Signal Processing

Speech Signal Processing

Speech Information Processing

Role of language

- **Semantics:** The meanings of words, and relations among them.
- **Syntax:** The order of words, role of function words.
- **Phonology:** Individual phonemic segments, features, stressed and unstressed vowels.
- For example,
  - What is the phonemic inventory of English?
  - How does it function? The concept of contrast (e.g., **pat** vs. **bat**).
  - Why do we believe that it is psychologically real?
  - Why does the same phoneme give rise to different, acoustic realizations in different utterances? (e.g., In fluent speech, "Joe ate his soup" loses the /h/ of "his", and the /t/ of "ate" doesn't look like a /t/ in "Tom".)
  - What are the principles that lead to modifications of segments in different environments?
  - How are phonemes usually described in terms of features translated into phonetic representations? (e.g., /z/ is + voiced, /s/ is - voiced; same relation for many pairs, like f-v; patterning of sounds is beautifully captured by feature concept.)

New terms that you will know..

Voiced, Unvoiced, pitch, intensity, timbre, formants, speech production, vocal tract, vocal cord, phonemes, manner and place of articulation, co-articulation, linear prediction, homo-morphic filtering, spectrograms, speech coding, speech enhancement, speech recognition, deep learning for speech, acoustic modelling, time/pitch scale modification, speech synthesis, Human auditory system, speech perception, speech quality measures (MOS, PESQ).....

