Dimensional Word Clouds to Show Cross-Cultural Differences in News Videos

Luvena Huo (lh2933)
Supervisor: Dr. John R. Kender
Columbia University

1. INTRODUCTION

With the continuous growth of globalization, any event could be reported by multiple news sources across different affinity groups in the world. This then results in potential differences in opinions from different affinity groups on the same topic or piece of news. We are studying the differences and similarities in the portrayals of and reactions to the same event across different affinity groups, specifically through video news.

In this report, we wanted to continue to explore news relevant to the initial COVID-19 outbreak to find discussions and differences between separate affinity groups [5]; more specifically, the potential similarities and differences in COVID-19 news in the U.S. versus in China. Our goal was to explore and analyze the different comments and reactions that people of different affinity groups might have when consuming similar pieces of news stories, and the potential underlying cultural variability they might expose that was previously uncaptured in our analysis.

2. PREVIOUS STUDIES

2.1 Continuing COVID-19 Outbreak Comments & Reactions

Previously, we switched to using COVID-19 news due to its large impact on a global audience [5]. We wanted to see how different affinity groups might have portrayed and responded to the same event, and we specifically focused on the initial outbreak of the virus in the respective groups (around January-February for China and around March-April for the U.S.) to provide a point of comparison. We continued our previous analysis of reactions to this news by building off of the circumplex model and explored how to capture differences that we had noticed.

2.2 Exploring Cultural Variation in Emotional Expression

Our previous research focused on finding a way to quantify and analyze the emotions that different affinity groups were experiencing and expressing online in response to the news that they consumed. We quantified emotionally charged words and phrases from comments in a user study for both English and Chinese audiences in response to news videos regarding the outbreak. We then used a circumplex model to display the averaged results for both countries, as shown in Figures 1 and 2. The circumplex model was able to pinpoint certain emotions and allow us to understand how different emotions relate to each other on a visual scale, which was in turn helpful for comparing differences between similar emotions cross-culturally.

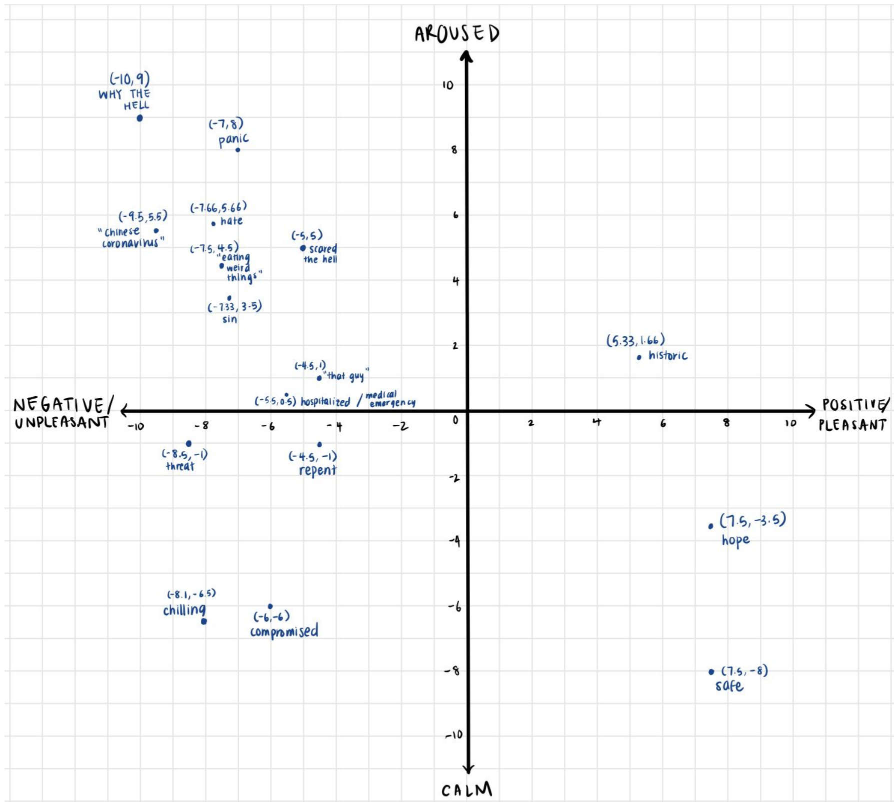

Fig. 1. U.S. Averaged Emotion Circumplex Plot[5]

