

Research paper

A Cross-Cultural Study of the Most Frequent Words in the News during COVID-19

Samira Hashemi

Novinborna Language Institute

Citation

Hashemi, S. (2023). A cross-cultural study of the most frequent words in the news during Covid-19.

Journal of new advances in English Language Teaching and Applied Linguistics, 5(1), 1232-1242.

 10.22034/jeltal.2023.5.1.12

Received

2022-12-11

Revised

2023-02-26

Accepted

2023-05-30

Keywords:

collocates,
cross-cultural study,
keyword nouns,
keyword verbs,
news discourse,
recurrent words

Abstract

Due to the global impact of the Corona pandemic, COVID-19-related news had become a top priority issue in the world news. This cross-cultural study aims to investigate the most common English and Persian keywords that co-occurred with the collocates employed in the virus-related news. Based on a 1,247,523-word corpus, this study examined how the news about COVID-19 had evolved throughout 2021. A total of 2400 COVID-19-related news, 1200 of which were in English and 1200 in Persian were randomly selected from news agency websites and analyzed using AntConc to determine the most recurrent keywords and their first collocations in both corpora. Frequent keyword nouns, first collocation of these nouns, and frequent keyword verbs were explored to understand the most important public's concerns that are reflected in news media. The results indicated what the most prevalent keyword nouns are and how they get along with collocates, besides the most recurrent keyword verbs during the pandemic in both corpora. This cross-cultural corpus-based descriptive and analytical study assists Iranian medical students, teachers, and researchers by providing a comprehensive understanding of English words related to pandemics.

*Corresponding Author: Samira Hashemi

Address: Novinborna Language Institute, Pnaje Azar Street, Esfahan, Iran Postal Code:8138995881,

Tel: (+98) 9371261765 **E-mail:** s.hashemi_1990@yahoo.com

Introduction

By January 2021, about 84 million had been infected with Covid-19 and almost 1.9 million people had died (World Health Organization, 2021). With the much damage this epidemic has caused to various businesses and many people, it is not surprising that under the influence of this pandemic, various news agencies, the mass media, social networks, and newspapers publish a huge number of COVID-19-related news daily. This is because it is the first time in history that technology has been used extensively during a pandemic to maintain safety, inform,

connect, and facilitate productivity (Igorovich, 2021). On one hand, studying online news has recently become more preferred and popular than other traditional ways due to their easier accessibility (Gottfried & Shearer, 2017). On the other hand, with the growth of digital technologies, news media organizations and public health agencies have less control over what will be published especially during the crisis (Kostkova, 2013). Since the official and government news agencies are mirrors of society, which manage the public's opinions and have made the Covid-19 related news widely covered in the world by the international press (Prieto-Ramos & Cheng, 2020) and because news is a crucial channel for the public to get information (Yu et al., 2020), therefore, using more reliable media or websites is preferable in these situations.

Researching the use of language during the pandemic had become very important and these studies were conducted in a variety of fields such as Saed et. al.'s (2022) list of the lemmatized words for ESP learners and journalists, Ghaffori's (2022) nonverbal politeness, Setiawati and Budiasih's (2022) strategies on teaching reading comprehension, Orts and Vargas-Sierra's (2022) critical discourse analysis study of newspapers.

This study attempts to investigate the most frequent words used on news agency websites to show how the Covid-19 related news had changed over a year, and what was the biggest concern of the press regarding this pandemic. The official and government news agencies inform the public with the daily top news to keep them safe. Some previous empirical studies were conducted to investigate how language is used to present COVID-19-related news. for example, COVID-19 lexemes in The Economist website (Katermina & Yachenko, 2020), collocations in a Malaysian online newspaper (Nor & Zulcafli, 2020), and the most frequent nouns and verbs in the media (Jiang & Hyland, 2022).

A significant aspect of this study is that it compares the most frequent words in English with that in Persian which has received relatively less research. Consequently, the objectives of the current study are as follows:

1. What are the most frequent words including nouns and verbs in 2400 news texts in English and Persian corpora?
2. What are the first collocates of the most recurrent nouns in both corpora?

