

Evaluating Collaborative Editing of

AI-Generated Live Subtitles

by Non-Professionals

in German University Lectures

# Einleitung

## Ausgangslage: Untertitel für Vorlesungen

- + Gehörlose und Schwerhörige, aber auch andere Studierende profitieren.
- Professionelle Untertitler sind teuer und nicht immer verfügbar.

## Herangehensweise: Einsatz von Spracherkennung

- + Automatisiert, kostengünstig und immer verfügbar.
- Unzureichende Genauigkeit (<98%) macht manuelle Korrektur erforderlich.

**Tool** für manuelle, kollaborative Korrekturen durch Studierende

- ▷ Nutzertest als Grundlage für die Entwicklung

# Methodik

## Forschungsfrage:

Können **Studierende** im Kontext einer **Vorlesung**  
**KI-generierte Untertitel kollaborativ** korrigieren?

## Nutzertest (“Wizard of Oz”):

- Studierende: Studierende und Non-Professionals
- Vorlesung: Zoom Video Konferenz
- KI: Microsoft Azure
- Kollaboration: Google Sheets

# Methodik: Setup

The image shows a Zoom meeting window on the left and a Google Sheet on the right. The Zoom window displays a CAST | Universal Design for Learning (UDL) grid with the following categories and options:

- Engagement:** Provide multiple means of engagement, Provide multiple means of representation, Provide multiple means of action & expression.
- Accommodations:** Provide options for sustaining interest, Provide options for perception, Provide options for physical action.
- Support:** Provide options for sustaining effort & persistence, Provide options for language & symbols, Provide options for expression & communication.
- Interventions:** Provide options for self-regulation, Provide options for comprehension, Provide options for executive functions.

The Google Sheet contains the following text:

Herzlich willkommen zum forza karrierefalle halt die Halle nutzt,  
wenn wir es schaffen, dass die Maßnahme der Barrierefreiheit allen Studierenden  
zugutekommen nicht nur denen, sei mal den relativ kleinen Teil der Studierenden,  
die jetzt irgendwo registriert sind doch offensichtlich eine Behinderung haben,  
wenn es das gelingt, dass er sozusagen die gesamte Studierendenschaft im Glück haben,  
dann wird es auch leichter werden, diese Maßnahme durchzusetzen und die Lehrenden  
und auch die des managemente Hochschulen dafür zu gewinnen und das ist die Frage,  
wie können wir das machen?

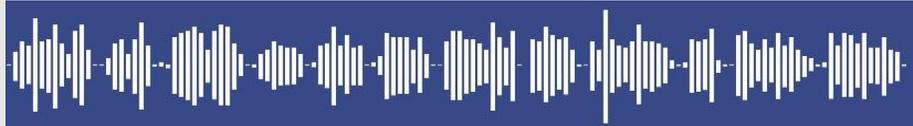
# Methodik: Bedingungen

Die Untertitel wurden entweder

- **simultan** oder
- realistisch **zeitverzögert**

eingefügt.

Audio:



Bed. 1:

Hallo, wir sind das SHUFFLE-Team

Bed. 2:

Hallo, wir sind das SHUFFLE-Team

# Ergebnisse

Yay, statistics!