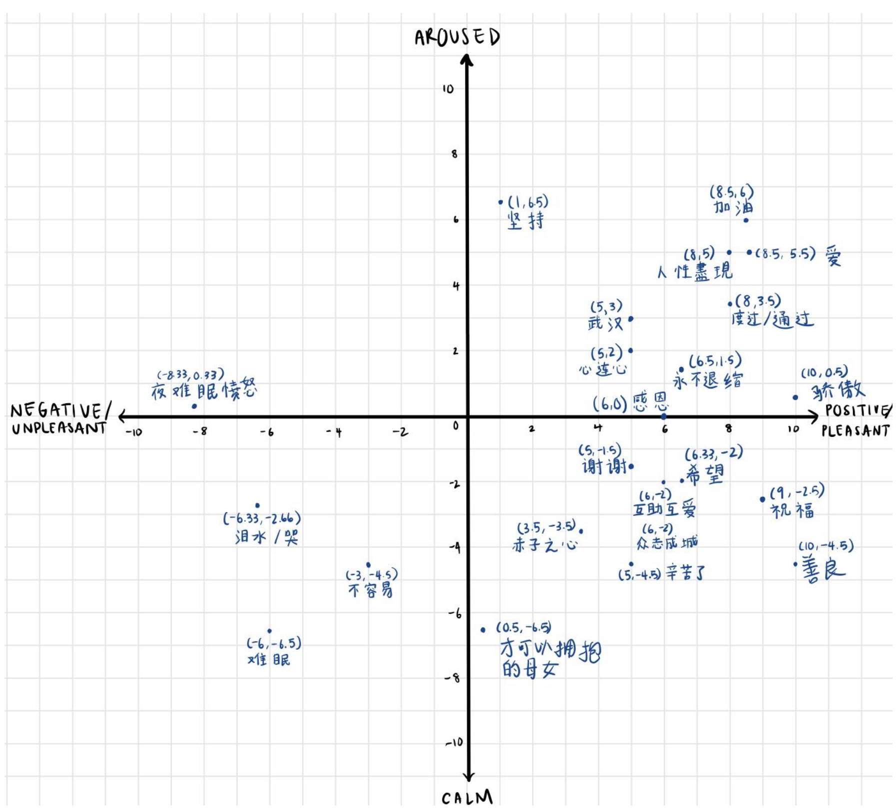

Fig. 2. China Averaged Emotion Circumplex Plot[5]

We can clearly see that American audiences had written more comments that fell on the negative end of the valence scale and Chinese audiences had written more comments that fell on the positive end.

However, there are still multiple pieces of information that we had noticed that weren't emphasized or fully captured in these circumplex graphs. This mainly included differences in how emotions were expressed in English versus Chinese. For example, more Chinese comments expressed their comments in terms of physical reactions, including "难眠", which can be translated to "unable to sleep", or "泪水 / 哭", which can be translated to "tears / crying". This is contrasted with English audiences that would directly express emotion using the word "sad" itself, as seen in the comments snippet from last semester's research report [5]. There are even some studies that mention this difference in emotional expression, citing that the Chinese language is more likely to use somatic or social words for expression (Tsai, Simenova, & Watanabe, 2004 [12]; Ning, 1995 [8]).

Another difference that is not immediately apparent is the emotional connections that were made. While American comments would bring up religion references such as "sin", Chinese comments would often mention patriotic sentiments such as "加油中国 加油武汉" ("good luck China, good luck Wuhan"), showing a difference in religious versus patriotic connections in emotional expression.

These differences, however, are not easily spotted without a deep dive into the circumplex graphs. We were curious about this potential cultural variability, and wanted to see if there was a way to both represent this and also understand more deeply if there are certain emotional expressions that one affinity group tends towards that are significantly different than that of the other group. In this report, we will focus on enhancing the circumplex model to capture additional interesting emotional expressions within the comments of those who commented under COVID-19 outbreak news, specifically using word clouds.

3. INITIAL EXPLORATIONS

3.1 Plutchik's Model for Capturing Cultural Variation

Initially, we were interested in exploring whether or not it would be useful to supplement the circumplex model with Plutchik's model (Plutchik, 1980 [9]), which depicts emotions using a wheel of 4 pairs of opposite emotions, including some variation of joy-sorrow, acceptance-disgust, fear-anger, and surprise-anticipation (see Fig. 3). These 8 emotions were considered the basic emotions and more subtle emotions could be created using different combinations of these emotions to create "dyads" (see Fig. 4). Primary dyads were generally formed when emotions next to each other on Plutchik's wheel were combined, secondary dyads formed when emotions skipping its immediate neighbor were combined, and tertiary dyads formed when skipping 2 neighbors. Primary dyads are defined as more commonly felt emotions than secondary dyads, and secondary dyads more common than tertiary dyads (Turner, 2000 [13]).

