Comment sentiment associations with linguistic features of educational video content

Lenart Motnikar
Faulty of Education,
University of Ljubljana,
Slovenia
lenart.motnikar@gmail.com

ABSTRACT

As people spend an increasing amount of time on social media, researchers are motivated to study the newly emerging communities and the interpersonal relationships within them. This study examines one such relationship, namely between the audiences of educational videos and its presenters. A dataset of sentiment-labeled comments from TEDx and TED-Ed YouTube videos was extended to include linguistic features of video content. It was revealed that the features significantly varied between animations and presentations, and in the latter case, the speakers' genders. A correlation analysis showed that sentiment depended on a number of features, where the most notable observations included associations between negative sentiment and negative emotional content, and between positive sentiment and (first person singular) personal pronouns.

Keywords

TED Talk, YouTube, comments, LIWC, sentiment analysis

1. INTRODUCTION

As social media platforms like YouTube became so prevalent in our daily lives [1], offering opportunities for interaction with wide audiences, educators and scholars are often motivated to participate with their own content [2]. The interactions on these platforms, however, are not always civil, and are frequently characterized by unwanted behavior [3]. In order to foster better online communities, recent research has focused on understanding contentious individuals and studying the effects of various design and moderation measures [4, 5]. Research has also suggested that individuals sharing content online mind the potential reactions of their audience and, motivated by not being badly perceived, adapt their behavior accordingly [6]. Little research has, however, been done on the specifics of these behavioral measures, or their effectiveness in terms of influencing the audience. A study examining vloggers, for example, found that they use a distinctive viewer-oriented speaking style, often characterized by explicit or implicit encouragements of desired behaviors (e.g. commenting, subscribing) [7]. Building upon these observations, this study, using a quantitative approach, explores potential ways for content creators to influence their audiences' behavior. By applying methods and theory previously unused in such a setting it explores associations between the language used in educational videos and the sentiments expressed in the comments, opening opportunities for future inquiries into the dynamics between individuals and large online audiences.

1.1 Lexical inquiry and word count (LIWC)

For linguistic analysis, the Linguistic Inquiry and Word Count (LIWC) program was used [8]. LIWC is a text analysis software which, using a predefined dictionary, measures the frequency of words across a variety of categories relating to grammar and psychological processes, and rates the text's manifestations of four underlying psychological dimensions - analytical thinking, authenticity, clout (expression of social status) and emotional tone. In the last two decades LIWC has become the most popular tool for automated text analysis in socio-psychological studies, as it helped illuminate how a person's choice of words reflects their mental states (for a review, see [9]). One of the most notable revelations stemming from LIWC research was the importance of function words in human social dynamics. Personal pronouns were shown to be particularly revealing as they, by conveying information about attentional focus, let us know how people relate to themselves and others, disclosing details ranging from one's social status to their emotional states.

1.2 Sentiment analysis

The research field of sentiment analysis or opinion mining aims to capture the public's feelings about various entities, be it products, people or ideas [10]. Due to the availability of a wide variety of tools and data, a significant portion of the field deals with the analysis of texts gathered from social media. The sentiment in this study was assessed with the SentiStrength [11] tool, which, using a lexical approach, identifies sentiment-related tokens and scores social web texts on a dual positive and negative scale.

1.2.1 Comment sentiment on TED YouTube videos.

The current study builds upon a dataset compiled by Veletsianos et al. [12]. The authors collected English-speaking educational YouTube videos posted on TEDx Talks and TED-Ed channels and investigated how presenter gender, video format and comment threading effect the sentiment expressed in the comments. They observed that presentations with female presenters, relative to those with male, exhibited greater polarity in positive and negative sentiment, and that animated videos were more neutral than presentations. These differences not only held for comments directed toward the video, but replies to the comments as well. The study also examined the relationship between sentiment and video topic by analyzing description and title keywords, and found that some topics exhibit more positive (e.g. beauty) and others more negative (e.g. cancer) sentiment.

2. METHOD

A modified »YouTube TED Talk Comment Sentiment Data« dataset [13] was used. The dataset contained positive (1 to 5) and negative (-1 to -5) sentiment scores of comments from 665 videos, information about whether the video was an animation or a presentation, and in the latter case, the information about presenter's gender. In this study, the dataset was extended to include LIWC scores of video subtitles. The subtitles were assessed using the LIWC2015 dictionary, scoring each subtitle track across 93 linguistic categories. As not every subtitle track featured all of the categories, occurrences where the score of a category equaled zero were ignored in the analysis.

