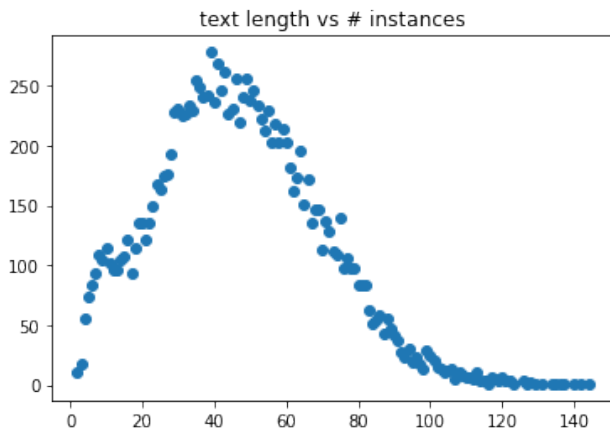# „Evolution" of german tts (thorsten) dataset
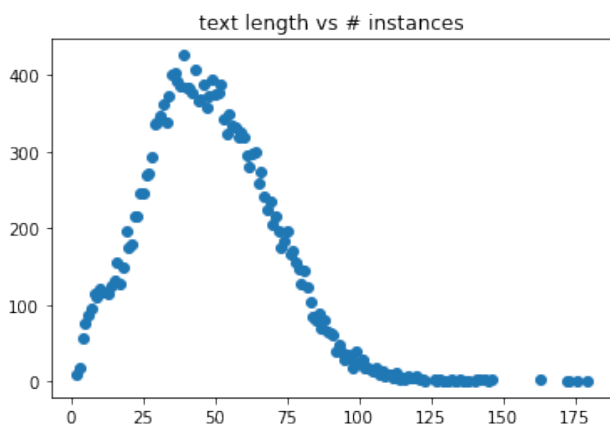
text length vs. # instances

„phase 1":   cheap usb microphone and
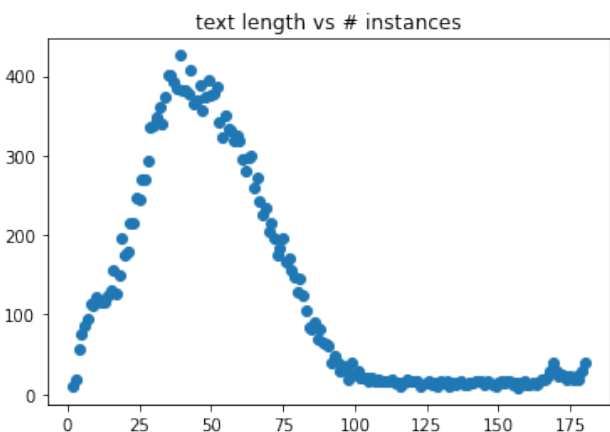not optimal recording room

Recording time: 2019-10-01 to 2020-01-07
Recorded phrases: 14.253
Phrase length (chars): 2 (min), 144 (max), 46.12 (avg)
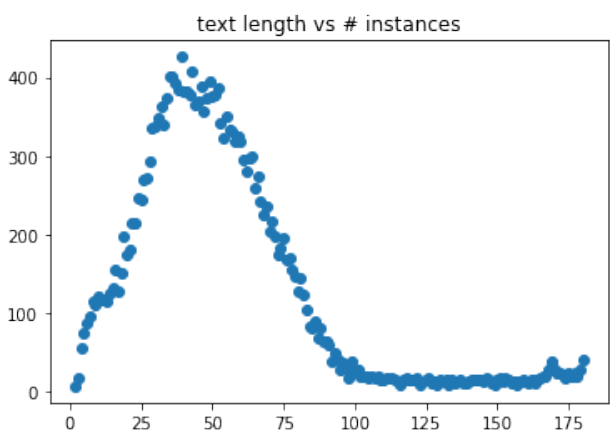
Problems with random noise and echos.

