

AESRC2020 Q&A

Q1: Will baseline experiments be released to participants for each track?

A1: Yes, we will release the accuracy of accent recognition and WER of accent speech recognition on the DEV set. For comparing, we suggest divide the DEV set according to the specific speakers.

Q2: Will accent label be released together with the speech in the test set?

A2: No. Test set doesn't contain accent label.

Q3: Are there any accents that are not covered by the training set in the test set?

A3: Yes.

Q4: What should I do with the punctuation marks in the test set?

A4: No punctuation marks except “ . ” And “ ’ ” in “Mr.”, “I’m”, “you’ll” etc. will appear in the standard transcription. Positions that may be in dispute will be dealt with specially to ensure fairness.

Q5: Which data is permitted to use in language model training:

A5: Only transcription of permitted speech data can be used in language model training. Any kind of other text data is forbidden.