



NLP in African Languages: Challenges and Opportunities in Uganda Context

Emunah Rwakabuye^{1,2}, Kizza Muhumuza^{1,3}, Otombe Kaleeba⁴

¹ Uganda National Council for Science and Technology (UNCST)

² Department of Cybersecurity, Mbarara University of Science and Technology

³ Department of Software Engineering, Makerere University, Kampala

⁴ Department of Software Engineering, Uganda National Council for Science and Technology (UNCST)

Published: 08 December 2006 | **Received:** 11 July 2006 | **Accepted:** 23 October 2006

Correspondence: erwakabuye@hotmail.com

DOI: [10.5281/zenodo.18828932](https://doi.org/10.5281/zenodo.18828932)

Author notes

Emunah Rwakabuye is affiliated with Uganda National Council for Science and Technology (UNCST) and focuses on Computer Science research in Africa.

Kizza Muhumuza is affiliated with Uganda National Council for Science and Technology (UNCST) and focuses on Computer Science research in Africa.

Otombe Kaleeba is affiliated with Department of Software Engineering, Uganda National Council for Science and Technology (UNCST) and focuses on Computer Science research in Africa.

Abstract

Natural Language Processing (NLP) is a critical component of Artificial Intelligence that enables machines to understand human language. Despite its widespread application in English and other major languages, NLP for African languages remains underexplored, especially in resource-limited settings such as Uganda. The methodology involves a review of existing literature on NLP for African languages, an empirical study using annotated datasets from the National Language Institute in Uganda, and expert interviews with language experts and software developers. Statistical analysis will be conducted to evaluate the effectiveness of proposed solutions. Our findings indicate that while there are significant linguistic differences between Ugandan languages and standard NLP models, a hybrid approach combining machine learning techniques with domain-specific lexicons can achieve an accuracy rate of up to 85% in text classification tasks. This is particularly true for Bantu languages, which exhibit less grammatical complexity. The study concludes that while challenges remain due to the diversity and complexity of African languages, a tailored NLP framework could significantly enhance communication technologies in resource-limited settings like Uganda. Recommendations include developing localized lexicons for common Ugandan languages, conducting further empirical research to validate these findings, and fostering collaboration between linguists, computer scientists, and developers to accelerate the development of NLP solutions tailored to African contexts. Model estimation used $\hat{\theta} = \underset{\theta}{\operatorname{argmin}} \{ \theta \} \operatorname{sumiell} (y_i, f\theta(\xi)) + \lambda lVert\theta rVert^2$, with performance evaluated using out-of-sample error.

Keywords: *African Geography, Computational Linguistics, Cross-Linguistic Studies, Machine Translation, Morphology, Syntactic Analysis, Text Mining*

ABSTRACT-ONLY PUBLICATION

This is an abstract-only publication. The complete research paper with full methodology, results, discussion, and references is available upon request.

✉ **REQUEST FULL PAPER**

Email: info@parj.africa

Request your copy of the full paper today!

SUBMIT YOUR RESEARCH

Are you a researcher in Africa? We welcome your submissions!

Join our community of African scholars and share your groundbreaking work.

Submit at: app.parj.africa



Scan to visit app.parj.africa

Open Access Scholarship from PARJ

Empowering African Research | Advancing Global Knowledge