„phase 2":   better audio microphone and
room situation
additional to „phase 1"

Recording time: 2020-01-28 to 2020-04-02
Recorded phrases: 21.589
Phrase length (chars): 2 (min), 179 (max), 47.49 (avg)

„phase 3":   Recorded more phrases between
100 and 180 chars
additional to „phase 2"

Recording time: 2020-05-17 to 2020-06-22
Recorded phrases: 22.672
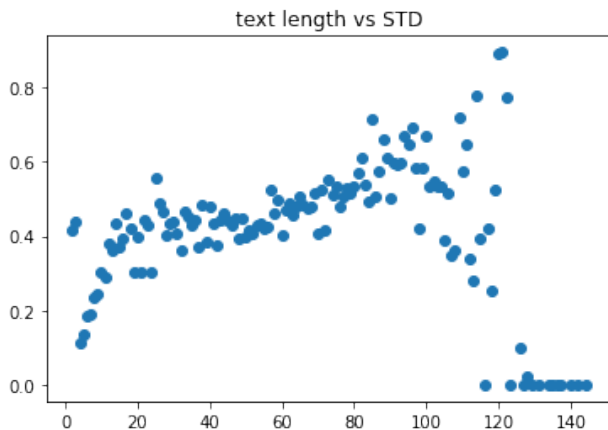Phrase length (chars): 2 (min), 180 (max), 52.35 (avg)

„final phase":   removed worse recordings and
optimized recordings on echo
and random noise
Recorded phrases: 22.668

- stereo to mono
- rnnoise
- highpass 100
- lowpass 7000
- samplerate 22050

Gladly donated to humanity by Thorsten Müller (https://github.com/thorstenMueller/deep-learning-german-tts)

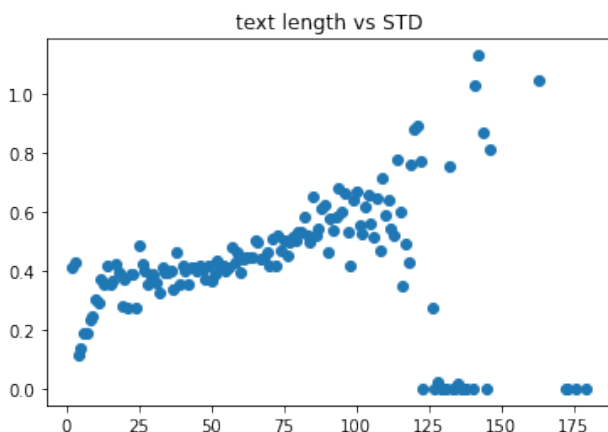# „Evolution" of german tts (thorsten) dataset

text length vs. STD

„phase 1":   cheap usb microphone and
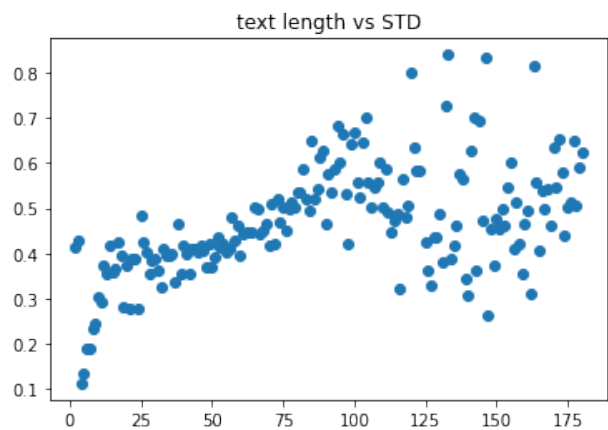not optimal recording room

Recording time: 2019-10-01 to 2020-01-07
Recorded phrases: 14.253
Phrase length: 2 (min), 144 (max), 46.12 (avg)
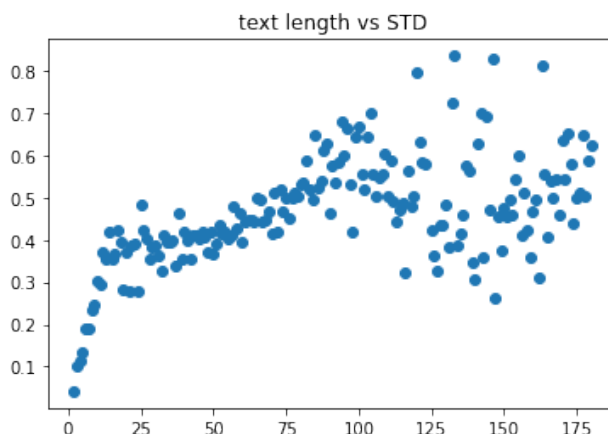
Problems with random noise and echos.