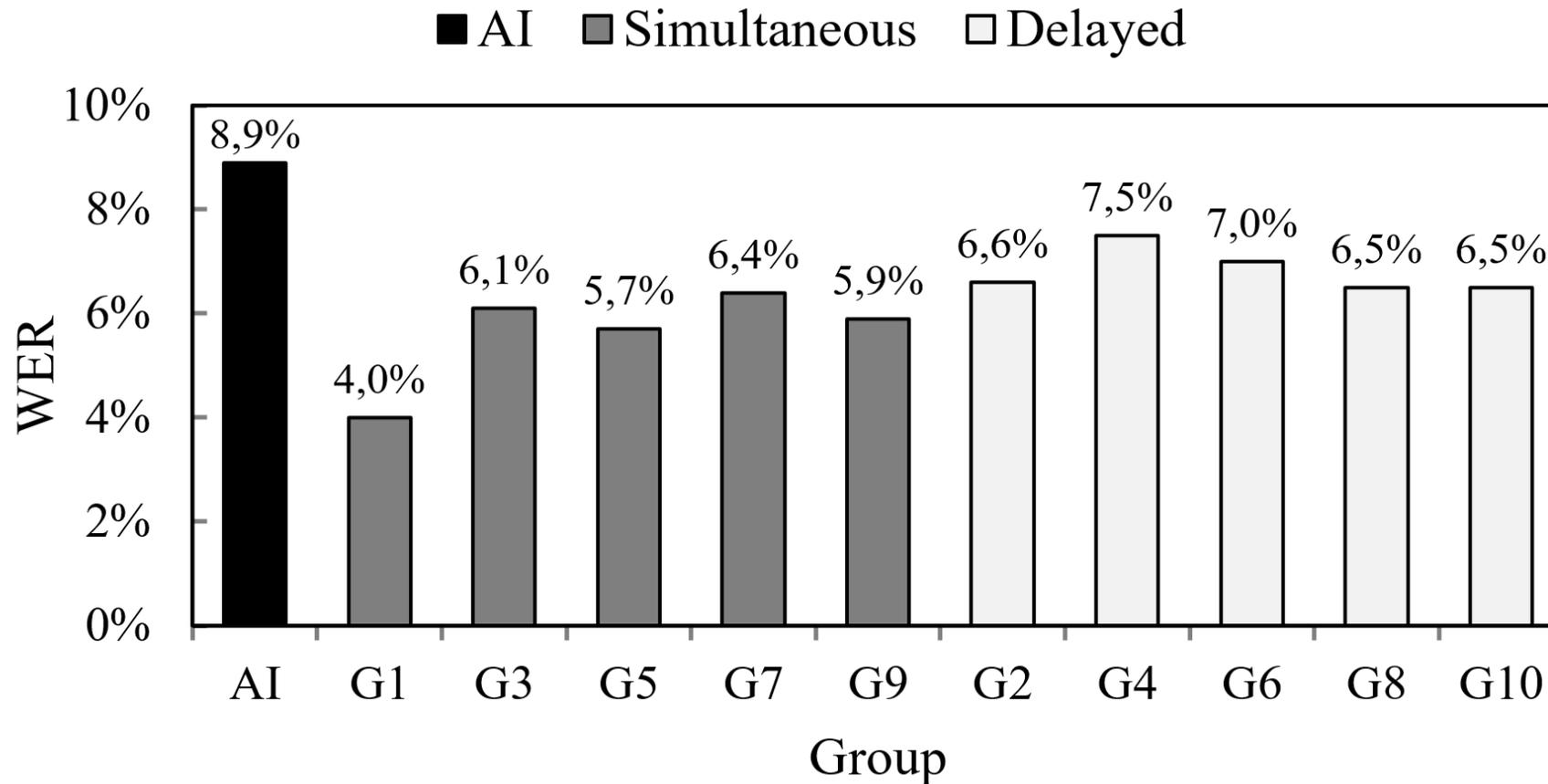
# 1. Forschungsfrage

Hat der Zeitpunkt des Einfügens der Untertitel  
Einfluss auf die Genauigkeit der manuellen Korrekturen?

## Word Error Rate

$$WER = \frac{\text{Ersetzungen} + \text{Löschungen} + \text{Einfügungen}}{\text{Anzahl der Wörter}_{\text{Referenz}}} \times 100$$

die manuellen Korrekturen positiv beeinflusst.



## 2. Forschungsfrage

Beeinflusst der Zeitpunkt des Einfügens der Untertitel die wahrgenommene kognitive Anstrengung der Teilnehmer?

