

Nautilus Speaker Characterization (NSC) Corpus

Laura Fernández Gallardo (laura.fernandezgallardo@tu-berlin.de)

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

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



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) and speaker speech (wav) and tags+timestamps (csv) - over the recording session

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,...

Self-assessed personality traits and vocal attractiveness

Annotations: interpersonal speaker attributions

34-item semantic differential rating scale

sympathisch / unsympathisch (likable / non-likable)
gefühlskalt / herzlich (cold / hearty)
kompetent / inkompetent (competent / incompetent)
erwachsen / kindlich (adult / childish)
...

For 20 "extreme" speakers: naïve voice descriptions