

BUSAYO AWOBADE

iambusayor.github.io ♦ [Google Scholar](#) ♦ +234(0) 703 650 2407 ♦ hawobade@gmail.com

Lagos, NG

RESEARCH INTERESTS

Multilingual speech and language models; African and low-resource NLP; speech evaluation and error diagnosis; data-centric AI; efficient and robust ML.

EDUCATION

Bachelor of Science in Computer Science, Federal University of Agriculture, Abeokuta 2017 - 2023
DeepLearning.AI: Neural Networks and Deep Learning; Structuring Machine Learning Projects; Convolutional Neural Networks; Sequence Models.

PUBLICATIONS

- **Busayo Awobade**, ..., Tobi Olatunji. *AfriVox-v2: A Domain-Verticalized Benchmark for In-the-Wild African Speech Recognition*. Interspeech 2026.
- **Busayo Awobade**, ..., Tobi Olatunji. *AfriVox: Probing Multilingual and Accent Robustness of Speech LLMs*. EACL 2026.
- Gabriel Zencha Ashungafac, Mardhiyah Sanni, **Busayo Awobade**, Alex Gichamba, Tobi Olatunji. *AfriSpeech-MultiBench: A Verticalized Multidomain Multicountry Benchmark Suite for African Accented English ASR*. IJCNLP-AAACL 2025.
- Chris Emezue, NaijaVoices Community, **Busayo Awobade**, et al. *The NaijaVoices Dataset: Cultivating Large-Scale, High-Quality, Culturally-Rich Speech Data for African Languages*. Interspeech 2025.

MANUSCRIPTS & WORKSHOP PAPERS

- **Busayo Awobade**, et al. *Beyond Aggregate WER*. Manuscript under revision for ACL ARR.
- Mardhiyah Oduwole, ..., **Busayo Awobade**, et al. *The Role of Data Augmentation in African Machine Translation*. LM4UC (AAAI) 2026.
- **Busayo Awobade**, Mardhiyah Oduwole, Steven Kolawole. *What Happens When Small Is Made Smaller? Exploring the Impact of Compression on Small Data Pretrained Language Models*. AfricaNLP (ICLR) 2024.

RESEARCH EXPERIENCE

Research Scientist | Intron

Nov. 2024 - Present

- **Lead research and engineering efforts in multilingual TTS**, with a focus on African languages, regional accents, and real-world speech.
- Engineered a **production-grade multispeaker TTS model** supporting 13 African languages and 10 regional accents, using pipelines designed for noisy, in-the-wild speech and legacy ASR corpora.
- Architected an end-to-end pipeline for processing **35k+ hours** of conversational and in-the-wild speech across 24 African languages, including a custom chunking algorithm optimized for multilingual speech.
- Designed a **language-agnostic framework for evaluating ASR and speech-language models** across read, conversational, and in-the-wild African speech.
- Led the development of **AfriVox**, an open benchmark for multilingual and accent-robust speech evaluation spanning 24 African languages, and extended the work through AfriVox-v2 with domain-verticalized, in-the-wild evaluation.
- Developed diagnostic analyses of ASR failure modes, including errors involving named entities, numbers, and other linguistically salient phenomena, to connect evaluation findings with targeted data interventions.

Research Engineer | Babs

Jan. 2024 - Jun. 2024

- Trained and evaluated **TTS models** for Yoruba, Swahili, and Hausa using the **VITS architecture** on publicly available African speech corpora.
- Extended **Meta's MMS model** for Yoruba and Hausa speech synthesis using Common Voice aligned corpora.
- Led research on **speech-to-speech translation with expressivity transfer**, following Meta's Seamless Expressivity work.
- Built a cascaded speech-to-speech translation system using ASR (**Whisper**), MT (**M2M-100**), and expressive TTS (**Bark**).

Founding Machine Learning Engineer | ASALytics (now NazariAI)

Mar. 2022 - Dec. 2022

- Awarded **\$115k grant** from the Algorand Foundation to build the ASALytics platform.
- Set up **CI/CD infrastructure** for model and application servers using GCP Cloud Build and Cloud Run.

Data Science Intern | Zummit Info Labs

Jun. 2021 - Oct. 2021

- Built an object detection and classification system for road traffic signs using YOLO and a deep learning classifier.
- Trained a FaceNet + LightGBM model for a research project on racial profiling and bias in airport security systems.

RESEARCH & COMMUNITY LEADERSHIP

Research Scientist / Research Lead | MLCollective

Nov. 2023 - Present

- Lead research activities for the Nigeria chapter, mentoring early-career researchers and organizing research talks, meetups, presentations, and webinars on NLP and machine learning.

Community Leadership | Masakhane / Deep Learning Indaba

- Co-organized the Masakhane EfficientNLP workshop at Deep Learning Indaba 2023.
- Previously led the AI community at the Federal University of Agriculture, Abeokuta.

SELECTED AWARDS

- **Best Poster**, Deep Learning Indaba 2023 — for *What Happens When Small Is Made Smaller?*
- **Algorand Foundation Grant**, \$115k, 2022 — for development of the ASALytics platform.
- **1st Runner-up**, DSN Data Science Hackathon, 2020 — among 750 participants.
- **4th Best AI University in Nigeria**, UmojaHack Africa University Challenge, 2021.

ACADEMIC SERVICE

- **Peer reviewer**: ACL Rolling Review (ARR, May 2026); EMNLP 2026; AfricaNLP@ICLR 2024; EMNLP 2023; TinyPapers@ICLR 2023.