source transfer with Cohen’s d effect sizes ranging from 0.80 to	✓	0.24
The study evaluates four allocation strategies across named entity recognition and sentiment analysis for three African	✓	0.35
The value of embedding similarity as a selection proxy is task-dependent, with random selection outperforming similarity	✓	0.35

## References

- <http://arxiv.org/abs/2309.05311v1>
- <http://arxiv.org/abs/2603.27651v1>
- <http://arxiv.org/abs/2307.01310v2>