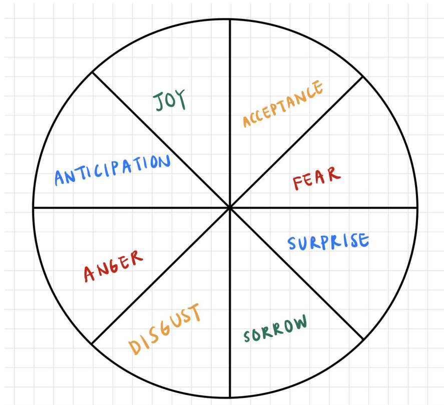

Fig. 3. Plutchik's model of emotion (Plutchik, 1980 [6]) displaying the 8 basic emotions. Pairs of opposite emotions lie opposite each other on the circular graph and are in the same color. Taken from last semester's paper[5].

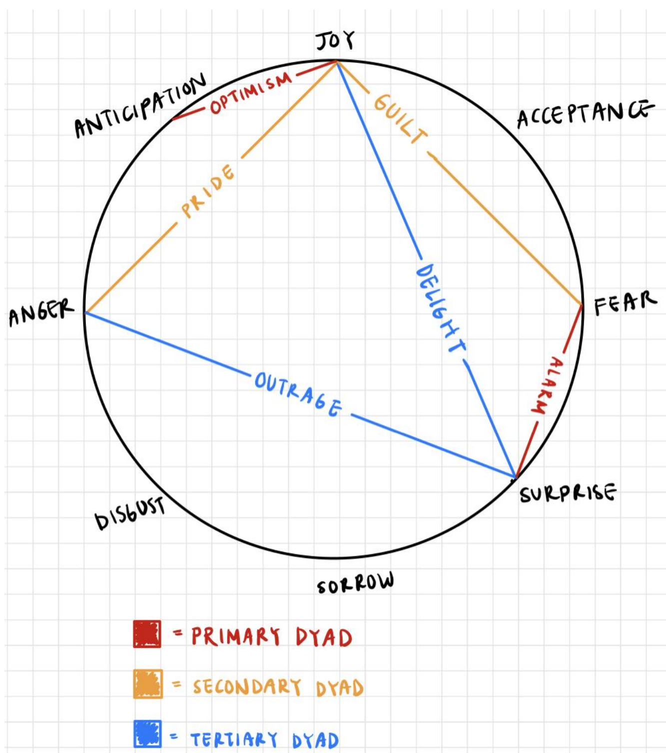

Fig. 4. Dyads formed through different combinations of basic emotions in Plutchik's model of emotion. Primary dyads are shown in red, secondary in orange, and tertiary in blue. Dyad combinations taken from Turner's analysis (Turner, 2000 [13])

We believed that since this model focuses on how emotions are interrelated, it could be helpful in understanding the nuances and differences in emotion buckets in Chinese comments versus English comments. It would also be interesting to see how Plutchik's model might interpret the emotions, and assign different mixes for similar emotions sourced from different affinity groups.

First, we wanted to know whether or not Plutchik's model was transferable to different cultures. We were able to find examples of applications of Plutchik's model cross-culturally - in a 2020 psychology study on the emotional effects of caregiving for a schizophrenia patient, emotions for American, Taiwanese, Canadian, and Japanese families were analyzed and compared using Plutchik's model (Shiraishi & Reilly, 2020 [11]). Shiraishi and Reilly explored the nuances of anxiety and worry across different cultures using the model, leading us to believe that at least in theory, Plutchik's model is applicable cross-culturally.

However, there were still a couple concerns we had about this approach in explaining cultural variability. We weren't certain that our data would always fall into these perfectly sorted and ordered buckets, and were worried that Plutchik's model might explain nuances better in theory than in practice. Additionally, not all of the combinations of basic emotions felt completely accurate - for example, fear-anger did not seem to be fitting opposites.

3.2 Third Dimension in Circumplex Model to Capture Cultural Variation

We also wanted to explore whether a third dimension in the circumplex model might be able to capture some of the differences we were seeing in Chinese emotional expression versus English emotional expression. Previously, we had discounted a third dimension because there was some debate over what the ideal third dimension would be to accurately capture emotions - some studies mentioned power or the level of control over the emotion felt (Jin & Wang, 2005 [7]), some mentioned degree of aggression (Bush, 1973 [1]), and some mentioned attention versus rejection (Schlosberg, 1954 [10]).

