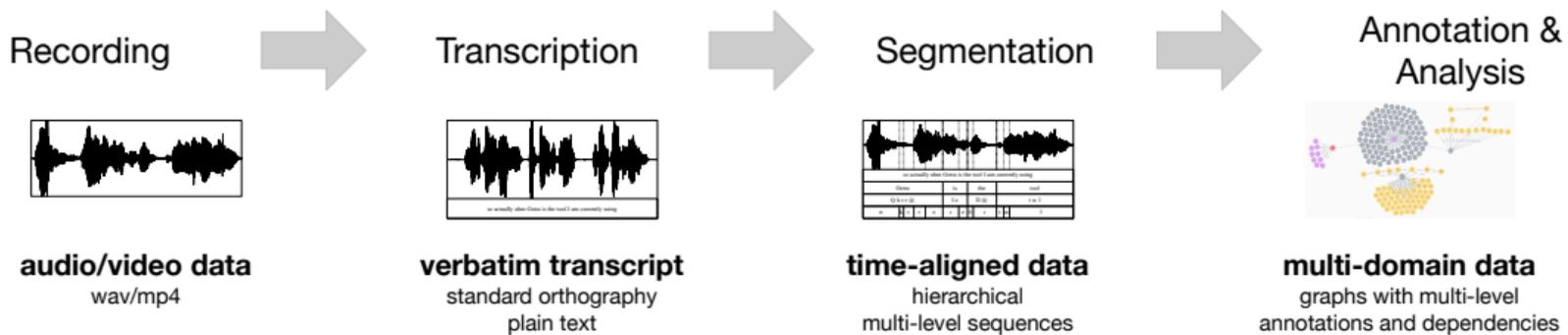


Perfect Speech Recordings: Why and How?

Christoph Draxler
Bavarian Archive for Speech Signals BAS
draxler@phonetik.uni-muenchen.de

Thanks to YYY for translating this presentation

Scientific Speech Processing



Processing Flow

In Spoken Language Processing, output of one task is input to the next – and recordings are the decisive first step!

- ▶ ALL other steps depend on the recordings
- ▶ recording errors CANNOT be repaired later

Recordings require technical skills

- ▶ can be learned
- ▶ experience and routine are important

Recordings are a *social* activity

Signal quality

Optimal signal quality is important!

- ▶ sample rate \geq 32 kHz, 16 bit linear, WAV-format
- ▶ microphone close to the speaker
 - ▶ lapel (or Lavalier) microphone
 - ▶ close-talk microphone
- ▶ be aware of environment noise

Microphone placement & turn-taking

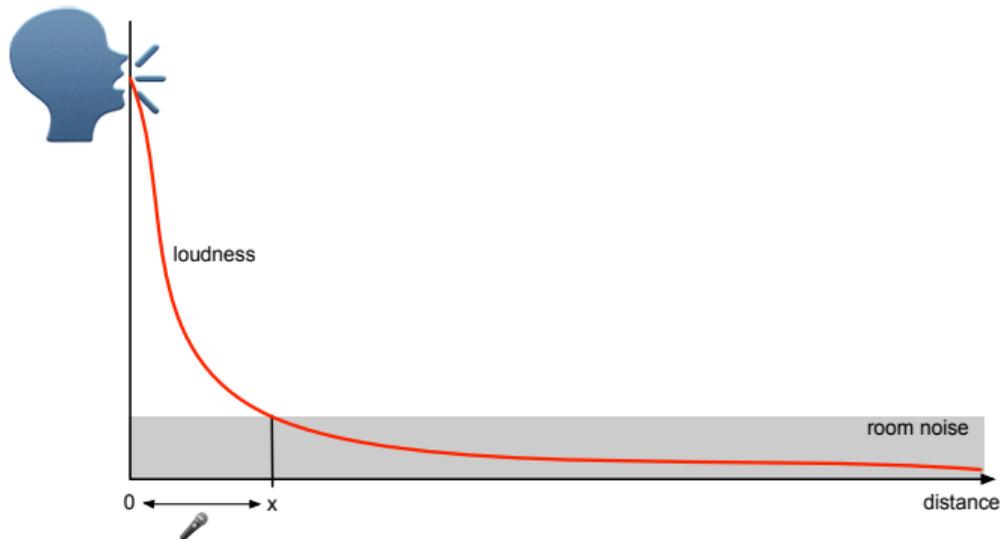
1. one microphone per speaker, close to the mouth, e. g.



