Phonetic matching for spoken words of rare languages in low-resource environments

By Auke Schuringa, Master Thesis, January 2025

Master thesis under supervision of dr. Anna Bon for the master Computer Science at the Vrije Universiteit Amsterdam and the University of Amsterdam.
Abstract: The thesis presents an energy-efficient method for matching recordings of spoken words from rare languages in low-resource environments.
Adapted from "Energy-efficient phonetic matching for spoken words of rare languages in low-resource environments".

The Situation

- Rare languages are often spoken rather than written
- There is a need for automation in handling of speech to increase efficiency and spur innovation:
  - Interacting with applications through speech

The Problem

- Little data of the language is available
- Rare languages (often) originate from low-resource environments:
  - Efficient use of energy is important
  - Less economic incentive to develop solutions compared to rich(er) areas
  - Solutions cannot be expensive

Existing Solutions

- Artificial neural networks to perform speech recognition:
  - Expensive to train
  - Not enough data available
  - Very accurate
- Manual matching:
 - Very labor intensive
  - Not scalable

Our Solution

- Low-resource environment friendly
  - Only pre-trained speech to token networks used
  - No training of a network required
- Energy-efficient:
  - One time processing of a recording for matching other recordings
- Sacrifices some accuracy compared to expensive specially trained models
  - Accuracy remains high on sufficiently phonetically contrastive words

Try it out!

- Create a new group in Telegram with only yourself
- Add the **audio-indexer** bot

Record
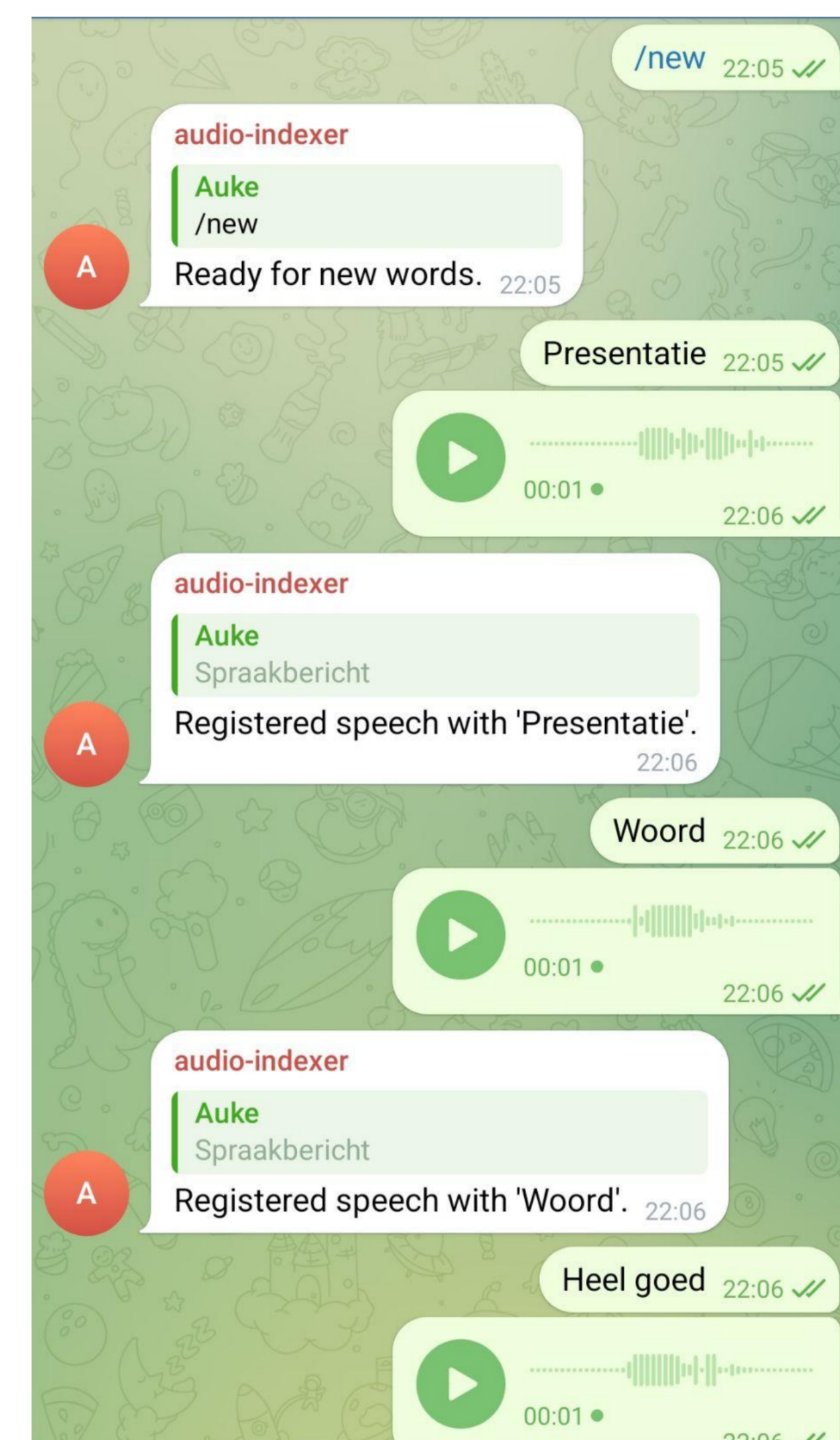
Match

Applications and Research

- Voice matching applications:
  - Package for developing Android applications that include interaction through speech
  - Shopping bots for speech-controlled buying and selling
  - Website integration for voice control
- Research and improvements:
  - Personalized speech matching
  - Language analysis by clustering recordings of words by similarities
  - Model fine-tuning on language specific acoustics
  - Automated data filtering
  - Sentence searching

https://euridice.eu

EURIDICE

VU — VRIJE UNIVERSITEIT AMSTERDAM

UNIVERSITY OF AMSTERDAM

W4RA