Literature Review

Analytical Approach

The current study attempts to concentrate on linguistic items by identifying keyword nouns and keyword verbs in monthly news about COVID-19. **Keywords** are word forms and clusters of words that have a statistically higher frequency than expected in a corpus (Hyland & Jiang, 2020), and they can demonstrate what the corpus is about by highlighting important themes and subjects as well (Scott & Tribble, 2006). Keywords also operate as the gist and carry the main idea and subject matter (Peng & Hu, 2022). About every fourth word in English is a **noun**, making it the most frequent word class (Jiang & Hyland, 2022). **Verbs** have a significant role in a sentence, syntactically, and show activities, states, or processes, semantically (Ruminda,

2016). The verbs are essential for news writing since they convey action and provide movement and momentum to the story (Handford, 2013). Finally, **collocations** as the tendency for words to join regularly in a language, provide naturalness of language since they can explain the whole meaning of the individual components (Rahmyani, 2022). “collocations are to be defined as the habitual and recurrent juxtaposition of semantically related words” (Bartsch & Evert, 2014, p. 48)

Previous Studies

In the pandemic, as a result of business closures, stock market declines, job losses, travel restrictions, lockdowns, working from home, and school closures, the use of media has been strengthened. In other words, news media as a significant platform for informing the public, helps people to stay up-to-date (Fletcher et al., 2020). Some previous studies (e.g., Katermina & Yachenko, 2020; Nor and Zulcafli, 2020; Jiang and Hyland, 2022) investigated lexemes, collocations, keyword nouns, keyword verbs, and noun phrases in different online newspapers as follows:

Katermina and Yachenko (2020) examined COVID-19 lexemes in the eight issues of The Economist website consisting of 84 pages on average, in which the word “COVID-19” is used 1216 times. They worked on word combinations, neologisms, collocations, and the contexts of the word *Covid-19* and its synonyms. They found that the words *coronavirus* and *Covid-19* were used frequently in the noun-noun type with such lexemes as *outbreak*, *pandemic*, *crisis*, etc., They also applied with such verbs as *to destroy*, *to attack*, *to rage*, etc., and with adjectives like *new* and *novelty* in the word combinations. Coronavirus and COVID-19 lexemes as productive stems formed new words such as *corohara* (short for coronavirus harassment)—an act of harassment of people coughing in public places in neologisms. The findings revealed that due to the COVID-19 pandemic, English culture has been affected, and new words and combinations have been added to English vocabulary to make new connotations with the lexemes that have already existed. In other words, this study showed how mass media discourse influences public awareness, opinion, language, and culture. Nor and Zulcafli (2020) analyzed the top frequent collocations in 1018 news articles on COVID-19 issues in The Star Online news. Their findings showed how Malaysians feel about the virus and different variants of Covid-19 in Malaysia, and how the Malaysian government deals with the virus as well. In addition, considerable use of top frequent nouns such as *outbreak*, *cases*, *patients*, and *tests* and top frequent verbs such as *spread*, *fight*, *prevent*, *died*, *combat*, and *battling* co-occurred with the word COVID-19 showed feelings of threat and fear of the virus. Jiang and Hyland (2022) investigated frequent keyword nouns, verbs, and noun phrases in 120,000 news texts from the NOW corpus (News on The Web), totaling 12.3 million words from January to December 2020. The most common words in their corpus were *symptom*, *guideline*, and *restriction* that showed the news in the early months of the outbreak was primarily focused on the symptoms of the virus with guidelines and protocols. The increased usage of noun phrases such as face mask, herd immunity, infection prevention, and contact tracing which were less common in the last decades or somehow new, had changed the focus of the news because of their frequent use, thereby showing how news tracked the pandemic from its beginnings. *Reopening* and *solution* were two frequent words after six months passed until the word *vaccine*

became increasingly prominent as the disease progressed. As a result, their study confirms this quotation of Oxford Languages (2020, p.3) "The English language, like all of us, has had to adapt rapidly and repeatedly this year."

Methodology

To explore the lexical coverage of Covid-19 related keyword nouns and keyword verbs in news, a 1,247,523-word corpus consisting of 679,205 words in English and 568,318 words in Persian, from 2400 news associated with Covid-19 including 1200 news in English and 1200 news in Persian (i.e., 100 English news and 100 Persian news per month) was created (Table 1).

