

The ARU speech corpus (University of Liverpool)

This corpus comprises single channel recordings of IEEE (Harvard) sentences (IEEE, 1969) spoken by twelve adult native British English speakers in anechoic conditions. The recordings were made in October and November 2017 in the Acoustics Research Unit (ARU) at the University of Liverpool as part of a research project on speech intelligibility by Dr Simone Graetzer, Dr Gary Seiffert and Professor Carl Hopkins.

All individuals included in this corpus have given consent for their recordings to be made available to the public. Further information about the corpus is provided below.

Talkers

The six female and six male talkers of between 21 and 47 years of age were required to speak British English as a first language and to have completed primary and secondary schooling in the UK. The county in which primary and secondary schooling took place is identified for each talker in Table 1. Talkers were also required to have accents that were not strongly regional, and accents that were close to Received Pronunciation were preferred. Talkers were screened to meet this requirement. In addition, talkers were required to be non-smokers who reported no history of hearing or speech impairments.

The hearing screening of all participants involved a pure tone audiometry test in an audiometric booth according to BS EN ISO 8253-1:2010. Talkers had thresholds of 20 dB HL or better (age adjusted) at frequencies ranging from 125 Hz to 8 kHz.

Table 1. Talkers providing the speech recordings.

ID	Gender	Age	County	Country
01	M	47	Avon	England
02	M	21	Ceredigion	Wales
03	F	23	Berkshire	England
04	F	35	Surrey and Middlesex	England
05	M	35	Denbighshire and Conwy	Wales
06	M	47	Kent	England
07	F	24	Norfolk	England
08	F	32	Merseyside	England
09	F	44	Wirral	England
10	M	29	Cheshire	England
11	F	45	East Sussex	England
12	M	32	Leicestershire	England

Recording setup

Twelve talkers were recorded by a microphone at 1m on axis while producing the 720 IEEE sentences in the ARU anechoic chamber (internal dimensions 5 x 4 x 2.6m). The sampling rate was 65536 Hz. A Brüel & Kjær (B&K) free-field half-inch microphone Type 4190 was used with a B&K preamplifier Type 2669 No. 3004348. This was connected to a B&K Nexus conditioning amplifier (Serial 2301697), which was itself connected to a B&K LAN-XI multipurpose generator module (Type 3160-A 4/2). This was connected via an ethernet switch to a personal computer running B&K Pulse Time Data Recorder v20.

Procedure

The talkers were seated in the ARU anechoic chamber and asked to produce the IEEE word list sentences with a normal vocal effort. The instruction was to '[s]peak normally, as you would in everyday conversation'. The sentences were spoken in a randomised order. If the talker hesitated or made an error, he or she was asked to repeat the sentence. The talkers were instructed to face the microphone when speaking, and the position of the talker's head was monitored via a video camera.

Signal processing

Signals were high-pass filtered to remove energy below 60 Hz (using a Finite Impulse Response filter with a Kaiser window method) and low-pass filtered to attenuate energy above 9 kHz (predominantly electrical background noise). To obtain the same active speech levels (ITU-T P.56, 2011), speech samples were normalised using the *activlev* function provided in VOICEBOX (Brookes, 2014-2016) before saving to wav format with 24 bits per sample.

Filenames

Each talker produced all 720 IEEE sentences divided into 72 word lists of 10 sentences each.

Talkers are identified with an ID number at the beginning of each filename, and the List field indicates the IEEE wordlist number.

How to cite

When you use the corpus, please cite this website in publications using the following information:

Hopkins, C., Graetzer, S., Seiffert, G. (2019). ARU adult British English speaker corpus of IEEE sentences (ARU speech corpus) version 1.0 [data collection]. Acoustics Research Unit, School of Architecture, University of Liverpool, United Kingdom. DOI: [10.17638/datacat.liverpool.ac.uk/681](https://doi.org/10.17638/datacat.liverpool.ac.uk/681).

References

Brookes, M. (2014-2016). VOICEBOX: Speech Processing Toolbox for MATLAB.
<http://www.ee.ic.ac.uk/hp/staff/dmb/voicebox/voicebox.html>

BS EN ISO 8253–1:2010 (2010). Acoustics: Audiometric test methods part 1: Basic pure tone air and bone conduction threshold audiometry. International Organization for Standardization.

IEEE (1969). Recommended practice for speech quality measurements, IEEE Transactions on Audio and Electroacoustics, 17 (3), 227-246.

ITU-T P.56. (2011). Objective measurement of active speech level. International Telecommunication Union.