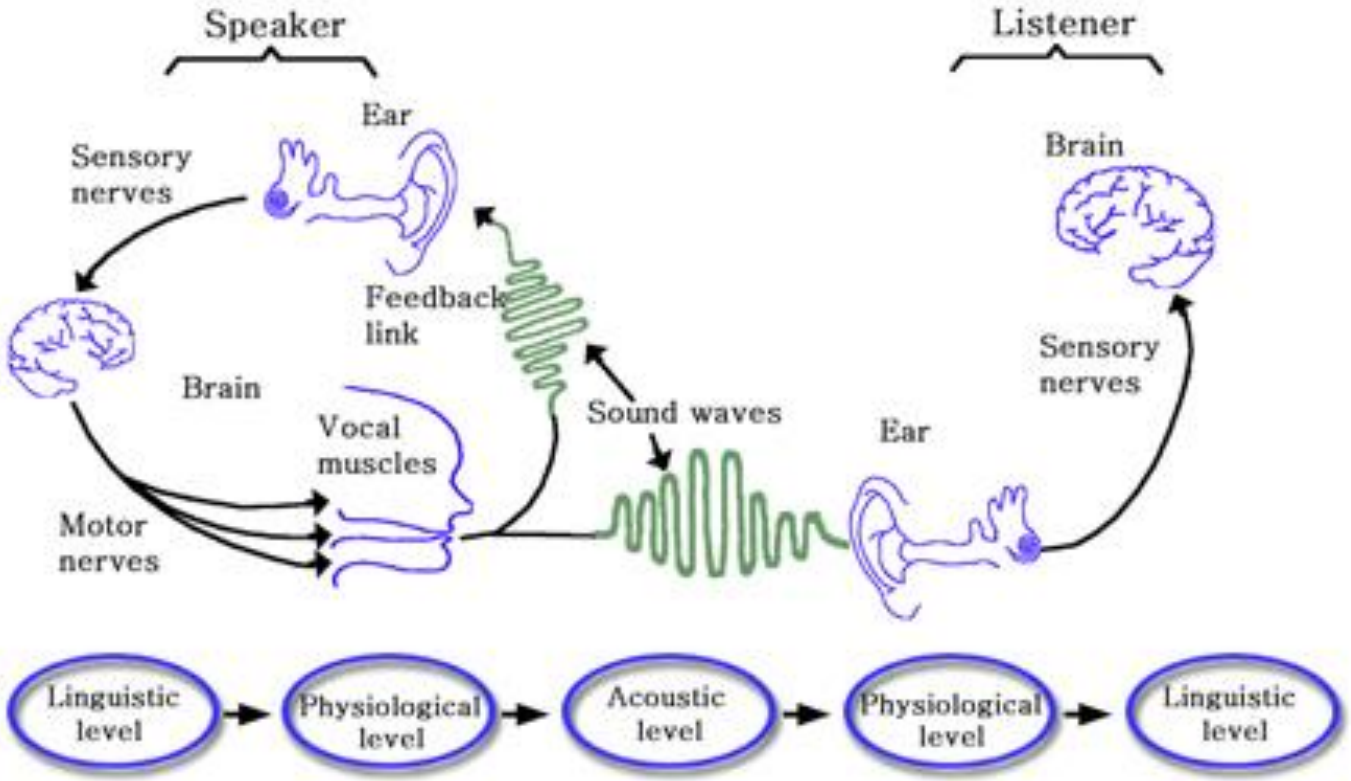
Speech Information Processing

Speech Information Understanding and Modeling

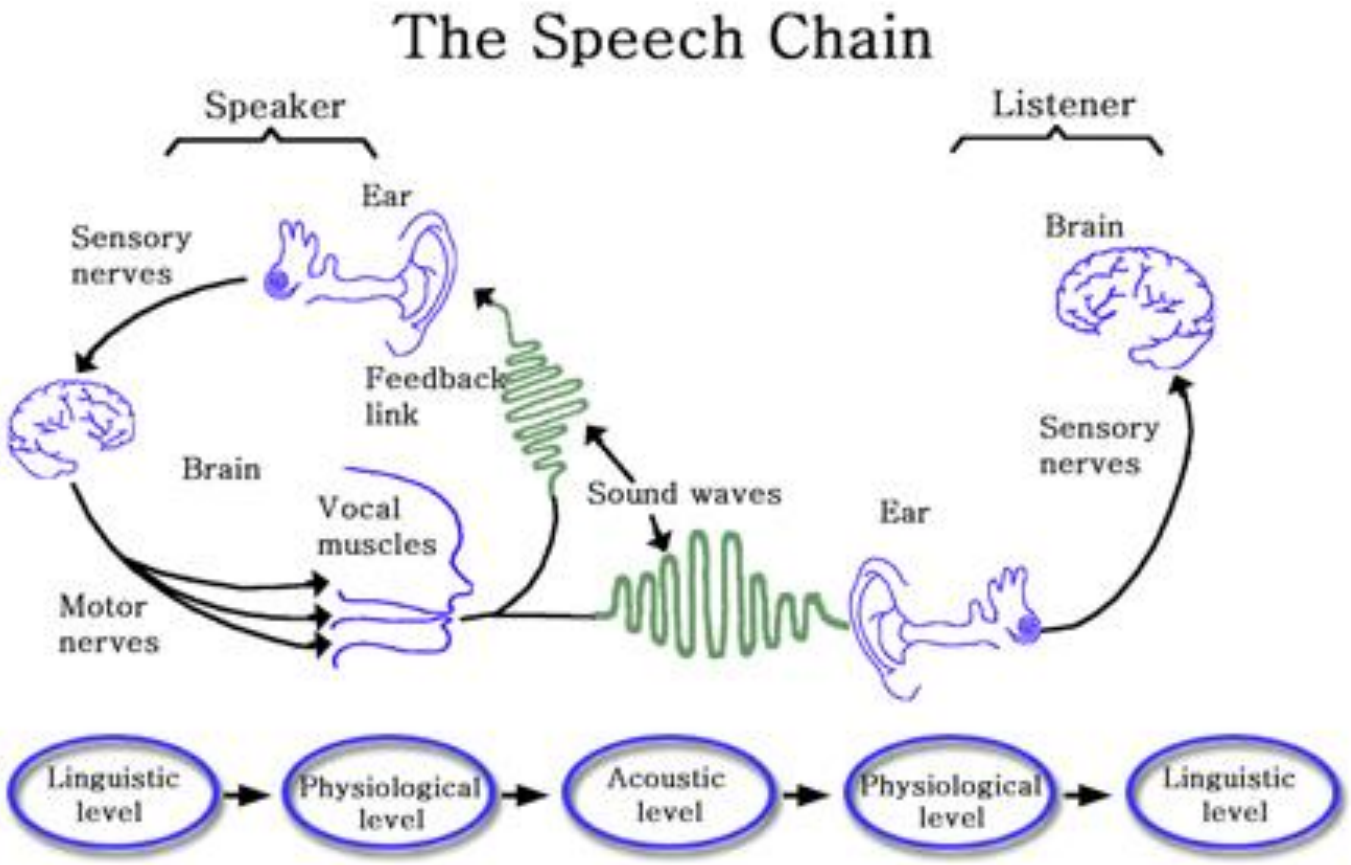
What is information in speech and how it is encoded?