Table 1.
Corpus information

Corpora	Number of news	Number of words
English corpus	1200	679,205
Persian corpus	1200	568,318
Total	2400	1,247,523

From January 2021 to December 2021, about one year had passed since pandemic and the discourse in news websites had become more different than a year before; because scientists and the public had become more aware about that novel disease. Consequently, the current study's data were randomly collected from four official and government news agency websites including two English news agency websites (i.e., <https://www.upi.com> and <https://www.reuters.com>) and two Persian news agency websites (i.e., <https://www.irna.ir> and <https://www.isna.ir>) consisted of 200 Covid-19 related news per month from January 2021 to December 2021 (Tables 2 and 3).

Table 2.
Frequency of news and words in the English corpus by month

Months	January	February	March	April	May	June
News	100	100	100	100	100	100
words	61089	57449	67149	59545	65916	55858
Months	July	August	September	October	November	December
News	100	100	100	100	100	100
words	46495	54344	56878	51590	47376	56516

Table 3.
Frequency of news and words in the Persian corpus by month

Months	Jan.(دی)	Feb.(بهمن)	Mar.(اسفند)	Apr.(فروردین)	May(اردیبهشت)	Jun.(خرداد)
News	100	100	100	100	100	100
Words	48783	47424	50068	42614	53002	47366
Months	Jul.(تیر)	Aug.(مرداد)	Sep.(شهریور)	Oct.(مهر)	Nov.(آبان)	Dec.(آذر)
News	100	100	100	100	100	100
Words	47342	53604	41637	44020	42499	49959

Data Analysis

In order to analyze the data, first, they were converted into text format to be entered into AntConc version 3.5.8. Next, all of the frequent grammar words were removed since the data

was about the news. Also, the words coronavirus, COVID, and COVID-19 were filtered out because the corpora were about this disease and it was obvious that those terms were the most frequent ones. Afterward, the top frequent words that appeared in a year in both English and Persian corpora separately were identified by this software, and they were analyzed with their left or right collocations. Then the most frequent keyword nouns and keyword verbs were recognized monthly to analyze them more accurately in the news since monthly analysis shows the changing of the topics month by month and this is why this research was carried out. For more reliability, the findings were checked by a rater manually to ensure whether they were meaningful and relevant or not. Finally, the Persian examples selected for the result section of the article were translated into English using Google Translate. Given that the author is fluent in both English and Persian language, the translated examples were edited by her and then they were adapted to English.

Result

In this section, first, the analysis of the most common words overall in one year used in news agency websites in English and Persian is presented. Since the keyword analysis may be used to compare and determine the words that are significantly more recurrent in two corpora (Kheovichai, 2022), the current study demonstrated that besides people, disease, and vaccine which are the three most recurrent keywords overall in the English corpus, people(مردم), (استان)province, and vaccine(واکسن) are the three most frequent ones in the Persian counterpart as well. In the following tables show the words which were analyzed in the text based on the words they frequently co-occurred with, and then checked and selected with their first left collocates.

Overall Distribution of the First Left Collocates in English Corpus

Table 4 displays the overall distribution of the first left collocates of people, disease, and vaccine in the English corpus. The findings indicate that *COVID-19*, *million*, and *severe* were the most frequently used collocates in English corpus, respectively.

