

Chapman University

Chapman University Digital Commons

Student Scholar Symposium Abstracts and Posters

Center for Undergraduate Excellence

Spring 5-3-2023

Analysis of Speech-to-Text Algorithms in Recognizing Down Syndrome Conversations

Kayla Anderson

Chapman University, kayanderson@chapman.edu

Cecilia Marie Abrahamsson

Chapman University, abrahamsson@chapman.edu

Yingying 'Yuki' Chen

Chapman University, yingchen@chapman.edu

Franceli L. Cibrian

Chapman University, cibrian@chapman.edu

Follow this and additional works at: https://digitalcommons.chapman.edu/cusrd_abstracts

Recommended Citation

Anderson, Kayla; Abrahamsson, Cecilia Marie; Chen, Yingying 'Yuki'; and Cibrian, Franceli L., "Analysis of Speech-to-Text Algorithms in Recognizing Down Syndrome Conversations" (2023). *Student Scholar Symposium Abstracts and Posters*. 575.

https://digitalcommons.chapman.edu/cusrd_abstracts/575

This Poster is brought to you for free and open access by the Center for Undergraduate Excellence at Chapman University Digital Commons. It has been accepted for inclusion in Student Scholar Symposium Abstracts and Posters by an authorized administrator of Chapman University Digital Commons. For more information, please contact laughtin@chapman.edu.

Analysis of Speech-to-Text Algorithms in Recognizing Down Syndrome Conversations.

Kayla Anderson, Cecilia Marie Abrahamsson, Yingying 'Yuki' Chen, Franceli L. Cibrian



Introduction

Speech-to-text software has become extremely prevalent today. For example, in supporting voice assistants.

However, much of the software is unreliable for the speech patterns of females, children, accents, and individuals with disabilities such as Down syndrome. **This raises the issue of unfair access to current leading technology.**

Particularly, Down Syndrome speech has not been investigated in depth. Such understanding is essential to inform voice-recognition algorithms and improve their performance.

I'm not sure I understand.

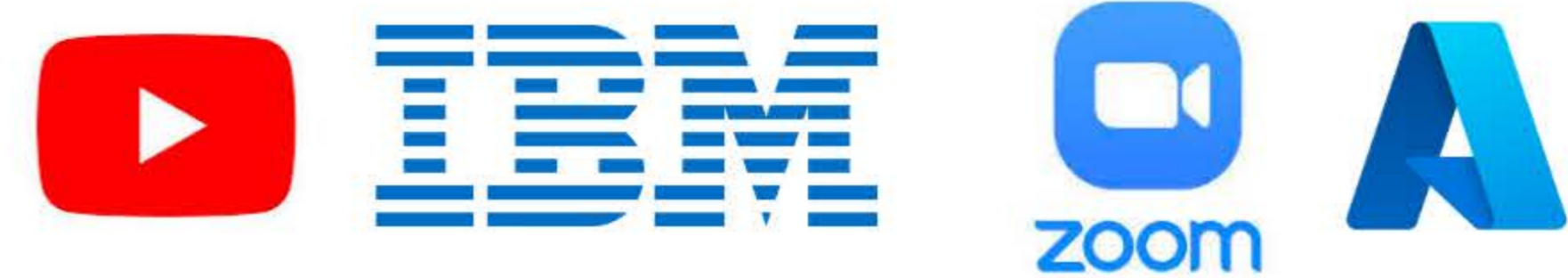


Methods

Aim: to analyze the accuracy of speech recognition systems for young adults with Down Syndrome.