„phase 2":   better audio microphone and
room situation
additional to „phase 1"

Recording time: 2020-01-28 to 2020-04-02
Recorded phrases: 21.589
Phrase length: 2 (min), 179 (max), 47.49 (avg)

„phase 3":   Recorded more phrases between
100 and 180 chars
additional to „phase 2"

Recording time: 2020-05-17 to 2020-06-22
Recorded phrases: 22.672
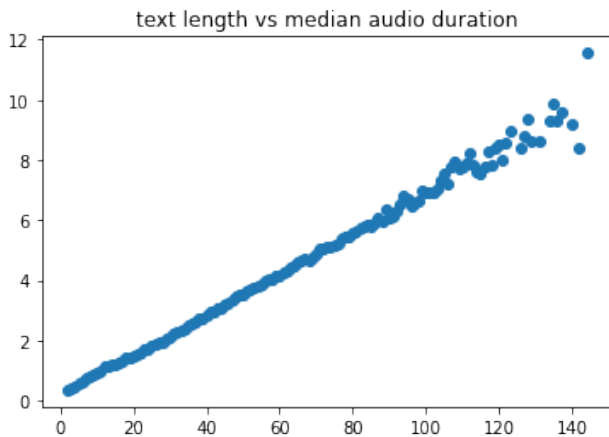Phrase length: 2 (min), 180 (max), 52.35 (avg)

„final phase":   removed worse recordings and
optimized recordings on echo
and random noise
Recorded phrases: 22.668

- stereo to mono
- rnnoise
- highpass 100
- lowpass 7000
- samplerate 22050

Gladly donated to humanity by Thorsten Müller (https://github.com/thorstenMueller/deep-learning-german-tts)

„Evolution" of german tts (thorsten) dataset
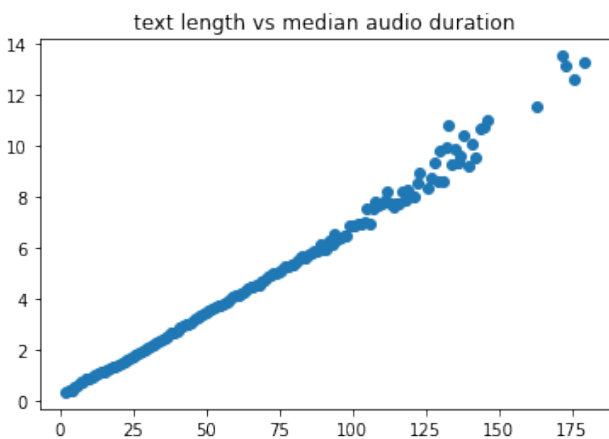
text length vs. median audio duration

text length vs median audio duration

„phase 1": cheap usb microphone and
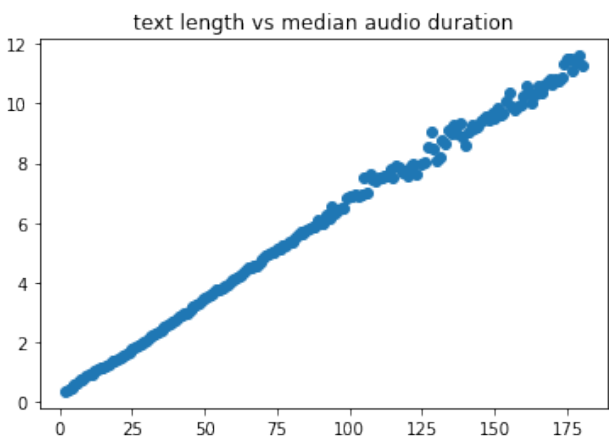not optimal recording room

