ARTICLE

<https://doi.org/10.1057/s41599-021-00798-7>

OPEN

Rise and fall of the global conversation and shifting sentiments during the COVID-19 pandemic

Xiangliang Zhang¹ , Qiang Yang¹, Somayah Albaradei¹, Xiaoting Lyu², Hind Alamro¹, Adil Salhi¹, Changsheng Ma¹, Manal Alshehri¹, Inji Ibrahim Jaber³, Faroug Tifratene¹, Wei Wang² , Takashi Gojobori¹, Carlos M. Duarte⁴ & Xin Gao¹

Social media (e.g., Twitter) has been an extremely popular tool for public health surveillance. The novel coronavirus disease 2019 (COVID-19) is the first pandemic experienced by a world connected through the internet. We analyzed 105+ million tweets collected between March 1 and May 15, 2020, and Weibo messages compiled between January 20 and May 15, 2020, covering six languages (English, Spanish, Arabic, French, Italian, and Chinese) and represented an estimated 2.4 billion citizens worldwide. To examine fine-grained emotions during a pandemic, we built machine learning classification models based on deep learning language models to identify emotions in social media conversations about COVID-19, including positive expressions (*optimistic*, *thankful*, and *empathetic*), negative expressions (*pessimistic*, *anxious*, *sad*, *annoyed*, and *denial*), and a complicated expression, *joking*, which has not been explored before. Our analysis indicates a rapid increase and a slow decline in the volume of social media conversations regarding the pandemic in all six languages. The upsurge was triggered by a combination of economic collapse and confinement measures across the regions to which all the six languages belonged except for Chinese, where only the latter drove conversations. Tweets in all analyzed languages conveyed remarkably similar emotional states as the epidemic was elevated to pandemic status, including feelings dominated by a mixture of *joking* with *anxious/pessimistic/annoyed* as the volume of conversation surged and shifted to a general increase in positive states (*optimistic*, *thankful*, and *empathetic*), the strongest being expressed in Arabic tweets, as the pandemic came under control.

¹Computational Biosciences Research Center (CBRC), Computer, Electrical and Mathematical Sciences and Engineering (CEMSE) Division, King Abdullah University of Science and Technology (KAUST), Thuwal, Saudi Arabia. ²School of Computer Science and Information Technology, Beijing Jiaotong University, Beijing, China. ³IT Department, King Abdullah University of Science and Technology (KAUST), Thuwal, Saudi Arabia. ⁴Red Sea Research Center (RSRC) and Computational Biosciences Research Center (CBRC), King Abdullah University of Science and Technology (KAUST), Thuwal, Saudi Arabia.

email: Xiangliang.Zhang@kaust.edu.sa; wangwei1@bjtu.edu.cn; Carlos.Duarte@kaust.edu.sa; Xin.Gao@kaust.edu.sa

Introduction

The coronavirus disease 2019 (COVID-19) pandemic, the first severe pandemic since the Spanish flu a century ago, has infected over 132 million people and killed about over 2.8 million as of April 8, 2021 (<https://covid19.who.int/>), thus causing colossal economic and employment losses globally (Fernandes, 2020) and confining ~58% of the world's population (Bates et al., 2020), forcing them to work from home while parenting their children, who have been unable to attend school. The convergence of economic losses, job insecurity, health risks, and confinement represents an extraordinary pressure on citizens worldwide, which has raised concerns about their well-being and the psychological impacts of confinement and concurrent difficulties. Social media has been instrumental in keeping people informed and in touch during their confinement, thereby providing a record of their impressions, interests, and sentiments that represents an important resource to assess which events among the multiple disruptions affected citizens and how these events and the ensuing sentiments varied across cultures. We conducted sentiment analysis to evaluate the global rise and fall of sentiments during the COVID-19 pandemic. We performed a mass examination of sentiments expressed in millions of social media posts (Twitter and Weibo) about COVID-19 in English, Spanish, Arabic, French, Italian, and Chinese languages using deep learning classification. The data extracted and analyzed in this study are available in Yang et al. (2020).

Sentiment analysis has emerged as a useful tool for monitoring individuals' attitudes and reactions shared through social media in response to various events (Zhang et al., 2018; Jaidka et al., 2020). In public health events, conducting sentiment analysis on people's reactions and concerns helps public health institutes design interventions (Zunic et al., 2020) such as promoting physical distancing and mask-wearing to reduce the risk of COVID-19 infection and protecting the public against misinformation (Tsao et al., 2021). Stronger negative emotions have been exhibited to increase citizen engagement through government social media (Chen et al., 2020). Additionally, public sentiment surveillance can be a vital tool in public health agencies' arsenal of interventions to promote accurate and reliable information for the public on social media (Tsao et al., 2021). Meanwhile, deep learning has automated sentiment analysis by assigning one or several sentiment labels to input raw text without the need for feature engineering and extraction, which are time-consuming. Most work on sentiment analysis has been focused on classifying reviews/comments into coarse-grained categories, with emotion labels such as positive, neutral, and negative or five scores ranking the degree of emotions. However, people's feelings during a pandemic are far more diverse than those expressed in reviews of movies, restaurants, and books often discussed in sentiment analysis. Even with a popular benchmark such as SemEval-2018 (Mohammad et al., 2018), which includes 11 categories, one cannot identify sentimental categories such as *joke* (irony) and *denial* of conspiracy theories, which are commonly observed in social media discussions on COVID-19. Because of the lack of relevant training data, sentiment analysis on COVID-19 tweets has been conducted either unsupervised (Xue et al., 2020; Kabir and Madria, 2020; Drias and Drias, 2020; Chen et al., 2020) or supervised but with limited training data (Kleinberg et al., 2020). In a previous study (Xue et al., 2020), Latent Dirichlet Allocation, which is a widely used topic modeling method, and NRC Lexicon were applied for English tweets to predict one of the several emotions (*anger*, *anticipation*, *fear*, *surprise*, *sadness*, *joy*, *disgust*, and *trust*) for each tweet. In a past analysis (Kabir and Madria, 2020), a real-time COVID-19 tweet analyzer was built to visualize topic modeling results and model human emotions in the U.S. where the sentiments are only

