



The 3rd International Conference on Artificial Intelligence in Education Technology (AIET 2022)

**TIM SCHLIPPE, KATRIN FRITSCHÉ, YING SUN, MATTHIAS WÖLFEL**

# **AI-BASED VISUALIZATION OF VOICE CHARACTERISTICS IN LECTURE VIDEOS' CAPTIONS**

Wuhan, China  
July 3, 2022

# AGENDA

---

**Introduction**

---

**Related Work**

---

**AI-based Visualization of Voice Characteristics in Captions**

---

**User Study**

---

**Conclusion and Future Work**

---

**1**

**2**

**3**

**4**

**5**

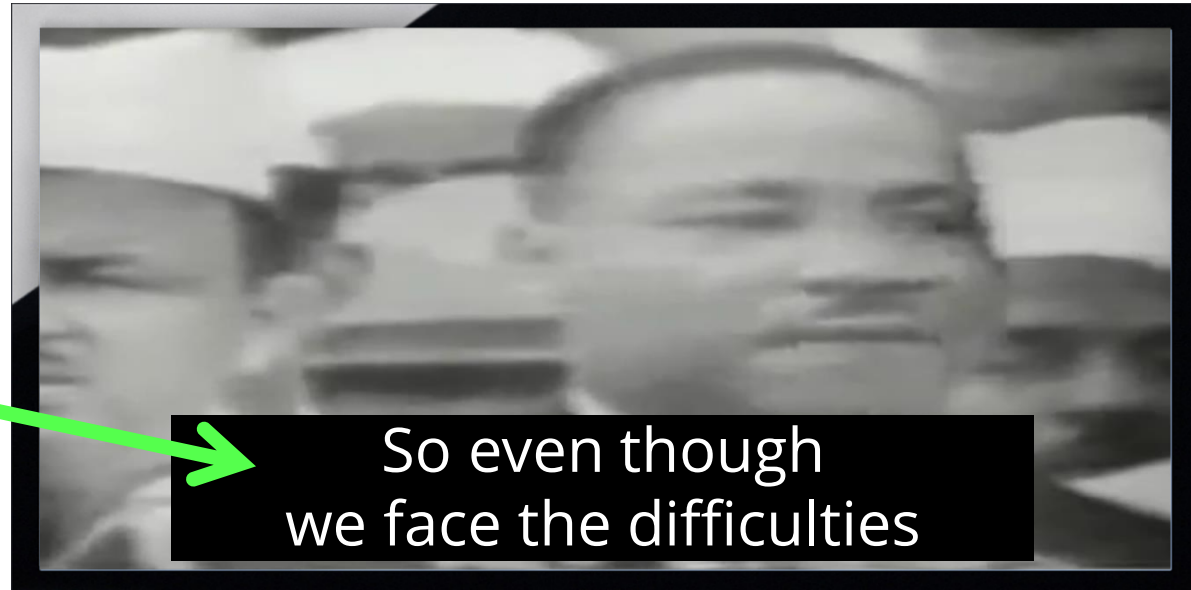
# 1

## INTRODUCTION

# CAPTIONING IN LECTURE VIDEOS

## 100+ STUDIES:

CAPTIONING a video improves  
+ COMPREHENSION of  
+ ATTENTION to  
+ and MEMORY for  
the video



# CAPTIONING IN LECTURE VIDEOS

## 100+ STUDIES:

**CAPTIONING** a video improves  
**+ COMPREHENSION** of  
**+ ATTENTION** to  
**+ and MEMORY** for  
the video

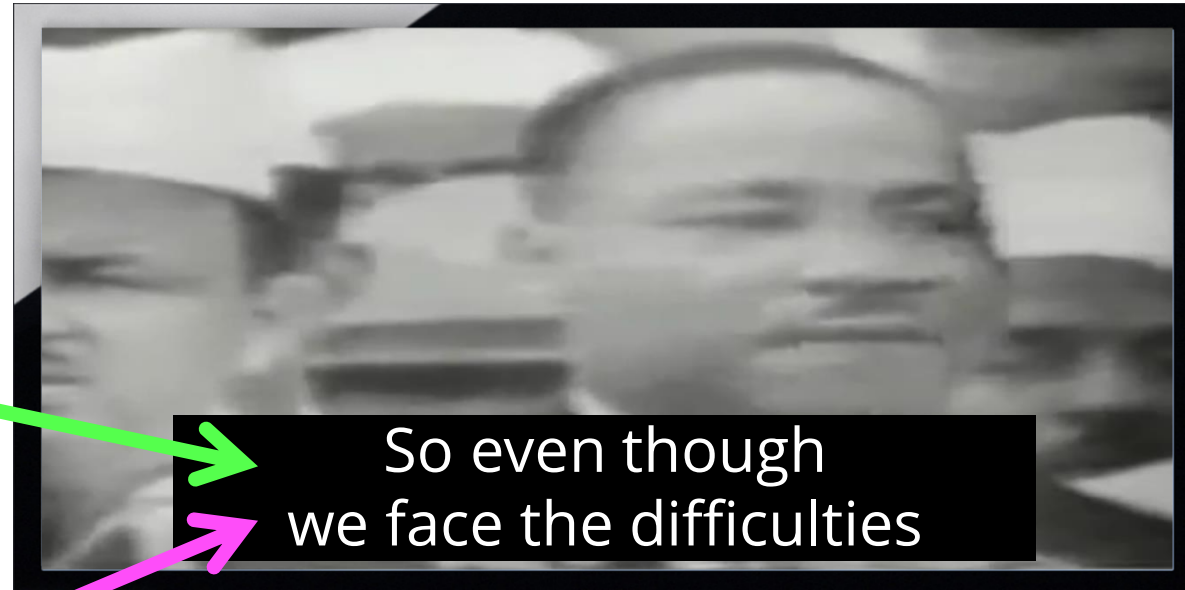
**Contribution to making messages clear  
in human communication:**

