

# Recording a High-Quality German Speech Database for the Study of Speaker Personality and Likability

*Laura Fernández Gallardo, Ph.D.*

*Quality and Usability Lab, Technische Universität Berlin, Germany*




1

# Outline

- Motivation for data recording
- Database material
- Description of the recording procedure
- Listening tests with this database
- Future directions
- Summary & sharing possibilities

# Outline

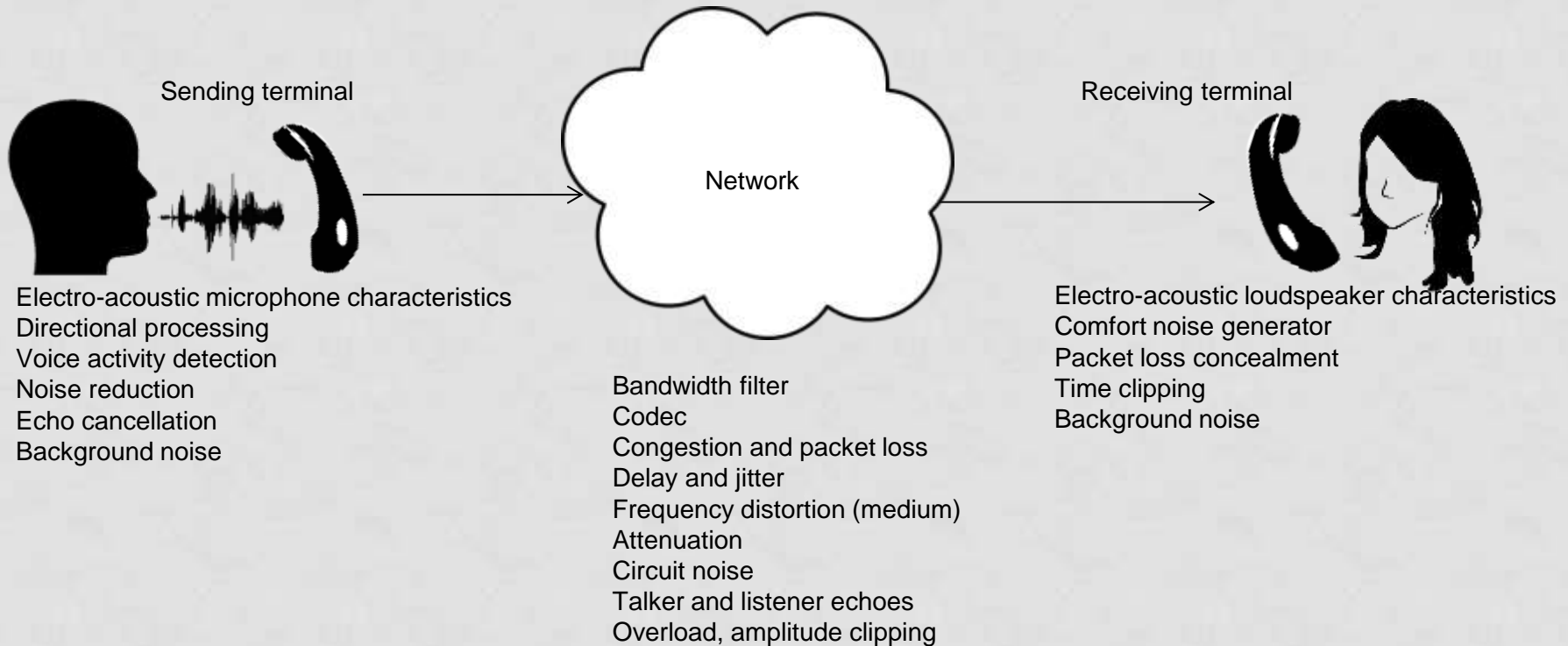
- Motivation for data recording 
- Database material
- Description of the recording procedure
- Listening tests with this database
- Future directions
- Summary & sharing possibilities

