

CORPUS INFORMATION FORM

Check (x) or fill in the blank (_____) if appropriate.

Name of the Corpus: TH CoSS-0 (Corpus of Speech Synthesis No. 0)

1. IPR Holder: HCI&MM/Tsinghua University

2. Corpus Type:

- Speech Corpus (x)
- Text Corpus ()

3. If it is a speech corpus:

- Purpose:
 - ASR ()
 - TTS (x)
 - Other, please specify _____
- Language:
 - Putonghua (SC) (x)
 - Mandarin in Taiwan (TC) ()
 - Cantonese in HK (TC) ()
 - Other, please specify _____
- Style:
 - Read speech (x)
 - Spontaneous speech ()
 - Conversational speech ()
 - Other, please specify _____
- Channel:
 - Close-talk Microphone (x)
 - Telephone ()
 - Mobile phone ()
 - Other, please specify _____
- Sampling Rate: 8KHz k Hz
- Sampling Precision:
 - PCM (x), _____ bits per sample
 - A-law ()
 - Miu-law ()
 - Other, please specify _____
- Corpus size: _____ hours 1 speakers 3,000 Chinese syllables
- SNR level: _____ dB
- Transcriptions:
 - Character tier (SC) (x)

- Character tier (TC) ()
- Canonical Pinyin tier ()
- Other canonical pronunciation tier, please specify _____
- Surface form IF tier ()
- Surface form IPA tier ()
- Surface form SAMPA-C tier ()
- Other surface form tier, please specify _____
- Other transcription, please specify _____
- Other transcription, please specify _____
- Other transcription, please specify _____

4. If it is a text corpus:

- Language:
 - SC ()
 - TS ()
 - Other, please specify _____
- Domain:
 - Culture ()
 - Economy ()
 - Military ()
 - News ()
 - Politics ()
 - Sciences ()
 - Sports ()
 - Other, please specify _____
 - Other, please specify _____
 - Other, please specify _____
 - Other, please specify _____
 - Other, please specify _____
 - Other, please specify _____
- Corpus size: _____ Mega characters
- Tag Information:
 - Word segmentation: ()
 - Part-of-Speech ()
 - Other, please specify _____
 - Other, please specify _____
 - Other, please specify _____
 - Other, please specify _____

5. A brief Description of the Corpus:

It is a database for Chinese TTS system, male voice.