Instead of using these metrics, we wanted to use the additional dimension to capture differences in emotional expression in separate affinity groups. Therefore, following this thought process, we attempted to find studies representing the differences we had noticed on two ends of a continuum. One possibility we explored was expressing emotion using somatic phrases, which had been mentioned previously to be characteristic of Chinese language, versus the mind, which was characteristic of English language. However, studies (Chan, Ho, & Chow, 2002 [2]) actually indicated that there is a strong belief in Chinese culture that the mind and body are interrelated, rather than placing an emphasis on emotional expression on the body only. In fact, Chan, Ho, and Chow mention that health is "perceived as a harmonious equilibrium that exists between… external sources of harm and the seven emotions (joy, sorrow, angger, worry, panic, anxiety, and fear)" - this could actually explain why many emotions in Chinese comments were expressed with somatic phrases.

Another possibility for a potential scale for an additional dimension was the difference that we saw in Chinese comments referencing patriotism versus English comments referencing religion. While this religion vs patriotism scale could be interpreted to be similar to the individualism vs collectivism scale (Gudykunst, 1997 [4]), which is widely considered to be a dimension of cultural variability, these differences were only present in a handful of phrases on each graph.

This brought into question whether or not it is sound to create a dimension on the graph to this information, as an additional dimension would be more helpful if it were contributing to the entire plot rather than just a few points on the graph. Additionally, we had concerns about whether or not it was possible to create a direct and definite connection between language and individualism vs collectivism. Therefore, neither of the potential observed differences were well suited to be on separate ends of a third dimension.

4. DIMENSIONAL WORD CLOUDS

We then decided to try a different approach where we went back to the root of the problem to break down what we are trying to capture. We knew that we were seeing differences not only in the emotions that are currently being expressed, which we can see through the circumplex graphs, but also the way in which the emotions are expressed. What we really want is to visualize or show the cultural variance and differences in expression that are currently missing in the circumplex graphs. Additionally, when capturing this information, we want to make sure that we are not taking away from the information expressed by the circumplex model, but rather building upon it. Therefore, rather than trying to quantify these cultural variances on a different axis or scale, we believed that it would make more sense to find a way to display the differences so that people could view and understand the differences more clearly themselves, while also still capturing the emotional differences that we are seeing from the circumplex model.

With these considerations in mind, we decided to create a multi-dimensional word cloud to represent and display cultural variability. Rather than replace the circumplex model, we wanted to use the circumplex model as a foundation for the word cloud, while also using the word cloud as a side-by-side visual aid to help enhance the understanding of cultural emotional differences at a glance.

4.1 Capturing the 2-Dimensional Circumplex Model Using Color

When building our word cloud, one of the first things we considered was how to create a point of connection between the word cloud and the circumplex model such that we were not losing any information from the circumplex model in our word cloud. We decided to use the 2-dimensional circumplex model (see Fig. 1 and Fig. 2) as the base system and reinforce it with hue and saturation as our two axes. In simpler terms, we want to capture a 2 dimensional coordinate, or the circumplex model, in the word cloud while using color to do so.

This was done by using hue to capture the x-axis, or the valence, and saturation to capture the y-axis, or the degree of arousal or intensity. This created a two dimensional gradient across the coordinate axis, with the color in each point of the coordinate plane being the color of the phrase that lay there within the word cloud. The colors of the words in the word cloud would then be able to capture the original position of that word or phrase within the circumplex model, therefore more accurately conveying where that emotion is in terms of valence and arousal.

4.1.1 Hue

We used hue to capture valence, or the x-axis in the circumplex model, by assigning one end to be pure blue (negative side) and the other end to be pure red (positive side). We then created a gradient that transitioned pure blue into different shades of purple by mixing red in, and ultimately into pure red from one side of the axis to the other.

We did this by assigning RGB (Red, Green, Blue) coordinates to each point, with the leftmost end of the coordinate system as (0, 0, 255), which is pure blue, and the rightmost end of the coordinate system as (255, 0, 0), which is pure red. We then began mixing the red and blue values across the coordinate system gradually by calculating the amount of red and blue that should be at each point on the x-axis (see Fig. 5).