Videos that did not have English subtitles available were excluded from the dataset (n = 57), reducing the sample size of videos and comments by 8.6% and 6.7%, respectively. Additionally, the analysis only included first-level comments representing 50% of the sample. Because comments on YouTube come in two general forms, posted directly under the video or as a reply to another comment, this study followed the interpretation that replies are directed towards other comments rather than the video itself.

Table 1: Descriptive statistics of videos and comments

Format/ gender	Videos *n*	Comment *n*	Comment *n* M	Comment *n* SD
Female	66	38572	584.42	782.40
Male	130	89642	689.55	1575.45
Animation	412	197385	479.09	873.51
	608	325599	535.52	1056.97

While the removal of videos minimally affected the reported differences between video formats and presenters, the exclusion of replies significantly increased both positive and negative average sentiment. The general trend that videos with female presenters exhibited greater polarity and that animations were the most neutral, however, still remained.

Table 2: Sentiment differences of comments by format and gender

Format/ gender	Positivity		Negativity	
	M	SD	M	SD
Female Speaker	2.16	0.98	-1.72	1.06
Male Speaker	1.96	0.95	-1.63	0.98
Animation	1.60	0.78	-1.62	0.94

Each video then received two aggregated sentiment scores by separately averaging the positive and negative sentiment of all its comments.

Table 3: Differences of aggregated sentiment scores by format and gender

Format/ gender	Positivity M	Negativity SD	Positivity M	Negativity SD
Female	2.23	0.21	-1.71	0.26
Male	2.02	0.25	-1.61	0.31
Animation	1.63	0.18	-1.58	0.27

This further increased the average positivity and negativity, reflecting the otherwise statistically insignificant trend that sentiment averages decrease as the number of comments on a video increases.

3. RESULTS

The data was tested for differences in LIWC scores between video formats and presenter gender. The Wilcoxon rank sum test revealed that the video formats significantly (p < 0.01) differed in 70 and genders in 26 of the 93 LIWC2015 categories. Differences in summary variables and language metrics showed that animations were more analytical and used longer words and sentences, whereas the presentations had a greater word count and exhibited more clout, authenticity and emotional tone. Similar differences could be observed between the genders, where videos with male presenters exhibited greater analytical thinking and those with female presenters more authenticity. Numerous differences in categories relating to style and content were also observed, a selection of which is shown in Table 4.

Table 4: Word category prevalence by format and gender

	Animation	–	Talk
Female	anxiety, body	*negative emotion, sadness, female referents, feeling, health*	pronouns, 1st person singular, regular verbs, conjunctions, negations, *affect, certainty*
–	prepositions, adjectives, comparatives, *death, anger, seeing, sexuality, ingesting, relativity, space, religion, friends, swearing*	3rd person, *tentativeness, differentiation, assent, home*	1st person plural, 2nd person, auxiliary verbs, adverbs, interrogatives, *positive emotion, social processes, insight, discrepancies, hearing, time orientation, drives, motion, work*
Male	articles, quantifiers, numbers	*money*	*informal speech, leisure*

Note. Content categories are presented in italic

Across the five (sub)samples, correlating positive and negative aggregated sentiments with LIWC scores revealed 302 significant (p < 0.05) correlations, of which 83 were stronger than |r| = 0.3. Because the correlations covered a large majority of the LIWC2015 categories, only the categories exhibiting correlations above |r| = 0.3 in at least two sentiment-sample pairings are reported in Table 5. In the sample containing all videos, correlations with three out of four summary variables could be observed. Positive sentiment was positively associated with authenticity and inversely with analytic thinking, while emotional tone positively correlated with both positive and negative sentiment (note that negative sentiment was represented by a value between -1 and -5). The association between emotional tone and negative sentiment, however, remained in all samples. Significant correlations with language metrics could also be observed. The percentage of words longer than six letters exhibited a general inverse correlation with negative sentiment, and in the case of videos with female speakers, positive sentiment as well.

was related more to content whereas positive sentiment to style and grammar, especially (first person singular) personal pronouns.

