

#CUNYTuesday

City University of New York

School of Professional Studies
Data Science Workshop

Duubar Villalobos Jimenez

MS in Data Science Graduate Student
CUNY School of Professional Studies

This work was made in conjunction with:
Dipika Shrestha
Graduate Student
Masters in International Affairs
Baruch Marxe School of Public and International Affairs

About me

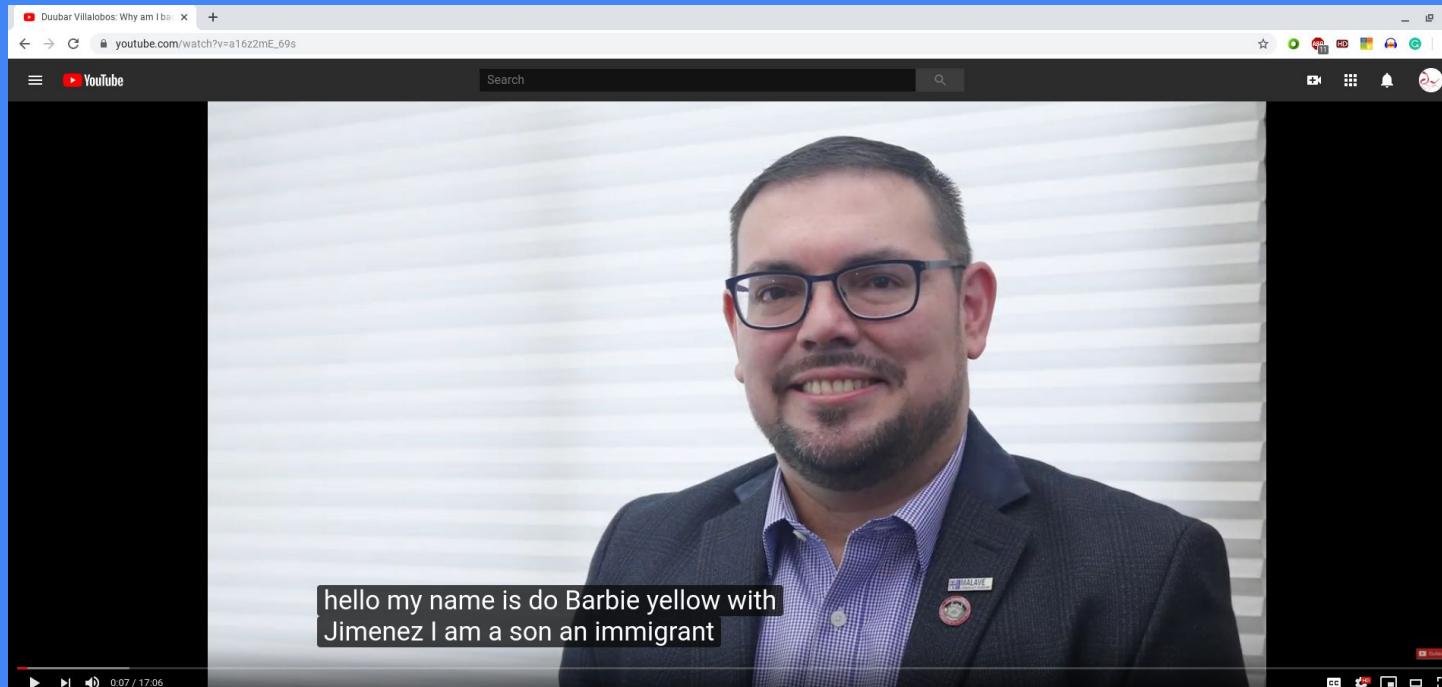
I'm passionate about building great relationships with social impact. I have over 16 years of experience strategizing and innovating digital experiences.

I am an avid data science student, and I'm excited to share my experiences with you!

About the Capstone Project

This capstone project represents highly sophisticated, but practical, solutions to address real problems.

Real world problem:



hello my name is do Barbie yellow with
Jimenez I am a son an immigrant

Duubar Villalobos: Why am I back to school after 16 years (Podcast Episode 63)

#DuubarVillalobos #ATheEndofTheDay #Podcast

344 views • Premiered Aug 4, 2019

13 0 SHARE SAVE ...

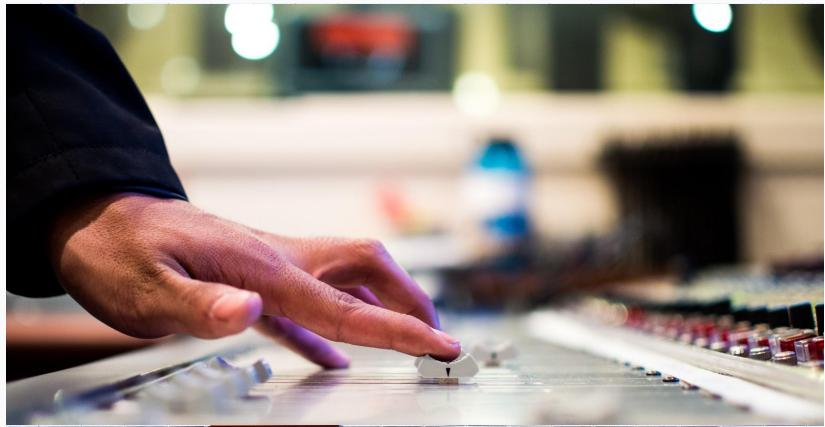
Chat Reply is disabled for this Premiere.