**WORDS**

**7%**

**TONE  
INTONATION  
VERBAL PACE**

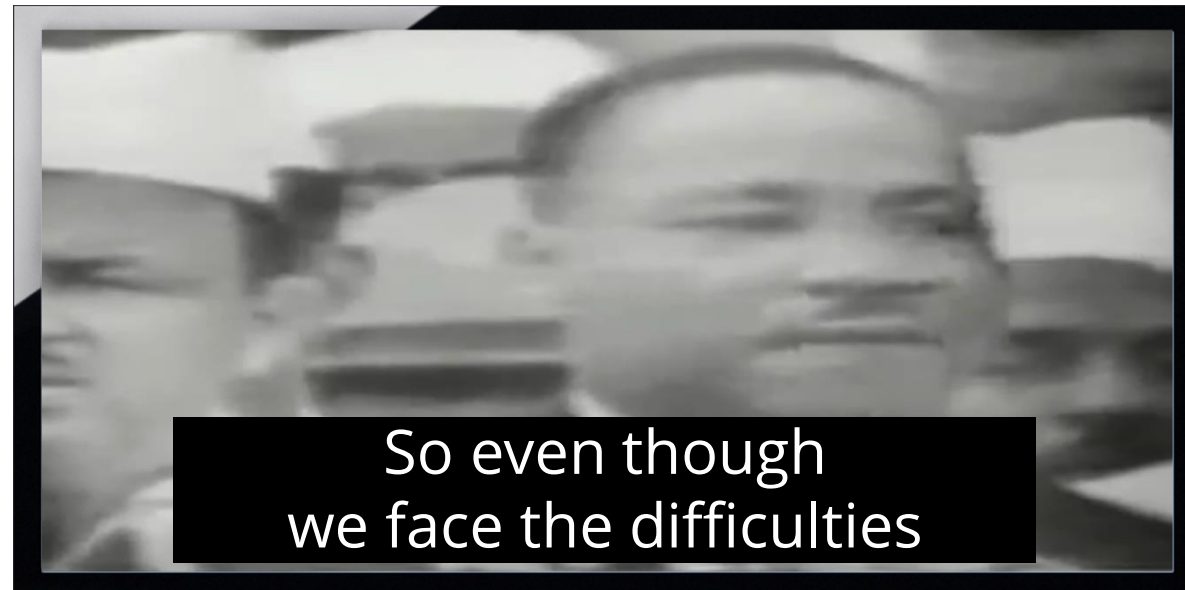
**38%**



# NO PROGRESS IN CAPTIONS & SUBTITLES



**for decades**

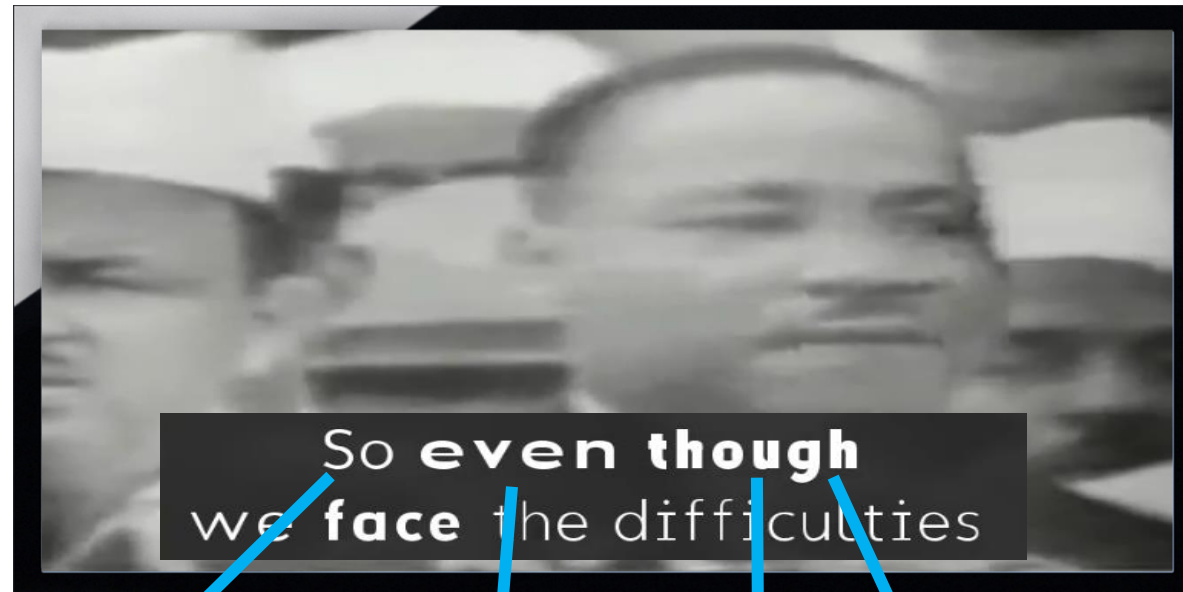


**limited**  
**boring**

No information  
from the voice

No emotions  
from the voice

unique & innovative



quieter slower faster **louder**

**informative**

Information  
from the voice

**emotional**

Emotions  
from the voice

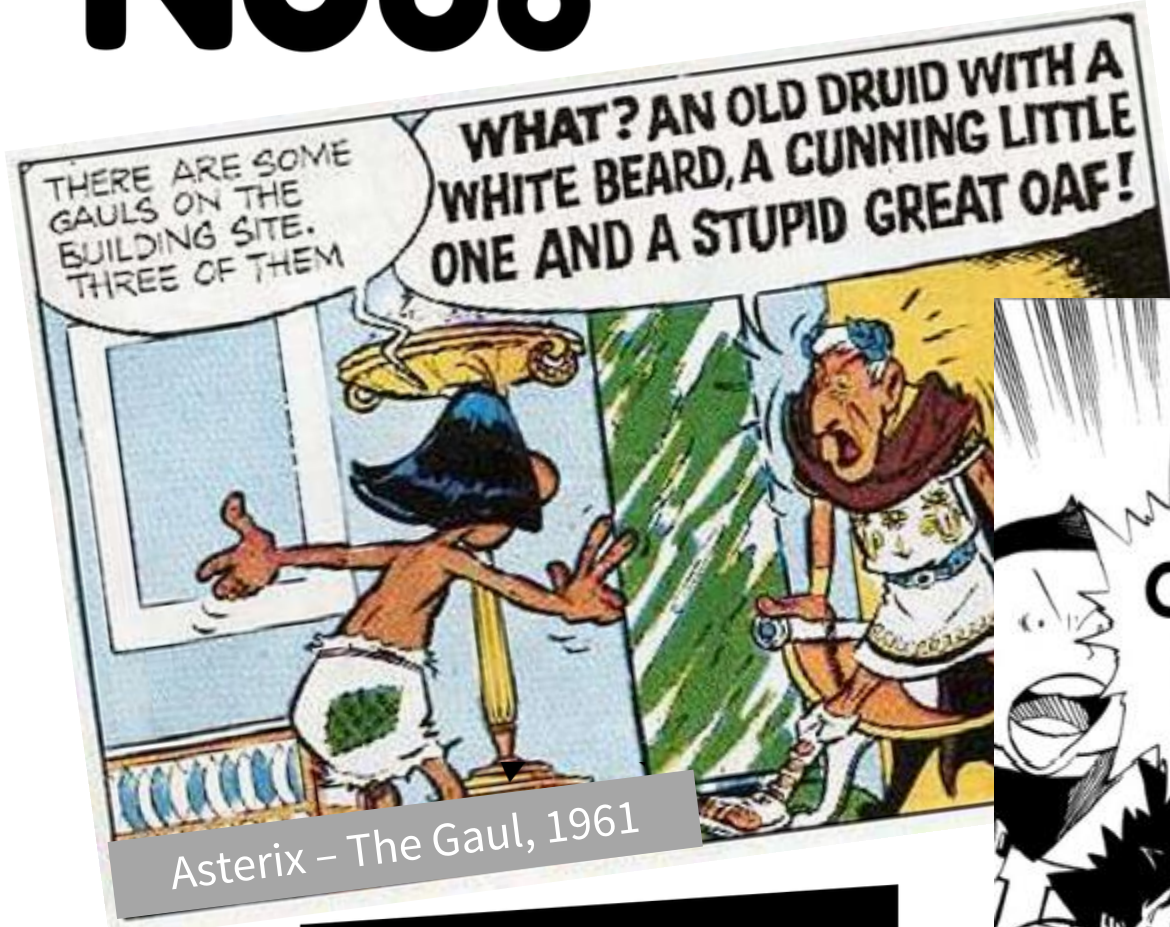


**I have a dream**

**that one day**



# Nooo



**AARGH!**



WEAK OR  
FADING SPEECH

# 2

## RELATED WORK

# RELATED WORK: Font Development and Representation



## Typography as stylish device



Dadaismus

Asterix – The Gaul