Lets give it a try

The Speech Chain



- Speech Research –
  - Speech Science
    - - Linguistics
    - - Physiology of Speech Production
    - - Acoustics
    - - Auditory Nervous System
    - - Psychophysics of Auditory System
    - - Cognitive Psychology
    - - Computer-based Algorithms
  - Speech Technology
    - Speech Recognition
    - Speaker Recognition
    - Speech Synthesis
    - Speech coding



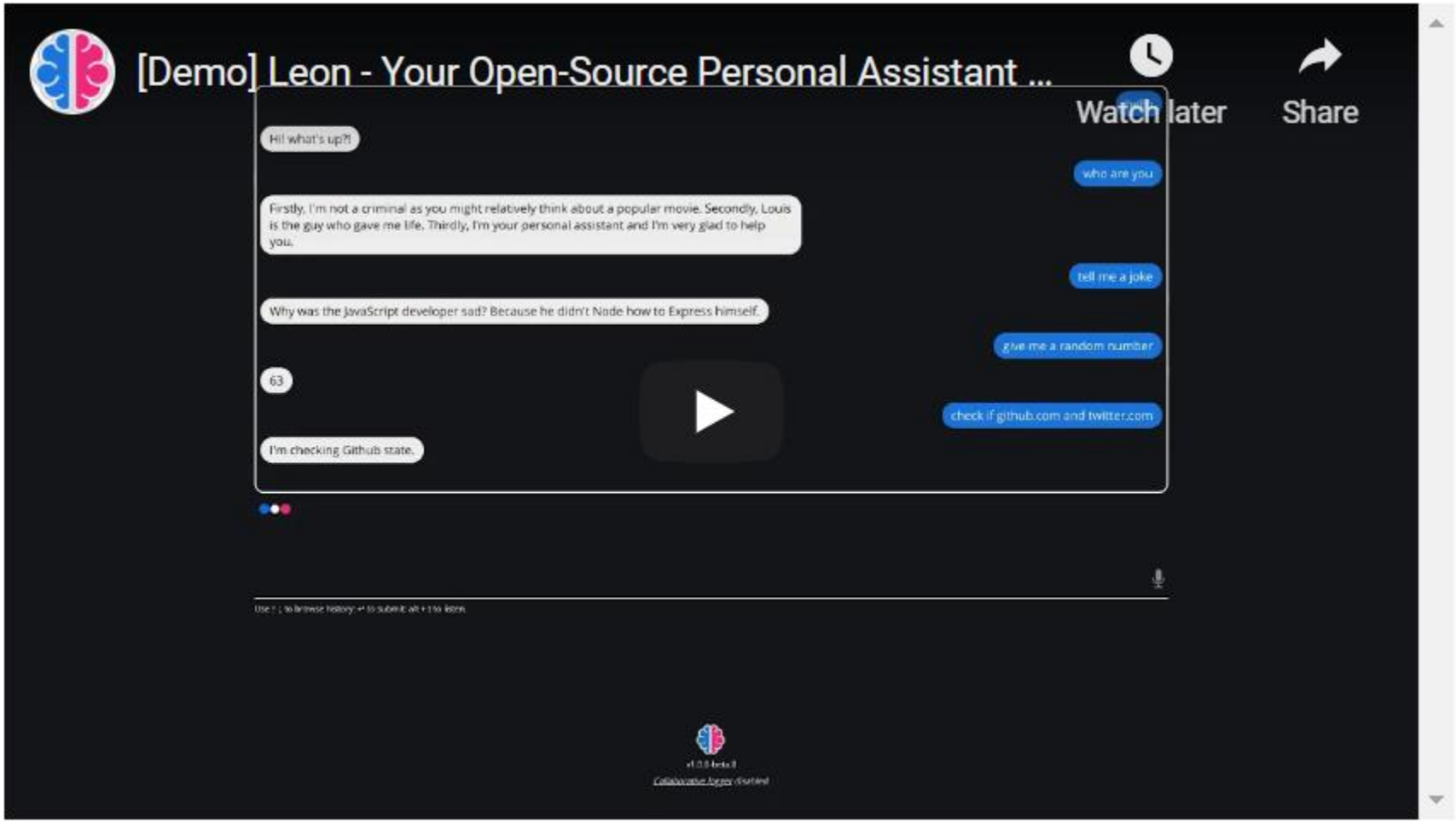
- Speech Coding
- Speech Synthesis
- Speech Recognition
- Speech Understanding
- Speaker Recognition
- Language Recognition