7:51



The guest pronounced

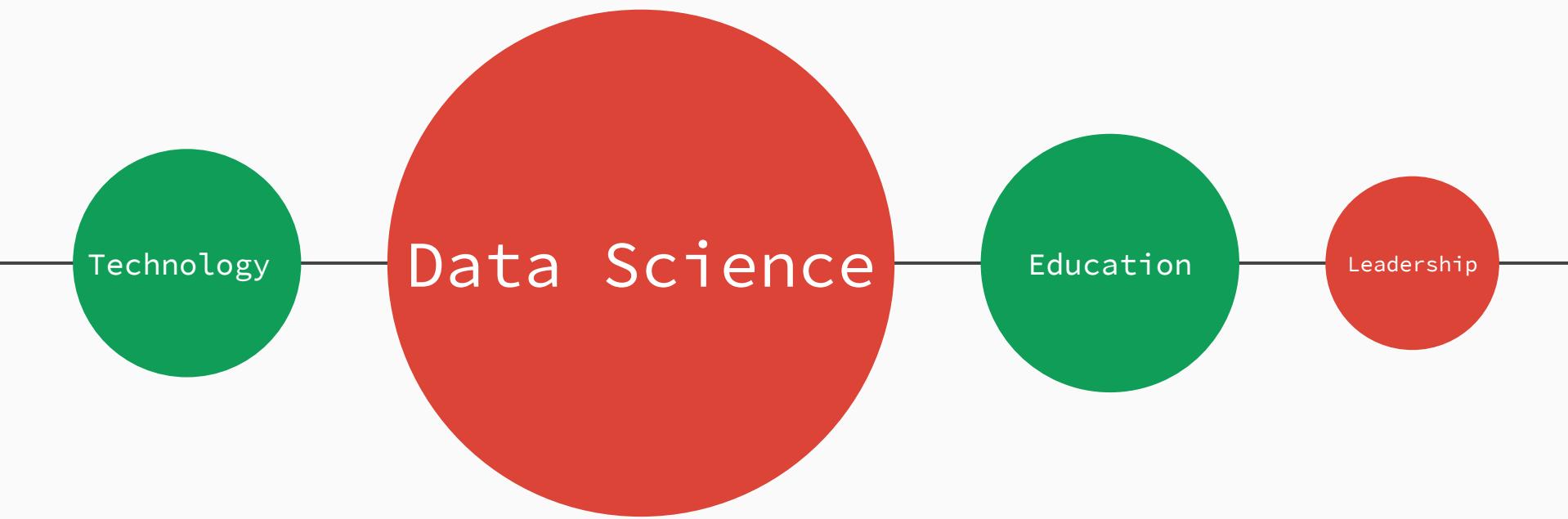
"Hello my name is Duubar Villalobos..."



YouTube's automated transcript

"Hello my name is do Barbie yellow with..."

Youtube's Automated Speech Recognition, Automated Captioning, and Sentiment Analysis from Non-native English Speakers with Foreign Accents.



#	Automated	Manual	Accurate	Words	Similarity	Meaning	Gender	Accent	in US
1	University Columbia needs to invite you	University Columbia University NYU	No	+	Different	Change	F	Rus	4-8
2	years ago of the bad baby from	years ago after graduating from	No	+	Different	Unclear	F	Sp	4-8
3	went to Technica is fine and and it was	Juventud Technica its Spanish and and it was	No	+	Different	Change	F	Sp	4-8
4	this is a validation that health	this organization that helps	No	+	Different	Change	F	Sp	4-8
5	listened in a vision they a lot of	listened Univision there a lot of	No	+	Different	Change	F	Sp	4-8
6	report when you to know everything we	want you to know everything want	No	+	Different	Mssng Wrds	M	Sp	4-8
7	proper no	proper	No	+	Similar	Unclear	M	Sp	16+
8	campus I never been to campus before no	campus I never been to campus before	No	+	Very Similar	Similar	M	Sp	16+

Question 1

How accurately does YouTube's transcription technology turn the spoken words into text captions for various English foreign accents?

Hypotheses

Null: Is one hundred percent accurate.

Alternative: Is not one hundred percent accurate.

Question 2

Does automatic YouTube transcription technology change the expressed sentiment of the spoken words?

