

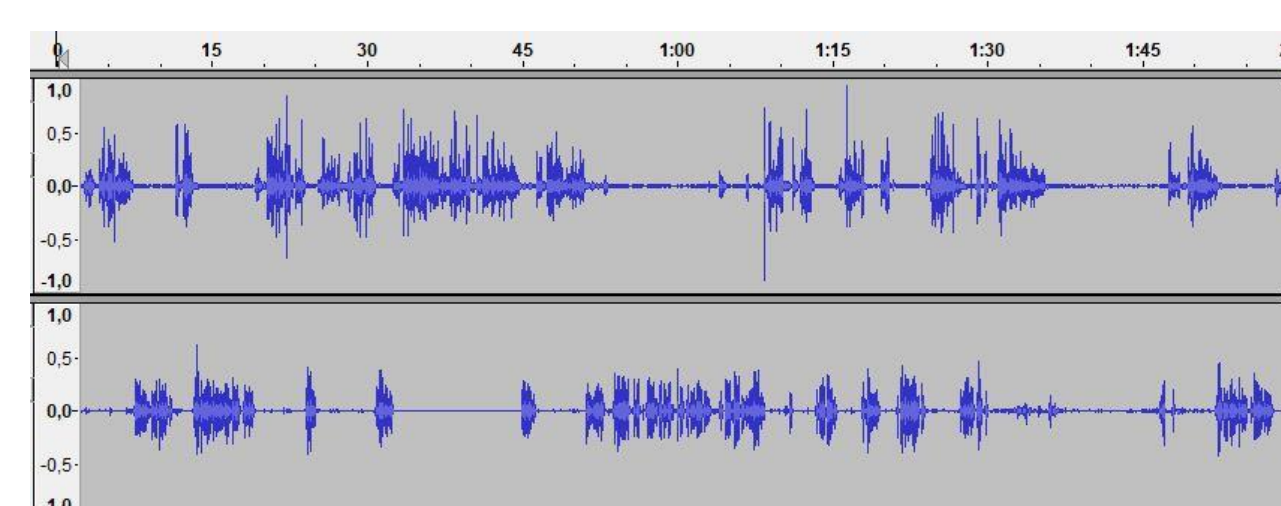
# Nautilus Speaker Characterization (NSC) Corpus

Laura Fernández Gallardo (laura.fernandezgallardo@tu-berlin.de)  
Benjamin Weiss (benjamin.weiss@tu-berlin.de)



Quality and Usability Lab, Technische Universität Berlin, Germany

<b>Purpose</b>	Speaker characterization (paralinguistics)
<b>Speakers</b>	300 speakers; 126 males, 174 females
<b>Age</b>	18 to 35 years old
<b>Language</b>	German (mother tongue)
<b>Speech type</b>	conversational; scripted, semi-spontaneous, and spontaneous
<b>Size</b>	50 GB, 155 hours of speech
<b>File format</b>	1-channel audio/wav, 48 kHz, 16 bit speakers: microphone AKG C 414B-XLS interlocutor: headset Sennheiser HMD 46



## Scripted dialogs

- Dialog 1: Health insurance
- Dialog 2: Mobile phone rate plan
- Dialog 3: Car rental inquiry
- Dialog 4: Real estate agency



## Semi-spontaneous dialogs

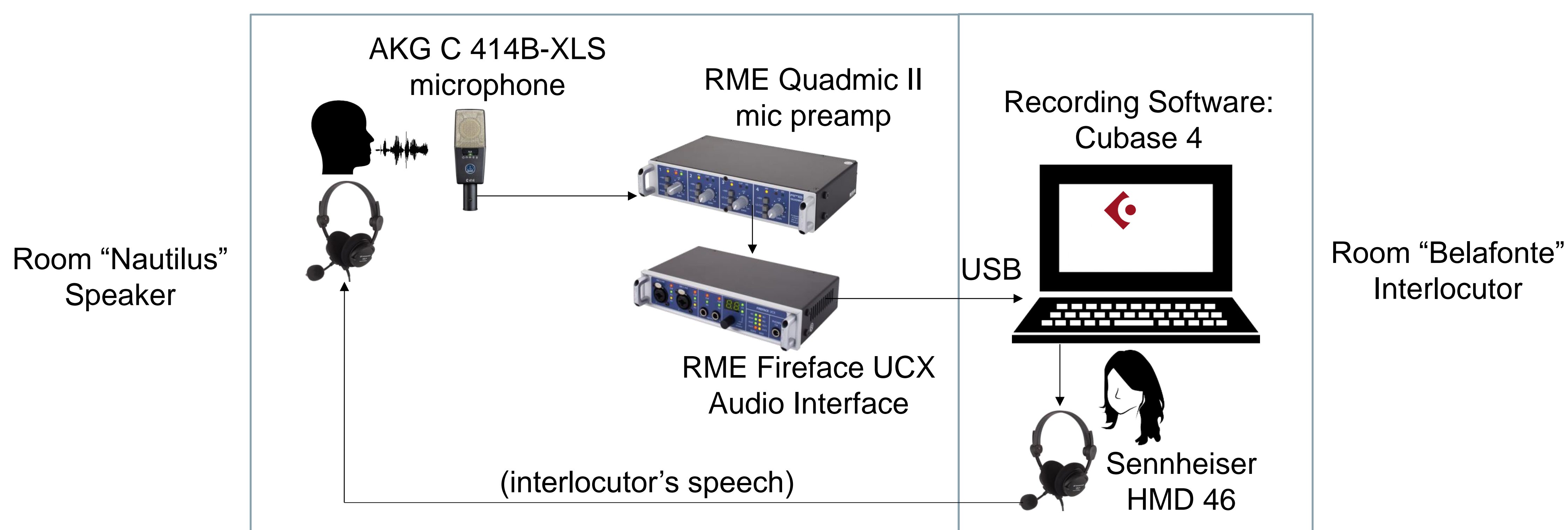
- Dialog 5: Car rental booking
- Dialog 6: Pizza order
- Dialog 7: Book from the library
- Dialog 8: Doctor's appointment