**1917**

**1961**



# RELATED WORK: Font Development and Representation



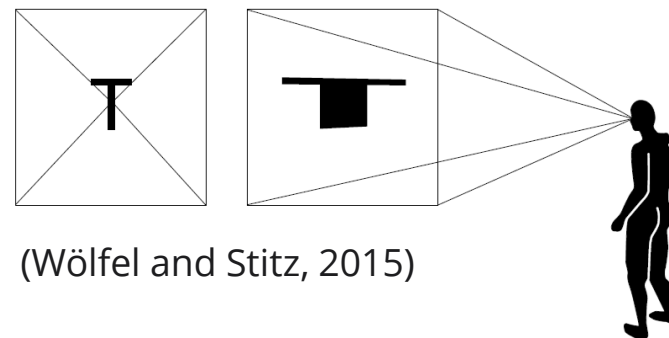
“Digital reading media [has to be treated] as a system of variable and dynamic design elements for text presentation and text accessibility in communicative spheres.” (Kuhn, 2017)

## Typography as stylish device



## Responsive Type

(Lee et al., 2006)  
(Rashid et al., 2008)



(Wölfel and Stitz, 2015)

Dadaismus

Asterix – The Gaul

1917

1961

2000

# RELATED WORK: Font Development and Representation



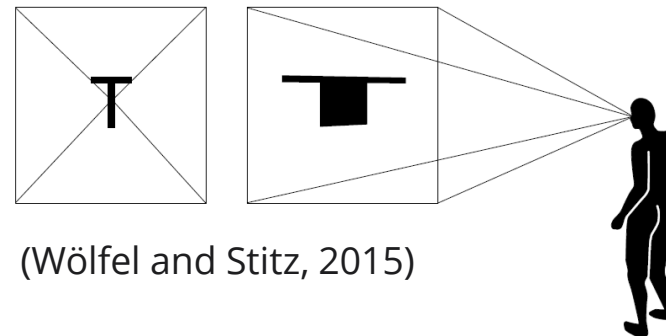
“Digital reading media [has to be treated] as a system of variable and dynamic design elements for text presentation and text accessibility in communicative spheres.” (Kuhn, 2017)

