

Corpus Analysis of Coronavirus Disease-19 (COVID-19)-Related Loneliness in Twitter

Chereen Shurafa, Wajdi Zaghouani

Submitted to: JMIR Mental Health
on: February 22, 2024

Disclaimer: © The authors. All rights reserved. This is a privileged document currently under peer-review/community review. Authors have provided JMIR Publications with an exclusive license to publish this preprint on its website for review purposes only. While the final peer-reviewed paper may be licensed under a CC BY license on publication, at this stage authors and publisher expressly prohibit redistribution of this draft paper other than for review purposes.

Table of Contents

Original Manuscript.....	5
---------------------------------	----------

Preprint
JMIR Publications

Corpus Analysis of Coronavirus Disease-19 (COVID-19)-Related Loneliness in Twitter

Chereen Shurafa^{1*} MS, PhD; Wajdi Zaghouni^{1*} PhD

¹Hamad bin Khalifa University Doha QA

*these authors contributed equally

Corresponding Author:

Chereen Shurafa MS, PhD
Hamad bin Khalifa University
Education City
Doha
QA

Abstract

Background: Loneliness is a complex mental health issue involving feelings of isolation, disconnectedness, and a lack of purpose in life. The coronavirus disease (COVID-19) pandemic and lockdowns significantly increased social isolation, contributing to a global rise in loneliness on social online platforms. However, research on mental health expressions in Arabic on social media is limited, exacerbating the lack of understanding of loneliness and mental health issues in these communities. Stigma and societal resistance in the Middle East hinder individuals from seeking mental health support, a trend which was particularly noticeable amidst the challenges of the pandemic.

Objective: This study aimed to conduct a corpus analysis of an open-source dataset to gain insights into the origin of how loneliness was expressed in Arabic on Twitter during the initial months of COVID-19 regulations. Additionally, it explored temporal trends during the same time period to enhance comprehension of the prevalence of loneliness and mental health struggles during the initial months of the pandemic.

Methods: Grounded in corpus linguistic methodology, the study also employs thematic analysis, explores temporal trends, and conducts Inter-Annotator Agreement (IAA) to comprehend the prevalence of loneliness and mental health struggles during the pandemic's early stages.

Results: Our analysis of Twitter data using AntConc revealed crucial insights into the origins of mental health challenges and loneliness during the early days of the COVID-19 pandemic. Employing sophisticated linguistic techniques, we identified over 18,000 instances of loneliness-related terms, which deepened our understanding of the issue. Emotions were expressed through various themes, with a significant focus on COVID-19 news and opinions. The advanced search features of AntConc were instrumental in this process, allowing us to gauge word frequency and locate relevant tweets. The association between loneliness and the pandemic was clear, especially with the word "corona" being the most frequent collocate. Thematic analysis identified the seven key themes that were despair in the face of uncertainty, grief for the loss of life, hope through faith, opinions on COVID-19, opinions on isolation, COVID-19 news, and fear of illness which characterized the emotional landscape during this time. Inter-Annotator Agreement highlighted the diversity of interpretations.

Conclusions: Our study demonstrates the crucial role social media plays in capturing and understanding human emotions during major global events like the COVID-19 pandemic. Our AntConc analysis offers insights into mental health and loneliness during the early COVID-19 pandemic. It highlights the Arabic-speaking online community's experiences and contributes to global understanding of mental health and social media dynamics. This approach can be applied to different languages, enriching our understanding of loneliness and mental health during crises.

(JMIR Preprints 22/02/2024:55970)

DOI: <https://doi.org/10.2196/preprints.55970>

Preprint Settings

1) Would you like to publish your submitted manuscript as preprint?

✓ Please make my preprint PDF available to anyone at any time (recommended).

Please make my preprint PDF available only to logged-in users; I understand that my title and abstract will remain visible to all users.
Only make the preprint title and abstract visible.

No, I do not wish to publish my submitted manuscript as a preprint.

2) If accepted for publication in a JMIR journal, would you like the PDF to be visible to the public?

✓ **Yes, please make my accepted manuscript PDF available to anyone at any time (Recommended).**

Yes, but please make my accepted manuscript PDF available only to logged-in users; I understand that the title and abstract will remain visible to all users.

Yes, but only make the title and abstract visible (see Important note, above). I understand that if I later pay to participate in <http://www.jmir.org/>



Original Manuscript

Corpus Analysis of Coronavirus Disease-19 (COVID-19)-Related Loneliness in Twitter

Chereen Shurafa ¹, Wajdi Zaghouani ²

¹ College of Humanities and Social Sciences, Hamad bin Khalifa University;
chsh34020@hbku.edu.qa

² College of Humanities and Social Sciences, Hamad bin Khalifa University;
wzaghouani@hbku.edu.qa

* Correspondence: chsh34020@hbku.edu.qa

Abstract

Background: Loneliness is a complex mental health issue involving feelings of isolation, disconnectedness, and a lack of purpose in life. The coronavirus disease (COVID-19) pandemic and lockdowns significantly increased social isolation, contributing to a global rise in loneliness on social online platforms. However, research on mental health expressions in Arabic on social media is limited, exacerbating the lack of understanding of loneliness and mental health issues in these communities. Stigma and societal resistance in the Middle East hinder individuals from seeking mental health support, a trend which was particularly noticeable amidst the challenges of the pandemic.