Table 5: Correlations between aggregated sentiments and LIWC categories

		Positive sentiment					Negative sentiment				
		All videos	Presentations			Animated videos	All videos	Presentations			Animated videos
	LIWC categories		All	Male	Female			All	Male	Female	
Summary variables	**Analytic thinking**	-.62***	-.24***	-.08	-.33**	-.08	.06	-.19**	-.23**	-.24*	.08
	Authenticity	.28***	.31***	.16	.45***	-.12*	.04	.01	.05	.06	.11*
	Tone	.27***	-.06	-.05	.02	.09	.28***	.44***	.40***	.49***	.30***
Language metrics	**Words >6 letters**	-.40***	-.17*	-.05	-.42***	.05	-.14***	-.23**	-.22*	-.30*	-.24***
	Dictionary words	.56***	.37***	.29**	.36**	.00	-.13**	.06	.09	.14	-.14**
Style and grammar	**Function words**	.57***	.32***	.20*	.35**	-.05	-.00	.16*	.15	.35**	.07
	Total pronouns	.66***	.34***	.20*	.44***	.11*	-.05	.16*	.16	.34**	-.01
	Personal pronouns	.67***	.45***	.29***	.56***	.12*	-.07	.11	.13	.24	-.06
	1st person singular	.63***	.40***	.19*	.60***	.23*	-.09	-.01	.04	.04	-.04
	Articles	-.46***	-.28***	-.09	-.40***	-.10	.15***	-.04	-.13	.01	.18***
	Regular verbs	.55***	.16*	-.02	.36**	.00	-.00	.19**	.22*	.25*	.05
	Quantifiers	-.23***	-.31***	-.17	-.42***	-.03	.15***	.02	-.01	-.06	.17***
Content	**Affect words**	.34***	.18*	.11	.18	.24***	-.39***	-.17*	-.10	-.30*	-.47***
	Negative emotion	.10*	.21**	.16	.12	.14**	-.53***	-.44***	-.35***	-.61***	-.58***
	Anger	-.08	.029	.07	-.01	.08	-.28***	-.19*	-.17	-.35**	-.35***
	Sadness	.06	.20**	.15	.16	.15*	-.36***	-.42***	-.29**	-.66***	-.36***
	Biological processes	-.04	.16*	.14	-.01	.04	-.20***	-.33***	-.33***	-.27*	-.19***
	Health	-.03	.15*	.17	-.03	-.05	-.34***	-.36***	-.37***	-.31*	-.35***
	Past focus	.35***	.31***	.23**	.41***	.05	-.07	-.05	-.06	.05	-.03
	Death	-.25***	.01	.07	.11	-.08	-.19***	-.44***	-.53***	-.22	-.18**

Note. For visualization purposes, the significant correlations are colored with a grey-to-black gradient, representing their strength.
*p<0.05, **p<0.01, ***p<0.001

The presentation subsamples also exhibited correlations between positive sentiment and the percentage of words caught by the dictionary. Regarding style and grammar, positive sentiment was associated with function words, particularly (first person singular) personal pronouns. In the female presenter subsample, associations with positive sentiment were observed between regular verbs, quantifiers and articles, while negative sentiment positively correlated with the percentages of function words, pronouns and verbs. Contentwise, a majority of significant correlations was with negative sentiment, most of which were inverse and related to negative affective processes like anger and sadness, or concerns like health and death. Positive sentiment exhibited fewer and weaker content related associations, except in the case of presentations expressing a greater focus on the past.

4. DISCUSSION AND CONCLUSIONS

An important caveat before delving into interpretations is that the videos included in this study had different audiences. In fact, more than 90% of commenters only commented on one or two videos, as different topics and formats invite different profiles of people. While this does not change the overall experience for the comment reader, it should be noted that the results would likely differ with a constant or randomized audience.

Nevertheless, the analysis returned some interesting results. A general pattern was observed, showing that negative sentiment

The association with content is not that surprising as it can at least partially be attributed to video topic, as has been reported in the original study. Additionally, the emotion tokens SentiStrength and LIWC used for analysis overlap to some degree. This explanation also holds for the association with emotional tone, as it merely combines the words from emotion categories.

From a socio-psychological perspective the association between positive sentiment and style is more intriguing. While the importance of function words in human social dynamics is well documented, it has so far been limited to studies of smaller groups of people, like couples or teams [14]. This is the first time that a reaction of a larger audience has been associated with a speaker's pronoun use. What this observation means in terms of social psychology is less clear. It should be noted that sentiment, as it was assessed here, is a theoretically unsound construct and a particularly crude measure of emotion (for a critique, see [15]). It only measures emotion on a dual positive/negative scale, and does not differentiate between the nuances of human emotional experience and expression. For example, on a video discussing suicide, a comment personally attacking the speaker might receive the same sentiment score as one where the commenter shares their experience with depression. The motivations for these behaviors are vastly different, as are the readers' reactions. For this reason, one should be careful when interpreting sentiment and take into account the variety of factors contributing to its manifestation.

These limitations considered, the observed associations still encompass some psycholinguistic information about speaker-audience interaction, and call for a deeper inquiry into the topic.

A question that still remains is why the sentiments were differently associated with content and style in the first place. The observation may reveal information about the social aspects of emotion processing. If we only focus on the clearest examples, negative emotion and first person singular, a general explanation could be that the former evokes more sympathy whereas the latter, which entails more self-focus, evokes cheer.