## Typography as stylish device



## Responsive Type

(Lee et al., 2006)  
(Rashid et al., 2008)



(Wölfel and Stitz, 2015)

## Voice Driven Type Design

(Wölfel et al., 2015), (Schlippe et al, 2018),  
(Bessemans et al., 2019)  
(de Lacerda Pataca and Costa, 2020)  
(de Lacerda & Dornhofer Paro Costa, 2022)



Dadaismus

Asterix – The Gaul

1917

1961

2000

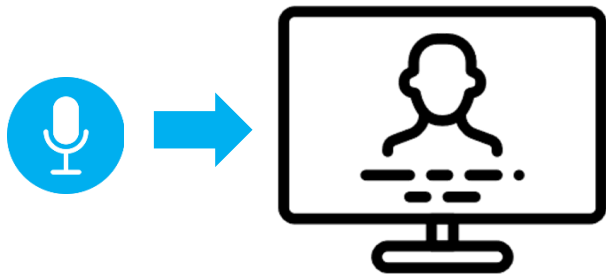
2015

# RELATED WORK: Captioning and Subtitling



## Computer assisted captioning

(Martone et al., 2004)  
(Boulianne et al., 2006)  
(Levin et al., 2014)  
(Tremblay et al., 2015)



**2000**

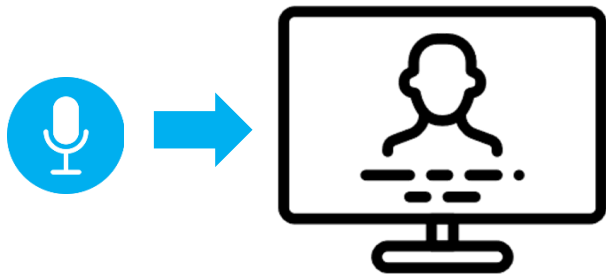
# RELATED WORK:

## Captioning and Subtitling



### Computer assisted captioning

(Martone et al., 2004)  
(Boulianne et al., 2006)  
(Levin et al., 2014)  
(Tremblay et al., 2015)



### Placement and design

(Vy and Fels, 2010)  
(Brown et al., 2015)  
(Fox, 2016)



**2000**

**2010**

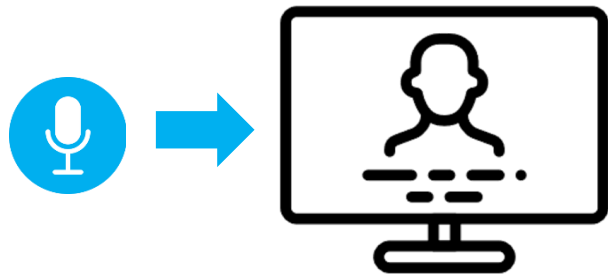
# RELATED WORK:

## Captioning and Subtitling



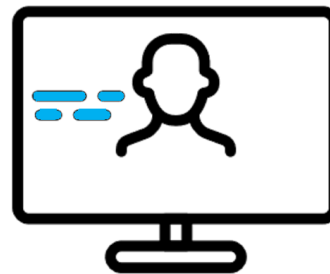
### Computer assisted captioning

(Martone et al., 2004)  
(Boulianne et al., 2006)  
(Levin et al., 2014)  
(Tremblay et al., 2015)



### Placement and design

(Vy and Fels, 2010)  
(Brown et al., 2015)  
(Fox, 2016)



### Emojis and emoticons

(El Taweel, 2016)



**2000**

**2010**

**2016**

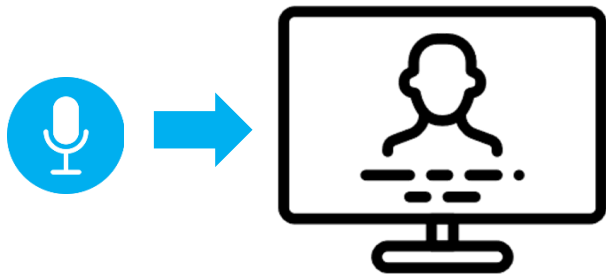


# RELATED WORK: Captioning and Subtitling



## Computer assisted captioning

(Martone et al., 2004)  
(Boulianne et al., 2006)  
(Levin et al., 2014)  
(Tremblay et al., 2015)



## Placement and design

(Vy and Fels, 2010)  
(Brown et al., 2015)  
(Fox, 2016)



## Emojis and emoticons

(El Taweel, 2016)



## WaveFont

(Schlippe et al., 2020)  
(Schlippe et al., 2022)