Objective: This study aimed to conduct a corpus analysis of an open-source dataset to gain insights into the origin of how loneliness was expressed in Arabic on Twitter during the initial months of COVID-19 regulations. Additionally, it explored temporal trends during the same time period to enhance comprehension of the prevalence of loneliness and mental health struggles during the initial months of the pandemic.

Methods: Grounded in corpus linguistic methodology, the study also employs thematic analysis, explores temporal trends, and conducts Inter-Annotator Agreement (IAA) to comprehend the prevalence of loneliness and mental health struggles during the pandemic's early stages.

Results: Our analysis of Twitter data using AntConc revealed crucial insights into the origins of mental health challenges and loneliness during the early days of the COVID-19 pandemic. Employing sophisticated linguistic techniques, we identified over 18,000 instances of loneliness-related terms, which deepened our understanding of the issue. Emotions were expressed through various themes, with a significant focus on COVID-19 news and opinions. The advanced search features of AntConc were instrumental in this process, allowing us to gauge word frequency and locate relevant tweets. The association between loneliness and the pandemic was clear, especially with the word "corona" being the most frequent collocate. Thematic analysis identified the seven key themes that were despair in the face of uncertainty, grief for the loss of life, hope through faith, opinions on COVID-19, opinions on isolation, COVID-19 news, and fear of illness which characterized the emotional landscape during this time. Inter-Annotator Agreement highlighted the diversity of interpretations.

Conclusion: Our study demonstrates the crucial role social media plays in capturing and understanding human emotions during major global events like the COVID-19 pandemic. Our AntConc analysis offers insights into mental health and loneliness during the early COVID-19 pandemic. It highlights the Arabic-speaking online community's experiences and contributes to global understanding of mental health and social media dynamics. This approach can be applied to different languages, enriching our understanding of loneliness and mental health during crises.

Keywords: loneliness; COVID-19; corpus analysis; thematic analysis; Twitter; Arabic

Introduction

Loneliness, a psychological state devoid of purpose, encompasses diverse dimensions, including the subjective perception of isolation and the experience of negative emotions intertwined with scarce social connections. Robert Stuart Weiss posits loneliness as a natural and universal phenomenon capable of affecting individuals irrespective of demographic characteristics [1]. The advent of the coronavirus-19 (COVID-19) pandemic dramatically altered the global landscape, introducing widespread lockdowns that precipitated increased social isolation. The surge in loneliness during the COVID-19 pandemic has become a poignant societal concern, with individuals grappling with isolation, social distancing measures, and disruptions to traditional social interactions [2].

Social media has emerged as a crucial platform for individuals to navigate and alleviate feelings of loneliness. While social media platforms like Facebook, Twitter, and Instagram can encourage excessive use and negative interactions, they can also serve as valuable outlets for expression [3]. The multilingual nature of social media platforms has enabled diverse voices to express the effects of isolation on mental well-being.

Twitter is a popular choice for text analysis due to its short length of expression (140 characters) and worldwide use of Twitter for expressing people's opinions on different issues [4]. Examining specific aspects of loneliness expression on Twitter provides valuable insights into how individuals articulate and communicate their feelings in the digital realm. Burke and Kraut [5] have shed light on the patterns and nuances of loneliness expression on Twitter. Individuals often share personal reflections, emotional states, and experiences related to isolation and solitude through tweets [6]. The COVID-19 pandemic has heightened emotional struggles among younger individuals, as social isolation and loneliness cause depression and anxiety symptoms [7].

The study aimed to demonstrate how corpus linguistics tools could be used to find expressions of mental health struggles and indications of the origin of loneliness due to COVID-19 within the corpus. It also aims to show how AntConc could be used to analyze an Arabic dataset to conduct a thematic analysis and analyze temporal trends to increase our understanding of the expression of mental health struggles and loneliness.

Methodology

Dataset collection

To conduct the corpus analysis, we used the open-source dataset ArCov-19, as described previously [8]. ArCOV-19, an extensive Arabic COVID-19 Twitter dataset, chronicling one year from January 27, 2020 to January 31, 2021, and comprising a substantial repository of approximately 2.7 million tweets. We sampled the dataset for the months of March, April, May, and June of 2020, as this timeframe corresponds with the months of the COVID-19 lockdown [9], [10], [11], [12]. Our dataset began on March 11, when COVID-19 was considered a pandemic by the World Health Organization (WHO) [13]. We collected a total of 889,306 tweets.

Corpus analysis techniques

We identified the dataset from March to June of 2020 as the target corpus and loaded the corpus into

AntConc for each month, along with the reference corpus. AntConc conducts corpus linguistics, or the analysis of electronic texts, to provide commonalities in language.

We used the Almaany Arabic Dictionary and the Cambridge Dictionary English to Arabic feature to create a list of words associated with loneliness in Arabic [14], [15]. We searched for words in English and selected the English to Arabic option on both sites. Upon searching for the words lonely, loner, and alone, we were provided with their Arabic equivalents. We looked at words associated with loneliness and searched and included the words that included results, such as lonely, loner, alone, and secluded. Searching the synonyms of loneliness through both dictionaries resulted in the words provided in Table 1.