Table 4.
The first left collocates of vaccine, people, and disease

	people			disease			vaccine		
Jan.	million	infected	young	heart	Coronavirus	severe	Covid-19	coronavirus	biontech
Feb.	million	Infected	younger	severe	coronavirus	heart	Covid-19	Astrazeneca	coronavirus
Mar.	million	vaccinated	homeless	severe	Coronavirus	heart	Covid-19	Astrazeneca	pfizer
Apr.	million	older	young	Coronavirus	severe	respiratory	Covid-19	Astrazeneca	mRNA
May	million	younger	healthy	Coronavirus	Severe	Rare	Covid-19	variant	rapid
Jun.	million	vaccinated	young	Coronavirus	critical	Severe	Covid-19	coronavirus	Astrazeneca
Jul.	million	vaccinated	unvaccinated	coronavirus	symptomatic	chronic	Covid-19	coronavirus	sinopharm
Aug.	pregnant	million	vaccinated	coronavirus	severe	autoimmune	Covid-19	pfizer	biontech
Sept.	pregnant	vaccinated	million	severe	coronavirus	heart	Covid-19	Moderna	pfizer
Oct.	million	young	pregnant	severe	preventable	moderate	Covid-19	pfizer	biontech
Nov.	million	vulnerable	unvaccinated	severe	milder	heart	Covid-19	pfizer	Moderna
Dec.	million	vaccinated	unvaccinated	severe	heart	kidney	Covid-19	pfizer	coronavirus

In the following section, some collocates along with some exemplifications from the English corpus are presented.

Exemplifications of the First Left Collocates in English: People, Disease, and Vaccine

Table 4 indicates the distribution of the most recurrent left collocates of *people*. The first word that catches the attention is *million*. It shows that millions of people were involved in this disease. Examples (1) and (2) indicate that other collocations in the first 2 months such as *infected* and *young* show that many young people were infected with the coronavirus.

- (1) *Adults between 18 and 24 years old account for nearly 60% of new COVID-19 cases among young people* (Dunleavy, 2021). (Jan.)
- (2) *At least half of COVID-19 transmission globally may have been caused by symptom-free infected people unknowingly spreading the virus to others* (Dunleavy, 2021).

As COVID-19 continued in the second six months of the year, collocates for *people* became more specific. Examples (3), (4), (5), and (6) indicate that as much as the press strongly suggested that people who had not been *vaccinated* get *vaccinated* as quickly as they could, it also encouraged *pregnant* and *vulnerable* people to get *vaccinated* as well.

- (3) *fully vaccinated people do not need masks outdoors or in most indoor spaces. It also said that fully vaccinated people will not need to physically distance in most places* (Reuters, 2021).
- (4) *CDC encourages all pregnant people or people who are thinking about becoming pregnant and those breastfeeding to get vaccinated to protect themselves from COVID-19* (Dunleavy, 2021). (Aug.)
- (5) *unvaccinated people are 29 times more likely to be hospitalized for COVID-19 than fully vaccinated people* (Norton, 2021).
- (6) *"AZD7442 ... can be used to protect vulnerable people from getting COVID-19 and can also help prevent progression to severe disease," he said* (Dunleavy, 2021). (Nov.)

Table 4 illustrates the most frequent first-left collocates of *disease* too. In Example (7) the word *severe* was a similarly prevalent term that co-occurred with *disease* throughout the year, besides, using the word *critical* with *disease* in example (8) shows that the news agencies tended to indicate the acuteness of this disease.

- (7) *The immunity offered by two doses of Pfizer's coronavirus vaccine drops off after about two months, although protection against severe disease, hospitalization, and death holds strong* (HealthDay News, 2021).
- (8) *"It's much easier to prevent severe critical disease, and I think you're much more likely to have longer lasting protection against severe critical disease," he said* (Thompson, 2021). (Jun.)
- (9) *The researchers also found that COVID-19 patients' rates of new diagnoses of respiratory disease, cardiovascular disease and diabetes were 27, 3 and 1.5 times greater, respectively, than those observed in the control group* (Mundell & Preidt, 2021). (Apr.)

- (10) *As many as two of every five Americans who've died from COVID-19 were suffering from diabetes, making the chronic disease one of the highest-risk conditions during the pandemic, an expert says (Thompson, 2021). (Jul.)*
- (11) *Young people with pre-existing conditions -- such as obesity, chronic kidney disease, heart disease or immune disorders -- are 25 times more likely to get severe COVID-19 than kids without pre-existing conditions (HealthDay News, 2021).*

In the above-cited examples (9), (10), and (11), the press attempted to inform people with underlying diseases such as respiratory, chronic, kidney, and heart diseases about the risk of catching COVID. Some other collocates that co-occurred with *disease* such as *preventable* in example (12), *moderate* in example (13), *symptomatic* in example (14), and *milder* in example (15) were employed in the news texts which show these words represent different aspects of the disease because the disease was new and the way to deal with was different.