classified as *positive*, *neutral*, and *negative*. Similarly in a previous study (Drias and Drias, 2020), NRC Lexicon was used to predict emotions based on word matching. In a previously performed examination (Chen et al., 2020), sentiment features and topic modeling were employed to reveal substantial differences between the use of controversial terms in tweets about COVID-19. A supervised learning-based study (Kleinberg et al., 2020) used only 5000 training texts to train linear regression models and predict emotional values (*anxiety*, *fear*, *sadness*, *worry*, etc.) based on frequency-inverse document frequency and part-of-speech features.

In addition, most contemporary sentiment analysis applications focus on a single language or a single country or region. In two previous studies (Barkur and Vibha, 2020; Venigalla et al., 2020), lexicon-based methods or emotion dictionaries were used to analyze people's emotions on India's nationwide lockdown because of the COVID-19 outbreak. In a previous analysis (Alhajji et al., 2020), naive Bayes models categorized Saudis' attitudes toward COVID-19 preventive measures as *positive*, *negative*, or *neutral*. In (Pastor, 2020), questionnaire survey and sentiment package were conducted to determine the Philippine students' sentiments on the synchronous online delivery of instruction as a result of the country's extreme community quarantine during the pandemic. In (Ziems et al., 2020), a simple logistic regression classifier was used with linguistic features, hashtags, and embedded tweets to identify anti-Asian hate and counter hate text. However, the nature of the pandemic since late February 2020 has broadened the interest and focus on COVID-19 to a global audience. We therefore present an analysis based on six languages (English, Spanish, Arabic, French, Italian, and Chinese) that represent ~2.4 billion citizens globally. We examined around 105+ million tweets that were collected between March 1 and May 15, 2020, and Weibo messages collected between January 20 and May 15, 2020, with a dominance of English posts (64%). After excluding neutral reports (e.g., media and government reports), we classified the content to reflect three sets of sentiments, each with several expressions, including positive (*optimistic*, *thankful*, and *empathetic*), negative (*pessimistic*, *anxious*, *sad*, *annoyed*, and *denial*), and a complicated expression (*joking*). We built the classification models using deep learning language models such as Bert (Devlin et al., 2018) for Spanish, French, and Italian; XLNet (Yang et al., 2019) for English; AraBert (Antoun et al., 2020) for Arabic; and ERNIE (Zhang et al., 2019) for Chinese.

Methods

Data collection. We collected the tweets using Twint (<https://github.com/twintproject/twint>), an open-source Twitter crawler and formed requests with specific parameters and scraped the resulting responses into JSON documents. We then used a unified query across the target languages: "COVID-19 OR coronavirus OR covid OR corona OR كورونا (Arabic for corona)." We launched 12 instances on 24 cores to download daily updates. Data rates slightly varied throughout the period, with an average of a little over a million tweets per day. We saved the tweets as JSON files and pooled them into a shared medium so that the language models can preprocess and consume them for sentiment analysis, which was conducted on a GPU server (GTX 1080ti GPU and 20 CPUs). Because of the limited number of Chinese posts on Twitter, COVID-19 sentiment analysis for China was performed through Sina Weibo, the largest social media platform in China. Weibo records were collected through Sina Weibo API by first gathering hashtags about COVID-19 and then extracting Weibo records that include these hashtags.

Table 1 Accuracy of classifiers (mean ± std of accuracies in the fivefold cross-validations).

	English (XLNet)	Arabic (AraBERT)	Spanish (BERT)	French (BERT)	Italian (BERT)	Chinese (ERNIE)
Accuracy	0.847 ± 0.004	0.905 ± 0.002	0.823 ± 0.001	0.824 ± 0.004	0.827 ± 0.002	0.880 ± 0.001

Sentiment annotation. Because of the lack of a sentiment benchmark to support our fine-grained emotion analysis, we randomly selected 10,000 English and 10,000 Arabic tweets for sentiment annotation. We chose these languages because they are two of the top five most popular languages worldwide. Additionally, English can be effectively translated into other languages when needed. Domain experts determined the sentiment category after a review of a tweet subset and several rounds of discussion. The final set of ten labels reflected complicated sentiments during the pandemic; these labels covered auxiliary emotions such as *optimistic* (representing *hopeful*, *proud*, and *trusting*), *thankful* for efforts to combat the virus, *empathetic* (including *praying*), *pessimistic* (*hopeless*), *anxious* (*scared*, *fearful*, etc.), *sad*, *annoyed* (*angry*), *in denial* toward conspiracy theories, *official report*, and *joking* (*ironical*). We recruited more than 50 experienced annotators, at least three of whom labeled each tweet. We provided them example tweets in advance with suggested categories. Each tweet was allowed to be assigned to multiple labels, in line with the convention and supporting the analysis of the complicated emotions during the pandemic. To measure the reliability of these sentiment annotations, we conducted a verification study on the annotated tweets and determined the final results of the labeled tweets via majority voting. We acquired the annotated data for Spanish, French, and Italian by translating the labeled English tweets using Google Translate (<https://translate.google.com/>), which has been considered as a reliable strategy in multilingual studies (Balahur and Turchi, 2013). We then evaluated the quality of these translations by calculating the BLEU score, wherein we compared A and A', where A' is translated back by, for example, A(En)→B(Es)→A'(En). The BLEU score in our study was 0.33 (note that the state-of-the-art machine translation model has a BLEU score of 0.39 using a tied transformer), indicating a good translation quality. After performing a manual check of a subset of the translated tweets, we were surprised by the high quality of translation. For Weibo, we examined the COVID-19 posts and annotated 21,173 posts under 7 sentiment categories: *optimistic*, *thankful*, *surprised*, *fearful*, *sad*, *angry*, and *disgusted*. The detailed statistical analysis of the annotated data can be found in (Yang et al., 2020).