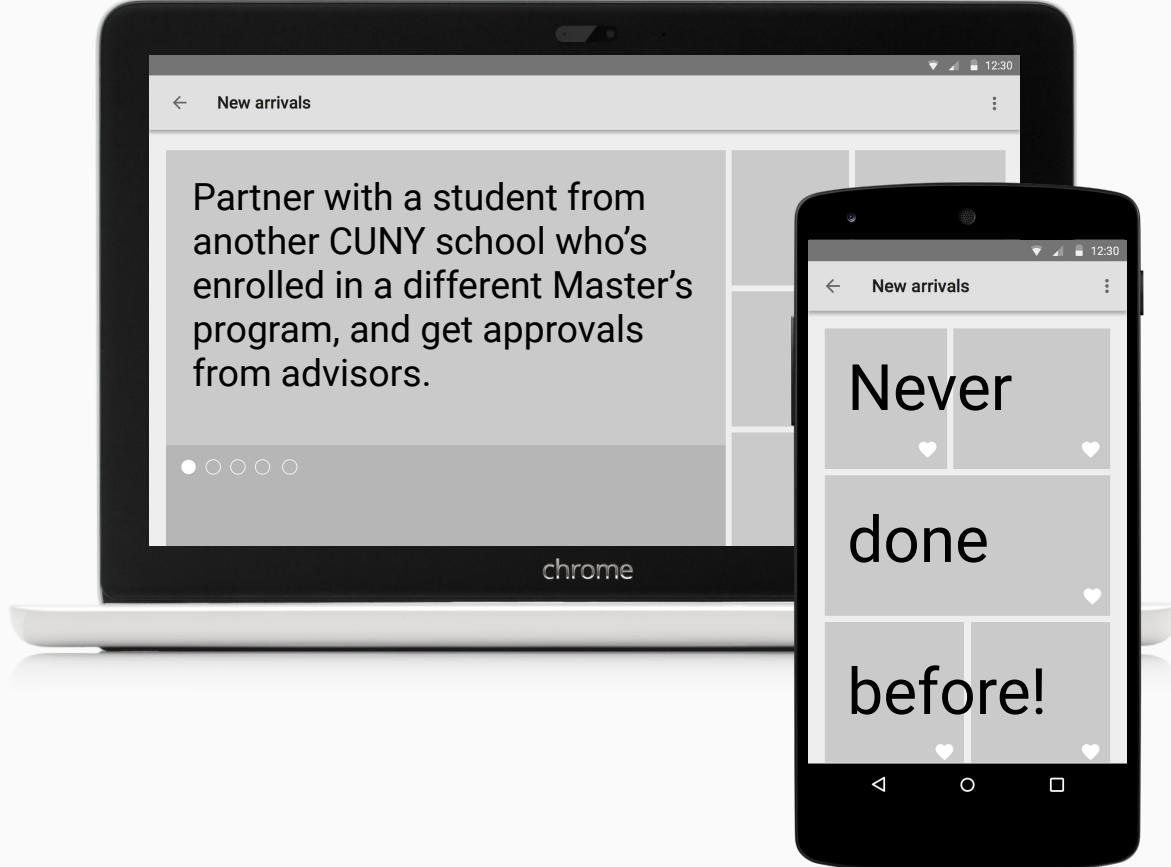
Hypotheses

Null: The sentiment does not change.

Alternative: The sentiment changes.

Methodology

1. Identify videos that fit the real world problem criteria.
2. Extract automated transcripts from YouTube videos.
3. Manually transcribe videos.
4. Extract metadata from videos.
5. Perform sentiment analysis on both transcripts.
6. Compare outputs and results.
7. Present results based on diverse comparisons.



Real World Problem Highlights

Jun 21, 2018

- More than **1.9 billion** logged-in users -worldwide- who come to YouTube **every month**.
- YouTube localized versions stretching across **90 countries and 80 languages**.

Deafness around the world

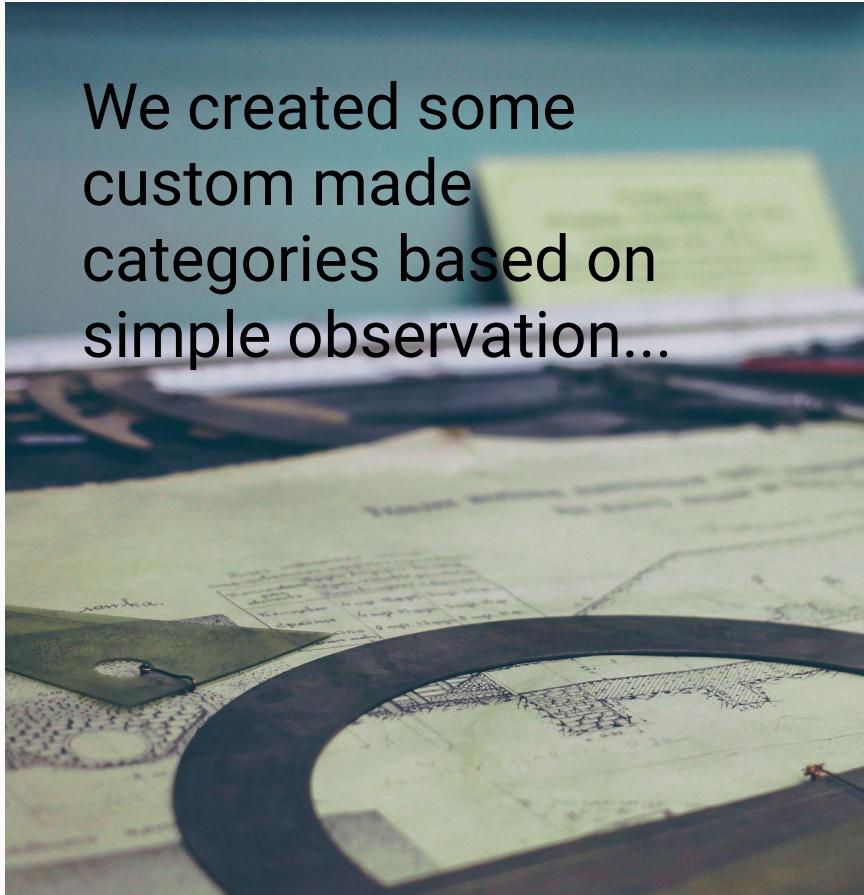
- 466 million: The number of people worldwide that have disabling hearing loss.
- 1.1 billion: The number of young people (aged between 12–35 years) at risk of hearing loss due to exposure to noise in recreational settings.
- One-third: The number of people over 65 years of age that are affected by disabling hearing loss.