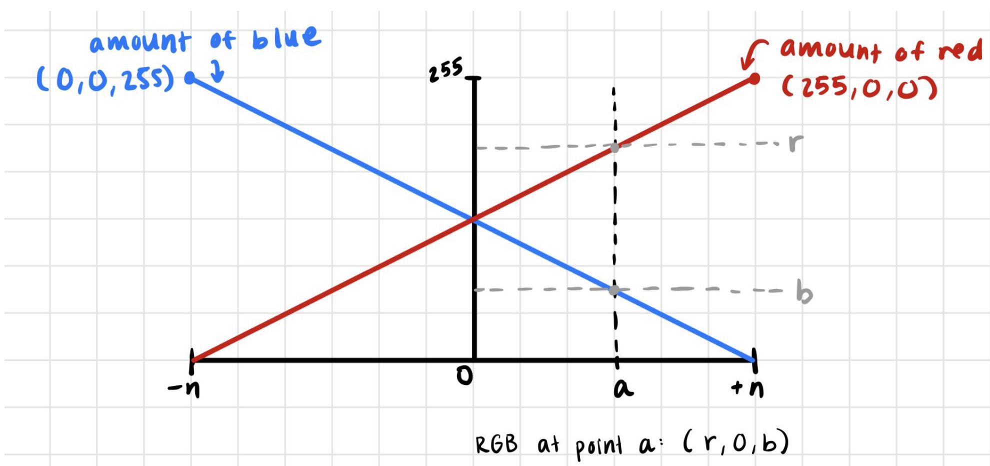

Fig. 5. Calculation for each RGB value at 100% saturation. Represents the valence value on the circumplex model, or the x-axis.

The hue of a specific section along the x-axis would then be able to indicate how positive or negative the emotional connotation of a word is. Pure blue would be the most negative and pure red would be the most positive, while a purple leaning slightly towards blue would be slightly negative.

4.1.2 Saturation

We then used saturation to capture the level of arousal or intensity of the emotion, or the y-axis in the circumplex model. The colors with 100% saturation would represent the highest intensity emotion felt, while lower saturation would represent lower intensity emotions. To do this, we simply started at 100% saturation at the highest arousal on the graph, and gradually decreased it in even increments. This created a gradient along the y-axis that represented the level of saturation at every point.

To calculate saturation, we simply converted the RGB coordinate number, which we have from calculating the hue, to a Hue Saturation Value (HSV) number. The pure mix of red and blue in the original RGB coordinate would always be 100% saturation, and we would slowly decrease the percentage saturation within the HSV value to 5% at the lowest saturation on the coordinate axis. We then converted the new HSV value back into an RGB value to input it into our coordinate axis.

4.1.3 Dead Zone

We then put the hue and saturation together over the coordinate axis to represent valence and level of arousal respectively (see Fig. 6).

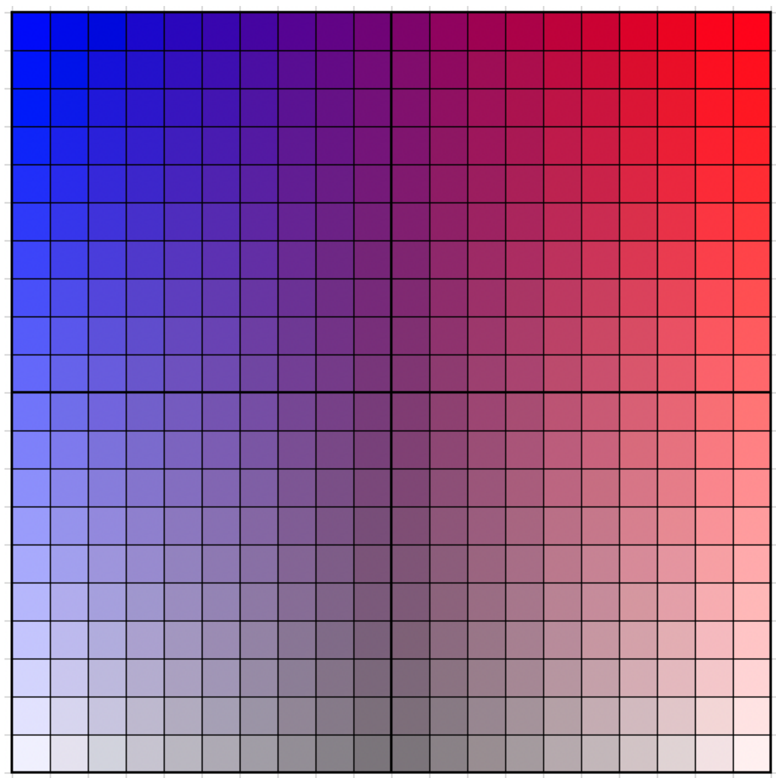

Fig. 6. Original colored graph overlaying the circumplex model, with hue and saturation representing valence and arousal, respectively.