Recording time: 2019-10-01 to 2020-01-07
Recorded phrases: 14.253
Phrase length: 2 (min), 144 (max), 46.12 (avg)

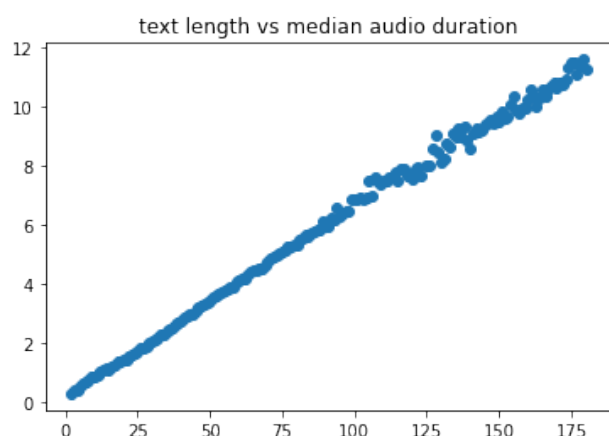Problems with random noise and echos.

text length vs median audio duration

„phase 2": better audio microphone and
room situation
additional to „phase 1"

Recording time: 2020-01-28 to 2020-04-02
Recorded phrases: 21.589
Phrase length: 2 (min), 179 (max), 47.49 (avg)

text length vs median audio duration

„phase 3": Recorded more phrases between
100 and 180 chars
additional to „phase 2"

Recording time: 2020-05-17 to 2020-06-22
Recorded phrases: 22.672
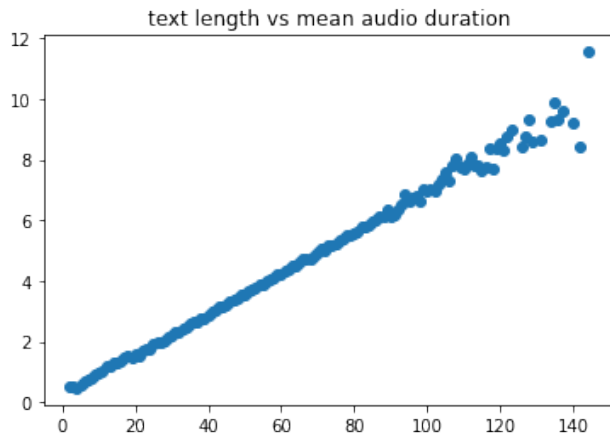Phrase length: 2 (min), 180 (max), 52.35 (avg)

text length vs median audio duration

„final phase": removed worse recordings and
optimized recordings on echo
and random noise
Recorded phrases: 22.668

- stereo to mono
- rnnoise
- highpass 100
- lowpass 7000
- samplerate 22050

Gladly donated to humanity by Thorsten Müller (https://github.com/thorstenMueller/deep-learning-german-tts)

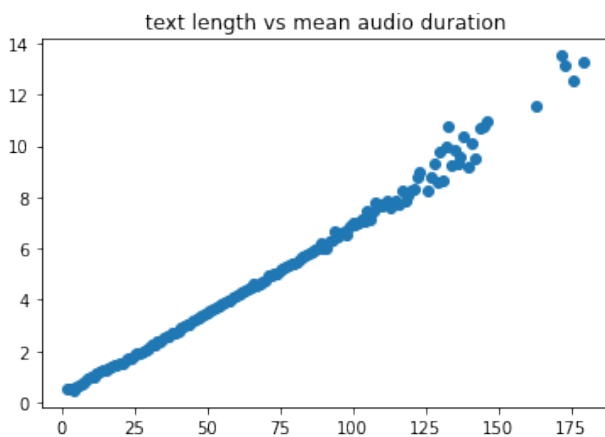# „Evolution" of german tts (thorsten) dataset