English Speakers around the world

- 565 million: Use English on the internet
- 385 million: Native English speakers
- 800 million: Non-native English speakers

We do not want to make another sentiment analysis study. We want to create scientific and social awareness.

We created some
custom made
categories based on
simple observation...



Custom categories for variations in transcripts

- No Difference
 - Missing Words
 - Similar meaning
 - Change in meaning
 - Opposite meaning
 - Unclear meaning
 - Offensive meaning
1. We compared the scripts in between them.
 2. We compared by **gender**.
 3. We compared by **language**.
 4. We compared by the **number of years residing** in the U.S.
 5. We compared **similarities based on word count**.
 6. We compared **similarities based in The Cosine approach**.
 7. We calculated **The Flesch** readability score and compared the results.
 8. We calculated **The Dale–Chall** readability level and then we compared the results.
 9. We performed Sentiment Analysis using **VADER** on both scripts.

Summaries

#	YouTube_ID	Length	Gender	Country of Origin	Accent	Years in US
1	7jff8Y9zvV0	0:23:30	Female	Kazakhstan	Kazakh	8-12
2	a16z2mE_69s	0:17:06	Male	Costa Rica	Spanish	16+
3	KBT2F0hdqQc	0:07:49	Female	Russia	Russian	4-8
4	uXJwQNEXuug	0:10:32	Male	Burkina Faso	French	4-8
5	xtNBDzgWt2I	0:18:11	Male	Cuba	Spanish	4-8
6	yl3k91L1hng	0:14:06	Female	Cuba	Spanish	4-8

Distribution of summaries

Gender	Percent
Female	50%
Male	50%

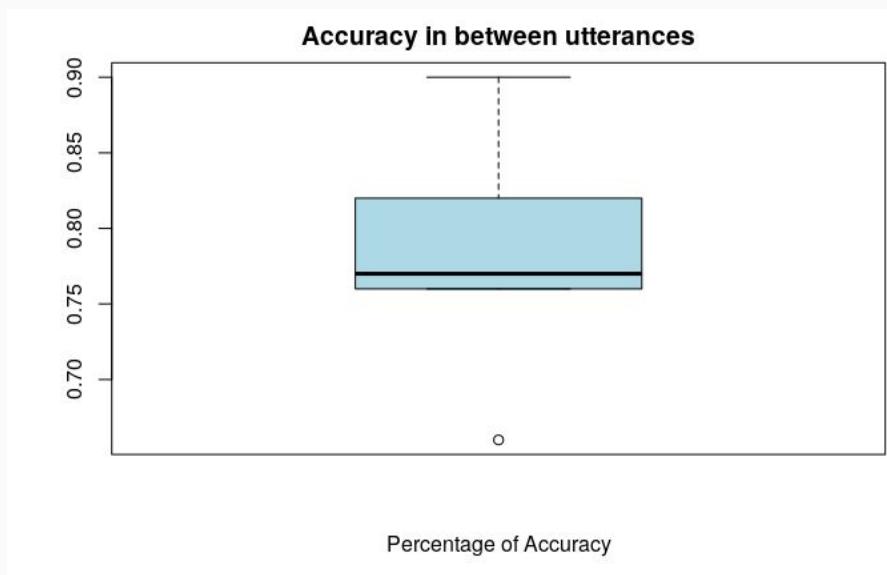
Accent	Percent
Spanish	50.0%
French	16.7%
Kazakh	16.7%
Russian	16.7%

Years in US	Percent
4-8	66.7%
16+	16.7%
8-12	16.7%

How accurately does YouTube's transcription technology turn the spoken words into text captions for various English foreign accents?

Based on the **simple comparison** of utterances, the **accuracy is less than 100%**.

		Accurate	
#	YouTube_ID	Yes	No
1	7jff8Y9zvV0	82.2%	17.8%
2	a16z2mE_69s	90.5%	9.5%
3	KBT2F0hdqQc	77.2%	22.8%
4	uXJwQNEXuug	66.4%	33.6%
5	xtNBDzgWt2I	75.9%	24.1%
6	yl3k91L1hng	77.0%	23.0%
Average		78.2%	21.8%



How accurately does YouTube's transcription technology turn the spoken words into text captions for various English foreign accents?

Based on the **visual comparison** of utterances, the **accuracy is less than 100%**.