While the two-dimensional gradient seems to look as we had expected, we can notice grayness in the bottom middle of the graph instead of the lighter purple that we were expecting. This made it difficult to recognize the hue of the shades that have grayness in them. To combat this, we removed the bottom four horizontal lines that contained the grayness, or the “dead zone”, and redistributed the rest of the saturation. Instead of decreasing saturation to 5%, we would decrease it to 25% at the lowest. We then recalculated the colors on the graph.

4.1.3 Putting It Together

After removing the dead zone from our hue and saturation graph, we are left with the final product (see Fig. 7). From this graph, we can generally map the color to the location of the word within the circumplex model.

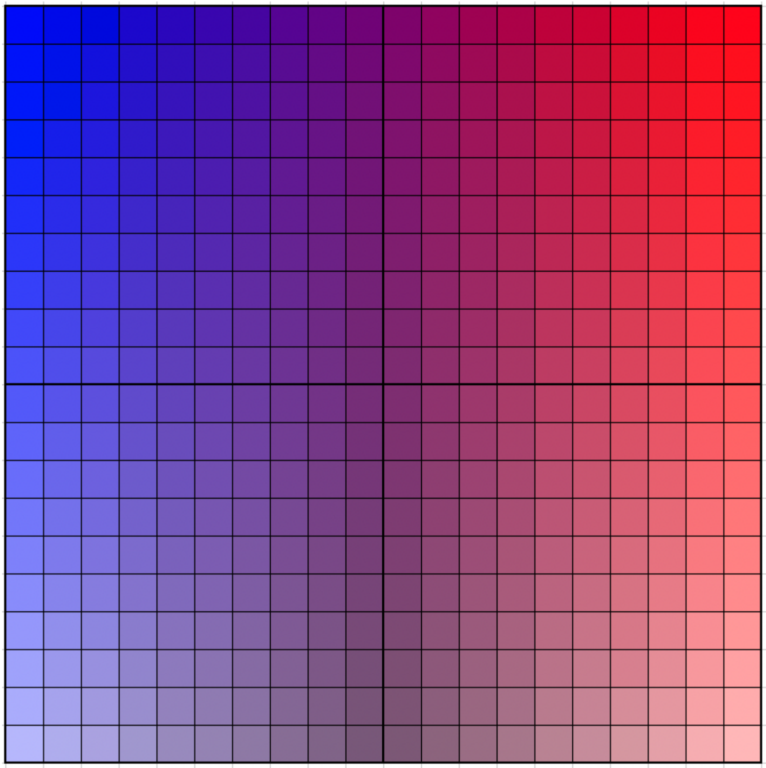

Fig. 7. Final colored graph overlaying the circumplex model, with hue and saturation representing valence and arousal, respectively.

4.2 Word Sizes Based on Occurrence

A variable that was not captured in the circumplex model was the frequency of the emotions that were expressed in the different comments, or basically how many people had felt the same way. A word cloud gives us a chance to relay that through the word sizes within the cloud.

4.2.1 Font Scaling Factor

Yang, J. Li, Lu, Chen, Xhang, and Y. Li conducted a study regarding the impact of the font scale on the semantic expression of a word cloud, and how it influenced the audience viewing the word cloud (Yang, Li, Lu, 2020 [16]). The equation presented was an exponential function as shown below:

$$F = nf^{x}$$

where F is the word's font size, n is the proportionality constant based on how large the entire word cloud is, f is the frequency of the word, and x describes the scaling factor, or how much of an impact we want the frequency to have on the word.

Based on their study's final analysis across audience response accuracy, time, and confidence, they were able to conclude that a scaling factor of around 1 or slightly above would be reasonable. Therefore, in our paper, we will generally use a scaling factor of 1, or a linear function for our word sizes.

4.2.2 Considerations

It is worth noting that there was one word - "加油" (good luck / you got this) - in the Chinese comments under COVID-outbreak videos that were repeated over 40 times, as it was in many comments. A linear scale for this word would not make much sense, as the word would be so much larger than the other words in the word cloud that we would not be able to see the full picture. Therefore, the scaling accuracy on this phrase specifically is not completely accurate; however, we will still make sure that it is significantly larger than its neighbors.

4.3 Semantic Positioning

By capturing a word's position within the two-dimensional circumplex model using color, we were essentially able to free up the positioning of words within our word cloud. We wanted to use the positioning of words within the word cloud to add semantic value by positioning similar or relevant words clustered together.