# Motivation

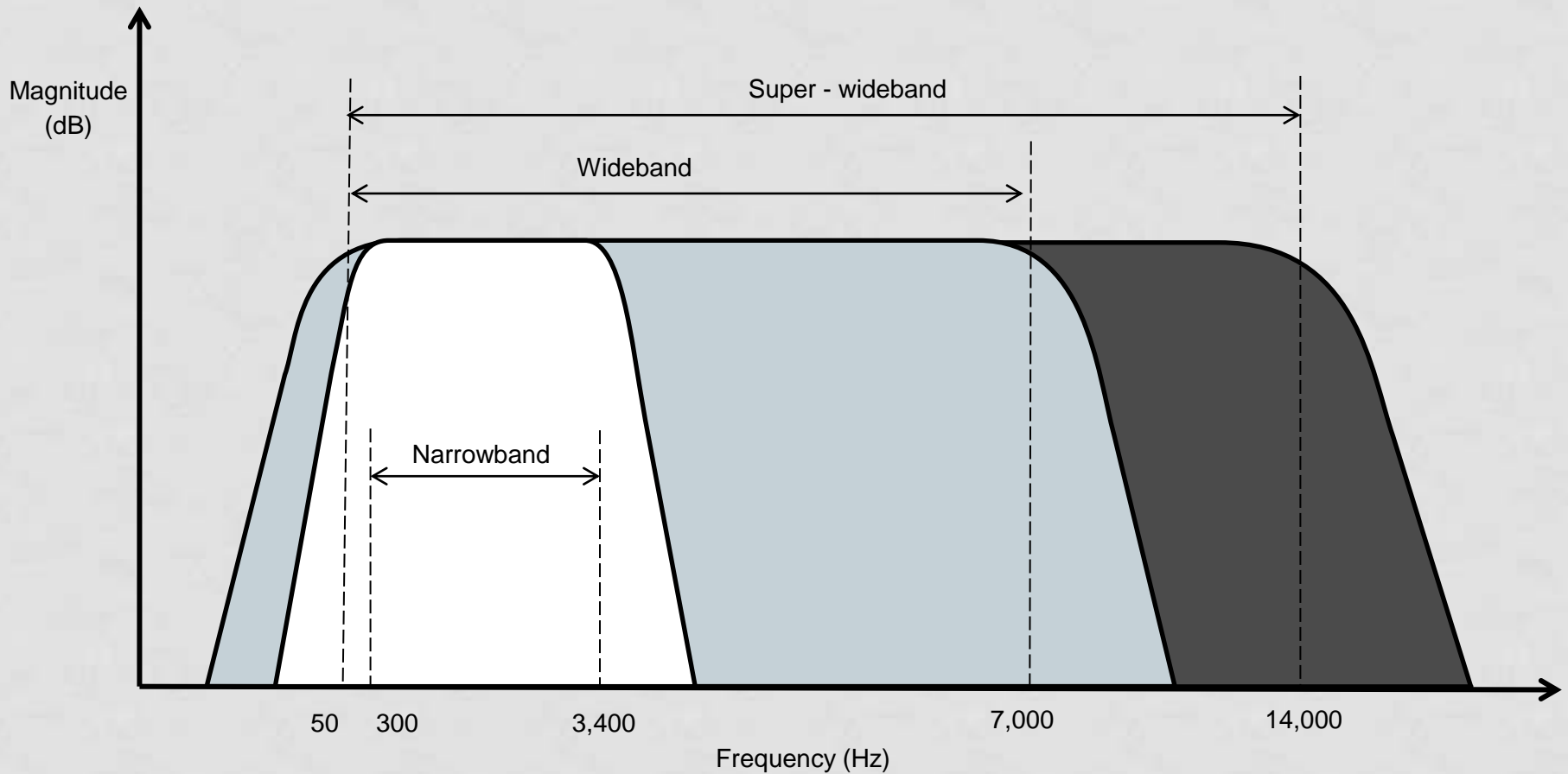
- Needed: speech database for speaker characterization
- Social characteristics:
  - Speaker personality
  - Speech likability
- Effects of transmission channels



# Motivation



# Motivation



# Motivation

The existing databases are not sufficient for the study of the influence of transmission channels on human and automatic speaker characterization

# Motivation


- Existing databases
  - Transmitted utterances or  $f_s < 16$  kHz (SPC, PersIA, EAR, aGender, SLD, VoiceClass, VeriDat)
  - Different microphones and recording environments (Voxforge)
  - Language is not German (SPC, PersIA, EAR, AMI, Maptraits)
  - Limited neutral speech or too little material (EmoDB, SmartKom, VAM, SEMAINE, OLLO)
  - Non-spontaneous speech (PhonDat)



# Motivation

- New high-quality database of German speech
  - 200 – 300 German speakers (gender-balanced)
  - scripted and spontaneous speech
  - clean conditions, 48 kHz sampling frequency
- Currently being recorded