				Meaning				
#	YouTube_ID	Accurate	Missing Words	Similar	Change	Unclear	Opposite	Offensive
1	7jff8Y9zvV0	82.2%	1.2%	9.7%	4.3%	2.2%	0.0%	0.4%
2	a16z2mE_69s	90.5%	0.3%	3.5%	2.7%	2.7%	0.0%	0.3%
3	KBT2F0hdqQc	77.2%	0.6%	8.3%	9.4%	2.8%	0.6%	1.1%
4	uXJwQNEXuug	66.4%	2.2%	13.4%	15.1%	3.0%	0.0%	0.0%
5	xtNBDzgWt2I	75.9%	1.6%	9.6%	7.1%	2.5%	2.7%	0.5%
6	yl3k91L1hng	77.0%	0.7%	10.0%	6.7%	4.4%	0.4%	0.7%
	Average	78.2%	1.1%	9.1%	7.6%	2.9%	0.6%	0.5%

How accurately does YouTube's transcription technology turn the spoken words into text captions for various English foreign accents?

Based on the **Cosine comparison** of utterances, the accuracy is *less than 100%*.

Based on Angle: Accurate = $0 < \text{Very Similar} \leq \pi/8$ Similar $\pi/8 < \pi/4 \leq \text{Different} < 3\pi/8 \leq \text{Very Different}$

#	YouTube_ID	Accurate	Vector			
			Very Similar	Similar	Different	Very Different
1	7jff8Y9zvV0	82.6%	2.0%	13.2%	1.8%	0.4%
2	a16z2mE_69s	90.5%	0.5%	6.8%	2.2%	0.0%
3	KBT2F0hdqQc	79.4%	2.2%	16.7%	1.7%	0.0%
4	uXJwQNEXuug	66.8%	3.9%	19.0%	7.8%	2.6%
5	xtNBDzgWt2l	77.3%	3.0%	15.6%	3.6%	0.5%
6	yI3k91L1hng	77.4%	1.5%	13.3%	7.0%	0.7%
Average		79.0%	2.2%	14.1%	4.0%	0.7%

How accurately does YouTube's transcription technology turn the spoken words into text captions for various English foreign accents?

Based on the **simple comparison** of utterances by Gender.

From the left table:

		Accurate	
#	Gender	Yes	No
1	Female	79.7%	20.3%
2	Male	79.2%	20.8%
	Average	79.5%	20.6%

- Automated transcription performed slightly better for utterances captured from Females speakers compared to utterances captured from Male speakers for all videos that were analyzed.

How accurately does YouTube's transcription technology turn the spoken words into text captions for various English foreign accents?

Group	Study	Accuracy
Compare	Side by side	79.5%
	Visual	79.5%
	Cosine	80.1%
	Number of words	90.4%
<hr/>		
Reading	Flesch Reading Ease	91.4%
	Dale–Chall readability	98.1%
	Sentiment Analysis	97.7%

How accurately does YouTube's transcription technology turn the spoken words into text captions for various English foreign accents?

Group	Study	Side by Side	Visual	Cosine	Word Count
Gender	Female	79.7%	79.7%	80.5%	91.7%
	Male	79.2%	79.2%	79.8%	89.0%
Accent	Kazakh	82.2%	82.2%	82.6%	92.9%
	Spanish	81.6%	81.6%	82.2%	91.2%
	Russian	77.2%	77.2%	79.4%	92.8%
	French	66.4%	66.4%	66.8%	79.3%
Years in US	16+	90.5%	90.5%	90.5%	94.8%
	8-12	82.2%	82.2%	82.6%	92.9%
	4-8	74.3%	74.3%	75.4%	87.6%

How accurately does YouTube's transcription technology turn the spoken words into text captions for various English foreign accents?

Group	Study	Side by Side	Visual	Cosine	Word Count
Sentiment	Accurate	81.3%	81.3%	82.0%	91.3%
	Not Accurate	0.0%	0.0%	0.0%	50.0%
Words	Accurate	88.7%	87.9%	88.7%	90.4%
	Decreased	0.0%	0.0%	0.0%	0.0%
	Increased	0.0%	0.0%	0.0%	0.0%
FRES	Accurate	86.9%	86.9%	87.4%	95.8%
	Not Accurate	0.0%	0.0%	2.4%	32.9%
Dale-Chall	Accurate	81.0%	81.0%	81.7%	91.6%
	Not Accurate	0.0%	0.0%	2.7%	29.7%

How accurately does YouTube's transcription technology turn the spoken words into text captions for various English foreign accents?

Group	Type	FRES	Dale-Chall	Sentiment
Gender	Female	92.20%	98.10%	97.60%
	Male	90.70%	98.00%	97.80%
Accent	Kazakh	92.70%	98.40%	98.60%
	Spanish	92.20%	98.40%	97.00%
	Russian	92.80%	97.20%	97.80%
	French	84.10%	96.60%	98.70%
Years in US	16+	94.80%	99.20%	99.20%
	8-12	92.70%	98.40%	98.60%
	4-8	89.60%	97.50%	96.80%

How accurately does YouTube's transcription technology turn the spoken words into text captions for various English foreign accents?