**2000**

**2010**

**2016**

**2020**

3

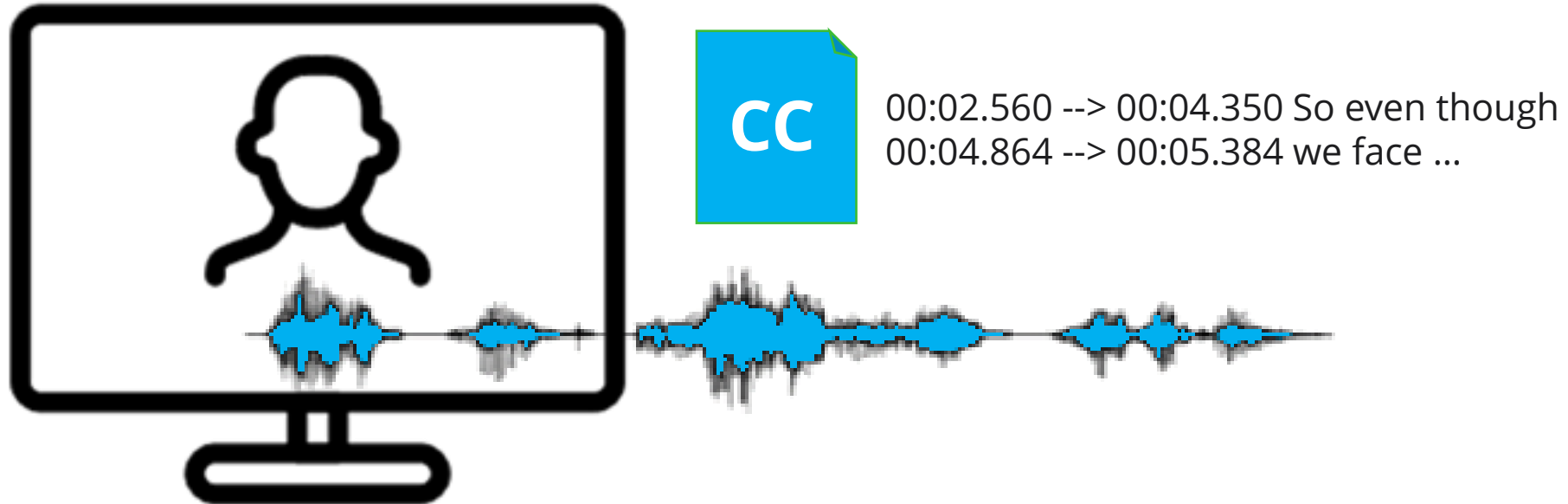
# AI-BASED VISUALIZATION OF VOICE CHARACTERISTICS

## IN CAPTIONS

# TECHNICAL IMPLEMENTATION: Visualization of Voice Characteristics



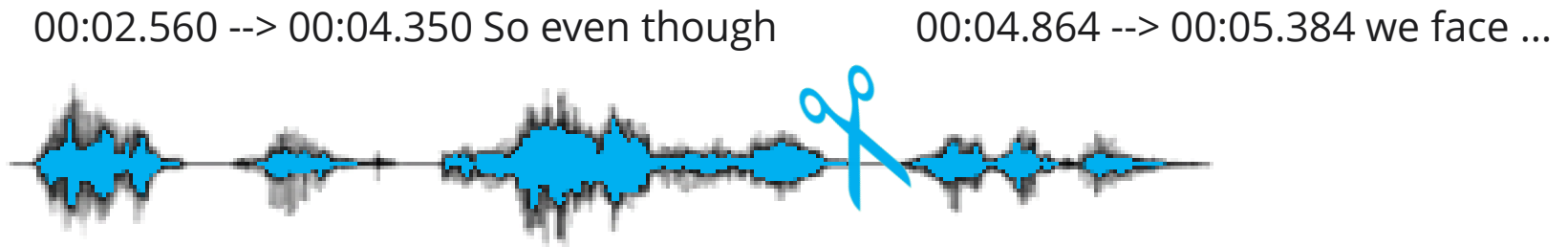
## 1. Extraction of the audio track from the video



# TECHNICAL IMPLEMENTATION: Visualization of Voice Characteristics



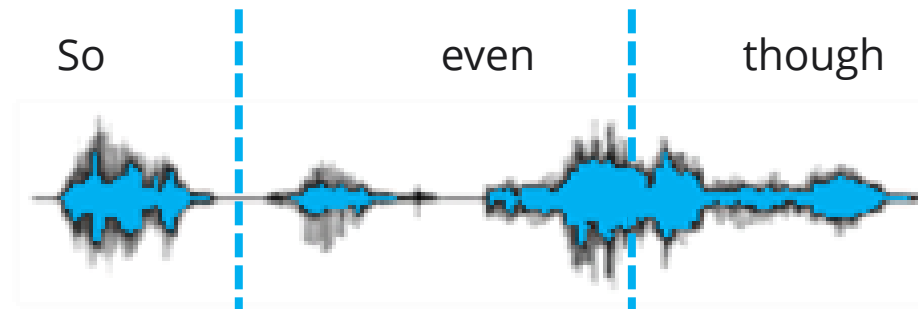
## 2. Segmentation into smaller audio files