**RATING SHEET**

MENTAL DEMAND

PHYSICAL DEMAND

TEMPORAL DEMAND

PERFORMANCE

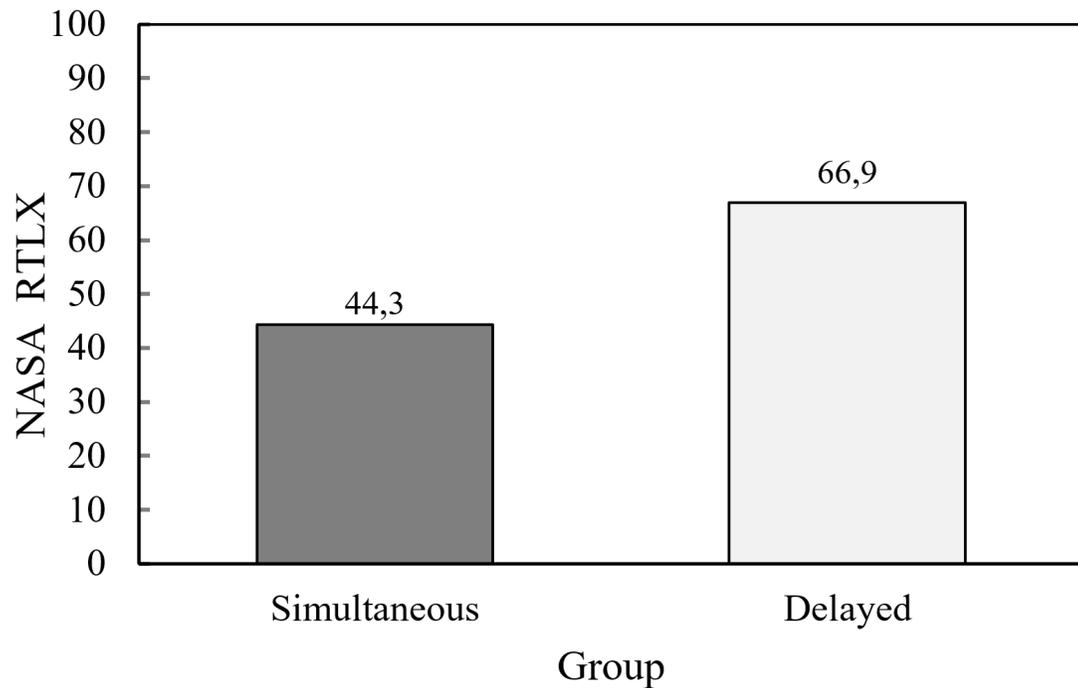
EFFORT

FRUSTRATION

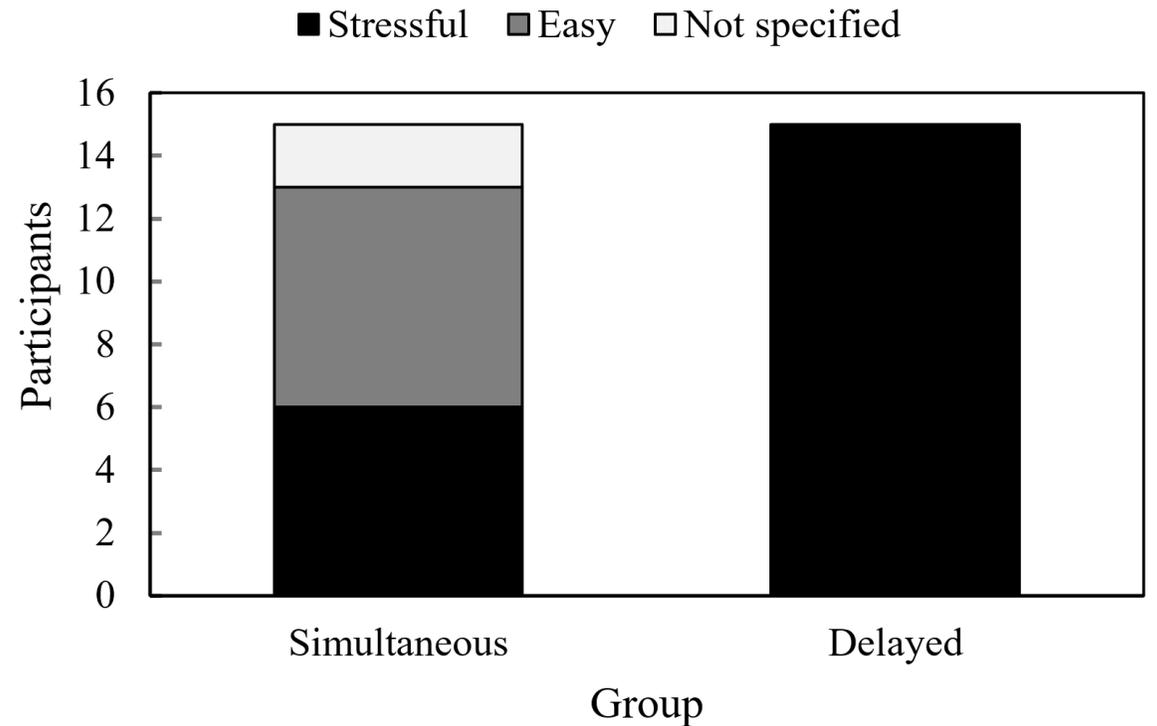
**NASA Raw Task Load  
Index (RTLX)**

wahrgenommene kognitive Belastung der Teilnehmer negativ.