text length vs. mean audio duration

„phase 1":   cheap usb microphone and
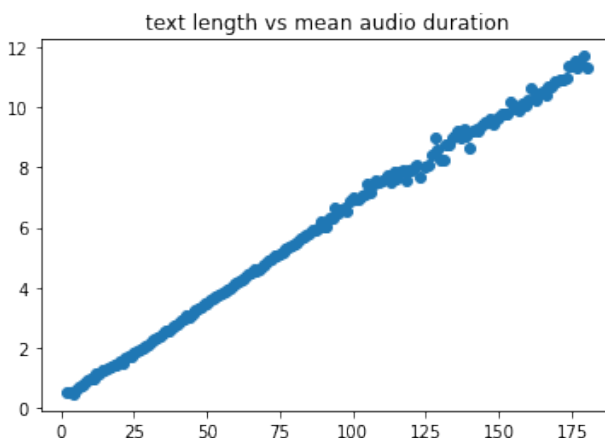not optimal recording room

Recording time: 2019-10-01 to 2020-01-07
Recorded phrases: 14.253
Phrase length: 2 (min), 144 (max), 46.12 (avg)

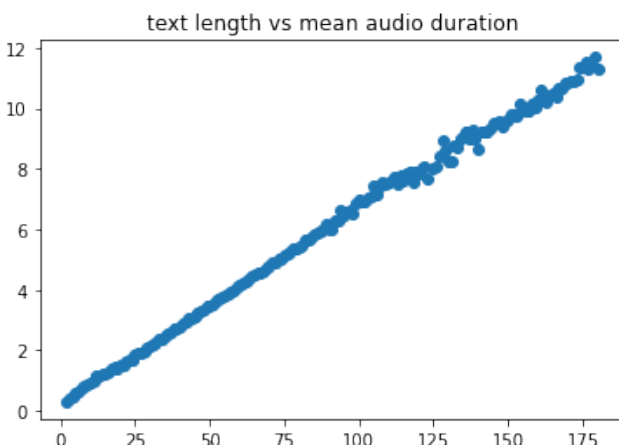Problems with random noise and echos.

„phase 2":   better audio microphone and
room situation
additional to „phase 1"

Recording time: 2020-01-28 to 2020-04-02
Recorded phrases: 21.589
Phrase length: 2 (min), 179 (max), 47.49 (avg)

„phase 3":   Recorded more phrases between
100 and 180 chars
additional to „phase 2"

Recording time: 2020-05-17 to 2020-06-22
Recorded phrases: 22.672
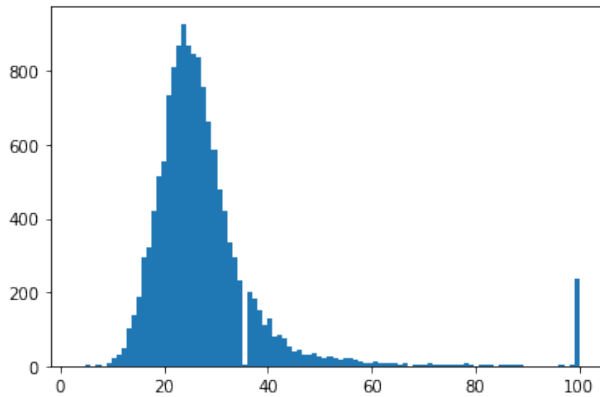Phrase length: 2 (min), 180 (max), 52.35 (avg)

„final phase":   removed worse recordings and
optimized recordings on echo
and random noise
Recorded phrases: 22.668

- stereo to mono
- rnnoise
- highpass 100
- lowpass 7000
- samplerate 22050

Gladly donated to humanity by Thorsten Müller (https://github.com/thorstenMueller/deep-learning-german-tts)