Group	Type	FRES	Dale-Chall	Sentiment
Sentiment	Accurate	92.30%	98.30%	97.70%
	Not Accurate	54.50%	88.60%	0.00%
Words	Accurate	96.90%	99.40%	98.70%
	Decreased	45.90%	90.60%	94.10%
	Increased	35.40%	81.80%	82.80%
FRES	Accurate	91.40%	99.40%	98.60%
	Not Accurate	0.00%	84.10%	87.80%
Dale-Chall	Accurate	92.60%	98.10%	97.90%
	Not Accurate	29.70%	0.00%	86.50%

How accurately does YouTube's transcription technology turn the spoken words into text captions for various English foreign accents? -- By Visual Comparison

Group	Type	Accurate	Missing Words	Similar	Change	Unclear	Opposite	Offensive
Gender	Female	79.7%	1.0%	9.5%	5.9%	3.0%	0.2%	0.6%
	Male	79.2%	1.2%	8.2%	7.4%	2.7%	1.0%	0.3%
Accent	Kazakh	82.2%	1.2%	9.7%	4.3%	2.2%	0.0%	0.4%
	Spanish	81.6%	0.9%	7.5%	5.4%	3.1%	1.1%	0.5%
	Russian	77.2%	0.6%	8.3%	9.4%	2.8%	0.6%	1.1%
	French	66.4%	2.2%	13.4%	15.1%	3.0%	0.0%	0.0%
Years in US	16+	90.5%	0.3%	3.5%	2.7%	2.7%	0.0%	0.3%
	8-12	82.2%	1.2%	9.7%	4.3%	2.2%	0.0%	0.4%
	4-8	74.3%	1.3%	10.3%	9.2%	3.2%	1.1%	0.6%

How accurately does YouTube's transcription technology turn the spoken words into text captions for various English foreign accents? -- By Visual Comparison

Group	Type	Accurate	Missing Words	Similar	Change	Unclear	Opposite	Offensive
Sentiment	Accurate	81.3%	0.9%	8.6%	6.0%	2.5%	0.5%	0.2%
	Not Accurate	0.0%	11.4%	18.2%	36.4%	15.9%	6.8%	11.4%
Words	Accurate	87.9%	0.0%	5.5%	4.0%	1.8%	0.4%	0.3%
	Decreased	0.0%	22.4%	43.5%	16.5%	10.6%	5.9%	1.2%
	Increased	0.0%	2.0%	37.4%	44.4%	14.1%	0.0%	2.0%
FRES	Accurate	86.9%	0.4%	6.8%	3.7%	1.7%	0.3%	0.2%
	Not Accurate	0.0%	8.5%	30.5%	38.4%	14.6%	4.3%	3.7%
Dale-Chall	Accurate	81.0%	1.1%	8.3%	5.9%	2.6%	0.6%	0.4%
	Not Accurate	0.0%	2.7%	35.1%	43.2%	13.5%	0.0%	5.4%

How accurately does YouTube's transcription technology turn the spoken words into text captions for various English foreign accents? -- By Cosine Comparison

Group	Type	Accurate	Very Similar	Similar	Different	Very Different
Gender	Female	80.5%	1.9%	13.9%	3.3%	0.4%
	Male	79.8%	2.3%	13.1%	4.0%	0.0%
Accent	Kazakh	82.6%	2.0%	13.2%	1.8%	0.4%
	Spanish	82.2%	1.7%	11.8%	4.0%	0.4%
	Russian	79.4%	2.2%	16.7%	1.7%	0.0%
	French	66.8%	3.9%	19.0%	7.8%	2.6%
Years in US	16+	90.5%	0.5%	6.8%	2.2%	0.0%
	8-12	82.6%	2.0%	13.2%	1.8%	0.4%
	4-8	75.4%	2.7%	16.0%	5.1%	1.0%

How accurately does YouTube's transcription technology turn the spoken words into text captions for various English foreign accents? -- By Cosine Comparison

Group	Type	Accurate	Very Similar	Similar	Different	Very Different
Sentiment	Accurate	82.0%	2.1%	12.6%	2.8%	0.5%
	Not Accurate	0.0%	2.3%	52.3%	40.9%	4.5%
Words	Accurate	88.7%	0.1%	9.6%	1.4%	0.2%
	Decreased	0.0%	21.2%	56.5%	18.8%	3.5%
	Increased	0.0%	21.2%	43.4%	30.3%	5.1%
FRES	Accurate	87.4%	1.4%	9.5%	1.5%	0.2%
	Not Accurate	2.4%	9.8%	56.1%	26.2%	5.5%
Dale-Chall	Accurate	81.7%	2.0%	12.9%	2.9%	0.5%
	Not Accurate	2.7%	8.1%	43.2%	40.5%	5.4%

How accurately does YouTube's transcription technology turn the spoken words into text captions for various English foreign accents? -- By Word Count Comparison

Group	Type	Accurate	Decreased	Increased
Gender	Female	91.7%	3.8%	4.5%
	Male	89.0%	5.1%	5.9%
Accent	Kazakh	92.9%	4.1%	3.0%
	Spanish	91.2%	4.1%	4.7%
	Russian	92.8%	3.9%	3.3%
	French	79.3%	7.3%	13.4%
Years in US	16+	94.8%	2.2%	3.0%
	8-12	92.9%	4.1%	3.0%
	4-8	87.6%	5.4%	7.0%

How accurately does YouTube's transcription technology turn the spoken words into text captions for various English foreign accents? -- By Word Count Comparison

Group	Type	Accurate	Decreased	Increased
Sentiment	Accurate	91.3%	4.3%	4.4%
	Not Accurate	50.0%	11.4%	38.6%
Words	Counts	90.4%	4.5%	5.2%
	-----	-----	-----	-----
FRES	Accurate	95.8%	2.2%	2.0%
	Not Accurate	32.9%	28.0%	39.0%
Dale-Chall	Accurate	91.6%	4.1%	4.3%
	Not Accurate	29.7%	21.6%	48.6%