Table 1. Variations of loneliness-related words in Arabic, according to Almaany and Cambridge dictionaries.

Word	Dictionary meaning
لوحدي	Variation of alone (Almaany)
وحيدا	Variation of lonely and loner (Almaany)
وحيد	Variation of lonely and loner (Cambridge dictionary)
لوحيدة	Variation of lonely (Cambridge dictionary)
واحد	Variation of lonely (Almaany)
الوحيد	Variation of lonely and alone (Almaany)
منعزل	Variation of lonely and secluded (Almaany)
معزول	Variation of lonely and secluded (Almaany)
لوحدي	Variation of lonely (Almaany)
متوحد	Variation of lonely (Almaany)
الوحيدة	Variation of lonely and alone (Almaany)

We used our list of words from Table 1 and ran an advanced and basic search. For the advanced search, we used the concordance tab to search for tweets and used the keyword in context (KWIC) feature. For a basic search, we found collocates and ran a keyword analysis to analyze the keyness of the words.

Concordance search

Upon loading the words in the advanced search option, we were directed to the Concordance tab. This feature provided us with the frequency of each word, which would assist us in finding relevant tweets. We listed the frequency of each word, which consisted of a combined frequency of 18,299. The prevalence of the selected loneliness-related words was identified within the target corpus. Additionally, through Table 2, we provided an example of a concordance search result for each word directly associated with loneliness, along with a translation of each tweet.

After loading the list of words from Table 1, we used AntConc's advanced search feature to run a KWIC search and advanced search and loaded words to get the number of hits (17367 hits). We searched each word separately, demonstrating that sometimes the words were in the context of loneliness, while other times, the context had nothing to do with loneliness.

Table 2. Examples of loneliness-related tweets for searched words found through a Concordance

search in AntConc.

Loneliness related word	Translation of word	Example and translation of tweet from Concordance search
لوحدي	Alone	<p>كورونا طلعت اسوء مافينا". ناس بتتعمد تاذي غيرها ماشيه بمبدأ لو مت مش هاموت لوحدي- وناس بتبعد عن اقرب الناس ليها لمجرد اصابتهم-</p> <p>"Corona has become the worst thing for us. -People deliberately hurt others, following the principle that if I die, I will not die alone. -And people stay away from the people closest to them just because they are infected</p>
وحيدا	Lonely and loner	<p>كورونا علمتنا درساً في غاية الاهمية عندما تبقى وحيداً ستنجو وعندما تتبع القطيع ستموت.</p> <p>Corona taught us a very important lesson: when you remain alone, you will survive, and when you follow the herd, you will die.</p>
وحيد	Lonely and loner	<p>"الملياردير البرتغالي أنطونيو فييرا صاحب بنك سانتاندير بالبرتغال توفي بسبب عدوى #كورونا الأربعة الماضي ابنته كتبت في إحدى وسائل التواصل الاجتماعي: نحن #عائلة ثرية لكن والدي توفي وحيداً ومختنقاً يبحث عن شيء مجاني هو الهواء والمال بقي في المنزل</p> <p>"Portuguese billionaire Antonio Vieira, owner of Santander Bank in Portugal, died of #Corona infection last Wednesday. His daughter wrote on one of the social media outlets: We are a #rich_family, but my father died alone and suffocated, looking for something free, which is air and money. He stayed at home."</p>
وحيدة	Lonely	<p>ما أفجعني اليوم أنها فارقت الحياة. نعم هي ماتت .. لكن بصورة لا يتمناها أيأ منكم .. توفيت وحيدة دون مرافق وتشخيصها #كوفيد19 في #العزل . أشعر بالحزن تجاه كبار السن ومن استهتار البالغين من الأبناء.</p> <p>What saddens me today is that she passed away. Yes, she died...but in a way that none of you would wish for...she died alone, without an escort, and was diagnosed with #Covid19 in #isolation. I feel sad for the elderly and the carelessness of adults towards their children.</p>
واحد	Lonely	<p>2020 "اخوة اثنين في سجن واحد هما وحيد والديهم الأم والأب يعيشون ظروف صعبة بلا اولادهم ، ويزيد الألم وقلقهم بهذه الظروف مع هذا الوباء</p> <p>2020 "Two brothers in one prison. They are the only ones whose parents are the mother and father. They live in difficult circumstances without their children, and their pain and anxiety in these circumstances increases with this epidemic."</p>
الوحيد	Lonely or loner	<p>. انا الوحيد اللي مو مهتم في مرض ذا ولا ادري وش الوضع</p> <p>I am the only one who is not interested in this disease and I do not know what the situation is.</p>
منعزل	Lonely or secluded	<p>"خلصنا من فلم #الحظر_الشامل وداخليين في مسلسل #الحظر_الجزئي ولا نزال نعاني من اكشن ودراما ال #فيروس_كورونا_المستجد في العالم ، وبعض الناس قلبها طرارة ومكافآت وكأنه بقوقعة بروحة منعزل عن البشر!!</p> <p>"We have finished the movie #Comprehensive_Lockdown and entered the series #Partial_Lockdown and we are still suffering from the action and drama of the new #Corona_virus in the world, and some people's hearts are filled with joy and rewards as if they are in a shell with a spirit isolated from humans!!</p>