- (12) *Soon, this virus will be a preventable disease if people choose to get vaccinated.*
- (13) *Booster shots are necessary because "vaccine effectiveness against mild and moderate disease appears to wane over time," which could lead to blood clots and "long COVID" in some cases.*
- (14) *Pfizer said it, BioNTech and Israel's health ministry examined the effects of the vaccine between January 17 and March 6 and found it to be 97% effective in preventing symptomatic disease, severe and critical disease and death.*
- (15) *The slow-motion pandemics of obesity and chronic illness have made the acute COVID-19 pandemic far worse than it needed to be, both among those hospitalized, and even among those with milder disease.*

As shown in Table 4, the most frequent left collocate of the *vaccine* was *COVID-19* from January 2021 to December 2021. It demonstrates that *COVID-19* vaccines were portrayed as something new and unpredictable in these 12 months, because of the novelty of the virus. Additionally, in examples (16), (17), (18), and (19) the vaccine term was collocated with *AstraZeneca*, *Sinopharm*, *Moderna*, and *Pfizer*, respectively more than other words, all of which show the process of making and injecting different types of vaccines, the level of effectiveness, and the comparison of some of them:

- (16) *The AstraZeneca vaccine is still under review for emergency use approval by the WHO but the organization said 35.3 million doses of the drug will be arriving to these countries in the first stage of shipments (Coote, 2021). (Feb.)*
- (17) *Hungary has approved a third COVID-19 vaccine shot that would be administered on a case-by-case basis, after local media reported Hungarians vaccinated with China's Sinopharm vaccine did not reach the required level of antibodies against the novel coronavirus (Shim, 2021). (Jul.)*
- (18) *The Moderna vaccine appears to be most effective at preventing serious COVID-19, but both it and Pfizer-BioNTech shot offer the most protection, the CDC says (Dunleavy, 2021). (Sep.)*

(19) Although the Pfizer vaccine appears to be weaker against Omicron, it still offers about 70% protection against severe COVID-19, the study said (Hughes, 2021). (Dec.)

The following section describes the distribution of collocations in the Persian corpus with some exemplifications.

Overall Distribution of the First Right Collocates in Persian Corpus

The overall distribution of the first right collocates of people (مردم), province (استان), and vaccine (واکسن) in Persian corpus are shown in Table 5. The findings show that injection (تزریق) cooperation (همراهی), and state (وضعیت) were the most recurrent collocates the press employed in Persian corpus.

Table 5.

The first right collocates of people (مردم), province (استان), and vaccine (واکسن) in the Persian corpus

	دی (Jan.)	بهمن (Feb.)	اسفند (Mar.)	فروردین (Apr.)	اردیبهشت (May)	خرداد (Jun.)	تیر (Jul.)	مرداد (Aug.)	شهریور (Sep.)	مهر (Oct.)	آبان (Nov.)	آذر (Dec.)
مردم (people)	همکاری	همکاری	همه	توسط	سلامت	همه	مشارکت	سلامت	همکاری	درصد	جان	درصد
	همراهی	همه	سلامت	همراهی	همراهی	نقش	عمومی	جان	عمومی	سلامت	واکسیناسیون	همراهی
	مشارکت	همراهی	همکاری	بیشتر	روان	عموم	همراهی	همراهی	سلامت	همه	همه	سلامت
استان (province)	مردم	سطح	وضعیت	وضعیت	دانش آموزان	مردم	مردم	وضعیت	جنوب	مردم	مردم	وضعیت
	وضعیت	هلال احمر	سطح	سطح	مرکز	وضعیت	مناطق	جنوب	مناطق	مردم	مناطق	وضعیت
	مدارس	وضعیت	مردم	کل	وضعیت	نخستین	جمعیت	فرهنگیان	فرهنگیان	مدارس	فرهنگیان	دانش آموزان
واکسن (vaccine)	تزریق	ساخت	محموله	تزریق	ساخت	تزریق	تولید	تزریق	تزریق	دوز	دوز	تزریق
	تزریق	تولید	تزریق	دوز	تولید	دوز	دوز	دوز	دوز	دوز	تولید	سوم
	دریافت	تهیه	دوز	انسانی	دریافت	اول	تزریق	دریافت	دریافت	دریافت	بالینی	دوم

In the following section, some collocates along with some translated examples from the Persian corpus are explained.