- Challenges to machine speech processing:
  - Definition of information content
    - Multiple levels of information
    - Subjectivity of the listener
  - Robustness to
    - Interfering signals
    - Partial information
  - Algorithmic complexity

Speech Recognition

Source: https://mycroft.ai/blog/deepspeech-update/?cn-reloaded=1

Speech Recognition



Source: https://github.com/leon-ai/leon

Speech Recognition

Source: https://www.techleer.com/articles/455-wav2letter-a-facebook-ai-researchfair-automatic-speech-recognition-toolkit/

Speech Recognition

annyang! SpeechRecognition that just works

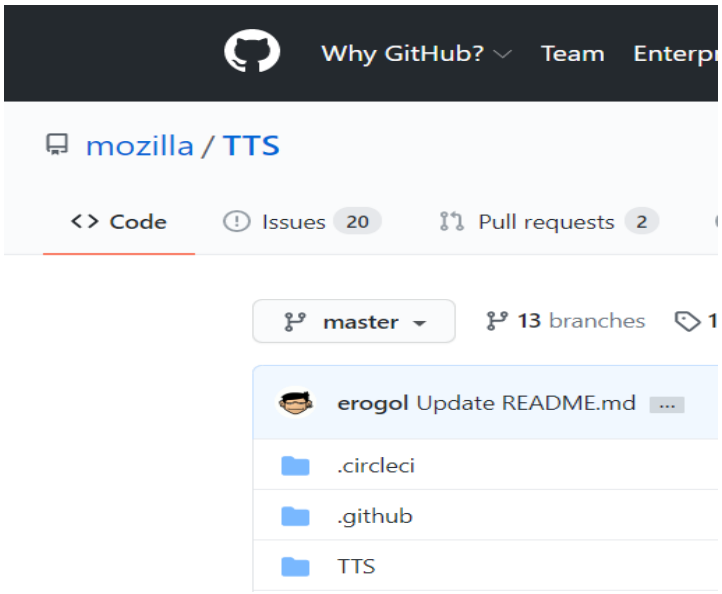
annyang is a tiny javascript library that lets your visitors control your site with voice commands.
annyang supports multiple languages, has no dependencies, weighs just 2kb and is free to use.