### RTLX



### Stress indications



# 3. Forschungsfrage

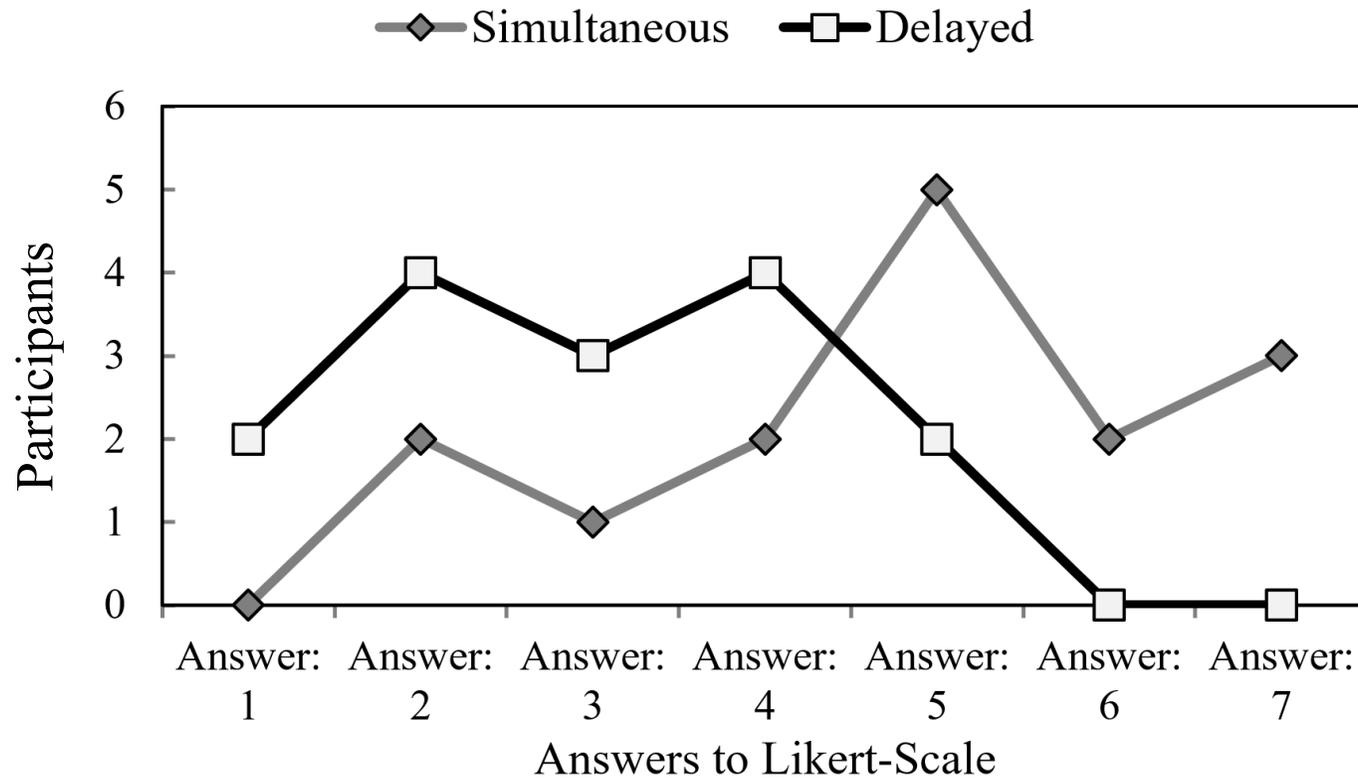
## Questionnaire

\*Were you able to follow the content of the video during the test?

Choose one of the following answers

- 1 - I could not follow the content at all
- 2
- 3
- 4
- 5
- 6
- 7 - I could follow the content fully

positiven Einfluss darauf, dem Inhalt folgen zu können.



# Zusammenfassung

Die Word Error Rate **verbessert** sich durch manuelles Editieren

Zeitversetzte Verbesserung erzeugt **mehr Stress** und ist **weniger effektiv**

# Einschränkungen / Offene Fragen

- Word Error Rate
  - Welche Genauigkeit ist für das **Textverständnis** wirklich notwendig?
  - Ist die WER für **deutsche Sprache** geeignet?
  - Was zeigen **andere Berechnungsmethoden**?
- User interface
  - Wie kann ein spezialisierter Editor **kollaboratives Editieren unterstützen**?
  - Was sind notwendige Zeitverzögerungen für **simultanes Editieren**?  
Ist das noch LIVE?

# Was kommt jetzt?

Entwicklung einer **Open-Source Untertitel-Software** "MELVIN"

-> **Weitere Studien** anhand des Editors

Wir sind auch als