Exemplifications of the First Right Collocates in Persian: People (مردم), Province (استان), and Vaccine (واکسن)

As shown in Table 5 the first right collocates of *people* (مردم). These words are cooperation (همراهی), all (همه), and health (سلامت). The following examples (20), (21), and (22) indicate that all people's cooperation is essential in order to keep them safe and healthy during the pandemic.

(20) برای خارج شدن از وضعیت قرمز کرونایی، تنها با همراهی مردم می توانیم موفق عمل کنیم.

'To get out of the red coronavirus situation, we can succeed only with the people's cooperation.'

(21) رعایت پروتکل های بهداشتی برای همه مردم الزامی است و حتماً باید برای حفظ سلامت یکدیگر در محل های تجمع از ماسک استفاده کنیم.

'Compliance with health protocols is mandatory for all people, and we must use masks to protect each other's health in gathering places.'

(22) عادی انگاری کرونا مهمترین تهدید سلامت مردم است.

'Taking corona as normal is the most important threat to people's health.'

Table 5 also displays the most frequently used first-right collocates of *province*(استان) too. In example (23) the use of the word *Status* (وضعیت) as the first frequent word that co-occurred with *province*(استان) over the year shows the Corona situations in different provinces with Iran's color-coded scale..

(23) باید توجه کنیم برگزاری برنامه‌های اربعین باعث تغییر رنگ وضعیت استان نشود تا به‌زودی بتوانیم وارد رنگ زرد شویم.

*We should pay attention that holding the Arbaeen programs does not change the color-coded COVID-19 of the **province status** so that we can enter the yellow color soon.*

(24) میزان مراجعه بیماران بستری در بیمارستان های استان رو به افزایش است و امید می رود که مردم استان با رعایت دستورالعمل ها و پروتکل های بهداشتی، در کنترل وضعیت به ما کمک کنند.

*'The number of hospitalized patients in the hospitals of the province is increasing and it is hoped that the **people of the province** will help us control the situation by following the guidelines and health protocols.'*

In the above example (24), the significant role of people in controlling critical situations is shown by the news agency authors. Table 5 demonstrates that injection (تزریق), get (دریافت), and production (تولید) are the most frequent collocates of the word vaccine (واکسن) throughout the year. In examples (25) and (26) the authors used injection (تزریق) and get (دریافت) as the first collocates of the term vaccine (واکسن) that show vaccination should be done as soon as possible to prevent the spread of COVID. It is important to know that in example (26), the word *get* (دریافت) is a verb in English and it is a noun in Persian.

(25) حضور دانش‌آموزان و دانشجویان در کلاس‌های درس منوط به تزریق واکسن است.

*'Attendance of students in classrooms depends on **vaccine injection**.'*

(26) از همه دانش‌آموزان و خانواده‌های‌شان درخواست می‌کنیم در اسرع وقت با مراجعه به مراکز واکسیناسیون نسبت به دریافت واکسن اقدام نمایند.

*'We request all students and their families to **get the vaccine** as soon as possible by visiting the vaccination centers.'*

Moreover, example (27) indicates that vaccine production faced many challenges during the pandemic because the virus was changing quickly.

(27) در خصوص تولید واکسن این بیماری برای کنترل اثربخشی یک واکسن باید فازهای مختلف بالینی طی شود و به نظر می رسد تا چندماه آینده واکسن در دسترس نخواهد بود به همین دلیل فعلا بهترین راهکار پیشگیری از ابتلا به ویروس کرونا استفاده از ماسک و رعایت فاصله گذاری اجتماعی است.

*'Regarding the **vaccine production** for this disease, in order to control the effectiveness of a vaccine, different clinical phases must be completed, and it seems that the vaccine will not be available for the next few months, for this reason, the best way to prevent the coronavirus is to use a mask and observe social distancing.'*

The following sections explain the distribution of the ten top keyword nouns in English and Persian corpus monthly.