Tweets/weibo preprocessing. We preprocessed the tweet data by removing user information, interactions (e.g., retweets and likes), emojis, and emoticons, although they are used to express emotions well because we focused only on textual information. Next, we filtered noisy symbols and texts, which cannot convey any meaningful semantic or lexical information and may even prevent the model from learning, such as user names with the mark “@”; the retweet symbol “RT”; hyperlinks; and some special symbols including line breaks, tabs, and redundant spaces. Unlike previous methods, which also removed hashtags from tweets, we retained these hashtags because they have meaningful semantics, such as in “Proud to be one of the few people who hasn’t texted their ex #COVID-19 #Quarantine #lockdown.” Moreover, we performed word tokenization, stemming, and tagging using the NLTK tool (<https://www.nltk.org/>) for English, Spanish, French, and Italian and Pyarabic for Arabic (<https://github.com/linuxscout/pyarabic>). For Weibo segmentation, we employed Jieba (<https://github.com/fxsjy/jieba>).

Multilabel sentiment classifiers. For each language, we built a multilabel sentiment classifier based on deep neural network language models because of their success in diverse natural language processing tasks (Yang et al., 2019; Safaya et al., 2020). An integration framework called simpletransformer (<https://simpletransformers.ai/>) helped fine-tune these pretrained models and train a customized classifier. We used the pretrained models Bert (Devlin et al., 2018) for Spanish, French, and Italian; XLNet (Yang et al., 2019) for English; AraBert (Antoun et al., 2020) for Arabic; and ERNIE (Zhang et al., 2019) for Chinese and added the last layer of a fully connected network with the sigmoid activation function. These pretrained models represent a tweet or a Weibo post as a 768-dimensional vector that is sent to the last fully connected layer to predict a probability between 0 and 1 for each label. We applied a threshold (e.g., 0.5) to determine if one label should be assigned to the input tweet/Weibo post. Therefore, we were able to assign multiple labels to one tweet when they are all predicted with high probabilities. Before reporting the classification findings, we evaluated the classification models via 5-fold cross-validations. Table 1 presents the accuracy results of different languages. With a validated performance (accuracy > 0.82), all sentiment classifiers were then trained by 10,000 labeled tweets; meanwhile, the Chinese sentiment classifier was trained using 21,173 Weibo posts. We then used the trained models to predict the sentiments of millions of COVID-19 tweets (March 1–May 15, 2020) for our analysis. The implementation details of these classifiers can be found in (Yang et al., 2020).

Results

Rise and fall in the COVID-19-related conversations. We first examined people’s attention to COVID-19, which we quantified as the volume of conversations observed on social media such as Twitter and Sina Weibo. Figure 1 shows the volume of COVID-19 tweets (normalized to the maximum number, with volume statistics in the insert) from March 1 to May 15, 2020, in English, Spanish, Arabic, French, and Italian, as well as the normalized volume of COVID-19 Weibo posts in Chinese from January 20 to May 15, 2020. The figure also shows the volumes of English tweets by topic reflecting major citizen concerns during the pandemic, such as *oil/stock price*, *economic stimulus*, *employment*, *herd immunity*, *medicine/vaccine*, and *working/studying from home*.

Our results across languages show similar patterns of a rapid increase followed by a gradual decline in global citizen conversations on COVID-19; however, Chinese posts reached a peak on January 22, 2 months earlier than the peak in all other languages, which was attained between March 12 and 21 (Fig. 1A). This aspect reflects a lag between the development of the epidemic, which was first detected in Wuhan and then spread to pandemic status. While the rise and fall in the volume of COVID-19 posts were remarkably correlated for English, French, Italian, Arabic, and Spanish (Fig. 1B), the increase occurred earlier in Italian, as Italy was the first Western nation to suffer from the pandemic (Fig. 1A). Remarkably, the time series of COVID-19 conversations on Twitter revealed a clear weekly signal with reduced volumes on weekends and higher activity during weekdays (Fig. 1A), which is consistent with past Twitter activity reports (Golbeck et al., 2010). This strong weekly signal exhibited that differences in habits between weekends and weekdays, which also include reported

Fig. 1 Rise and fall of the global conversation of COVID-19 on social media. **A** Volume of COVID-19 tweets per day from March 1 to May 15, 2020 (normalized to the maximum number, with volume statistics in the insert), in English (En), Spanish (Es), Arabic (Ar), French (Fr), and Italian (It), as well as the normalized volume of COVID-19 Weibo posts per day in Chinese (Zh) from January 20 to May 15. The vertical lines show Sundays for guidance. Discussions and attention quickly reached their peak when important decisions were made, such as the announcement of a state of alarm; the imposition of national lockdowns; and the closure of leisure, business, and religious establishments, as marked out on the peaks of curves. Additionally, people's attention diminished with time. High-correlation coefficient values (**B**) indicate that populations speaking different languages are responding similarly. **C** Volumes (without normalization) of English tweets per day by topic reflecting major citizen concerns during the COVID-19 pandemic, such as oil/stock prices, economic stimulus packages, employment, herd immunity, medicine/vaccine, and working/studying from home (normalized daily volume by maximum shown in **D-H**, respectively).

variations in topics and sentiments (Golbeck et al., 2010), were maintained despite confinement (e.g., lack of workplace transport) somewhat blurring these differences.

