

CORPUS INFORMATION FORM

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

1. Name of the Corpus: BIT-MobileTalk

(Mobile Phone Conversational Speech Corpus for Travel)

2. IPR Holder:

Modern Communication Lab of Beijing Institute Technology

3. Corpus Type:

- Speech Corpus (x)
- Text Corpus ()

4. If it is a speech corpus:

- Purpose:

- ASR (x)
- TTS ()
- Other, please specify speaker segmentation , speaker recognition and key-words spotting

- Language:

- Putonghua (SC) (x)
- Mandarin in Taiwan (TC) ()
- Cantonese in HK (TC) ()
- Other, please specify _____

- Style:

- Read speech ()
- Spontaneous speech ()
- Conversational speech (x)
- Other, please specify _____

- Channel:

- Close-talk Microphone ()
- Telephone (x)
- Mobile phone (x)
- Other, please specify _____

- Sampling Rate: 8 k Hz

- Sampling Precision:

- PCM (x), 8 bits per sample
- A-law ()
- Mu-law ()

- Other, please specify _____
- Corpus size: _____ 3.7 hours _____ 30 speakers _____ 107MB
- SNR level: _____ 15~27 _____ dB
- Transcriptions:
 - Character tier (SC) ()
 - 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 _____

5. If it is a text corpus:

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

6. A brief Description of the Corpus:

This is a conversational speech corpus by 30 speakers, half male and half female. The speakers are divided into groups and each group consists of 6 speakers (3 males and 3 females). In each group, one speaker has initiated calls to the other 5 speakers using his/her mobile phone and received calls from the other 5 speakers using a fixed wired-telephone. So any two speakers in one group talked twice and there are $5 \times 6 = 30$ conversations in one group. And we have collected 6 groups (i.e. 180 conversations) by now.

The corpus has some particular characteristics:

- 1) The telephone networks contain not only PSTN but also GSM and CDMA; and the

terminal equipments contain various mobile telephones with different types from different manufacturers.

- 2) The conversation between speakers is goal-oriented, that is, we have assigned a topic which is a travel site to each conversation. And the length of each conversation is about one minute and a half. Moreover, we have provided some tips, such as how to go there and the admission charge, to the speakers. So there are many kinds of keywords.
- 3) The conversations are spontaneous and contain variant phenomena of spoken language.

The conversational corpus can be used for speaker segmentation, speaker recognition and KWS. Moreover, the Mobile Phone Speech Corpus for Traffic Information Query in which the speakers are the same as this corpus can be used to train speaker models when applied in speaker recognition. This corpus is being updated and extended furtherly.