معزول	Lonely or secluded	"#الشروق حرب طبقية وعالم معزول.. سيناريوهات للمستقبل بعد انتهاء أزمة كورونا #Al-Shorouk Class war and a secluded world...scenarios for the future after the end of the Corona crisis
لوحدي	Alone	أنا مكنتش بهتم بموضوع كورونا خالص وسابها لله وأخري كحول ومبسلمش وبتاع لحد ما واحد صاحبي كانو شاكين فيه وبعد عشنا يوم رعب لحد ما التحليل طلع والحمد لله سليبي .. بس الرعب بجد انك تكون تعبان وعديت حد من اهلك ! أنا من ساعتها مبروحش ف حنة وبفطر لوحدي والموضوع مش سهل زي ما احنا فاكرين I didn't care about the Corona issue at all, and I left it to God and left it with alcohol and didn't give it to anyone. One of my friends was suspicious of it, and we lived through a terrifying day until the test came out negative, thank God. But what's really scary is that you're tired and you infect someone in your family! Since then, I have gone to breakfast alone, and it's not as easy as we think.
متوحد	Lonely	عادي قولوا عليا متوحدة بس منجد مرة مبسولة في ذا الوضع.. ما ابغى الحياة ترجع طبيعية، هي " كذا مرة تمام، بعيد عن البشر بعيد عن الالتزامات الاجتماعية.. ممكن. ذل الشيء لما يخلص، ما راح يتكرر ثاني، فعيشوا التجربة عادي يعني It's okay to say that I'm lonely, but was once happy in this situation. I don't want life to go back to normal, it's just like that every once in a while, far from people, far from social obligations. Maybe when this thing is over, it won't happen again, so live the experience normally.
الوحيدة	Lonely and alone	انا الوحيدة اللي في حجر صحي حق كورونا حقيقي مامعاي ولا حد بين اربع حيط في السرير اليوم كلو كورونا زاتو حا تعفيني I am the only one who is in quarantine. Corona is real, and I do not have anyone among the .four walls in the bed. Today, all of Corona has come, and it will make me exempt

The word **الوحيدة** is defined as a variation of lonely and alone, according to Almaany. It could simply mean that something or someone is the only thing or only one, not necessarily alluding to loneliness. The KWIC search assisted us in filtering through 905 hits, some of which did not have a context of loneliness and differentiating them from loneliness-related tweets. For example, **الوحيدة** is not used in the context of loneliness but as “the only thing” in the search result that provided the right context of **اللي هنتفتح وكورونا لسه في ازدياد لكن مفيش حل** (It will open while COVID-19 is still increasing, but there is no solution), and the left context **عمرو أديب: "الرهان على وعي الشعب فاشل.. وأحنأ البلد** (Amr Adeeb: “Betting on the people’s awareness is a failure... and we are the country).

Collocate search

We ran a separate search for the collocates of each word through a basic search. **كورونا** (corona) was the most frequent collocate for the words **لوحدي** (alone), **وحيد** (lonely), **وحيدة** (lonely and loner), **واحد** (lonely), **الوحيدة** (lonely), and **الوحيد** (lonely and alone). The collocate for the words **وحيدا** (lonely or loner) and **منعزل** (lonely or secluded) was the word **في** (of). For the word **معزول** (lonely or secluded), the collocate was the word **بعد** (after). For the word **متوحد** (lonely), across the four lockdown months, the most frequent collocate was the word **البيت** (the house).

While the collocates for each searched word, along with frequency, we looked at the words **الوحيد** (lonely, masculine) and **الوحيدة** (alone or lonely, feminine) and their top collocate **كورونا** (corona).

The results for الوحيد (alone or lonely, masculine) and الوحيدة (alone or lonely, feminine), along with the top collocate كورونا (corona) of both words.

Keyword analysis

We ran a keyword analysis and listed the top 20 keywords in Table 3 according to keyness. Of the 20 keywords, many of the words were closely related and associated with COVID-19, such as كورونا (corona), فيروس (virus), كوفيد (COVID), بفيروس (with the virus), and الكورونا (The corona). A separate cluster of words closely related to public health included the words الصحة (health), حالة (case), المستجد (update), إصابة (infected), حالات (cases), and وفاة (death). The last word on the list is the word بالبيت (at home), commonly used with the twelfth keyword خليك (stay) as a hashtag # بالبيت - خليك (#stayathome) to express staying at home or informing others to stay at home, due to lockdown, isolation, or to not infect others. The top keyword was كورونا (corona).

Table 3. Keyword analysis of target corpus according to keyness.