# TECHNICAL IMPLEMENTATION: Visualization of Voice Characteristics



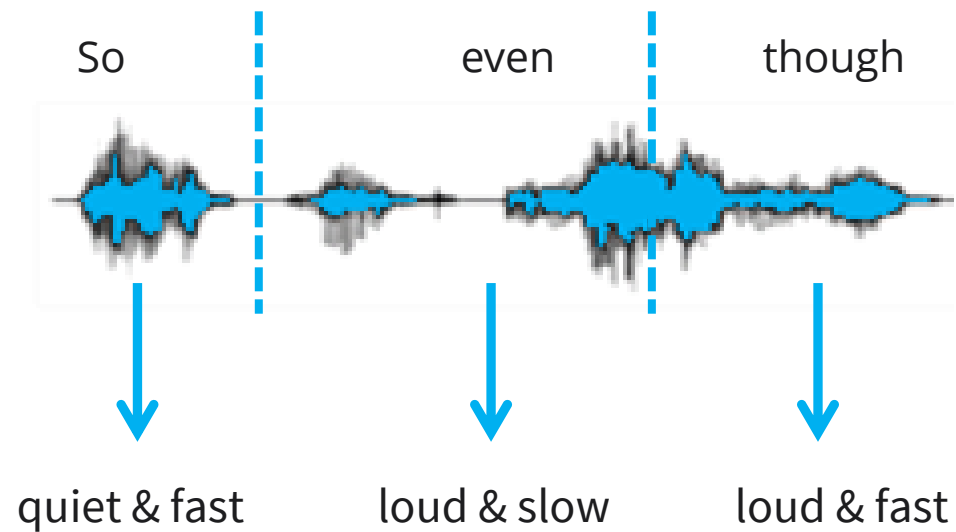
## 3. Automatic forced alignment



# TECHNICAL IMPLEMENTATION: Visualization of Voice Characteristics

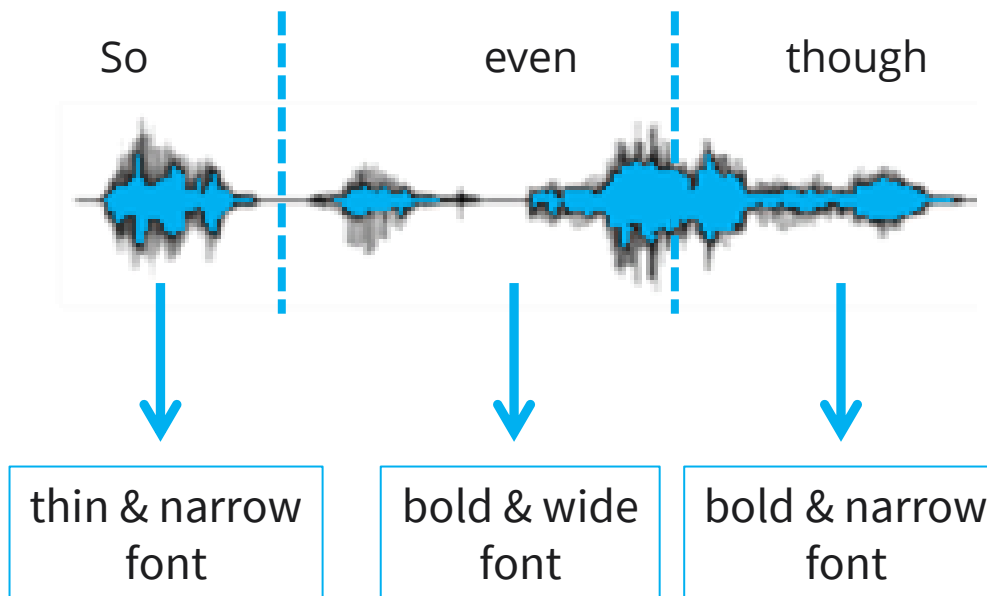


## 4. Acoustic feature extraction



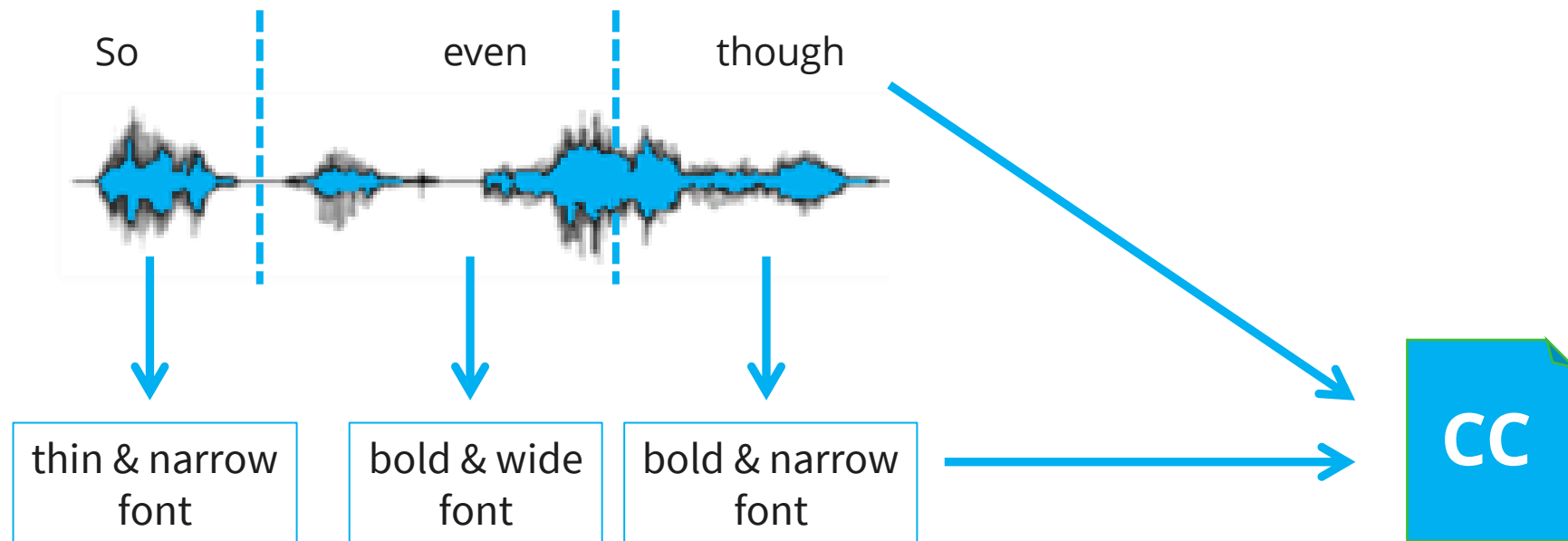
# TECHNICAL IMPLEMENTATION: Visualization of Voice Characteristics