This would create essentially what we can call a "storytelling cloud" (Wang et al., 2018 [15]), where different clusters within clouds tell stories, which would allow us to create a word cloud that is more representative of the main emotions that people from each affinity group are experiencing from the same piece of news. There are some studies that create these clusters through trees (Gambette & Véronis, 2010 [3]), and some that just simply use spatial proximity to define the clusters. One way to generate relationships between different words is to calculate word relevancy with one other with co-occurrence matrices and similarity vectors - this has been explored by multiple studies (Cui, Wu, Liu, Wei, Zhou, & Qu, 2010 [14] ; Xu, Tao, & Lin, 2016 [6]).

4.4 Superimposing Chinese and English Word Clouds

Finally, we explored whether it made sense to superimpose the Chinese word cloud and the English word cloud such that they were one cohesive whole, or if the word clouds would be more clear when presented separately.

Logically, it is possible to superimpose the two word clouds without too much of a clash: they both use the same linear occurrence scale, and Chinese characters and English words are different enough that we can immediately differentiate between the two on a screen. The benefit of superimposing the two word clouds would be that we are more clearly able to immediately compare and contrast the emotional differences at a glance. Additionally, if we position the words together semantically, we can better understand the nuances between the emotions felt between the two languages. For example, “hope” was a word that was commonly used in both English and Chinese comments. When semantically positioning them within the same word cloud, we can have a better understanding of what “color” or circumplex position they have, and how much they differ.

However, on the flip side, the benefit of having separate word clouds for the English emotional comments and Chinese emotional comments is that before making comparisons, we would be able to see the emotions emphasized by each language more clearly. Word clouds can also easily be placed side by side for comparison, and do not necessarily need to be mixed. To know definitively what would generate the most clarity in understanding would require a user test of both to be certain.

5. WORD CLOUD RESULTS USING COVID-OUTBREAK NEWS

5.1 English and Chinese Word Clouds

We used the data from and built off of the previous analysis done on COVID-19 outbreak news, where we generated a circumplex model to understand the similarities and differences in reactions of the two affinity groups to the news of the COVID-19 outbreak. We generated word clouds using EdWordle software (Wang et al., 2018 [15]). We converted the RGB values of the word cloud into HEX values that we then used in EdWordle, adjusted the word sizes, and adjusted the positioning of the words in the cloud to have semantic value.

Below is the English word cloud:

Fig. 8. English Word Cloud

We then did the same thing to the Chinese comment data, and generated a Chinese word cloud:

Fig. 9. Chinese Word Cloud

As we can see, there are several apparent differences between the English and Chinese word clouds when compared separately. We can immediately notice the differences that we had seen in the circumplex model - the colors in the Chinese word cloud tend towards red, while the words in the English word cloud tend towards blue, which indicate that valence-wise, Chinese comments are more often positive and English comments are more often negative.

The size of the words show which emotional words had been repeated in multiple comments in either group, with “加油” (or good luck) taking the clear lead in the Chinese word cloud, and “hope” taking the lead in the English word cloud. This information adds more nuance to the graphs that we are seeing - while most English words are negative, we can see that the most frequent word actually has a positive connotation.

Additionally, from the clustering, we are able to more clearly see different “subtopics” within each affinity group’s commentary. For example, in the English word cloud, we can see that there is a comments cluster regarding the general panic and fear around the pandemic, a cluster with more hate-focused comments, and a cluster with religious connotations. Within the Chinese word cloud, we have a cluster representing hope and perseverance on the left, love and gratitude in the bottom middle, pride and cheering in the top and right, and sadness and insomnia on the bottom right.

5.2 Superimposed Word Cloud

We also juxtaposed the two word clouds into the same word cloud with semantic positioning, as seen below:

Fig. 10. Superimposed English and Chinese Word Cloud

The superimposed wordle allows people to view by color, then see the English and Chinese breakdowns within that specific color cluster, as our eyes are more naturally drawn by color than the difference in language. For example, we can see that within the red-tinged words are mostly Chinese, and blue-tinged words are more often English.

We can also see that despite the word clouds being superimposed, some of the semantic clusters still separate into Chinese and English with little overlap. This is because the topics of the clusters are largely different between those in the Chinese comments versus those in the English comments, and therefore create a divide.

Additionally, we can notice other small details - for example, we can see on the left that hope in English and hope in Chinese (希望) are similar but different colors, as “希望” is slightly more saturated and therefore more intensely rated than “hope”. However, with a superimposed word cloud, it is more difficult to make conclusions about the Chinese emotional comments and English emotional comments separately.