Rank	Keyword	English	Keyness
1.	كورونا	Corona	2484.103
2.	فيروس	Virus	864.639
3.	كوفيد	Covid	241.580
4.	الصحة	Health	232.589
5.	حالة	Case	222.321
6.	الله	God	204.463
7.	المستجد	Update	160.517
8.	الكويت	Kuwait	145.347
9.	إصابة	Infected	141.837
10.	الجديد	New	127.127
11.	بفيروس	With the virus	107.719
12.	خليك	Stay	102.227
13.	عاجل	Urgent	101.479
14.	البيت	The house	94.713
15.	حالات	Cases	90.678
16.	الكورونا	The corona	89.572
17.	جديدة	New	85.742

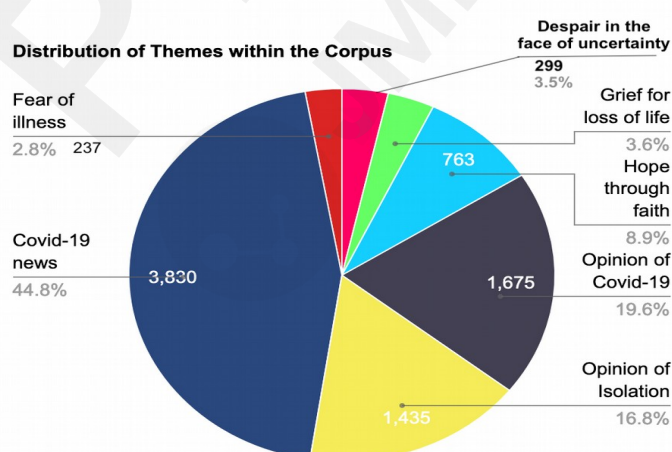
18.	وفاة	Death	82.464
19.	كلنا	Everyone	81.042
20.	بالبیت	At home	78.395

Theme identification

We analyzed prevailing thematic elements within the target corpus. The analytical process involved an analysis of 8,000 tweets, with a particular analysis of 100 tweets per month for each keyword. We selected 10 random tweets from each page per keyword to accumulate 100 tweets for each keyword for each month of the preliminary four lockdown months. The concordance feature was employed for viewing multiple hits and an in-depth review of the tweets identified by the keywords. This synergy of actions was enhanced, maximizing our ability to extract meaningful insights from the data.

Our analysis discerned prominent themes, such as despair in the face of uncertainty, grief for the loss of life, hope through faith, opinions on COVID-19, opinions on isolation, COVID-19 news, and fear of illness. These themes emerged through our meticulous examination of tweets within the corpus. Following categorizing 8,000 tweets based on the identified themes, we observed overarching patterns. The themes ranked as follows: opinions on COVID-19, opinion of isolation, hope through faith, grief for the loss of life, despair in the face of uncertainty, and fear of illness. The theme of COVID-19 news dominated 80% of the keywords, with exceptions for الله (God), البيت (the house), خليك (stay), and كلنا (everyone). The primary themes for the latter three keywords were Hope through faith and opinions on COVID-19. The distribution of themes within the corpus according to the 7 themes is presented in Figure 1.

Figure 1: Distribution of themes within the corpus according to the 7 themes.



Temporal trends

We discerned a gradual upswing in the articulation of loneliness over the ensuing weeks. The

identification of tweets associated with loneliness markedly intensified in May and June 2020. A noteworthy semantic distinction emerged in our examination, as the utilization of the term كورونا (corona) in tweets conveyed a more comfortable sentiment compared to the more serious tone invoked by the term كوفيد ١٩ (COVID-19) associated with news-related tweets. This observation was exemplified by contrasting tweets from June 25, 2020, where "COVID-19" was featured in a news context, while "corona" was humorously employed.

Individual expressions of loneliness during the pandemic exhibited temporal expansion over four months. Despair in the face of uncertainty was most pronounced in the initial two months when information about the pandemic was scarce, and awareness of the escalating loss of lives was heightened. A conspicuous rise in COVID-19-related news underscored the enduring thematic prevalence of pandemic-related content in public discourse. COVID-19-related news consistently held the highest thematic prevalence. Opinions on COVID-19 and isolation surged throughout the dataset, a trend attributed to lockdown restrictions and the escalating coverage of COVID-19 news. Fear of illness, albeit less prevalent, manifested in individuals who expressed concerns about contracting COVID-19, the potential loss of loved ones, the apprehension of an unending pandemic, and the fear of life never returning to pre-pandemic normalcy. Hope through faith emerged as a resilient theme, encapsulated in tweets that sought strength through religious prayers, extended prayers for isolated individuals or those who succumbed to COVID-19, and expressed hopefulness through prayer for a brighter post-pandemic future.

Inter annotator agreement (IAA)

Two annotators with graduate backgrounds in psychology and over 12 years of training in counseling were asked to annotate 200 tweets each according to the previously mentioned themes using Table 7 as a reference. There was an 82% average agreement score between the annotators. As an outcome, annotator 1 had an agreement of 84%, and annotator 2 had an agreement of 80%. Of the 200 tweets, the highest agreement stemmed from COVID-19 news, followed by hope through faith. Annotators agreed the least on despair in the face of uncertainty, followed by an opinion of isolation.