# Outline

- Motivation for data recording
- Database material 
- Description of the recording procedure
- Listening tests with this database
- Future directions
- Summary & sharing possibilities

# Database Material

- 4 scripted and 4 spontaneous dialogs
  - Information requests over the telephone

Tabelle 1: *Recording material for our database.*

scripted speech	Dialog 1: Health insurance
	Dialog 2: Mobile phone rate plan
	Dialog 3: Car rental—inquiry
	Dialog 4: Real estate agency
spontaneous speech	Dialog 5: Car rental—booking
	Dialog 6: Pizza
	Dialog 7: Book from the library
	Dialog 8: Doctor's appointment

# Database Material

## scripted dialogs

### Krankenversicherung

A: Guten Tag, DKS Versicherungen. Mein Name ist Heinmüller, was kann ich für Sie tun?

**B: Guten Tag, mein Name ist Schmidt. Ich hab eine Frage bezüglich meiner Krankenversicherung.**

A: Geben Sie mir bitte Ihre Vertragsnummer durch.

**B: Meine Vertragsnummer ist die 4035.**

A: Einen Moment bitte. Alles klar, wie lautet Ihre Frage?

**B: Ich hab von meinem Hausarzt ein Rezept über 10 Physiotherapie-Anwendungen bekommen. Deshalb wollte ich mal fragen, ob Sie die Kosten in vollem Umfang übernehmen.**

A: Ich verstehe. Ich leite Sie eben zu Ihrem persönlichen Sachbearbeiter weiter.

**B: Alles klar, Dankeschön!**

## spontaneous dialogs

### Autovermietung

Ihr Name: Pfeifer



1 Kleintransporter für einen Umzug



für ein Wochenende (Samstag und Sonntag),  
mindestens 1,80 m Ladehöhe,  
möglichst günstig



Modell : \_\_\_\_\_

Ladrtiefe : \_\_\_\_\_

Preis : \_\_\_\_\_ €/Tag



Reservierung : Eurocard  
Kreditkartennummer: 8823 7453 6592 0023  
gültig bis 12/19  
Führerschein Klasse 3



Ab wieviel Uhr kann ich den Wagen morgens abholen?

# Database Material

- Example Dialog 1 (scripted speech)



- Example Dialog 5 (spontaneous speech)




- Dialogs held with recording assistant
  - female, 26 years old, native German speaker with no accent

# Database Material

- Metadata:
  - Age
  - Gender
  - Place of birth
  - chronological places of residence and duration of stay
  - place of birth of the mother and of the father
  - highest education level
  - educational background
  - main occupation
  - past occupation(s) if any
  - years of work experience
  - self-assessed personality traits (by filling in BFI questionnaires)

# Database Material

- Speakers sign a consent form:
  - All data are pseudo-anonymized
  - The database may be publicly released under license agreement for exclusive use in scientific studies
  - The participation is voluntary and it is possible to withdraw from the study

**Technische Universität Berlin** 

Einwilligung des Teilnehmers

Ich erkläre, dass

- ich die Informationen zur Studie „Sprachaufnahmen zur Unterstützung sprachwissenschaftlicher Forschung“ gelesen habe und bereit bin, an der Studie teilzunehmen.
- meine Fragen ausreichend und verständlich beantwortet wurden.
- ich darüber informiert wurde, dass von mir heute folgende Daten erhoben werden: Audioaufnahmen und soziodemographische Daten.
- ich darüber informiert wurde, dass sämtliche Daten gespeichert und pseudonymisiert für wissenschaftliche Auswertungen genutzt werden.
- ich mir bewusst bin, dass die Audioaufnahmen und die Metadaten Teil einer öffentlichen Datenbank werden, die auch anderen Wissenschaftlern zu Forschungszwecken zur Verfügung gestellt werden soll. Alle Daten, die von mir für diese Studie erfasst werden, werden vertraulich behandelt.
- ich mir bewusst bin, dass die Teilnahme freiwillig ist und ich sie jederzeit ohne Angaben von Gründen widerrufen kann, ohne dass mir daraus irgendein Nachteil entsteht.
- ich eine Kopie der Informationen zur Studie und dieser Einwilligung erhalten habe.

Zudem ist mir bekannt, dass alles, was ich während der heutigen Sitzung gesehen und gehört habe, vertraulich ist und erkläre, dass ich diese Informationen nicht weitergeben werde.