Results also suggest a relationship between sentiment and language metrics, specifically the percentages of words longer than six letters and that of words caught by the dictionary. As the dictionary encompasses some 6000 words and stems in common use, this observation might indicate a relation to the simplicity or commonality of language used in the video. This could be interpreted in a way that people prefer simpler language, or that the use of more complex language encourages more sentiment-neutral conversation.

Lastly, the results shed light on the originally reported gender and format differences in sentiment. The groups varied in content and style, which might entail that some of the primarily observed discrepancies were due to the differences in topics the content makers chose, or the ways in which they were expressed. This considered, this explanation likely accounts only for a portion of the difference as there was still notable variation in correlation strengths between the samples, with the female subsample exhibiting the strongest correlations in most categories. For example, in the female subsample, but not the other two, positive sentiment exhibited an inverse correlation with articles and quantifiers. While this could still be due to the chosen topics, or some other confounding factor, another explanation for the phenomenon may lay in the fact that these words are mostly used in conjunction with concrete nouns, indicating a relation to concreteness or abstractness of a presentation. Why this relation would be only specific to female presenters, remains an open question.

Taken together, this study was mostly exploratory in nature, providing more avenues for research than solid findings. In order to thoroughly answer the questions emerged, future research should use more sound measures of behavior and mental states, as well as look into different communities and platforms where similar interpersonal interactions take place.

5. REFERENCES

[1] Smith A. and Anderson M. 2018. Social Media Use in 2018. *Pew Research Center* (March, 2018).

[2] Gruzd, A., Haythornthwaite, C., Paulin, D., Gilbert, S. and del Valle, M. 2016. Uses and Gratifications factors for social media use in teaching: Instructors’ perspectives. *New Media & Society*, 20, 2 (February, 2018), 475-494.

[3] Duggan M. 2017. Online Harassment 2017. *Pew Research Center* (July, 2017).

[4] Cheng, J., Bernstein, M., Danescu-Niculescu-Mizil, C. and Leskovec, J. 2017. Anyone Can Become a Troll. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing* (Portland, Oregon, February 25 - March 01, 2017). CSCW '17. ACM, New York, NY, 1217-1230.

[5] Cheng, J., Dansecu-Nisulescu-Mizil, C. and Leskovec, J. 2015. Antisocial behavior in online discussion communities. In *Ninth International AAAI Conference on Web and Social Media* (Oxford, UK, May 26 - 29, 2015). ICWSM '15. AAAI, Palo Alto, California, 61-70.

[6] Barasch, A. and Berger, J. 2014. Broadcasting and Narrowcasting: How Audience Size Affects What People Share. *Journal of Marketing Research* 51, 3 (June, 2014), 286-299.

[7] Frobenius, M. 2014. Audience design in monologues: How vloggers involve their viewers. *Journal of Pragmatics* 72, (October, 2014), 59-72.

[8] Pennebaker, J.W., Boyd, R.L., Jordan, K. and Blackburn, K. 2015. *The development and psychometric properties of LIWC2015*. University of Texas at Austin, Austin, TX.

[9] Tausczik, Y. R. and Pennebaker, J. W. 2010. The psychological meaning of words: LIWC and computerized text analysis methods. *Journal of language and social psychology*, 29, 1 (March, 2010), 24-54.

[10] Cambria, E. 2016. Affective Computing and Sentiment Analysis. *IEEE Intelligent Systems* 31, 2 (March, 2016), 102-107.

[11] Thelwall, M. 2017. Heart and soul: Sentiment strength detection in the social web with SentiStrength. In *Cyberemotions: Collective emotions in cyberspace*, J. Holyst, ED. Springer, Berlin, 119-134.

[12] Veletsianos, G., Kimmons, R., Larsen, R., Dousay, T. and Lowenthal, P. 2018. Public comment sentiment on educational videos: Understanding the effects of presenter gender, video format, threading, and moderation on YouTube TED talk comments. *PLOS ONE* 13, 6 (2018), e0197331.

[13] Kimmons, R. 2019. YouTube TED Talk Comment Sentiment Data. *BYU ScholarsArchive*, https://scholarsarchive.byu.edu/data/3.

[14] Kacewicz, E., Pennebaker, J., Davis, M., Jeon, M. and Graesser, A. 2013. Pronoun Use Reflects Standings in Social Hierarchies. *Journal of Language and Social Psychology* 33, 2 (March, 2014), 125-143.

[15] Paolillo, J. 2019. Against 'Sentiment'. In *Proceedings of the 10th International Conference on Social Media and Society* (Toronto, ON, Canada, July 19 - 21. 2019). SMSociety '17. ACM, New York, NY, 41-48.