Results

Corpus analysis

Our corpus analysis through AntConc demonstrated multiple components about the origin of mental health struggles and loneliness during the initial emergence of COVID-19. Our study employed advanced linguistic analysis techniques. Users expressed diverse forms of loneliness on Twitter, employing various keywords and synonyms. Concordance analysis uncovered 18,241 occurrences of loneliness-related words and enriched our thematic understanding. Emotions were conveyed through themes, and COVID-19 news and opinions were prominently featured.

Synonyms of loneliness and the multiple tools available on AntConc led us to direct expressions of mental health struggles and expressions of loneliness. Loneliness was eminently expressed from the start of the pandemic. The expressions of loneliness manifested directly and diversely in Arabic social media discourse during the initial months of the COVID-19 pandemic. The advanced search option in AntConc directed the study to the Concordance tab, offering the frequency of each word and aiding in finding relevant tweets. The combined frequency of loneliness-related words in Table 2

was 18,299. Selecting the word **لوحدي** (alone) led to tweets associated with the term. A sample of results included tweets expressing post-COVID plans and reflections on the impact of the pandemic. The word **الوحيد** (lonely and alone) had a frequency of 2968, with associated tweets illustrating feelings of isolation and unique experiences. Table 3 presented concordance search results for each loneliness-related word and translated tweet examples, enhancing the understanding of their contextual usage.

كورونا (corona) emerged as the most frequent collocate for words like **لوحدي** (alone), **وحيد** (lonely), **وحيدة** (lonely and loner), **واحد** (lonely), **الوحيدة** (lonely), and **الوحيد** (lonely and alone). Those for **الوحيد** (alone or lonely, masculine) and **الوحيدة** (alone or lonely, feminine) highlighted the association between the words and the pandemic, revealing expressions of concern, indifference, and personal experiences related to COVID-19. The top 20 keywords listed in Table 3 highlighted clusters related to COVID-19 and public health (Table 3). **كورونا** (corona) emerged as the top keyword, reflecting the COVID-19 focus. A separate cluster included words related to public health, and the hashtag **#باليك - خليك** (#stayathome) emphasized staying at home due to lockdown or isolation. The analysis demonstrated the association between keywords and loneliness expression.

Based on AntConc's advanced search, all words were loaded simultaneously to get an overall idea (17367 hits). Each word was searched separately. **الوحيدة**, defined as a variation of lonely, could also mean "the only thing." KWIC search helped filter 905 hits, revealing instances where words were used without a context of loneliness. This approach differentiated loneliness-related tweets from those unrelated, providing a nuanced analysis. Individuals employed expressions related to loneliness and its synonyms to articulate their emotional experiences during the early stages of lockdown measures in different ways. Our study was focused on the origin of loneliness during the pandemic. Within the first four months of the lockdown period, we found the mental health expression within the dataset consisted of humor, sarcasm, sorrow, grief, and longing.

Thematic analysis and IAA

Our study identified seven overarching themes, each contributing to the multifaceted emotional landscape during this period. We inferred from the identified themes that COVID-19-related news consistently dominated discussions, indicating the sustained impact of the pandemic on public awareness. While grief for the loss of life was not the primary focus, the staggering global mortality reported by the WHO emphasized the profound impact. Fear of illness, though less prevalent, manifested in concerns about contracting COVID-19, the potential loss of loved ones, and the fear of a prolonged pandemic. Hope through faith emerged as a resilient theme, expressed in tweets seeking strength through religious prayers and anticipating a brighter post-pandemic future. The temporal trends illuminated complex emotional responses during the initial pandemic phase and highlighted social media's adaptability in capturing diverse human emotions amid unprecedented global events.

The IAA revealed the potential for varied interpretations among annotators, but a consensus was apparent, particularly in understanding opinions on COVID-19 and COVID-19-related news posts. Through our analyses, expressions of loneliness were prevalent from the onset of the pandemic, persisting throughout the initial lockdown months. Different Arabic variations of the term loneliness were employed, resulting in various tweets and expressions.

Discussion

Studies have examined mental health challenges and loneliness amidst crises and pandemics in a broader light. Our study conducted a more in-depth analysis by delving deeply into understanding the articulation of loneliness and exploring how individuals express their feelings during isolation. We derived insightful inferences from thematic elements associated with COVID-19-related loneliness within the context of the Arabic-speaking population.

The increased usage of social media during a global health crisis provided useful resources for emotional, informational, and peer support [16]. The concordance feature is pivotal in facilitating thematic analysis, providing access to tweet segments for examination based on keywords and keyness levels. Building on this, our research sought to address a gap in the existing corpus analysis by exploring loneliness in Arabic, a language for which no such analysis was available online. Through this investigation, we gained valuable insight into the experiences of individuals in the Arabic-speaking community during the pandemic, identifying areas for further research and potential avenues for support and intervention.

There is a correlation between the duration of time spent in social isolation and the onset of loneliness. Spending more than 75% of one's time alone was associated with significantly higher loneliness scores [17]. Our study discerned that manifestations of loneliness in Arabic social media discourse during the initial months of the COVID-19 pandemic were evident from the commencement of lockdown measures, demonstrating a unique and diverse array of expressions. Users exhibited a lack of conformity to a singular form of expression and employed various keywords and multiple identified synonyms associated with loneliness to articulate their thoughts and emotions during this critical lockdown period.