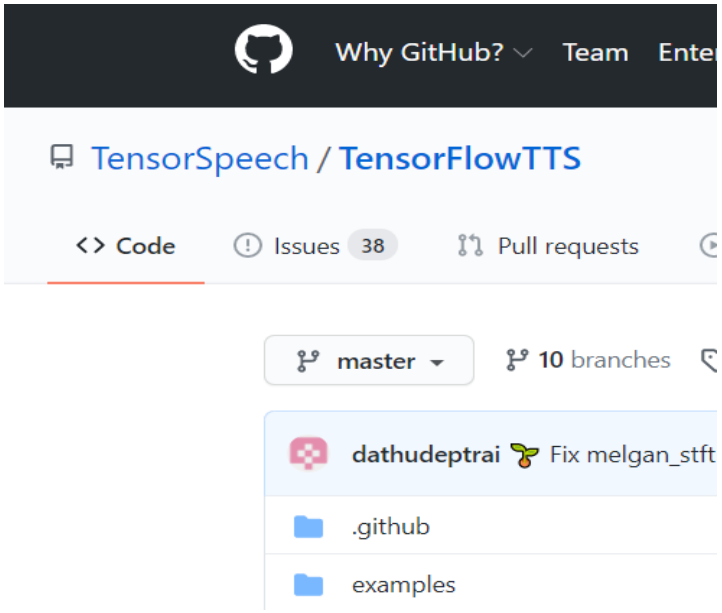
Source: https://github.com/TalAter/annyang

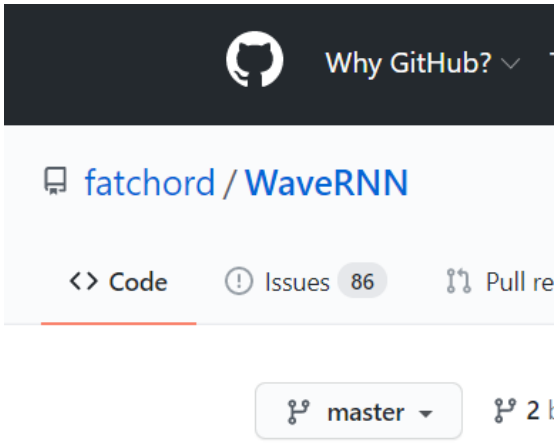
Speech Recognition

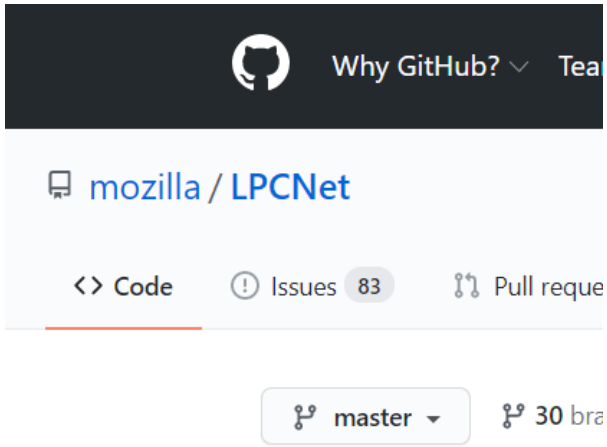
Source: https://pypi.org/project/SpeechRecognition/