## Spontaneous speech

casual neutral, emotional, question and statement utterances

## Interactions

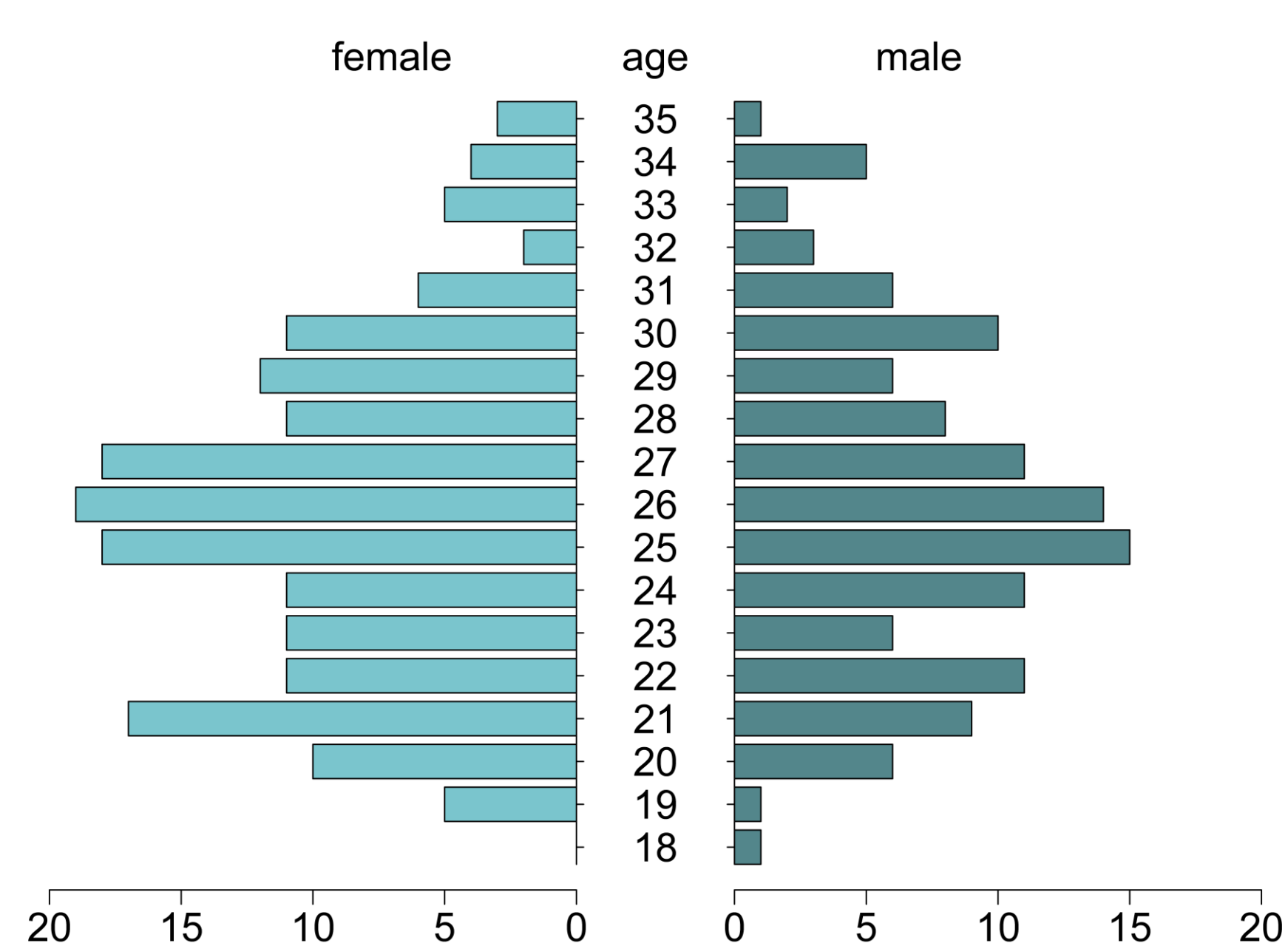
- interlocutor speech (wav)
- speaker speech (wav)
- tags+timestamps (csv)



## Self-assessed personality traits and vocal attractiveness

### Socio-demographic data

Age, gender, place of birth, chronological places of residence and duration of stay, place of birth of the mother and of the father,...



## Interpersonal speaker attributions (34 items)

- likable ----- non likable
- insecure ----- secure
- unattractive ----- attractive
- sympathetic ----- unsympathetic
- characterful ----- characterless
- competent ----- incompetent

(up to 34 items)

Reduced to 5 factors of speaker characteristics that can be subjectively assessed in listening-only and zero-acquaintance scenarios:  
**warmth, attractiveness, confidence, compliance, maturity**

### For 20 "extreme" speakers: naïve voice descriptions

- sonorous ----- flat
- low ----- high
- nasal ----- not nasal
- blunt ----- sharp
- even ----- uneven
- accented ----- without accent

(up to 34 items)

- Freely accessible, exclusive for **non-commercial scientific research**
- BAS CLARIN, ELRA, LDC repositories
- More information: <http://www.qu.tu-berlin.de/?id=nsc-corpus>