Two key factors drove the surge in tweet volume: government enforcement of confinement measures and the collapse of the stock market and the global economy. First, the effect of the government's confinement rulings on the volume of social media messages was clearly reflected in the peak in Chinese posts (January 22, 2020) when the decision to lock Wuhan down was made (Figs. 1A and S1) and Italian tweets (March 12, 2020) when the lockdown was tightened in Italy, closing all commercial and retail businesses (Figs. 1A and S2). English, French, and Spanish posts peaked on March 13 (Figs. 1A and S3–S5), with the U.K.'s initial announcement of the consideration of relaxed social distancing to promote herd immunity, Spain's imposition of a national lockdown and declaration of a state of alarm, and France's closure of all museums and culture and leisure areas. From that point, Europe became the center of the pandemic. In North America, on that day, the Canadian prime minister's wife tested positive, the U.S. declared a national state of emergency, and the U.S. president also tested positive. Meanwhile, the peak volume of Arabic tweets was reached a week later, with the closure of the two holy cities (Mecca and Medina) (March 21, 2020) (Figs. 1A and S6), the speech of Saudi Arabia's king on April 19, and the introduction of a popular hashtag (#البقاء في المنزل، #كُلُّنا مسؤول، #سلاحنا الأقوى لمواجهة الفيروس، #We are all responsible, staying at home is our most powerful weapon against the virus).

Second, the sharp uptick in tweet volume across languages, except Chinese, started on March 9, 2020, which preceded lockdown and other measures in Europe and corresponded to the stock market collapse and the oil price crash on the same day (Fig. 1C, D) and foretold the true catastrophic impact of COVID-19 and was identified as the main driver of the rise in social media conversations. The breakdown of the stock market was followed by increased discussions on unemployment and the need for economic stimulus packages (peaking on March 26, 2020) (Fig. 1C, D, and E). As the cost of such aim in human lives became lucid, the early surge in herd immunity interest quickly dissipated and was replaced by a focus on vaccines, which remained high throughout the study period (Fig. 1C, F, and G). The promotion of confinement measures increased the volume of conversations regarding the challenge of working and studying from home, peaking on March 17, 2020 (Fig. 1C, H).

The decline in the volume of conversations started earlier in China, reaching lower levels between February 25 and March 25 when the infection was controlled and decreasing further after March 25 when China started reopening (Figs. 1A and S1). The decrease in tweet volume after the March 13 peak followed a remarkably similar slope across non-Chinese languages (Fig. 1A). The volume of English tweets was dominated by discussions on therapies and vaccines, followed by employment and working and studying from home, and then conversations on economic stimulus programs and economic impacts (Fig. 1C).

Emotional imprint of the COVID-19 pandemic. Figure 2 presents the sentiment analysis results, which includes variations in the sentiment distribution in six languages over time and significant sentiments identified as sentiment percentage values above the 95% confidence interval ($>\mathcal{N}^{-1}(0.95; \mu, \sigma)$, where μ and σ are the mean and standard deviation of the daily values of one sentiment, respectively). Figure 3 statistically compares the nine sentiment categories (*denial*, *annoyed*, *sad*, *anxious*, *pessimistic*, *joking*, *empathetic*, *thankful*, and *optimistic*) across different languages with respect to the mean, standard deviation, and the increase/decrease slope in the daily sentiment sequences.

Consistent with their correlated trajectories in the conversation volume, English, Spanish, French, and Italian tweets expressed remarkably similar emotional states as the epidemic worsened to pandemic status, including feelings dominated by a mix of *joking* and *anxious/pessimistic/annoyed* as the conversation volume surged, the pandemic spread, and the stock market collapsed between March 1 and 15, 2020 (Fig. 2A–D). In contrast, *anxious* states, alongside *denial* and *empathetic (praying)*, prevailed in Arabic during that period (Fig. 2E). The drop in tweet volume after March 15 was followed by a general trend toward an increase in positive states (*optimistic*, *thankful*, and *empathetic*), which was the strongest in Arabic tweets (Fig. 2E(b)) but also present in English tweets (Fig. 2A(b)) and was conveyed with a combination of *annoyed*, *sad*, and *pessimistic* in Spanish, French, and Italian tweets (Fig. 2B(b)–D(b)). When Europe began its reopening plan in May, *anxious* states emerged in Spanish, French, and Italian tweets (Fig. 2B(b)–D(b)).