Our study harnessed AntConc's concordance tab to examine the frequency of loneliness-related words in the target corpus. A total of 18,241 occurrences were complemented by examples showcasing tweets associated with specific words and offering insights into the diverse contexts surrounding expressions of loneliness. Concordance search results for each word contributed valuable insights into the prevalence and nuances of loneliness expressions within the dataset. This analytical feature facilitated the identification of pertinent tweets and enriched our understanding of the multifaceted contexts in which loneliness was articulated on social media during the specified timeframe.

Users employed expressions related to loneliness and its synonyms to convey emotional experiences during the early stages of lockdown measures. Expressions spanned a spectrum from humor and fear to sarcasm and sorrow. The lack of expression was observed through retweeting or sharing news updates. Themes played a pivotal role, with the perspective transforming from mere COVID-19 news to opinions when added to a news update or a retweet. The demonstration of user expression through the KWIC search highlights the prevalence of emotional expression by utilizing diverse AntConc features.

We conducted a thematic analysis of themes to identify, analyze, and interpret patterns of meaning [18]. Our analysis revealed distinct themes that exhibited varying prevalence within 8,000 tweets. A predominant theme permeating the corpus is COVID-19 news, indicating the pervasive influence of pandemic-related information. The theme of opinions on COVID-19 emerged prominently, and Twitter serves as a platform to articulate users' thoughts and expressions regarding the emergence and comprehension of the virus.

The process of theme identification offers valuable insights into pertinent patterns within the data, especially during the lockdown period of COVID-19, demonstrating that while there is a prevalence

of mental health struggles expressed during this early stage of the pandemic, it was done through different themes. This analytical approach deepens our comprehension by elucidating that individuals while expressing thoughts, emotions, or news related to COVID-19, imply that users' experiences of loneliness were intricately influenced by these overarching themes. Users whose themes were characterized as despair in the face of uncertainty (3.5%), grief for the loss of life (3.6%), or fear of illness (2.8%) exemplified a tangible connection between prevailing themes and the articulation of loneliness in their online expressions. This refined perspective sheds light on the complex interplay between broader thematic elements and the manifestation of loneliness within the unique context of the COVID-19 lockdown.

Limitations and Future Research

While our research presents a comprehensive analysis of loneliness within the target corpus, several limitations offer avenues for further enrichment within this domain. Notably, the absence of corpora in the Arabic language specifically focused on loneliness, both pre and post the Covid-19 pandemic, poses a significant constraint. While our aim was to contribute a prominent research perspective, the scarcity of similar research in this area is evident. Although our study demonstrates the prevalence of loneliness on social media, the refined understanding of word and phrase usage is lacking. Additionally, our study focuses exclusively on Arabic Twitter data, which may limit its generalizability to other cultures or languages.

One limitation pertains to the necessity for extensive data cleanup before analysis, as failure to do so significantly impacts results within AntConc. Cleaning the corpus involves standardizing text and removing irrelevant characters, a meticulous process undertaken manually to ensure data accuracy. Despite focusing our data analysis on the months of March to June 2020, the sheer volume of data made this a time-consuming task, crucial for preventing skewed results. Temporal specificity could also be considered a limitation as our study is confined to the early stages of the pandemic, and its findings may or may not apply as the situation evolves.

In our research, we grappled with the semantic range, reflecting the intricate variations in word meanings based on context. Utilizing a list of designated words introduced a challenge during corpus analysis, where words were inconsistently defined within the specific context of loneliness. The employment of AntConc's advanced search feature for Key Words in Context (KWIC) searches allowed for nuanced exploration of contextual word usage, revealing instances where words extended beyond loneliness. Additionally, the textual analysis tools and techniques used, while robust, might miss nuanced expressions of loneliness.

One data limitation for methodological constraints is the reliance on a specific dataset (ArCov-19), which might introduce biases or limitations in the types of tweets captured. Moreover, due to time constraints, our research scope was limited to analyzing the theme of 8,000 tweets within the target corpus, whereas a more extensive analysis would have provided a deeper understanding of prevalent themes. Additionally, an emotion analysis of loneliness-related tweets would have added depth to our responses to RQ1 and RQ2, offering insights into emotional patterns and mental health struggles within the dataset.

While recognizing these limitations, we underscore their significance in shaping future research opportunities. The identified constraints pave the way for more extensive corpus analyses on loneliness related to Covid-19 in the Arabic language, contributing to a richer understanding of this complex phenomenon.

Conclusion

Our AntConc comprehensive analysis has yielded significant insights into the emergence of mental health struggles and loneliness during the initial phases of the COVID-19 pandemic. Our study offers a profound understanding of how individuals articulated and encountered loneliness within the Arabic-speaking online community during the early phases of the COVID-19 pandemic. This analysis contributes valuable insights to the broader scientific discourse on mental health, social media dynamics, and the evolving emotional terrain in response to global crises. By adapting similar methodologies to diverse linguistic contexts, scholars can contribute to a more comprehensive understanding of both universal and culture-specific aspects of loneliness, advancing our collective knowledge of the pandemic's impact on mental well-being worldwide.