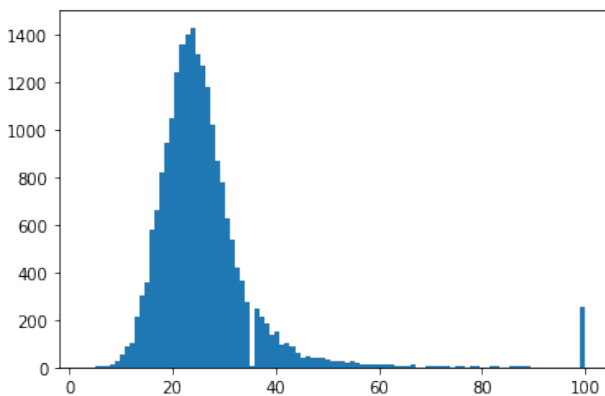# „Evolution" of german tts (thorsten) dataset

snr (signal noise ratio)

„phase 1":   cheap usb microphone and
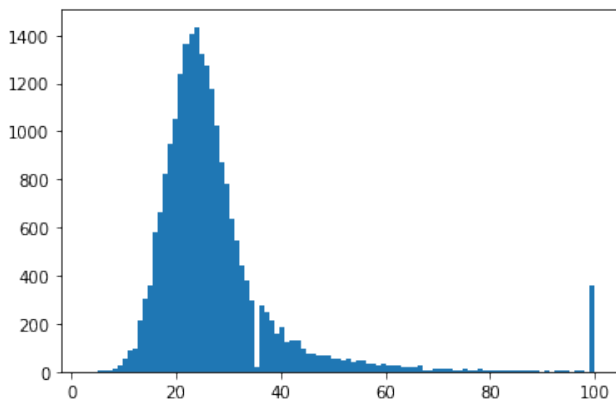   not optimal recording room

Recording time: 2019-10-01 to 2020-01-07
Recorded phrases: 14.253
Phrase length: 2 (min), 144 (max), 46.12 (avg)
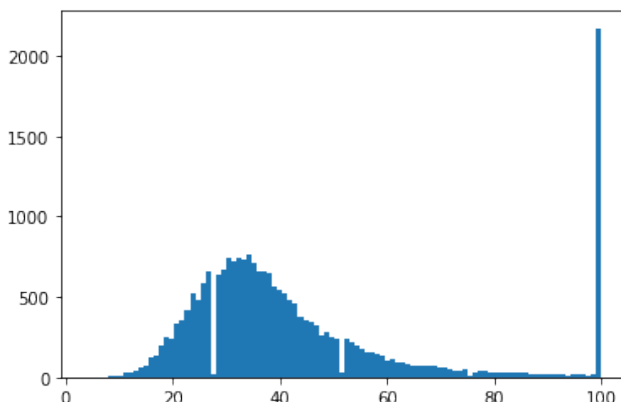
Problems with random noise and echos.

„phase 2":   better audio microphone and
             room situation
             additional to „phase 1"

Recording time: 2020-01-28 to 2020-04-02
Recorded phrases: 21.589
Phrase length: 2 (min), 179 (max), 47.49 (avg)

„phase 3":   Recorded more phrases between
             100 and 180 chars
             additional to „phase 2"

Recording time: 2020-05-17 to 2020-06-22
Recorded phrases: 22.672
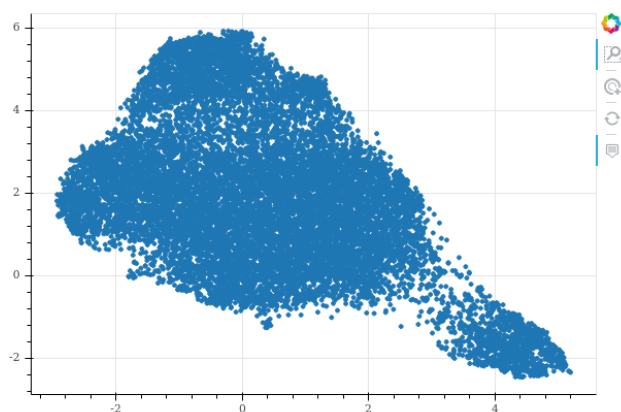Phrase length: 2 (min), 180 (max), 52.35 (avg)