This general tendency was disrupted by events eliciting strong but transient sentiment states and were largely language-specific. Two dates annoyed English Twitter users: April 19, when allegations of government negligence emerged because the U.S. COVID-19 death toll rose to 40,000, and May 10, when such annoyance was reflected by Barack Obama's comment: "*White House response to coronavirus has been absolute chaotic disaster*" (Fig. 2A(b) and S7). Spanish tweets also expressed a dominant *annoyed* state on March 22, when a family member of a civil guard agent succumbed to COVID-19, and on the weekend of March 28, 2020, when the European Union failed to agree on a stimulus package, and conveyed a *sad* state about the large death toll (16,972 on that date, the third highest tally in the world after the U.S. and Italy) (Fig. 2B(b) and S8). On May 5, people became *anxious* when a paper published on *bioRxiv* reported that "*COVID-19 has mutated to a more contagious version*" (Korber et al., 2020). This article immediately triggered strong reactions from other researchers, who challenged its conclusions. On May 9–10, the *sad* state significantly increased after reports on the deaths of children and cleaners in hospitals were received. *Sad* remained a significant sentiment for French Twitter users from March 24 to 29, with a spike in *denial* on April 17–18, 2020, associated with conspiracy theories regarding the origin of COVID-19 (Fig. 2C(b) and S9). The same *denial* of a laboratory as a source of the virus was observed in Spanish tweets on April 18–19 (Fig. 2B(b)). Reports on April 29 about possible links between COVID-19 and symptoms of a rare childhood disease caused *anxious* states (Fig. 2C(b) and S9). Most languages saw a slow emergence of positive states after their peak, with Arabian tweets showing the strongest increase in positive states, triggered by a positive tweet by Saudi Arabia's minister of health (#We are all responsible, staying at home is our most powerful weapon against the virus) (Figs. 2E and S10). Consequently, Arabic tweets showed higher proportions of positive states, especially *empathetic* and *thankful* (Fig. 3G, #H) and the strongest increase in *optimistic* (Fig. 3I), compared with those of other languages (Fig. 3). In contrast, Italian tweets reflected a mix of *pessimistic* and *sad* states throughout April, leading to the increase in *anxious* and *sad* states in May (Fig. 2D(b) and S11).

According to the previous study (Yigitcanlar et al., 2020), social media data (e.g., tweets) can guide authorities' interventions and decisions during a pandemic, and the effective use of government social media channels (e.g., Twitter) can facilitate the public's compliance with measures and restrictions. Findings in previous research (Chen et al., 2020) showed that stronger negative emotions lead to higher citizen engagement through government social media. This study's emotion analysis results in different languages also supported these conclusions. Our fine-grained sentiment classification can help us understand people's reactions

Fig. 2 Shifting sentiments in COVID-19 related conversation. **A-F**(a) Variations in the sentiment distribution (as percentage) of tweets in the languages analyzed over time [English (En), Spanish (Es), French (Fr), Italian (It), Arabic (Ar), and Chinese (Zh)]; **A-F**(b) Significant sentiments identified as sentiment percentage values above the 95% confidence interval ($>\mathcal{N}^{-1}(0.95; \mu, \sigma)$, where μ and σ are the mean and standard deviation of the daily values of one sentiment, respectively).

Fig. 3 Comparison of sentiment variations across languages. A–I Mean of daily sentiment in each category for each language. The symbol size is proportional to the standard deviation (amplified ten times to be visible) attached with a line indicating increase (pointing right) or decrease (pointing left) over time. The length of the line is proportional to the steepness of the linear line fitting one sentiment curve over days, i.e., the slope in % per day (amplified 30 times to be visible).

and concerns, which would benefit public health institutes when designing interventions.

More than 50% of Weibo posts in Chinese showed the prevalence of a *fearful* state when human-to-human transmission was confirmed on January 20, 2020 (Figs. 2F and S12), continuing until January 22, when Wuhan was locked down. This aspect, along with the Chinese New Year celebration, led to a growth in positive sentiment (*optimistic* and *thankful*) on January 23–25. *Sad* states were particularly significant on February 4, when the total confirmed cases reached 20,000; on February 7, when Dr. Wenliang Li (the coronavirus whistleblower) died; and April 4, when China held a nationwide memorial for COVID-19 victims. Meanwhile, an increase in *optimistic* states was observed when Wuhan eased its restrictions on April 8. The ensuing sentiments expressed a mix of *thanks* to health-care workers and *fear* of the spread from a few new cases, with spikes of *angry* states on specific events, followed by a rise in *optimistic* and *surprise* states after the country reopened (Fig. 2F(b) and S12).

The overall prevalence of different emotional states was similar among languages (Fig. 3), with the important component of positive states, such as *optimistic* and *joking* (about 20% of tweets each, Fig. 3F, I), but also negative ones, such as *annoyed* and *anxious* (Fig. 3B, D). The increase and decline of sentiments in Arabic and Chinese are more significant than those in other languages, driven by various cultural/religious events, holidays, and government announcements. Across all languages, the *optimistic* and *sad* states tended to increase over time, whereas *joking* tended to decrease (Fig. 3C, F, I). Remarkably, Arabic speakers revealed the highest *empathetic* sentiments (Fig. 3G), possibly because the pandemic overlapped with Ramadan, a period that particularly promotes empathy toward those in need. Another big portion of sentiment is *annoyed* (constituting 20% of all languages) (Fig. 3B). Although it seems stable over time, it showed sudden spikes, as shown in Fig. 2A(b)–C(b), caused by individual COVID-19-related events such as deaths, economic and political effects, and others (with detailed annotation and discussion in Figs. S7–S9).