How accurately does YouTube's transcription technology turn the spoken words into text captions for various English foreign accents? -- FRES Levels of Comparison

Based on **The Flesch Reading Ease score** confusion matrix, the **Accuracy = 91.4%**

Automated	Manual Reference								Total
	0	1	2	3	4	5	6	7	
0	27.4%	0.4%	0.3%	0.0%	0.0%	0.0%	0.0%	0.0%	28.0%
1	1.1%	13.8%	0.7%	0.2%	0.1%	0.1%	0.0%	0.0%	15.9%
2	0.5%	0.7%	21.4%	0.3%	0.2%	0.1%	0.0%	0.0%	23.1%
3	0.2%	0.1%	0.3%	5.3%	0.2%	0.1%	0.0%	0.0%	6.1%
4	0.0%	0.3%	0.2%	0.1%	6.2%	0.2%	0.1%	0.0%	7.1%
5	0.1%	0.1%	0.2%	0.3%	0.3%	7.0%	0.2%	0.0%	8.1%
6	0.1%	0.1%	0.1%	0.2%	0.2%	0.3%	6.2%	0.0%	7.0%
7	0.0%	0.0%	0.1%	0.1%	0.0%	0.0%	0.5%	3.9%	4.5%
Total	29.3%	15.5%	23.1%	6.4%	7.1%	7.7%	7.0%	3.9%	100.0%

How accurately does YouTube's transcription technology turn the spoken words into text captions for various English foreign accents? -- Dale-Chall Levels of Comparison

Based on **The Dale–Chall readability score labeling** confusion matrix.

The Accuracy = 98.1%

Difficulty		Manual Reference		
Level	Automated	0	1	Total
0	High School	81.6%	0.7%	82.3%
1	College	1.2%	16.5%	17.7%
	Total	82.8%	17.2%	100.0%

Question 1: How accurately does YouTube's transcription technology turn the spoken words into text captions for various English foreign accents?

Based on the above analysis performed on the utterances, we could conclude as follows:

We reject the Null hypothesis and accept the Alternative hypothesis.

Alternative: It is not one hundred percent accurate.

Does automatic YouTube transcription technology change the expressed sentiment of the spoken words?

Sentiment Analysis performed using **VADER** on individual utterances.

Sentiment	Transcript		
	Manual	Automated	Change
Negative	6.1%	6.9%	-0.8%
Neutral	63.5%	62.8%	0.7%
Positive	30.4%	30.3%	0.1%
Total	100.0%	100.0%	0.0%

Does automatic YouTube transcription technology change the expressed sentiment of the spoken words?

Based on **Sentiment Analysis** confusion matrix, the **Accuracy = 97.7%**

Automated	Manual Reference				Total
	Negative	Neutral	Positive		
Negative	5.9%	0.1%	0.1%		6.1%
Neutral	0.7%	62.2%	0.7%		63.6%
Positive	0.3%	0.5%	29.6%		30.4%
Total	6.9%	62.8%	30.4%		100.0%

Does automatic YouTube transcription technology change the expressed sentiment of the spoken words?

Sentiment Analysis by Gender and Source Type.

		Sentiment		
Gender	Type	Negative	Neutral	Positive
Female	Manual	6.2%	63.7%	30.1%
Female	Automated	6.8%	63.1%	30.1%
Male	Manual	6.0%	63.3%	30.7%
Male	Automated	6.9%	62.5%	30.6%
Female	Difference	-0.6%	0.6%	0.0%
Male	Difference	-0.9%	0.8%	0.1%

Does automatic YouTube transcription technology change the expressed sentiment of the spoken words?

		Sentiment		
Mother Language	Type	Negative	Neutral	Positive
French	Manual	4.7%	62.5%	32.8%
French	Automated	4.7%	62.5%	32.8%
Kazakh	Manual	5.3%	69.6%	25.2%
Kazakh	Automated	5.5%	70.4%	24.1%
Russian	Manual	7.8%	59.4%	32.8%
Russian	Automated	7.8%	58.3%	33.9%
Spanish	Manual	6.5%	61.5%	32.0%
Spanish	Automated	7.9%	59.9%	32.2%
French	Difference	0.0%	0.0%	0.0%
Kazakh	Difference	-0.2%	-0.8%	1.1%
Russian	Difference	0.0%	1.1%	-1.1%
Spanish	Difference	-1.4%	1.6%	-0.2%

Does automatic YouTube transcription technology change the expressed sentiment of the spoken words?

Years in US	Type	Sentiment		
		Negative	Neutral	Positive
4-8	Manual	6.9%	60.6%	32.5%
4-8	Automated	8.0%	59.2%	32.8%
8-12	Manual	5.3%	69.6%	25.2%
8-12	Automated	5.5%	70.4%	24.1%
16+	Manual	4.9%	63.6%	31.5%
16+	Automated	5.4%	62.8%	31.8%
4-8	Difference	-1.1%	1.4%	-0.3%
8-12	Difference	-0.2%	-0.8%	1.1%
16+	Difference	-0.5%	0.8%	-0.3%

Does automatic YouTube transcription technology change the expressed sentiment of the spoken words?