**Tech-Demo** in Lecco dabei.

The screenshot shows a video player interface. At the top, there is a title bar with a logo 'M', the text 'Dangerous superficial knowledge', and a menu icon. Below the title bar is a waveform visualization of the audio. The main content area is divided into two parts. On the left, there is a presentation slide titled 'ENTWICKLUNG' with a sub-heading '2.'. The slide contains three colored boxes: a green box for 'Anstrengung und Ausdauer', a purple box for 'Sprache und Symbole', and a blue box for 'Ausdruck & Kommunikation'. Below these boxes, there is a question: 'Dann kommt die Ebene der Entwicklung: Wie bleiben Lernende im Lernprozess? Wie'. On the right, there is a transcript of the video content, with each line of text followed by a timestamp. The transcript starts with 'Herzlich willkommen zum Vortrag "Barrierefreiheit, die allen nützt - Innovative Maßnahmen in der Hochschullehre im Sinne von Universal Design for Learning".' and continues with 'Ich bin Gottfried Zimmermann, Professor an der Hochschule der Medien in Stuttgart und bin der fachliche Leiter des Kompetenzzentrums für digitale Barrierefreiheit und auch Leiter des Forschungsprojekts SHUFFLE.' and 'In der Hochschullehre haben wir alle zumindest seit Corona irgendwie Erfahrungen mit digital'. The video player interface also includes a 'LIVE' indicator, playback controls (play, pause, stop, search), and a progress bar showing '04:30 / 06:48'.