

YembaEGRA — Building an AI Speech-Recognition Corpus to Diagnose Reading Deficiencies in Cameroon's Yemba Language

AI in Education · Good practice · Cameroon

Quality score: 33/100

Evidence strength: Descriptive / self-reported

Summary

Researchers at Université de Yaoundé I built YembaEGRA, a 69-speaker speech corpus in Bamendou, Cameroon, to train AI models diagnosing which reading sub-skill a Yemba-speaking child lacks, addressing shortages of multilingual assessors across Cameroon's 200+ languages.

Key dimensions

Dimension	Score
Evidence of learning impact	0/3
Equity & inclusion	2/3
Transparency, ethics & data protection	1/3
Teacher capability & pedagogy	1/3
Scalability & sustainability	1/3

Evaluated by C-NAPSE

Data sources

[C-NAPSE web research](#) — accessed 2026-09-16

[Université de Yaoundé I, funded by the EU Horizon 2020 MSCA-RISE 'Esperanto' project](#) — accessed 2026-09-16

[EdTech Hub library: Using AI for assessment in African languages learning in primary education](#) — accessed 2026-09-16

[Mendeley Data: YembaEGRA dataset](#) — accessed 2026-09-16

Cite this — Evidoria. *YembaEGRA — Building an AI Speech-Recognition Corpus to Diagnose Reading Deficiencies in Cameroon's Yemba Language*. Persistent ID: 37c644f9-a8d6-4223-aaf4-35d220391f05.