6. FUTURE WORK

6.1 Implementing Semantic Positioning

Due to the limited amount of time, we were not able to start computationally implementing the semantic positioning and were only able to do it manually. As mentioned in the previous sections, there is prior work in implementing this computationally, and we should be able to follow a similar approach using similarity vectors.

6.2 User Testing for Word Cloud Reception

Further user testing is needed to understand how audiences who are looking at the word cloud for the first time would interpret the information. Things that we would want to test for are whether or not the user is able to identify the relative position of a word or phrase on the circumplex model based on the color with reasonable accuracy (i.e., getting the correct quadrant of the coordinate plane), and whether or not the cloud is able to "storytell" with its clusters in a way that would allow users to form their own conclusions. Finally, we would like to see whether separate word clouds is preferable or if it would make more sense to superimpose the word clouds to form a more comprehensive narrative.

7. REFERENCES

[1] Bush, Lynn Ellison II. (1973). Individual differences multidimensional scaling of adjectives denoting feelings. *Journal of Personality and Social Psychology, 25*(1), 50–57. https://doi.org/10.1037/h0034274

[2] Cecilia Chan PhD, RSW, PhD, RSW, Petula Sik Ying Ho & Esther Chow MSW, RSW (2002) A Body-Mind-Spirit Model in Health, Social Work in Health Care, 34:3-4, 261-282, DOI: 10.1300/J010v34n03_02

[3] Gambette P., Véronis J. (2010) Visualising a Text with a Tree Cloud. In: Locarek-Junge H., Weihs C. (eds) Classification as a Tool for Research. Studies in Classification, Data Analysis, and Knowledge Organization. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-10745-0_61

[4] Gudykunst, William B. "Cultural Variability in Communication." Communication Research, vol. 24, no. 4, 1997, pp. 327–348., https://doi.org/10.1177/009365097024004001.

[5] Huo, Luvena. Cross-Cultural Emotive Responses to News Videos. http://www.cs.columbia.edu/~jrk/NSFgrants/videoaffinity/Interim/21y_Luvena.pdf

[6] J. Xu, Y. Tao and H. Lin, "Semantic word cloud generation based on word embeddings," *2016 IEEE Pacific Visualization Symposium (PacificVis)*, 2016, pp. 239-243, doi: 10.1109/PACIFICVIS.2016.7465278.

[7] Jin, Xuecheng & Wang, Zengfu. (2005). An Emotion Space Model for Recognition of Emotions in Spoken Chinese. 397-402. 10.1007/11573548_51.

[8] Ning Yu. (1995). Metaphorical Expressions of Anger and Happiness in English and Chinese, Metaphor and Symbolic Activity, 10:2, 59-92, DOI: 10.1207/s15327868ms1002_1

[9] Plutchik, Robert. (1980). *Emotion: A Psychoevolutionary Synthesis*. Harper & Row, Publishers.

[10] Schlosberg, H. (1954). Three dimensions of emotion. *Psychological Review, 61*(2), 81–88. https://doi.org/10.1037/h0054570

[11] Shiraishi, N., Reilly, J. Content analysis of the emotions affecting caregivers of relatives with schizophrenia. *Curr Psychol* (2020). https://doi.org/10.1007/s12144-020-01185-2

[12] Tsai JL, Simeonova DI, Watanabe JT. Somatic and social: Chinese Americans talk about emotion. Pers Soc Psychol Bull. (2004). doi: 10.1177/0146167204264014. PMID: 15359024.

[13] Turner, Jonathan H. (2000). *On the Origins of Human Emotions: A Sociological Inquiry into the Evolution of Human Emotions.* Stanford University Press, California.

[14] W. Cui, Y. Wu, S. Liu, F. Wei, M. X. Zhou and H. Qu, "Context preserving dynamic word cloud visualization," *2010 IEEE Pacific Visualization Symposium (PacificVis)*, 2010, pp. 121-128, doi: 10.1109/PACIFICVIS.2010.5429600.

[15] Y. Wang et al., "EdWordle: Consistency-Preserving Word Cloud Editing," in IEEE Transactions on Visualization and Computer Graphics, vol. 24, no. 1, pp. 647-656, Jan. 2018, doi: 10.1109/TVCG.2017.2745859.

[16] Yang, L., Li, J., Lu, W. et al. The influence of font scale on semantic expression of word cloud. J Vis 23, 981–998 (2020). https://doi.org/10.1007/s12650-020-00678-3