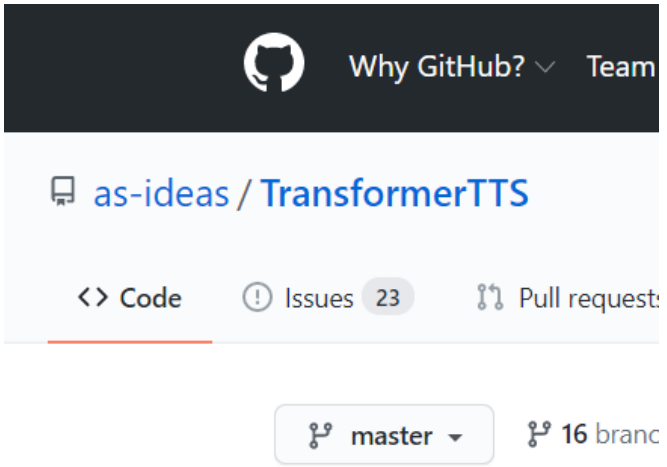
Text-to-Speech Synthesis



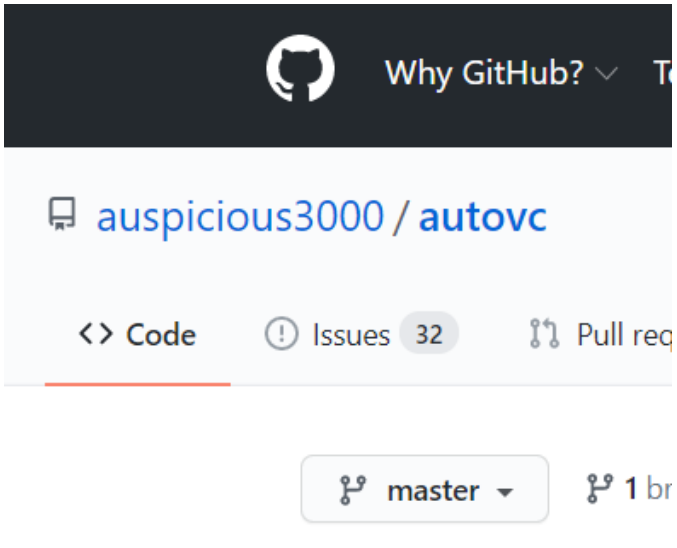


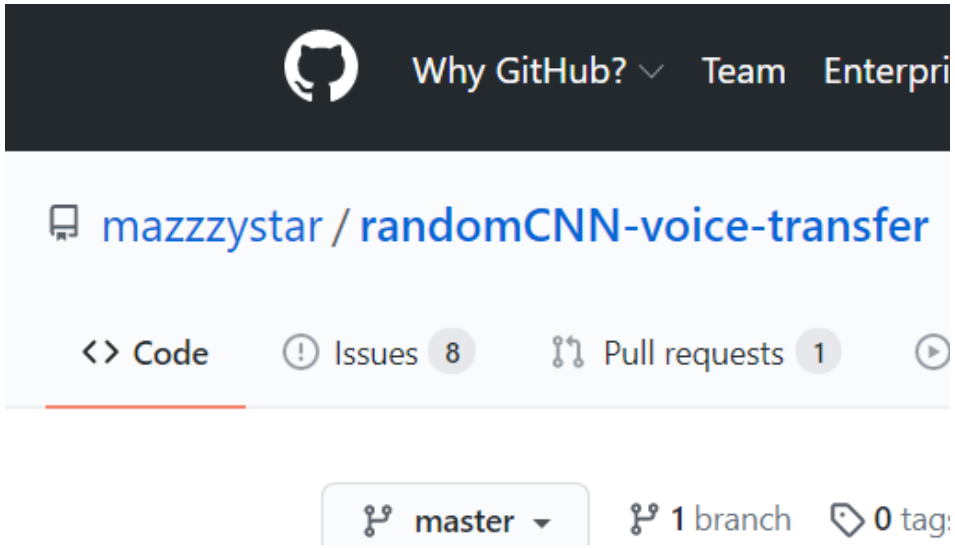


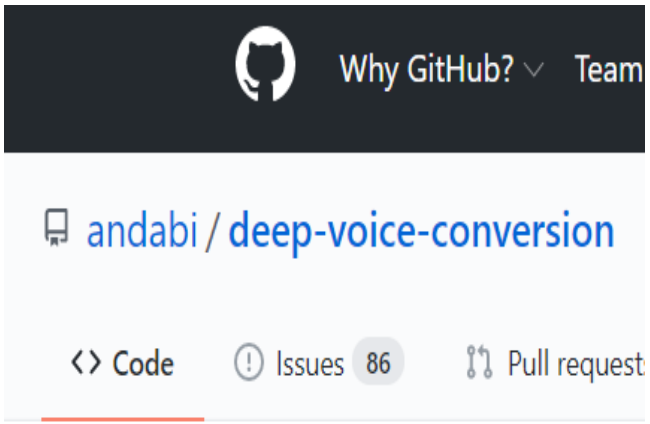


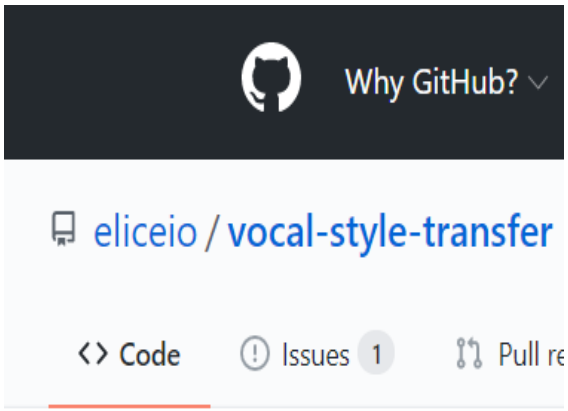


