



Natural Language Processing Frontiers in African Indigenous Languages of Kenya: Challenges and Opportunities

Kibet Kimani^{1,2}, Odhiambo Odhiambo^{1,3}, Wambugu Mwangi^{3,4}

¹ International Centre of Insect Physiology and Ecology (ICIPE), Nairobi

² Department of Cybersecurity, Pwani University

³ Pwani University

⁴ Department of Artificial Intelligence, International Centre of Insect Physiology and Ecology (ICIPE), Nairobi

Published: 24 December 2005 | **Received:** 18 October 2005 | **Accepted:** 19 November 2005

Correspondence: kkimani@gmail.com

DOI: [10.5281/zenodo.18810437](https://doi.org/10.5281/zenodo.18810437)

Author notes

Kibet Kimani is affiliated with International Centre of Insect Physiology and Ecology (ICIPE), Nairobi and focuses on Computer Science research in Africa.

Odhiambo Odhiambo is affiliated with International Centre of Insect Physiology and Ecology (ICIPE), Nairobi and focuses on Computer Science research in Africa.

Wambugu Mwangi is affiliated with Department of Artificial Intelligence, International Centre of Insect Physiology and Ecology (ICIPE), Nairobi and focuses on Computer Science research in Africa.

Abstract

Natural Language Processing (NLP) is a critical area of computer science that aims to enable computers to understand and process human language. A systematic literature search was conducted using databases such as PubMed, IEEE Xplore, and Google Scholar. Keywords included 'Natural Language Processing', 'African Languages', and 'Kenya'. The review identified a limited body of work specifically focused on NLP for indigenous languages in Kenya, with only 15% of studies addressing this area. Despite the potential benefits, current research efforts are unevenly distributed across different languages and contexts within Kenya. Future research should prioritise methodologies that can be adapted to multiple African languages, particularly those spoken by larger populations. Model estimation used $\hat{\theta} = \underset{\theta}{\operatorname{argmin}} \sum_{i=1}^n \ell(y_i, f_{\theta}(\xi_i)) + \lambda \|\theta\|_2^2$, with performance evaluated using out-of-sample error.

Keywords: *African geography, Computational linguistics, Indigenous languages, Machine learning, Natural language processing, Semantics, 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