Name: \_\_\_\_\_

Datum: \_\_\_\_\_

Unterschrift: \_\_\_\_\_

Versuchsleiterin:  
Laura Fernández Gallardo, Ph.D.  
Quality and Usability Lab, Telekom Innovation Laboratories, Technische Universität Berlin  
Ernst-Reuter-Platz 7, D-10587 Berlin, Germany  
+49 175 2629110  
Laura.Fernandez-Gallardo@telekom.de

# Outline

- Motivation for data recording
- Database material
- Description of the recording procedure 
- Listening tests with this database
- Future directions
- Summary & sharing possibilities



# Recording Procedure

- Speakers' recruitment
  - “Probandenportal” of the Technische Universität Berlin
  - mailing lists
  - announcements
- Requirements:
  - Native German speaker without dialect or accent
  - 18 – 35 years old
  - No hearing or speech disorders
  - No sickness or cold at the time of the recording
- Speakers were paid for their participation

# Recording Procedure

- Three different microphones



Stand-up microphone:  
AKG C 414B-XLS  
(frequency range 20–20,000 Hz)



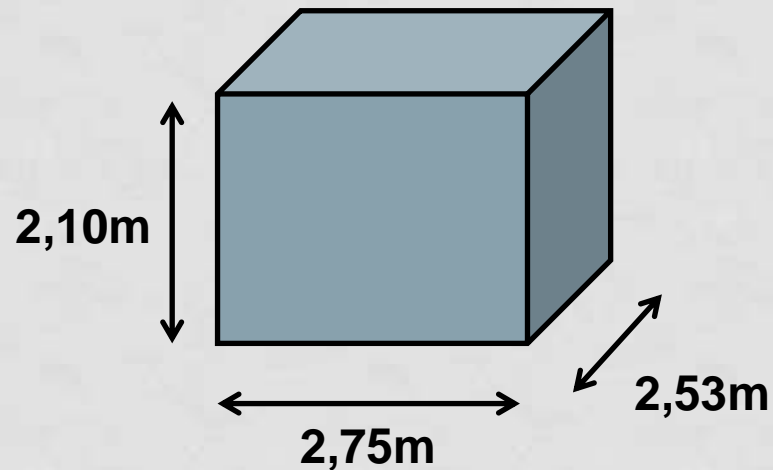
Table microphone:  
Audio-Technica U851R  
(frequency range 30–20,000 Hz)



Headset microphone:  
Sennheiser HMD 46  
(frequency range 100–12,000 Hz)

# Recording Procedure

- Room “Nautilus” (acoustically-isolated)