To quantitatively analyze the social patterns shared by the rise and fall of global sentiment, we measured the correlation coefficients of the average daily sentiment distribution among

the five languages (English, Spanish, French, Italian, and Arabic), as their results were all based on tweets. The presentation of these correlation coefficients (Fig. 4F) is similar to the form of volume correlation in Fig. 1B. We observed that the Arabic sentiment is less relevant to others, while Spanish, French, and Italian sentiments are highly similar (with correlation coefficients 0.97–0.99). English sentiment was closer to Spanish and Italian than to French sentiment. To further understand such a similarity, we applied principal component analysis and then t-distributed stochastic neighbor embedding (t-SNE) to visualize the daily sentiments of these different languages in a two-dimensional space, as shown in Fig. 4A–E. In each visualization figure, several samples were annotated with their corresponding dates. Several interesting findings were generated. First, Arabic sentiments (Fig. 4E) show two clusters, one mostly with date samples in March 2020 and the other with all other days. Notably, sentiments on March 21 were extremely close to those on May 15 rather than those on other days in March. March 21 is when the volume peaked, followed by a strong increase in positive states, thus the similarity with dates in May (Figs. 2E and S11). Second, English, Spanish, and French sentiments revealed a small cluster for days before the volume peak on March 12–13, which is consistent with our results in Fig. 2. The sentiments when the pandemic began were different from those after confinement measures were imposed on human mobility. Finally, the overall visualized sentiments in English, Spanish, French, and Arabic shared similar distributions, as quantitatively verified with correlation coefficients >0.89.

We also analyzed the sentiments of English tweets by topics reflecting major citizen concerns during COVID-19, such as *oil/stock price*, *economic stimulus*, *employment*, *herd immunity*, *medicine/vaccine*, and *working/studying from home*. Figure 5 summarizes the results, and Figs. S13–19 in the supplementary information presents details for each topic. The results showed that in relation to COVID-19, *working/studying from home* led with expressions of *optimism* and *thankfulness* in contrast to *herd immunity*, which elicited negative states of *denial* and *anxiety* (Figs. 5 and S13–S14). Conversations around *economic collapse* led to the strongest *annoyed* (stock market collapse), *pessimistic* (collapse of oil prices), and *sad* (high unemployment rate) states

Fig. 4 Visualization of daily sentiment compositions. **A–E** Visualization by principal component analysis and then t-distributed stochastic neighbor embedding (t-SNE) for daily sentiments in English (En), Spanish (Es), French (Fr), Italian (It), and Arabic (Ar) (each point presents the sentiment of one day). Sampled points were annotated with the corresponding dates. **F** The correlation coefficients of the average daily sentiment distribution among five languages (English, Spanish, French, Italian, and Arabic) analyzed on tweets.

(Figs. 5 and S15–S17). The sentiment in conversations about economic stimulus packages showed stronger *joking* and *annoyed* states than other emotions (Figs. 5 and S18). Discussions on drugs/vaccines showed the second strongest expressions of *optimism*, with significant *denial* and *annoyed* sentiments on events such as the poaching of vaccines from other countries and the hyped usage of unproven drugs (Figs. 5 and S19).

Discussion

Our analysis, based on over 105 million social media posts in six languages spoken by 2.4 billion citizens worldwide, revealed that the global conversation on COVID-19 followed the spread of the pandemic, with the rise of the volume in China preceding those in Western and Arabic languages by 2 months. These conversations, regardless of language, showed a remarkably similar pattern of a rapid rise and then a slow decline over time across all nations, and the surge being affected by economic collapse and confinement decisions in Western nations. The global conversation had strongly negative reactions to specific topics, such as herd immunity strategies, which involve the acceptance of massive losses of lives among the elderly, and fake news designed to scare the public or spread conspiracy theories. *Optimistic* and positive sentiments increased over time and in parallel to a decrease in the volume of conversations, as policies were implemented to release the population from confinement and reactivate the economy as a result of lower infections and deaths.

Our investigation provided an unprecedented application of sentiment analysis to multiple languages, including English, Spanish, French, Italian, Arabic, and Chinese, as opposed to most

studies that conducted monolingual research (Alamoodi et al., 2020), as they addressed sentiments related to events in a particular region and cultural setting. However, the COVID-19 pandemic has affected the entire world nearly synchronously except for the earlier effects in China, therefore providing an unprecedented opportunity to examine and compare sentiments across languages and over time.

Social psychology has shown that negative emotions such as *anger*, *sadness*, and *fear* often arise during stressful events, including those involving health risks (Ferrer et al., 2017). Hence, the pandemic, which caused stress in both health and the economy, was poised to generate an increase in overwhelmingly negative emotions. In contrast, positive emotions such as *empathy* and *optimism* can help overcome negative emotions and cope with stressful situations (Folkman, 2018). Indeed, positive emotions provide defense mechanisms against stressful events and help manage negative emotions, producing resilient individuals (Armstrong et al., 2011; Cohn et al., 2009; Tugade and Fredrickson, 2004). Our findings showed a rapid rise in negative emotions, signaling growing stress and rebalance, and sometimes, rise in positive emotions often associated with specific encouraging events. Additionally, the extent of positive responses was greater in Arabic Twitter conversations about COVID-19, which may reflect a cultural difference in coping with stress and crises and/or a more resilient society resulting from a more recurrent exposure to adverse situations. For instance, the Middle East, particularly the Arabic Peninsula, experienced a recent SARS epidemic (the Middle East respiratory syndrome coronavirus, MERS-CoV), which it successfully overcame and may have helped build greater resilience in the society when facing new pandemics (Algaissi et al., 2020).