2. as few interruptions or overlap as possible

#1 gives you a single speaker in high audio quality in separate channels in your audio file, #2 reduces transcription complexity

Microphone Distance (in rooms)



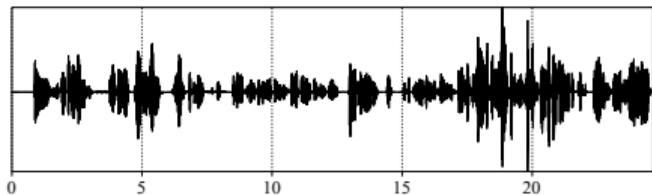
red loudness of speech – rapid decrease

gray constant level of noise and reverberation in a room

x distance from mouth where loudness of speech and room noise are equal, typically 30-40 cm

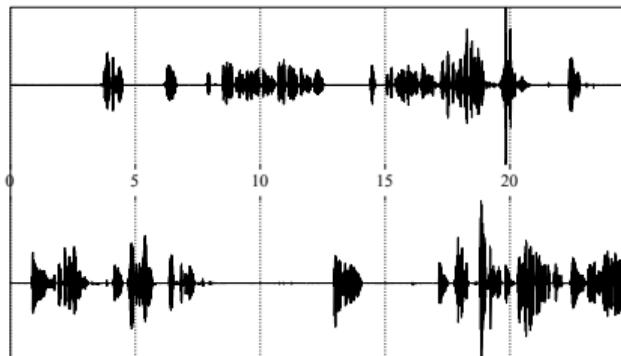
The microphone must be at most 40 cm from the speaker's mouth. Close-talk or lapel microphones are the best choice!

Spot the difference!



Two speakers, one microphone, mono

- ▶ no separation of speakers
- ▶ difficult to transcribe



Two speakers, two close-talk
microphones, stereo

- ▶ very good separation of speakers
- ▶ easy to transcribe
- ▶ perfect synchronisation

Take-home Messages

- ▶ Test signal quality before recording session
- ▶ Microphone close to the speaker!
- ▶ One microphone per speaker

A good microphone will not make good recordings in a bad recording situation.
It will make better recordings in a good recording situation.

Useful hints

- ▶ Ask the speaker to say time, date, and location
- ▶ Record his or her agreement on being recorded
- ▶ Document the recording situation, e. g. photo or sketch
- ▶ Have all device manuals available, e. g. as PDFs on your computer or tablet

Recommended Equipment in the Field I



Saramonic Blink 500/900

- ▶ 2 lavalier microphones
- ▶ 2 transmitters with Bluetooth
- ▶ 1 receiver connected to phone, tablet or computer via USB-C

Rechargeable batteries allow > 6 h recording time.

However, **no stereo** recordings are made with Blink 500!

Recommended Equipment in the Field II



Rode Wireless Go II

- ▶ 2 transmitters with Bluetooth
- ▶ 1 receiver connected to phone, tablet or computer via USB-C

Rechargeable batteries allow > 7h recording time.

Finally, stereo recordings are possible with the Rode Wireless Go II!

Scripted Recordings in the Field



USB headset Sennheiser SC60

images/speechrecorder_128

SpeechRecorder is a software for scripted recordings. Every item is saved into a separate file. This reduces the amount of signal editing dramatically.

