

Application Form for ATCSpeech Corpus

All the rights of the ATCSpeech corpus are reserved by:

- National Key Laboratory of Air Traffic Control Automation System Technology, Sichuan University. No.24 South Section 1, Yihuan Road, Chengdu, China, 610065.
- Wisesoft Co. Ltd. No 7, East Wuke 1st Road, Chengdu, China, 610045.

GENERAL INFORMATION

Organization category	<input type="checkbox"/> University <input type="checkbox"/> Educational institution <input type="checkbox"/> Enterprise <input type="checkbox"/> other (please specify)_____
Official organization name	
Contacts	
Email address	
Telephone number	
Other information	

DATA USAGE

Research interest	
Application reason	
Expected outcome	
Other information	

AUTHORIZED SIGNATURE

Applicant: _____

Date: _____

ADDITIONAL COMMENTS

- This corpus can only be used to academic research for university or educational institution.
- In any case, no data transfer, sell or any other commercial usage are allowed.
- An acknowledgment is required for any research outcome based on this corpus.
- Any correction or suggestion for this corpus will be highly appreciated.
- If you are interested in our work or make further research based on it, please cite our paper (<https://arxiv.org/abs/1911.11365>) as:

"Yang B, Tan XL, Chen ZM, Wang B, Li D, Yang ZP, Wu ZP, Lin, Y, ATCSpeech: a multilingual pilot-controller speech corpus from real Air Traffic Control environment [J], <https://arxiv.org/abs/1911.11365>"

If you have any concern about this corpus, please feel free to contact us at:

boyang@scu.edu.cn or yilin@scu.edu.cn.

ATCSpeech-Data-Feature

1 Data Description

The total duration of ATCSpeech corpus is about 57.8 hours, including 39-hour Chinese speech and 18.8-hour English speech. All the speeches are the pilot-controller conversations of air traffic control in China, concerning three aerodrome tower areas and one approach area. The raw speeches were collected from real air traffic control environment in China, which were recorded and stored by the MDSL device of Wisesoft Co. Ltd. The sample rate of the speeches in this corpus is 8000Hz. Over 10 frontline air traffic controllers contribute to developing the annotation standard for ATCSpeech corpus. More than 40 people participate in the corpus annotating and correcting.

2 Text

The speech corpus is collected from air traffic control environment, in which the air traffic related terminologies, code-switching and common communication procedure are implied in the transcription. Moreover, some special rules are also performed on the transcriptions, as shown below:

- 1) the texts concerning certain air traffic control areas or centers are removed.
- 2) all the symbols or punctuations are removed.
- 3) only the Chinese Character and English letter are allowed in the transcriptions.
- 4) the texts are formulated in one line with word segmentation.

3 Annotation

The transcription is generated by the actual human hearing under the following rules:

- 1) The transcription must be exactly same with that of in the speeches, no more, less or wrong words allowed.

- 2) All the digitals must be converted into Chinese character instead of Arabic numerals.
- 3) All the English speeches are denoted by English letters in lower case.

4 Attribution

In addition to the transcription, the following information of the speeches in the ATCSpeech corpus are also provided:

- a) Speaker gender: male (M) and female (F).
- b) Speaker role: pilot (P) and controller (C).
- c) Speech quality: clean (C) and noise (N).
- d) Flight phase: ground (GND), aerodrome tower (TWR), approach (APP) and en-route area control center (ACC).

5 Data organization

The architecture of the directory is shown below:

```
| ATCSpeech-Data-Feature .docx
| README.txt
| transcription.csv
└─ data
    | └─CN/EN
    |   | └─train/dev/test
    |   |   | └─TWRx
    |   |   |   | └─xxx.wav/feat
```