## 5. Mapping of acoustic features to font classes



# TECHNICAL IMPLEMENTATION: Visualization of Voice Characteristics

## 6. Type design





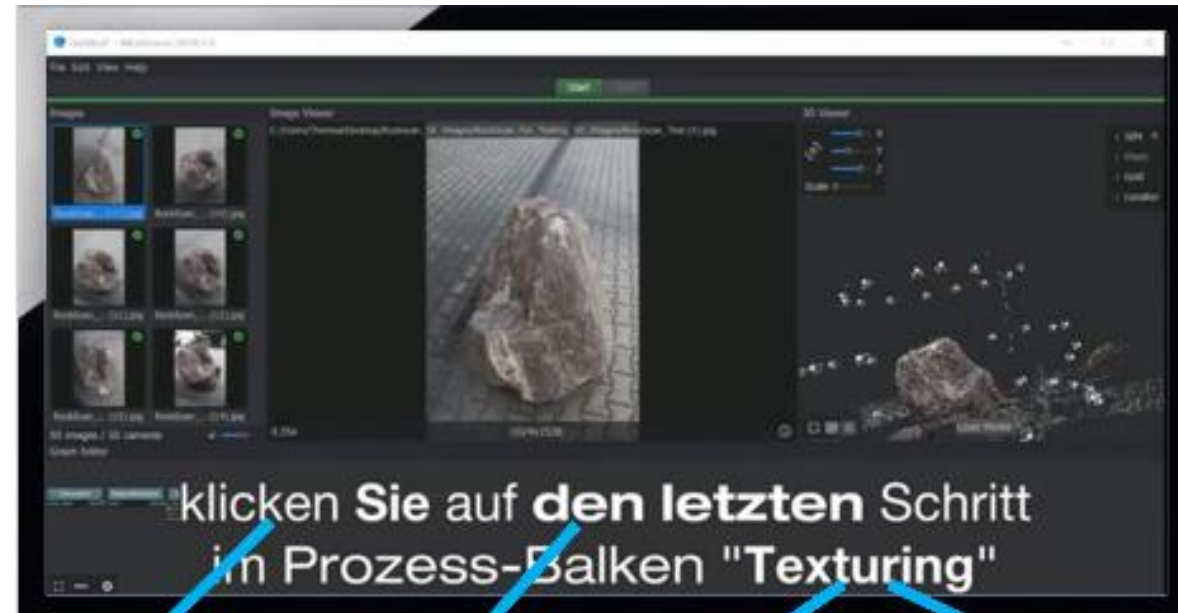
4

**STUDY**

# WAVEFONT IN LECTURE VIDEOS' CAPTIONS



**unique & innovative**



quieter slower faster louder

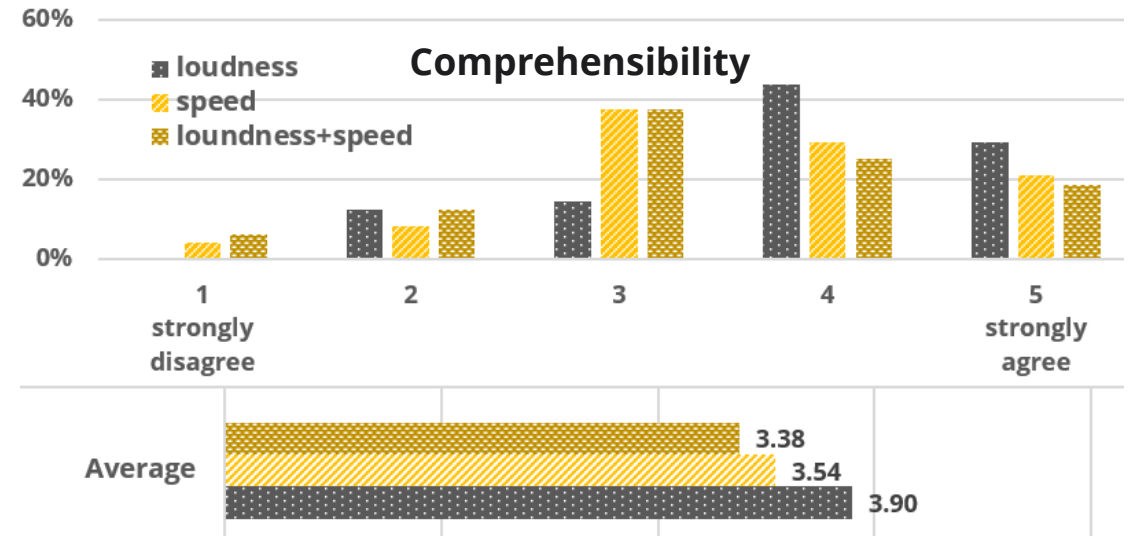
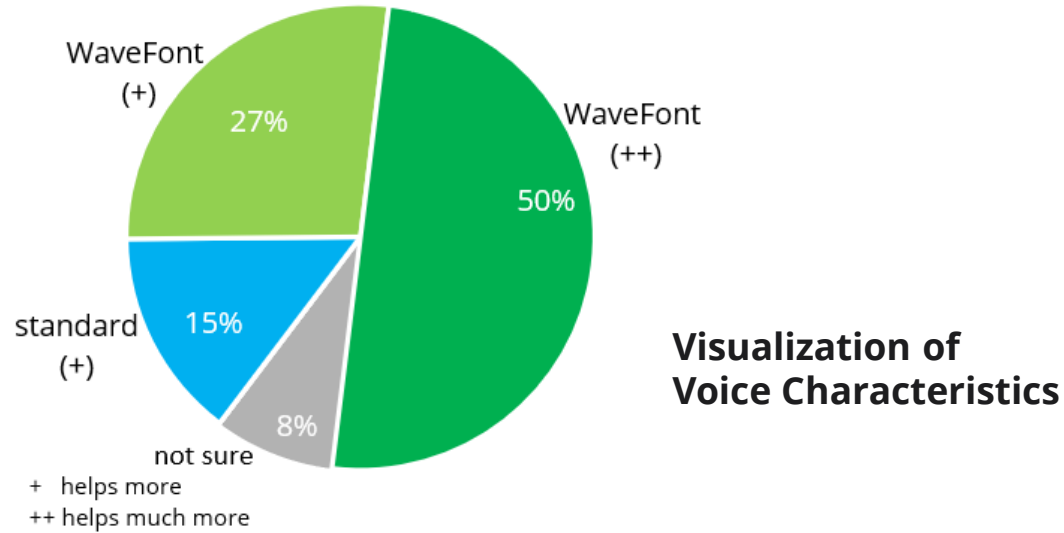
## STUDY

