UK Speech Conference 2023

Analysis of the Communication Rate Gap for Users of AAC Systems

Hussein Yusufali, Stefan Goetze, Roger K. Moore
Department of Computer Science - University of Sheffield
hsayusufali1, s.goetze, r.k.moore @sheffield.ac.uk

Motivation

- Over 70 million people worldwide face communication difficulties
- Communication rate gap between Augmentative and Alternative Communication (AAC) and typical speaking individuals is high
- Spoken conversational rates typical 140 – 150 WPM, whereas AAC systems have rates of 10 – 20 WPM
- Text input rate can increase by prediction system
- The poster proposes a text prediction user interface using fine-tuned BERT and RoBERTa language models

- Fine-tuned language models were tested across four communication scenarios as open-ended target tasks:
  - Scripted Dialogue (Ordering a coffee)
    - Fully scripted, users must copy the sentences provided to them
  - Half-prompted Dialogue (Small talk at work)
    - Half scripted, users can give own response with prompted topic
  - Open-ended question
    - Users are asked a question and able to give any response
  - Picture Description
    - Users are prompted with a picture and asked to describe

Large Language Model Fine-Tuning

- BERT and RoBERTa language models were fine-tuned on conversational data for AAC systems
- These models were selected for their high accuracy in predictive language generation

Corpora

Corpus	Description	Size
TV Corpus	TV episodes and shows	325M
Switchboard Corpus	Telephone conversations	300 hours
COCA Corpus	American English texts	1B
Wikipedia Corpus	Vast collection of texts	2B
AAC Corpus	AAC messages	6,000
Reddit Corpus	Reddit text collection	256M
Daily Dialog	Annotated dialogues	13k

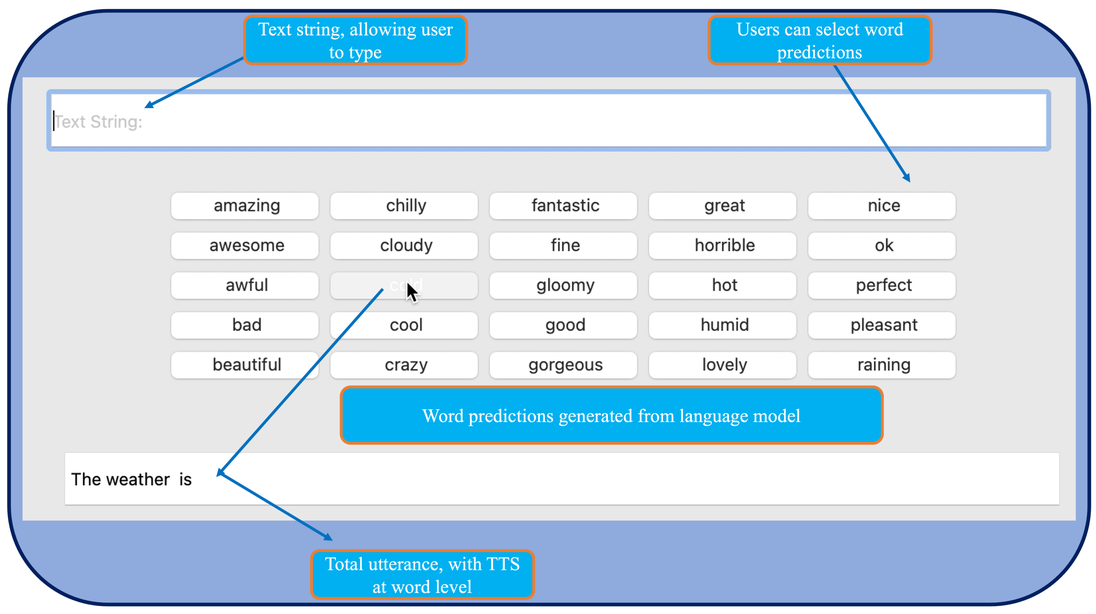
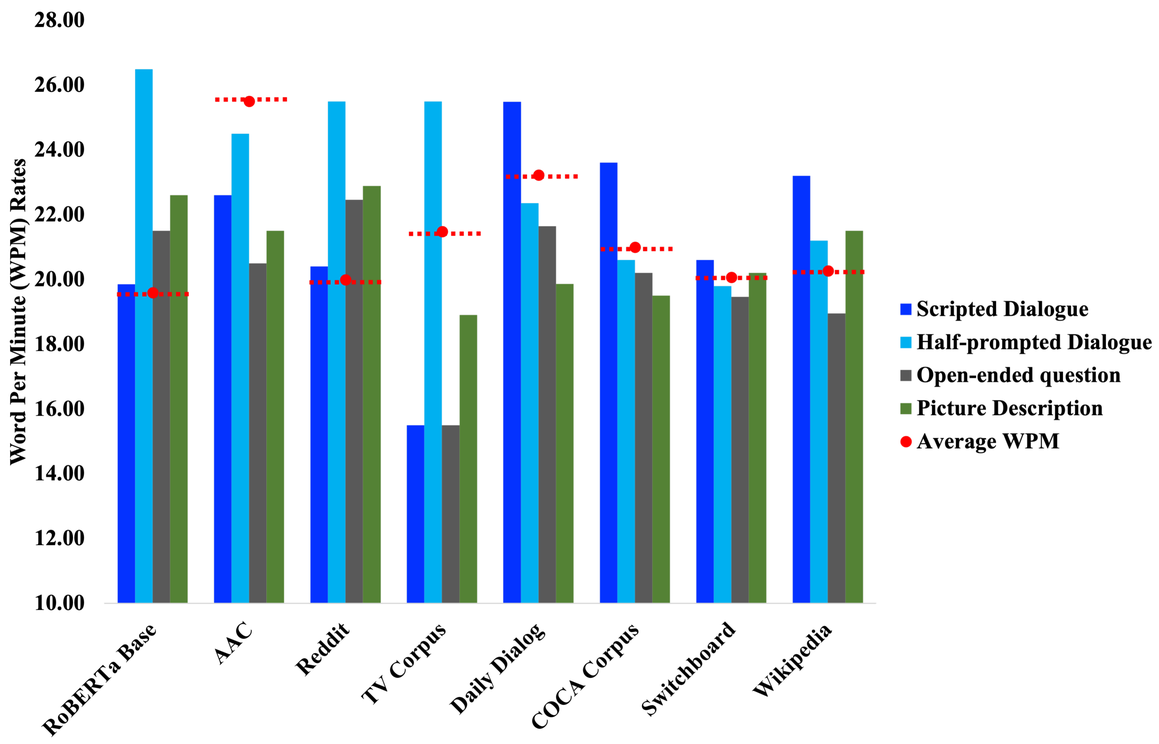


Figure 1: A user interface developed to display word predictions to the user in the form of sequential text input for use with a TTS system.

Results



Conclusions

- Conversational data fine-tuning enhances AAC system accuracy and satisfaction
- RoBERTa models outperform BERT models in accuracy
- Task-specific corpora, conversational data, improve accuracy and words per minute (WPM) compared to generic corpora. AAC corpora performed highest
- Alphabetical ordering of word predictions based on the first character proves advantageous

– This work was supported by the Centre for Doctoral Training in Speech and Language Technologies (SLT) and their Applications funded by UK Research and Innovation [grant number EP/S023062/1]: This work was also supported by Apple Inc