<http://www.speechrecorder.org/>

Distributed Recordings via the Internet

With WikiSpeech, you can do recordings even in places where you can't go

- ▶ register via mail to wikispeech@phonetik.uni-muenchen.de
- ▶ invite speakers by sending them a link
- ▶ speakers may use (almost) any device: phone, tablet, laptop, desktop...
- ▶ monitor incoming speech and the project progress without delay

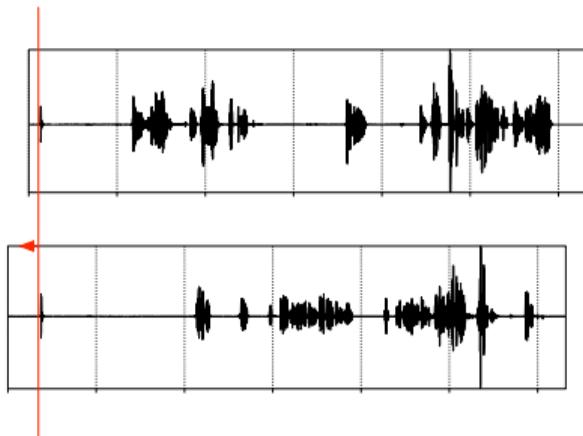
Free of charge for academic partners!

Easy, Cheap and Flexible!

1. every participant uses his or her own smartphone with a cable-bound headset, one participant acts as the master
2. wearing the headset, open the smartphone's built-in recording app, e.g. *Voice Memos* 📱 on iOS or *Voice Recorder* 🎙️ on Android, and start recording. Put the phone in your pocket or on the table in front of you
3. the recording master makes a loud and sharp noise, e. g. by hitting two pebbles together or by using a clapperboard – later, this noise will help synchronise the recordings
4. then, perform your recording session or interview
5. when done, stop the recording on every device and name the audio file with the corresponding speaker's name and date, e.g. Christoph_2022-03-25
6. transfer all recording files to the master device or a shared network folder for synchronisation and transcription

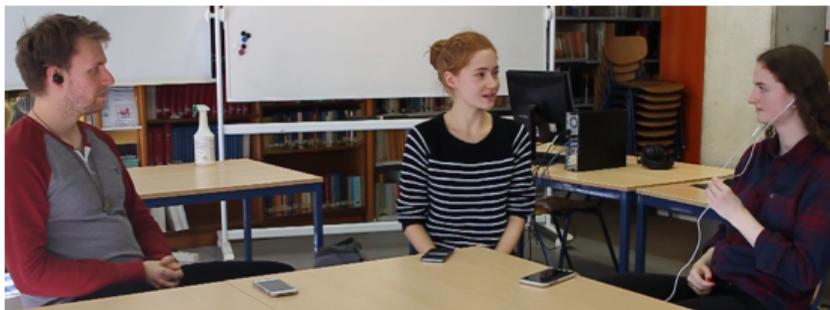
Synchronise audio files

The sharp noise made at the beginning of the recording session serves to synchronise the recordings.



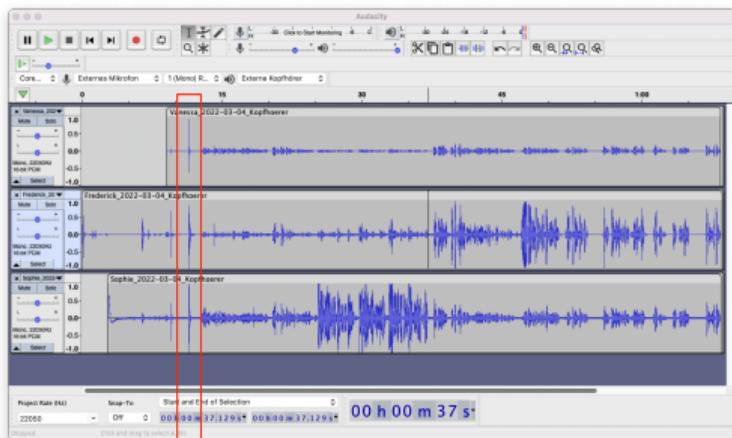
Using e. g. Audacity () , align the two files visually and audively until the oscillograms of the noise overlap perfectly and you do not hear an echo of the noise anymore.

Real-world Example: A Discussion in our Library



It's not perfect, but good!

- ▶ top track was recorded with the built-in microphone, so all speakers are almost equally loud.
- ▶ middle track is the left speaker, the bottom track is the right speaker. Both used a headset: here, the separation of speakers is much clearer.
- ▶ with the smartphone in their pockets, speakers could even move around!



align the
synchronisation
marks

Thanks to Vanessa, Frederick and Sophie!