

Inclusive Voices: Advancing Language Technology for People with Impaired Speech in Local Languages

Cave R., Ayoka G., Barbareschi G., Tomanek K., Wiafe I., Holloway C.

Richard.Cave.17@ucl.ac.uk - CHI 2025 workshop "Speech AI for All: Promoting Accessibility, Fairness, Inclusivity, and Equity"



PROBLEM DESCRIPTION

- **25%-50%** of people with disabilities globally face communication challenges.
- Impaired speech can lead to stigmatization and marginalization.
- Limited assistive technologies and SLT services in Low-Resource Language (LRL) communities in the Global South worsen the issue.
- Automatic Speech Recognition (ASR) for non-standard speech (NSS) can help maintain communication without altering speech.
- Lack of research and development for ASR in African languages, especially for non-literate speakers.
- Current ASR for African languages often targets standard speech for banking, ignoring NSS needs.
- Absence of primary data (atypical speech samples) in LRLs hinders inclusive ASR model development.
- **Existing datasets are mainly in English.**



The Centre for Digital Language Inclusion (CDLI) has eight principles:

1. Partner with local Universities & NGOs for ethical data collection.
2. Collect 50+ hours of recordings of impaired speech per language from diverse speakers.
3. Host and open-source speech data locally.
4. Provide open-source tools for data collection, transcription, and functional ASR apps.
5. Develop an open-source 'Cookbook' of best practices for collecting, transcribing NSS, and building ASR models.
6. Provide technical training and mentoring in data collection, transcription, AI ASR model development for University students, and co-design with people with non-standard speech.
7. Provide expert mentoring for innovators developing end-user ASR tools for non-standard speech.
8. Sponsor hackathons to encourage local innovation and co-design with individuals with impaired speech, offering mentorship and seed funding.

CDLI is led by the Global Disability Innovation Hub in partnership with UCL, and Keio Media Design. The University of Ghana, Talking Tips Africa Foundation,

VISION AND GOALS

Centre for Digital Language Inclusion

In January 2025 we launched the new Centre for Digital Language Inclusion (CDLI) <https://www.cdli-inclusion.com/>.

Our vision is to

- **Democratize speech recognition technology** by shifting the from 'standard speech only, top-down prioritization'.
- **Empower local communities** speaking LRLs to develop their own non-standard speech datasets, AI-driven ASR models, and user applications.
- **Create technologies** that support people with atypical speech in their local language and culture, starting with African languages.
- Develop **community-led user-centric research practices** for more available and useful applications.
- **Collect, validate, and make open-source** a large corpus of impaired speech in LRLs.
- **Build local, autonomous, and sustainable skills** in data collection, transcription, and AI-based ASR model and app development.



Low Resource Languages – starting with 10 in Africa

In Ghana we are collecting recordings of impaired speech in five languages: Akan, Ewe, Ga, Dagbani and Dagaare. These are the most spoken of the 80 indigenous languages in Ghana. We are soon to collect recordings in Kenya, starting with Swahili.

1st Low Resource Language Hackathon: Ghana 26-30 May 2025

All Universities in Ghana have been invited to submit teams.

Over 150 applications received from Ghana and elsewhere in Africa.

All teams required to have male and female participants and at least one person living with speech difference.

Hackathon Aims: To design and develop apps and services that support individuals with impaired speech to communicate effectively and inclusively in Ghanaian languages.

1. Integrate Text-to-Speech (TTS) or ASR for Ghanaian languages.
2. Develop an intuitive and user-friendly interface tailored to the unique needs of individuals with impaired speech, prioritizing ease of use, cultural relevance, and inclusivity.
3. Address accessibility standards, enabling individuals with varying levels of impairment.
4. Scalability of outputs to other low-resource languages.



QR Code to view the Hackathon website

Akan data collection and ASR modelling: lessons learned

A community-driven approach is vital for data collection, fostering participation from local stakeholders like linguists, speech therapists, and community organizations.

Culturally relevant prompts are essential. Image prompts proved more suitable than text prompts, reflecting diversity of literacy and cognitive skills. Describing images allowed for more comfortable and natural speech recordings.

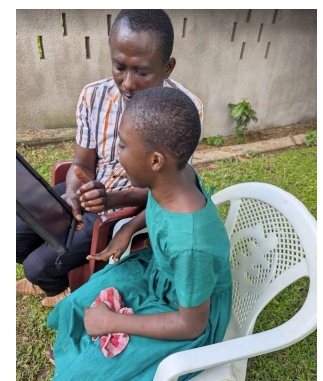
In-person data collection is the most effective method for personalized assistance

Recruitment and training facilitators who are proficient in the target language, digitally literate, and capable of supporting needs of participants with impaired speech. This ensures consistency and adherence to recording protocols.

The rules for written Akan are still developing. It is crucial for linguists to agree on transcription rules, particularly for code-switching and neologisms. Two-round transcription process incorporating double-blind cross-validation is important for identifying and correcting errors, enhancing the reliability of transcriptions.

Fine-tuning pre-trained multilingual ASR models, such as Whisper, using a dataset of NSS in Akan led to improved recognition accuracy.

Given the still-developing rules for written Akan (amongst other factors), overall WER of caption generation was significantly higher than Error Analysis of Meaning retention. We will be publishing a paper on this soon.



Led By



In Partnership with



Funded by



CDLI website