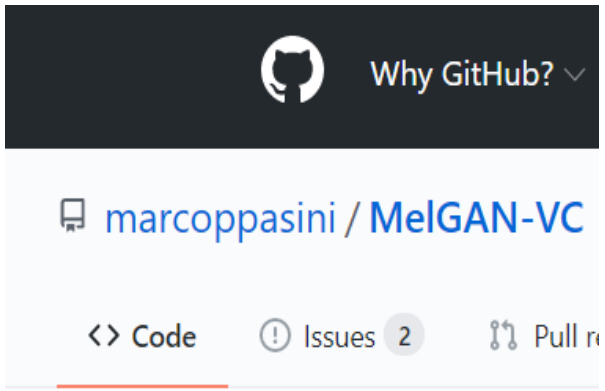
Voice conversion/ Style Transfer



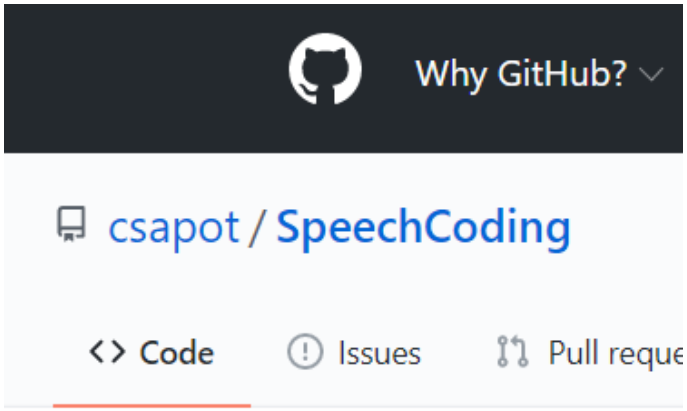


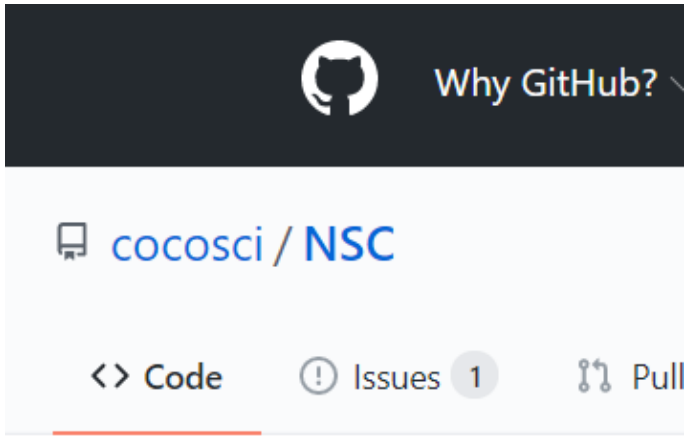


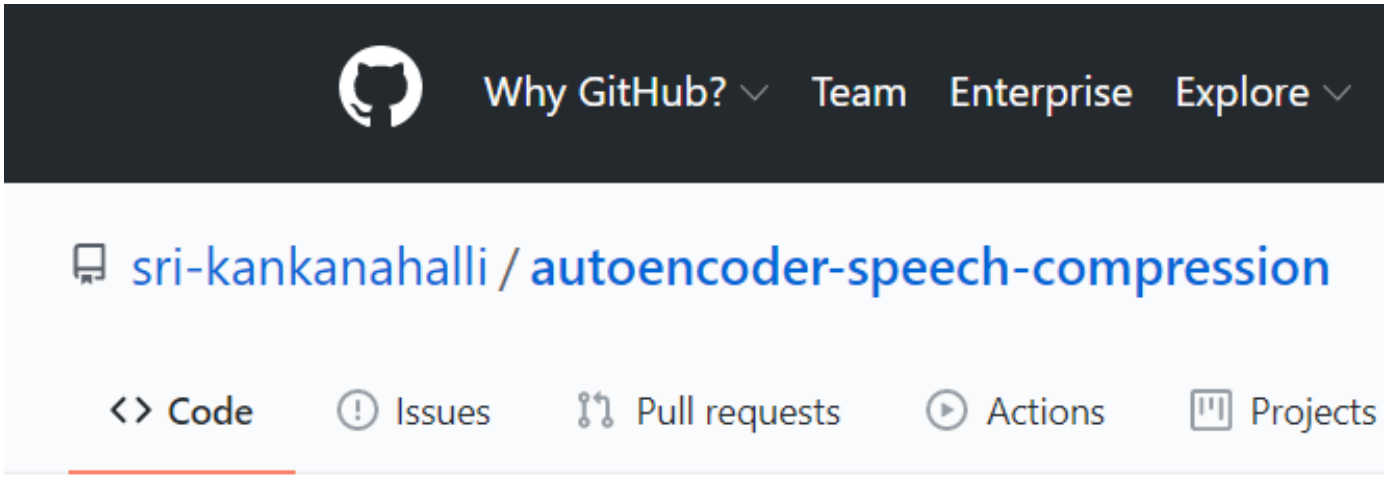




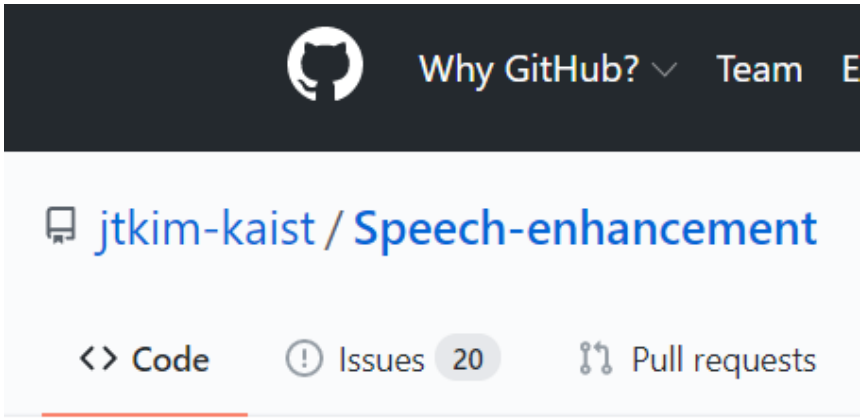
Speech Coding







Speech Enhancement



And so on...

Voice activity detection
Speaker recognition
Speaker diarization
Emotion recognition
Conversational systems
Voice cloning
Pronunciation evaluation
Language identification
……