TEXT-TO-SPEECH SYNTHESIS
for Constructed Languages

Oregon State University Cascades

What Are Conlangs?

Constructed languages, also known as *conlangs*, are artificial languages whose phonology, grammar and vocabulary that have been developed artificially, rather than being evolved naturally over time.

- **Artlangs** are conlangs created for a fictional world, such as for a book, film, or video game.
- **Auxlangs** are conlangs created to facilitate real-world communication between people who do not speak any other languages in common.
- **Proto-languages** are hypothesized ancestral languages to extant human languages that have been reconstructed by comparative linguistics.

Phonetic Alphabet

The International Phonetic Alphabet (IPA) is a written alphabetic notation representing standardized pronunciation of possible sounds in human speech.

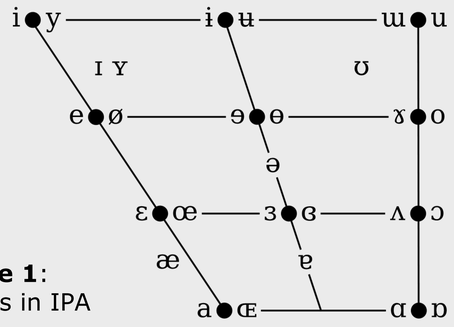


Figure 1: Vowels in IPA

How do we use the IPA in our system?

- **match** text in conlangs is converted to IPA
- **map** IPA symbols to spoken phonemes
- **concatenate** phonemes into audio files

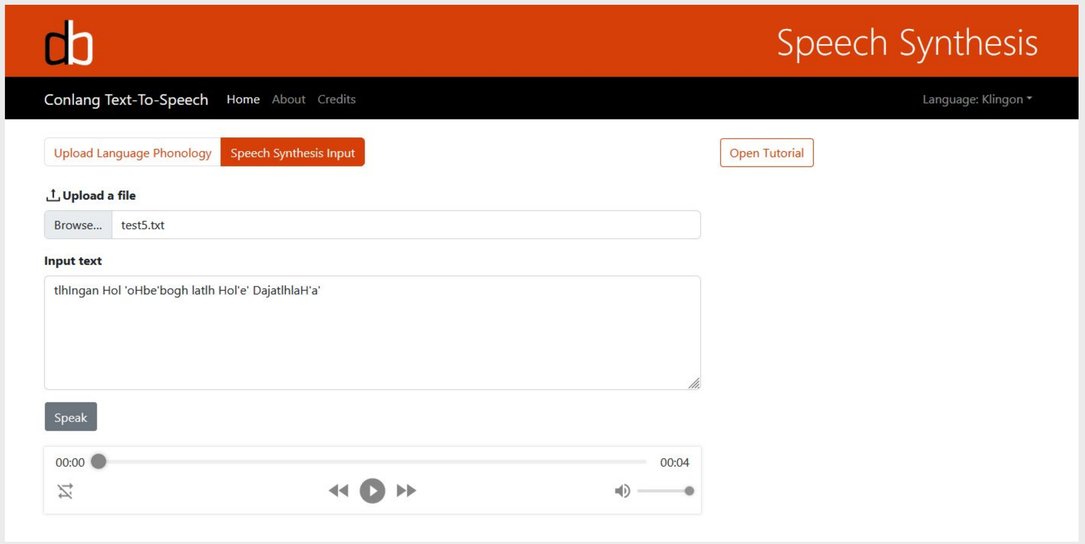
Synthesis Methods

Modern speech generation approaches rely on deep-learning and large datasets of spoken examples of the language. Absent this data, such methods are generally not appropriate for synthesizing most constructed languages.

This novel tool proposes a unique approach to synthesize hypothetical pronunciations without requiring large amounts of data. To generate speech clips for languages without any speakers or audio examples, we developed a innovative method to approximate this speech from the constitute phonetic sounds extracted from spoken examples of other languages. Our approach uses a combination of concatenative and formant synthesis techniques.

To realize words in a hypothesized language, we combine together short speech clips (one syllable or less) from a large database generated using an open-source software called `espeak-ng`, drawing from over 100 extant languages. This speech synthesizer creates speech sounds directly, which we categorize by original language, type of sound, corresponding IPA symbol, as well as context of the previous and following sounds in the source word.

Figure 2: Screenshot demonstrating the generation of speech.



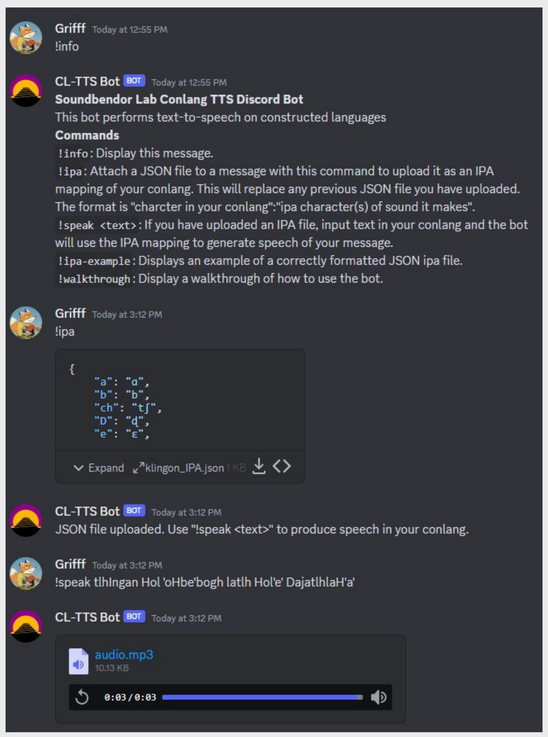


Figure 3: Example use of our Discord bot.

ACKNOWLEDGEMENTS

Team Members

- Nicholas Burrell
- Zachary Chand
- Jungmin Lee
- Gavin Lesher
- Avery Mao
- Griffin Riley
- Kevin Tran

Project Partner

Patrick Donnelly, Computer Science