References

- [1] J. Yanguas, S. Pinazo-Henandis, and F. J. Tarazona-Santabalbina, "The complexity of loneliness," *Acta Bio Medica Atenei Parm.*, vol. 89, no. 2, pp. 302–314, 2018, doi: 10.23750/abm.v89i2.7404.
- [2] L. G. E. Smith and S. Gibson, "Social psychological theory and research on the novel coronavirus disease (COVID-19) pandemic: Introduction to the rapid response special section," *Br. J. Soc. Psychol.*, vol. 59, no. 3, pp. 571–583, Jul. 2020, doi: 10.1111/bjso.12402.
- [3] A. Dhir, "Online social media fatigue and psychological wellbeing—A study of compulsive use, fear of missing out, fatigue, anxiety and depression," *Int. J. Inf. Manag.*, vol. 40, Mar. 2018, doi: 10.1016/j.ijinfomgt.2018.01.012.
- [4] K. Sailunaz, M. Dhaliwal, J. Rokne, and R. Alhajj, "Emotion detection from text and speech: a survey," *Soc. Netw. Anal. Min.*, vol. 8, no. 1, p. 28, Apr. 2018, doi: 10.1007/s13278-018-0505-2.
- [5] M. Burke and R. E. Kraut, "The Relationship Between Facebook Use and Well-Being Depends on Communication Type and Tie Strength," *J. Comput.-Mediat. Commun.*, vol. 21, no. 4, pp. 265–281, Jul. 2016, doi: 10.1111/jcc4.12162.
- [6] M. De Choudhury, M. Gamon, S. Counts, and E. Horvitz, "Predicting Depression via Social Media," *Proc. Int. AAAI Conf. Web Soc. Media*, vol. 7, no. 1, Art. no. 1, 2013, doi: 10.1609/icwsm.v7i1.14432.
- [7] L. Wilkialis *et al.*, "Social Isolation, Loneliness and Generalized Anxiety: Implications and Associations during the COVID-19 Quarantine," *Brain Sci.*, vol. 11, no. 12, Art. no. 12, Dec. 2021, doi: 10.3390/brainsci11121620.
- [8] F. Haouari, M. Hasanain, R. Suwaileh, and T. Elsayed, "ArCOV-19: The First Arabic COVID-19 Twitter Dataset with Propagation Networks." arXiv, Mar. 13, 2021. doi: 10.48550/arXiv.2004.05861.
- [9] A. Alqarni and A. Rahman, "Arabic Tweets-Based Sentiment Analysis to Investigate the Impact of COVID-19 in KSA: A Deep Learning Approach," *Big Data Cogn. Comput.*, vol. 7, no. 1, Art. no. 1, Mar. 2023, doi: 10.3390/bdcc7010016.
- [10] C. N. Ngonghala, E. A. Iboi, and A. B. Gumel, "Could masks curtail the post-lockdown resurgence of COVID-19 in the US?," *Math. Biosci.*, vol. 329, p. 108452, Nov. 2020, doi: 10.1016/j.mbs.2020.108452.
- [11] A. B. Said, A. Erradi, H. Aly, and A. Mohamed, "A deep-learning model for evaluating and predicting the impact of lockdown policies on COVID-19 cases." arXiv, Sep. 11, 2020. doi: 10.48550/arXiv.2009.05481.
- [12] A. Shanableh *et al.*, "Effects of the COVID-19 lockdown and recovery on People's mobility and air quality in the United Arab Emirates using satellite and ground observations," *Remote Sens. Appl. Soc. Environ.*, vol. 26, p. 100757, Apr. 2022, doi: 10.1016/j.rsase.2022.100757.
- [13] D. Cucinotta and M. Vanelli, "WHO Declares COVID-19 a Pandemic," *Acta Bio Medica Atenei Parm.*, vol. 91, no. 1, pp. 157–160, 2020, doi: 10.23750/abm.v91i1.9397.
- [14] "Cambridge English–Arabic Dictionary." Accessed: Nov. 16, 2023. [Online]. Available: <https://dictionary.cambridge.org/dictionary/english-arabic/>
- [15] "Almaany Arabic dictionary." Accessed: Nov. 21, 2023. [Online]. Available: <https://www.almaany.com/>
- [16] J. Abbas, D. Wang, Z. Su, and A. Ziapour, "The Role of Social Media in the Advent of COVID-19 Pandemic: Crisis Management, Mental Health Challenges and Implications," *Risk Manag. Healthc. Policy*, vol. 14, pp. 1917–1932, May 2021, doi: 10.2147/RMHP.S284313.
- [17] A. F. Danvers *et al.*, "Loneliness and time alone in everyday life: A descriptive-exploratory study of subjective and objective social isolation," *J. Res. Personal.*, vol. 107, p. 104426, Dec. 2023, doi: 10.1016/j.jrp.2023.104426.
- [18] V. Clarke and V. Braun, "Thematic analysis: The Journal of Positive Psychology: Vol 12, No 3." Accessed: Feb. 03, 2024. [Online]. Available: <https://www.tandfonline.com/doi/full/10.1080/17439760.2016.1262613>

Preprint
JMIR Publications