- ✓ visualization of voice characteristics,
- ✓ understanding the content,
- ✓ linguistic understanding,
- ✓ following the content,
- ✓ identifying important words

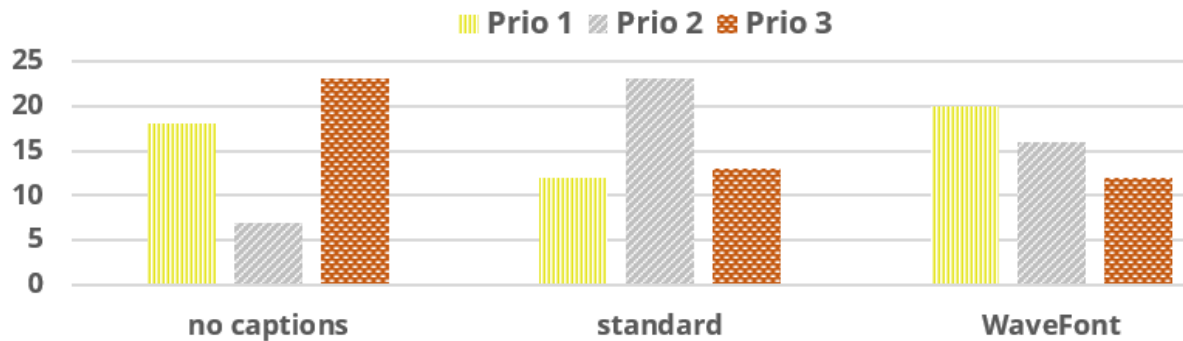
# WAVEFONT IN LECTURE VIDEOS' CAPTIONS



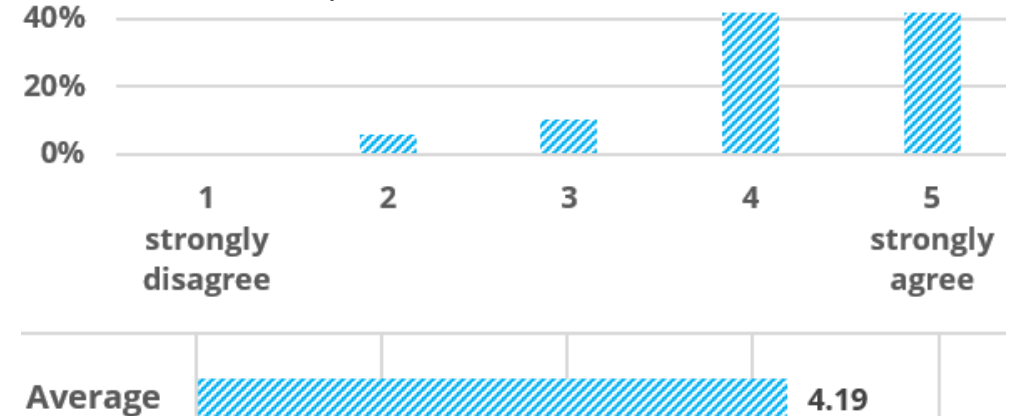
## VISUALIZATION OF VOICE CHARACTERISTICS



## Preferences in terms of captions in lecture videos (1st, 2nd, 3rd priority).



## Assumption of improved WaveFont comprehension with more practice



# 5

## CONCLUSION & FUTURE WORK

# CONCLUSION AND FUTURE WORK

## Conclusion

- AI-based visualization of voice characteristics in captions helps students improve the viewing and learning experience in lecture videos
- *WaveFont* technology processes the speech signal and intuitively displays loudness, speed, and pauses in the subtitle font
- In our survey, the AI-based visualization of speech features outperformed standard captions.
- Participants agree that they assume to understand *WaveFont* with more practice even better

## Future Work

- Analyze learning effect in more detail
- Investigate effect of language acquisition for non-native speakers
- Analyze the effect of *WaveFont* captions on learning styles

**THANK YOU**

Tim Schlippe

 [tim.schlippe@iu.org](mailto:tim.schlippe@iu.org)