Fig. 5 Comparison of sentiment variations across conversation topics. A–H Mean of daily sentiment in each category under each topic, excluding *empathetic* because of its small value (at the 10^{-4} level) under all topics. Symbol size is proportional to the standard deviation (amplified by ten times to be visible), attached with a line indicating increase (pointing right) or decrease (pointing left) over time. The length of the line is proportional to the steepness of the linear line fitting one sentiment curve over days, i.e., the slope in % per day (amplified by 60 times for visibility).

Sentiment analysis provides a previously unconsidered tool for policymakers to track societal states that can guide interventions on health and other types of issues, as well as emergency responses. Sentiment analysis applied to public health crises supports action in monitoring, discovery, news sharing, and policy formulation and evaluation (Alamoodi et al., 2020; Jang et al., 2021). Large-scale analysis of emotions expressed via social media provides a tool for business and policy decisions and responses to certain events, which is beyond the scope of this paper. However, we can assume that the sentiments expressed on social media, well above the “background” noise, somehow reached policymakers, whether formally or informally, and stimulated a discussion of where and how the pandemic may have occurred.

COVID-19 is the first pandemic in a globalized, online world, which both relays information to citizens and allows for the tracking of the volume of sentiments expressed through social media. The three-pronged impact on health, movement, and the economy affected citizens across languages. We found remarkable similarities in the global COVID-19 conversation in terms of volume, sentiments, and their triggers despite broad social and cultural differences inherent to the different languages used by the citizens monitored in this study. Global media provides a vehicle for the spread of fake news, but our analysis showed users’ ability to identify these and respond with strong negative sentiments. Our results provided evidence of citizens’ globalized reactions to the constraints and risks imposed by the first pandemic in the internet era, with the growth in optimism over time foretelling a desire to seek together a reset for an improved COVID-19 world.

Data availability

Our Tweets and Weibo data are available at <https://github.com/gitdevqiang/SenWave>.

Received: 21 October 2020; Accepted: 26 April 2021;
Published online: 17 May 2021

References

- Alamoodi A, Zaidan B, Zaidan A, Albahri O, Mohammed K, Malik R, Almahdi E, Chyad M, Tareq Z, Albahri A, Hameed H (2020) Sentiment analysis and its applications in fighting COVID-19 and infectious diseases: a systematic review. *Expert Syst Appl* 167:114155
- Algaissi AA, Alharbi NK, Hassanain M, Hashem AM (2020) Preparedness and response to COVID-19 in Saudi Arabia: Building on MERS experience. *J Infect Public Health* 13:834–838
- Alhajji M, Khalifah A, Aljubran M, Alkhalifah M (2020) Sentiment analysis of tweets in Saudi Arabia regarding governmental preventive measures to contain COVID-19. Preprint at <https://doi.org/10.20944/preprints202004.0031.v1>
- Antoun W, Baly F, Hajj H (2020) AraBERT: transformer-based model for Arabic language understanding. In Proceedings of the 4th workshop on open-source arabic corpora and processing tools, with a shared task on offensive language detection, 11–16 May, 2020
- Armstrong AR, Galligan RF, Critchley CR (2011) Emotional intelligence and psychological resilience to negative life events. *Pers Individ Differ* 51 (3):331–336
- Balahur A, Turchi M (2013) Improving sentiment analysis in twitter using multilingual machine translated data. In: Proceedings of the International Conference Recent Advances in Natural Language Processing (RANLP), 9–11 September, 2013
- Barkur G, Vibha G (2020) Sentiment analysis of nationwide lockdown due to COVID-19 outbreak: evidence from India. *Asian J Psychiatr* 51:102089
- Bates AS, Schlegel P, Roberts RJ, Drummond N, Tamimi IF, Turnbull R, Zhao X, Marin EC, Popovici PD, Dhawan S, Jamasb A et al. (2020) Complete connectomic reconstruction of olfactory projection neurons in the fly brain. *Curr Biol* 30(16):3183–3199
- Chen L, Lyu H, Yang T, Wang Y, Luo J (2020) In the eyes of the beholder: Sentiment and topic analyses on social media use of neutral and controversial terms for COVID-19. Preprint at <https://arxiv.org/abs/2004.10225>
- Chen Q, Min C, Zhang W, Wang G, Ma X, Evans R (2020) Unpacking the black box: how to promote citizen engagement through government social media during the COVID-19 crisis. *Comput Human Behav* 110:106380