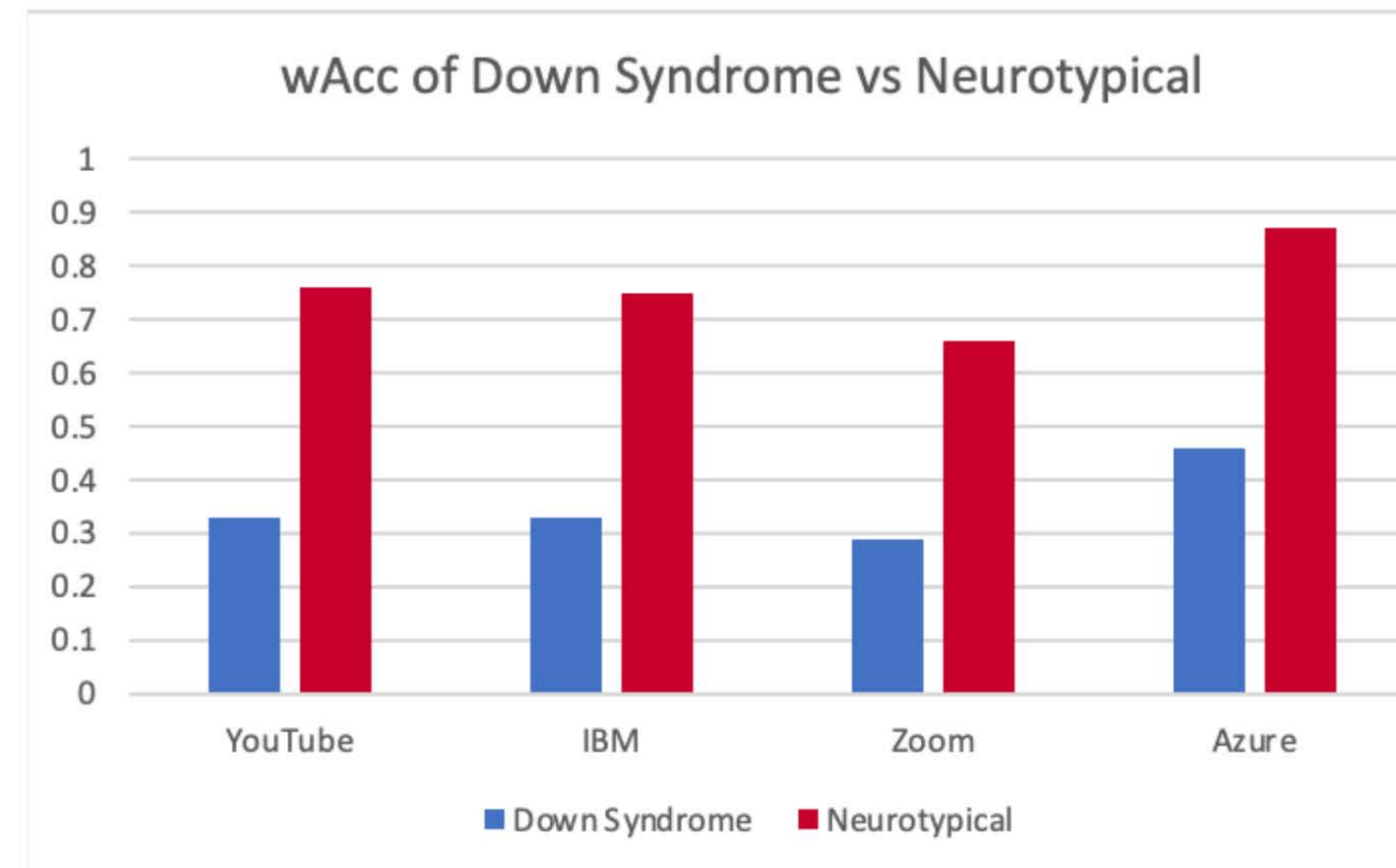
Procedure:

- Extract audio from videos
 - 15 individuals with Down Syndrome
 - 6 neurotypicals.
- Generate **transcripts** with speech-to-text **algorithms**
 - YouTube
 - IBM
 - Zoom
 - Azure
- Data collection
 - **3428** words from 330 dialogues for Down Syndrome
 - **3110** words and 302 dialogues for neurotypicals
- Python script to compute different metrics based on the transcripts
 - Mismatch
 - Deletion
 - Insertion
 - wAcc (weighted average of accuracy)
- Statistical Analysis to determine the significance of results



Results

RQ1) What is the difference in the accuracy of speech recognition algorithms when transcribing speech from neurotypical users and young adults with Down Syndrome?



Weighted average of accuracy (wAcc) for YouTube, IBM, Zoom, Azure for neurotypicals vs. Down Syndrome

RQ2) What are the most common types of errors that occur in speech recognition for young adults with Down Syndrome?

	YouTube	IBM	Zoom	Azure
Mismatch	1,281	1,614	1,228	1,301
Deletion	631	588	1,181	452
Insertion	330	382	142	318
Total	2,242	2,584	2,551	2,071

Mismatch, delete, and insert for all four algorithms for Down syndrome participants

Conclusion

- All algorithms perform around **40%** better for neurotypical participants
- **Azure is the best-performing** algorithm for both neurotypicals and people with Down syndrome
- **Mismatch** is the most common type of error, followed by deletion and insertion
- **Zoom** has the least amount of insertions since it generates transcripts in real time

Future Work

Our next step will be:

- Performing phonetic analysis to identify key sounds that prove difficult to detect in each algorithm,
- Include two new generative AI algorithms AssemblyAI and OpenAI in the transcripts.

Acknowledgement

The first three authors equally contributed to this research. Additional thanks to Dr. Vivian Genaro Motti for her contributions to this research.