„final phase":   removed worse recordings and
             optimized recordings on echo
             and random noise
Recorded phrases: 22.668

             - stereo to mono
             - rnnoise
             - highpass 100
             - lowpass 7000
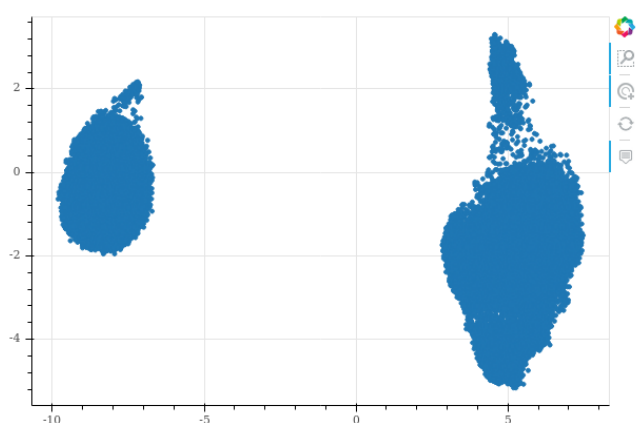   - samplerate 22050

„Evolution" of german tts (thorsten) dataset

resemblyzer bokeh

„phase 1":   cheap usb microphone and
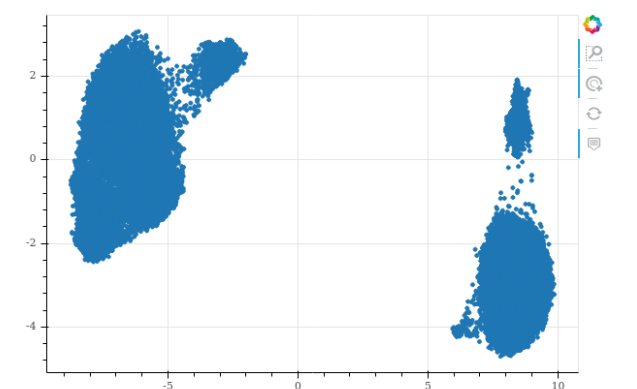not optimal recording room

Recording time: 2019-10-01 to 2020-01-07
Recorded phrases: 14.253
Phrase length: 2 (min), 144 (max), 46.12 (avg)
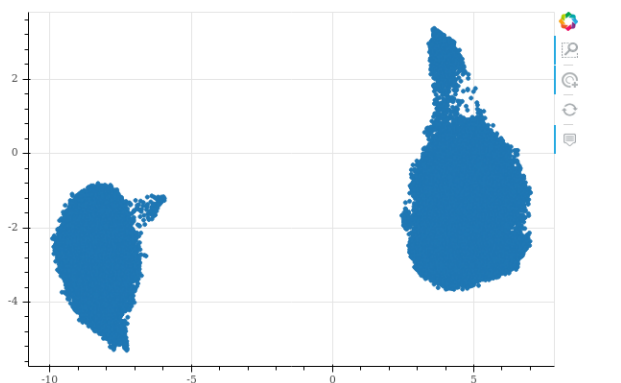
Problems with random noise and echos.

„phase 2":   better audio microphone and
room situation
additional to „phase 1"

Recording time: 2020-01-28 to 2020-04-02
Recorded phrases: 21.589
Phrase length: 2 (min), 179 (max), 47.49 (avg)

„phase 3":   Recorded more phrases between
100 and 180 chars
additional to „phase 2"

Recording time: 2020-05-17 to 2020-06-22
Recorded phrases: 22.672
Phrase length: 2 (min), 180 (max), 52.35 (avg)

„final phase":   removed worse recordings and
optimized recordings on echo
and random noise
Recorded phrases: 22.668

- stereo to mono
- rnnoise
- highpass 100
- lowpass 7000
- samplerate 22050

Gladly donated to humanity by Thorsten Müller (https://github.com/thorstenMueller/deep-learning-german-tts)

Phonemes Distribution

Phoneme coverage on dataset thorsten-de created on 2020-06-30_11-32