Sentiment Analysis results by simple comparison.

		Sentiment Changed
Transcript Accurate	No	Yes
Yes	100.0%	0.0%
No	88.8%	11.2%

In this case, we conclude that YouTube's transcription technology **change** the expressed **sentiment** of the spoken words when the transcript of the utterances is not accurate.

Does automatic YouTube transcription technology change the expressed sentiment of the spoken words?

Sentiment Analysis results by simple observation.

Meaning	Sentiment Changed	
	No	Yes
Accurate	100.0%	0.0%
Similar	95.3%	4.7%
Change	87.4%	12.6%
Unclear	87.0%	13.0%
Missing Words	76.2%	23.8%
Opposite	75.0%	25.0%
Offensive	44.4%	55.6%

Does automatic YouTube transcription technology change the expressed sentiment of the spoken words?

Sentiment Analysis results by Cosine Comparison.

Similarity	Sentiment Changed	
	No	Yes
Accurate	100.0%	0.0%
Very Similar	97.5%	2.5%
Similar	91.1%	8.9%
Very Different	83.3%	16.7%
Different	74.3%	25.7%

Does automatic YouTube transcription technology change the expressed sentiment of the spoken words?

Sentiment Analysis results by The Flesch Reading Ease score labeling.

		Sentiment Changed	
Difficulty Change		No	Yes
No	98.6%	1.4%	
Yes	87.8%	12.2%	

Does automatic YouTube transcription technology change the expressed sentiment of the spoken words?

Sentiment Analysis results by The Dale–Chall readability score labeling.

Difficulty Change	Sentiment Changed	
	No	Yes
No	97.9%	2.1%
Yes	86.5%	13.5%

Does automatic YouTube transcription technology change the expressed sentiment of the spoken words?

Sentiment Analysis results for the selected videos.

		Sentiment Changed	
#	YouTube ID	No	Yes
1	a16z2mE_69s	99.2%	0.8%
2	uXJwQNEXuug	98.7%	1.3%
3	7jff8Y9zvV0	98.6%	1.4%
4	KBT2F0hdqQc	97.8%	2.2%
5	xtNBDzgWt2I	95.9%	4.1%
6	yI3k91L1hng	95.6%	4.4%
Average		97.6%	2.4%

Question 2: Does automatic YouTube transcription technology change the expressed sentiment of the spoken words?

Based on the above analysis, we could conclude as follows:

We reject the Null hypothesis and accept the Alternative hypothesis.

Alternative: The sentiment does change.

Thank you

Feedback or Questions?

Contact

-Dipika Shrestha

dipika.shrestha@baruchmail.cuny.edu

-Duubar Villalobos Jimenez

Duubar.Villalobos.Jimenez@spsmail.cuny.edu



Summary

		Transcript Accurate			
Simple comparison	Yes	No	Sentiment Change		
	79.5%	20.5%	Accurate	Yes	No
	Yes	0.0%	Yes	0.0%	100.0%
	No	11.2%	No	88.8%	11.2%

Summary

Visual comparison	Transcript Accurate		
	Yes	No	
	79.5%	20.5%	Sentiment Change
		Meaning	Yes
		Accurate	0.0%
		Similar	4.7%
		Change	12.6%
		Unclear	13.0%
		Missing Words	23.8%
		Opposite	25.0%
		Offensive	55.6%
			44.4%

Summary

Cosine comparison	Transcript Accurate		Sentiment Change
	Yes	No	
	80.1%	19.9%	
Similarity		Similarity	Yes
		Accurate	0.0%
		Very Similar	2.5%
		Similar	8.9%
		Very Different	16.7%
		Different	25.7%
		No	100.0%
			97.5%
			91.1%
			83.3%
			74.3%

Summary

	Transcript Accurate		
Number of words	Yes	No	
	90.4%	9.6%	Sentiment Change
Word counts	Word counts	Yes	No
	Accurate	1.3%	98.7%
	Decreased	5.9%	94.1%
	Increased	17.2%	82.8%

Summary

	Transcript Accurate		
Flesch Reading Ease	Yes	No	
	91.4%	8.6%	Sentiment Change
Score Change	Yes	No	
	No	1.4%	98.6%
	Yes	12.2%	87.8%

Summary

Dale–Chall readability	Transcript Accurate			
	Yes	No	Sentiment Change	
	98.1%	1.9%	Level Change	Yes
			No	2.1%
			Yes	13.5%

Summary

Description	Transcript Accurate			Sentiment Change	
Gender	Yes	No		Yes	No
Female	79.7%	20.3%		2.4%	97.6%
Male	79.2%	20.8%		2.2%	97.8%

Summary

Description	Transcript Accurate			Sentiment Change	
	Yes	No		Yes	No
Kazakh	82.2%	17.8%		1.4%	98.6%
Spanish	81.6%	18.4%		3.0%	97.0%
Russian	77.2%	22.8%		2.2%	97.8%
French	66.4%	33.6%		1.3%	98.7%

Summary

Description	Transcript Accurate			Sentiment Change	
	Yes	No		Yes	No
16+	90.5%	9.5%		0.8%	99.2%
8-12	82.2%	17.8%		1.4%	98.6%
4-8	74.3%	25.7%		3.2%	96.8%