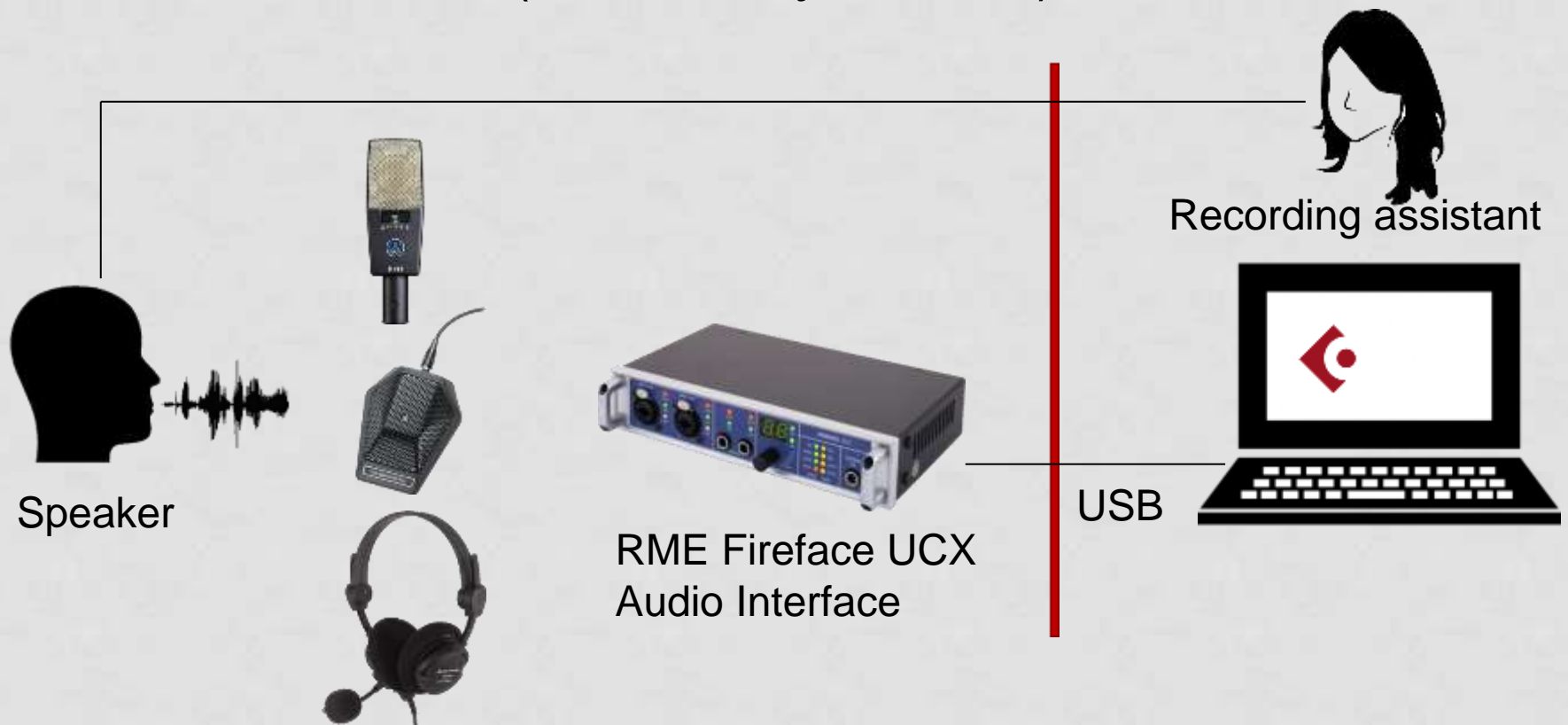
RT60 = 0.08 s at 2 kHz



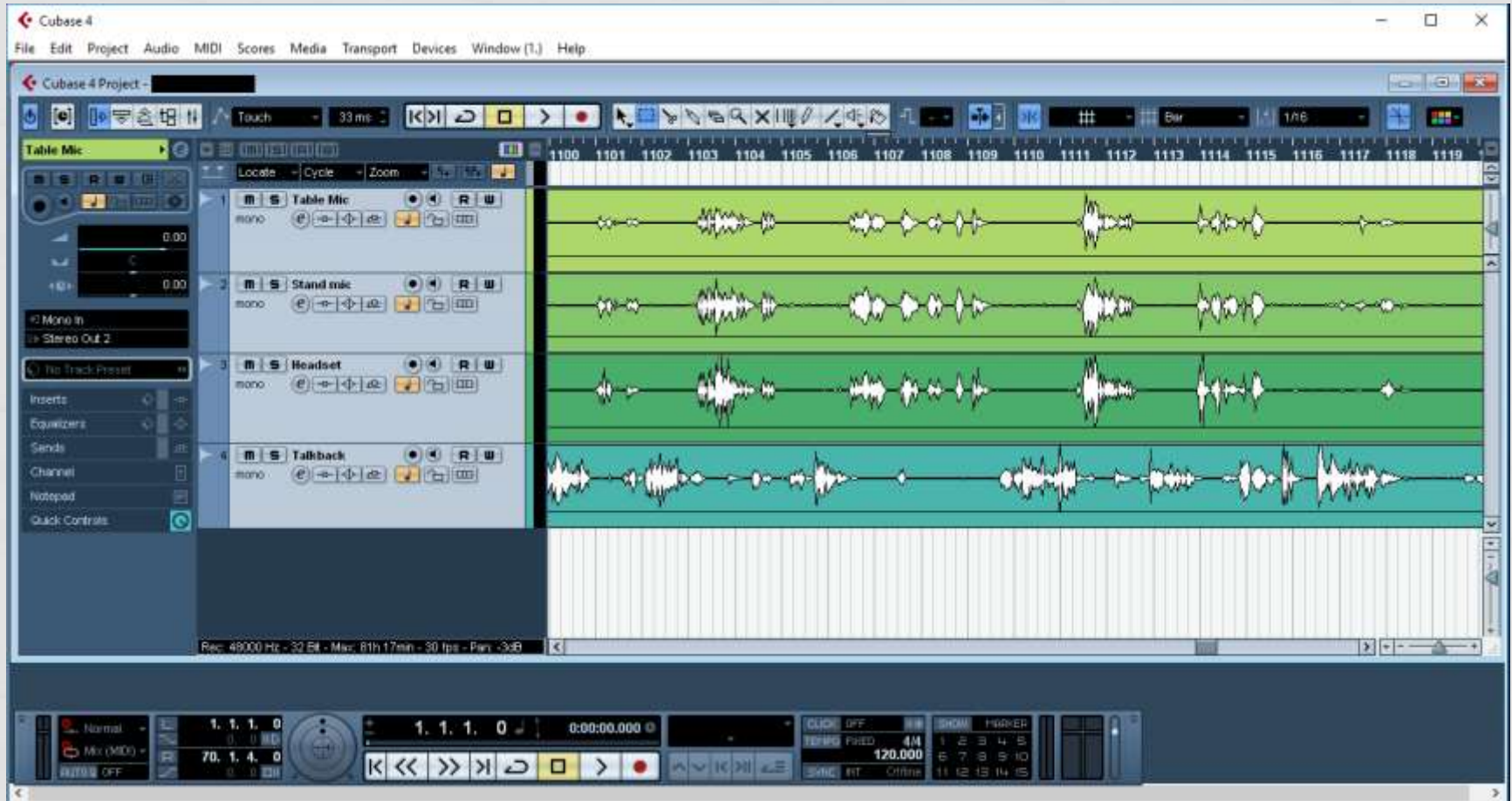
(In the position of a speaker is Laura Fernández Gallardo).