- Cohn MA, Fredrickson BL, Brown SL, Mikels JA, Conway AM (2009) Happiness unpacked: positive emotions increase life satisfaction by building resilience. *Emotion* 9(3):361–368
- Devlin J, Chang MW, Lee K, Toutanova K (2018) Bert: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of NAACL-HLT 2019*, 2–7 June 2019
- Drias HH, Drias Y (2020) Mining Twitter data on COVID-19 for sentiment analysis and frequent patterns Discovery. *medRxiv*. <https://doi.org/10.1101/2020.05.08.20090464>
- Fernandes N (2020) Economic effects of coronavirus outbreak (COVID-19) on the world economy. Available at SSRN: Fernandes, Nuno, Economic Effects of Coronavirus Outbreak (COVID-19) on the World Economy (March 22, 2020). IESE Business School Working Paper No. WP-1240-E, Available at SSRN: <https://ssrn.com/abstract=3557504>
- Ferrer RA, Klein WMP, Graff KA (2017) Self-affirmation increases defensiveness toward health risk information among those experiencing negative emotions: Results from two national samples. *Health Psychol* 36(4):380–391
- Folkman S (2018) The case for positive emotions in the stress process. *Anxiety Stress Coping* 21(1):3–14
- Golbeck J, Grimes JM, Rogers A (2010) Twitter use by the US Congress. *J Am Soc Inform Sci Tech* 61(8):1612–1621
- Jaidka K, Giorgi S, Schwartz HA, Kern ML, Ungar LH, Eichstaedt JC (2020) Estimating geographic subjective well-being from Twitter: a comparison of dictionary and data-driven language methods. *Proc Natl Acad Sci USA* 117(19):10165–10171
- Jang H, Rempel E, Roth D, Carenini G, Janjua NZ (2021) Tracking COVID-19 discourse on Twitter in North America: Infodemiology study using topic modeling and aspect-based sentiment analysis. *J Med Internet Res* 23(2): e25431
- Kabir M, Madria S (2020) CoronaVis: A real-time COVID-19 tweets analyzer. Preprint at <https://arxiv.org/abs/2004.13932>
- Kleinberg B, van der Vegt I, Mozes M (2020) Measuring emotions in the COVID-19 real world worry dataset. In *Proceedings of the 1st Workshop on NLP for COVID-19 at ACL 2020*, 9–10 July 2020
- Korber B, Fischer WM, Gnanakaran S, Yoon H, Theiler J, Abfalterer W, Foley B, Giorgi EE, Bhattacharya T, Parker MD, Partridge DG, Evans CM, Freeman TM, de Silva TI, on behalf of the Sheffield COVID-19 Genomics Group, LaBranche CC, Montefiori DC (2020) Spike mutation pipeline reveals the emergence of a more transmissible form of SARS-CoV-2. Preprint at *bioRxiv* <https://doi.org/10.1101/2020.04.29.069054>
- Mohammad S, Bravo-Marquez F, Salameh M, Kiritchenko S (2018) Semeval-2018 task 1: affect in tweets. In: *Proceedings of the 12th international workshop on semantic evaluation*, June 2018
- Pastor C (2020) Sentiment analysis on synchronous online delivery of instruction due to extreme community quarantine in the Philippines caused by COVID-19 pandemic. *Asia J Multidiscip Stud* 3(1):1–6
- Safaya A, Abdullatif M, Yuret D (2020) Kuisail at semeval-2020 task 12: Bert-cnn for offensive speech identification in social media. In: *Proceedings of the International Workshop on Semantic Evaluation (SemEval)*, 13–14 September, 2020
- Tsao SF, Chen H, Tisseverasinghe T, Yang Y, Li L, Butt ZA (2021) What social media told us in the time of COVID-19: a scoping review. *Lancet Digit Health* 3(3):e175–e194
- Tugade MM, Fredrickson BL (2004) Resilient individuals use positive emotions to bounce back from negative emotional experiences. *J Person Soc Psychol* 86(2):320–333
- Venigalla A, Vagavala D, Chimalakonda S (2020) Mood of India during COVID-19—An interactive web portal based on emotion analysis of twitter data. In *Proceedings of the Conference Companion Publication of the 2020 on Computer Supported Cooperative Work and Social Computing*, 17–21 October 2020
- Xue J, Chen J, Chen C, Zheng C, Li S, Zhu T (2020) Machine learning on big data from Twitter to understand public reactions to COVID-19. Preprint at <https://arxiv.org/abs/2005.08817>
- Yang Q, Alamro H, Albaradei S, Salhi A, Lv X, Ma C, Alshehri M, Jaber I, Tifratene F, Wang W, Gojobori T, Duarte C, Gao X, Zhang X (2020) Senwave: monitoring the global sentiments under the covid-19 pandemic. Preprint at <https://arxiv.org/abs/2006.10842>
- Yang Z, Dai Z, Yang Y, Carbonell J, Salakhutdinov RR, Le QV (2019) XLNet: generalized autoregressive pretraining for language understanding. In *Proceedings of the 33rd Conference on Neural Information Processing Systems (NeurIPS 2019)*, 8–14 December 2019
- Yigitcanlar T, Kankanamge N, Preston A, Gill PS, Rezayee M, Ostadnia M, Xia B, Ioppolo G (2020) How can social media analytics assist authorities in pandemic-related policy decisions? Insights from Australian states and territories. *Health Inf Sci Syst* 8(1):1–21
- Zhang Z, Han X, Liu Z, Jiang X, Sun M, Liu Q (2019) ERNIE: enhanced language representation with informative entities. In *Proceedings of the 57th Conference of the Association for Computational Linguistics (ACL)*, 28 July–2 August, 2019
- Zhang L, Wang S, Liu B (2018) Deep learning for sentiment analysis: a survey. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* 8(4)
- Ziems C, He B, Soni S, Kumar S (2020) Racism is a virus: Anti-asian hate and counterhate in social media during the Covid-19 crisis. Preprint at <https://arxiv.org/abs/2005.12423>
- Zunic A, Corcoran P, Spasic I (2020) Sentiment analysis in health and well-being: systematic review. *JMIR Med Inf* 8(1):e16023

Acknowledgements

The research reported in this publication was supported by funding from Computational Bioscience Research Center (CBRC), King Abdullah University of Science and Technology (KAUST), Saudi Arabia, with award numbers FCC/1/1976-17-01, FCC/1/1976-18-01, FCC/1/1976-19-01, FCC/1/1976-23-01, FCC/1/1976-24-01, FCC/1/1976-25-01, FCC/1/1976-26-01 and FCC/1/1976-31-01. We thank the KAUST writing support team for the proofread of our manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1057/s41599-021-00798-7>.

Correspondence and requests for materials should be addressed to X.Z., W.W., C.M.D. or X.G.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021