



Grassroots Radio FACTSHEET #03

TEXT-TO-SPEECH

amplifying the voice of communities

The Grassroots Radio project is creating a game changing network of inclusive digital platforms for citizen engagement, community deliberation, and the free flow of information by piloting solutions for community owned and operated radio across Europe. Our participatory approach includes methods for setting up stations and services along with an **emphasis on text-to-speech (TTS) technology to support the curation of audio content**, thus turning data into media.

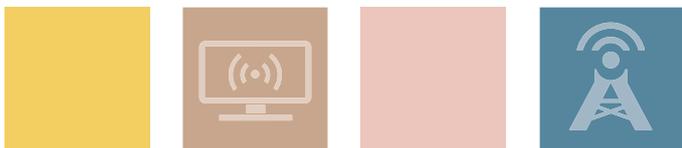
WHAT IS TTS?

Text-to-Speech (TTS) is a type of technology that provides artificial production of human speech. A TTS system **converts normal language text into speech**. TTS can take words on a computer or other digital device and convert them into audio. It works with nearly every personal digital device, including computers, smartphones and tablets. **All kinds of text files can be read** aloud, including Word and Pages documents. Even online web pages can be read aloud. The voice in TTS is computer-generated, and reading speed can usually be sped up or slowed down. **Voice quality varies**, but some voices can sound quite human and have **human-like features, like an accent** or an identifiable persona, such as a voice that sounds like a child speaking. Speech synthesis allows **environmental barriers to be removed for people with a wide range of disabilities**. TTS has long been used in screen readers for

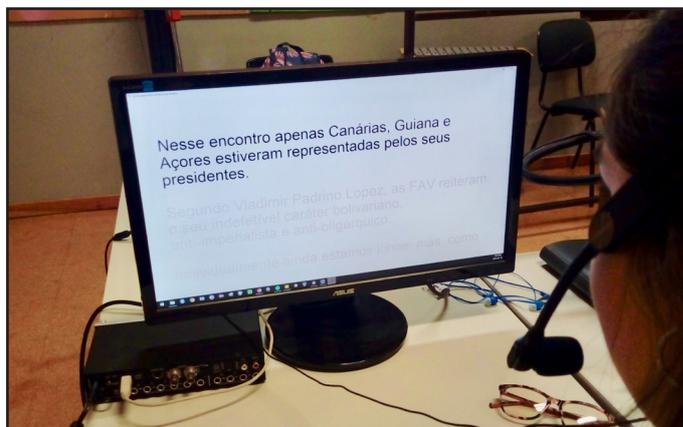
people with visual impairment, but it is now commonly used by people with dyslexia and other reading difficulties as well as by pre-literate children. They are also frequently employed to aid those with severe speech impairment usually through a dedicated voice output communication aid.

TTS IN GRASSROOTS RADIO

The TTS component of Grassroots Radio supports the curation of audio content. Integrating TTS technologies into the RootIO architecture means **opening up vast reserves of both community produced and internet based text data for broadcast on the community-led micro-stations**. Content that has already been written, from news to local history to local weather updates, can be automatically converted to audio data and added to the schedule. There is also the **opportunity to encourage listener interaction** with the stations and the TTS capabilities; messages sent by SMS or email, such as song dedications or classified-type ads, can instantly become audio files ready to play on the air. In terms of practical usage, **TTS capability for potentially under-staffed stations provides a method for semi-automated content generation** and is a way for community members to contribute to the station even when they are unwilling or unable to put their own voice on the air. From a **research perspective**, this provides an opportunity **to investigate the synthesis**



Waveform of a speech recording



Voice recording session with a volunteer in Madeira

of various forms of unstructured text, as well as explore the impact of combining synthetic voices with the local and intimate format of community radio. The TTS component of this project is not an add-on; we need to convert text to audio to be able to generate local, round-the-clock content and merge this with other networked community radio content as a means of offering community members a voice, and bringing people together around that voice. The aim is not to replace presenters but rather to offer support with extra automatically generated content.

TTS RESEARCH AREAS

Within Grassroots Radio, project partner CereProc will be assisting in research into two aspects of TTS.

LANGUAGE RESOURCES

While there are many low cost TTS systems available for dominant languages (e.g. with conventional accents), local accents and dialects for less resourced languages such as Irish and even for dominant languages such as Romanian many of the requirements for a TTS system have not yet been created. Furthermore, if resources do exist, they are not freely available. As part of the project we are setting up the open use Living Audio Dataset (LADS) which will make resources freely available both for commercial and open source TTS systems. A key research area is to explore how communities might be able to crowd build some of these resources such as pronunciation dictionaries as well as exploring how a resource like this could adapt and develop in the same way as Wikipedia.

APPLICATIONS OF TTS

Using TTS to support radio presenters is a completely novel application area for this technology. It is crucial that the technology does not get in the way of the enga-

gement and flow of a live presenter and there are many questions we need to answer to successfully use TTS in local radio.

For example:

- Will listeners accept the use of TTS in their radio stations?
- Can we design a TTS voice style to better support radio presentation?
- Would a local accent be preferred to higher quality synthesis?
- How should we fix / control (e.g. remove swearing, moderate) the content of TTS?
- Can TTS create a significant amount of content and/or income for community radio?

OUTCOMES

In the course of the project some useful resources will be developed both for local radio and for the TTS community in general.

- Integration of TTS into RootIO and standalone applications to support radio use of TTS.
- Idlak (an open source TTS engine which can be used free of charge) will be developed into a full end to end synthesiser.
- An open source Apache licensed RESTful API and server will be released for Idlak which allows engineers to integrate TTS into their applications using a server.
- The crowd-building web applications and program for creating phonetically balanced recording scripts will be made available.
- Language resources for Irish, Madeiran Portuguese, and Romanian TTS systems will be released under Apache license.
- Recording scripts for English, Irish, Madeiran Portuguese, and Romanian will be made available.
- Voice audio for English, Irish, Madeiran Portuguese, and Romanian which can be used to build a TTS system will be released.