# Recording Procedure


- Room “Nautilus” (acoustically-isolated)



# Recording Procedure



# Outline

- Motivation for data recording
- Database material
- Description of the recording procedure
- Listening tests with this database 
- Future directions
- Summary & sharing possibilities

# Conducted Listening Tests

- [15] L. Fernández Gallardo, L. and Weiss, B., "Speech Likability and Personality-based Social Relations: A Round-Robin Analysis over Communication Channels," in Interspeech, 2016.
- Round-Robin setup with 30 participants
- Personality and likability rated from speech
- Speech transmitted through narrowband and wideband
- The Social Relations Model to analyze mutual perceptions




# Conducted Listening Tests

- [17] Fernández Gallardo, L., "A Paired-Comparison Listening Test for Collecting Voice Likability Scores," in Informationstechnische Gesellschaft im VDE (ITG) Conference on Speech Communication, 2016.
- Paired-comparison vs. continuous scale to rate likability






# Outline

- Motivation for data recording
- Database material
- Description of the recording procedure
- Listening tests with this database
- Future directions 
- Summary & sharing possibilities

# Future Directions

- September 2016: ~170 speakers recorded
  - ~ 59% females
  - 26,2 years old on average (range: 19 - 35)
- Goal: 300 speakers
- Data will be labelled in terms of likability and personality
- The effects of communication channels will be studied

# Outline

- Motivation for data recording
- Database material
- Description of the recording procedure
- Listening tests with this database
- Future directions
- Summary & sharing possibilities 

# Summary



~300 speakers

fs = 48 kHz

Tabelle 1: *Recording material for our database.*

scripted speech	Dialog 1: Health insurance
	Dialog 2: Mobile phone rate plan
	Dialog 3: Car rental—inquiry
	Dialog 4: Real estate agency
spontaneous speech	Dialog 5: Car rental—booking
	Dialog 6: Pizza
	Dialog 7: Book from the library
	Dialog 8: Doctor's appointment

- + personality ratings
- + likability ratings
- + speakers' socio-demographic data

# Summary

- We are very open to various forms of collaboration
- Our data will be freely released to the scientific community under a license agreement
- Also individual judges' ratings, speech features or other analyses can be shared with interested researchers
  
- Author:
  - Laura Fernández Gallardo, Ph.D. (laura.fernandezgallardo@tu-berlin.de)
- Cooperating with this study:
  - Dr. Benjamin Weiss (benjamin.weiss@tu-berlin.de)
  - Prof. Dr.-Ing. Sebastian Möller (sebastian.moeller@tu-berlin.de)
  
- Supported by DFG, Grant FE 1603/1-1 to Laura Fernández Gallardo

# Thank you for your attention!

*Laura Fernández Gallardo, Ph.D.  
laura.fernandezgallardo@tu-berlin.de*



# Questions?



*Laura Fernández Gallardo, Ph.D.  
laura.fernandezgallardo@tu-berlin.de*