Changing of the Topics by Keyword Nouns in the English and Persian News

The use of keyness shows that the occurrence of a particular item is more recurrent than what would be assumed by coincidence, thus, it assists in identifying the primary topic or fundamental concepts of a set of texts (Jiang & Hyland, 2022).

Keyword Nouns in the English News

Table 6 indicates the keyword nouns for each month of 2021. As previously mentioned, the news was dominated by *people*, *disease*, and *vaccine* across the year. These words together with *study* and *case* appear in the top ten keywords nearly every month, in example (28) the word *study* was used to express that scientists had been studying this disease and the way they could prevent the growth of this virus during that year. Vaccinated or Covid-19 *Cases* and *deaths* were regularly reported as it is shown in example (29). However, the number of deaths was lower announced by the press in some months.

- (28) *The vaccine appears safe and effective since throughout the study period there were no significant cases of COVID-19 infection in the study population (Dunleavy, 2021).*
- (29) *Since the start of the pandemic, the United States has recorded more than 21 million cases and 357,385 deaths (Jacobson, 2021).*

Table 6.
Significant noun keywords in the English corpus on a monthly basis 2021

Jan.	Feb.	Mar.	Apr.	May	Jun.	Jul.	Aug.	Sep.	Oct.	Nov.	Dec.
Cases	Vaccine	Vaccine	Vaccine	Pandemic	Vaccine	Vaccine	Vaccine	Vaccine	Vaccine	Vaccine	Cases
Vaccine	Cases	People	Disease	Disease	People	People	People	People	People	People	Omicron
People	People	Health	Year	People	Study	Health	Cases	Study	Study	Health	Variant
Disease	Disease	Disease	Health	Vaccine	Cases	Disease	Delta	Dose	Children	Cases	Deaths
Health	Health	Cases	People	Year	Disease	Cases	Health	Disease	Booster	Pfizer	Vaccine
Deaths	Dose	Pandemic	Pandemic	Cases	Pandemic	Study	Variant	Health	Health	Booster	Health
Dose	Deaths	Study	Dose	Government	Health	Dose	Disease	Pfizer	Deaths	Pandemic	People
Pandemic	Pandemic	Virus	Month	Company	Year	Pandemic	Study	Researchers	Pandemic	Study	Study
Year	Study	Year	First	Dose	Deaths	Vaccination	Data	Data	Data	Deaths	Disease
Study	Year	Dose	Study	Work	Dose	Death	Dose	Pandemic	Pfizer	Variant	Booster

Examples (30), (31), (32), and (33) indicate the frequent use of the keyword vaccine in the first four months of the year that show vaccines were in their clinical trials.

- (30) *Branche and colleagues at the University of Rochester Medical Center in New York were running a clinical trial for a vaccine against the coronavirus, which kills Black people at three times the rate it kills whites - yet it was mostly whites signing up. (January)*
- (31) *Researchers will test the efficacy of the vaccine through a clinical trial of 300 volunteers, ages 6 to 17, to determine if it produces a strong immune response (Brokaw, 2021). (February)*
- (32) *Clinical trials have shown that vaccines made by Pfizer-BioNTech and Moderna are both about 95% effective -- but only after two full doses. (March)*

(33) *Pfizer and partner BioNTech announced Thursday that six months after getting the second dose of its vaccine, volunteers in a clinical trial remained 91% protected against COVID-19 and were better protected against severe forms of the illness. (April)*

Example (34) expresses that in April, the *first* dose of some vaccines was injected and it is clear that the news agency websites attempted to persuade people to get the first dose by the end of April.

(34) *Around 20% of the population will have been given at least a first dose by the end of April and all adults should have been offered a shot by the end of summer (Reuters & Copley, 2021).*

Examples (35), (36), and (37) show that eventually, new *variants* of COVID-19 became increasingly prominent, and that is why the news had been informing people about receiving the *booster* dose vaccine in the last months of the year due to the emergence of new *variants* such as *Delta* and *Omicron*. The example is as follows:

(35) *Delta is serious, and we have every reason to believe this is the Delta variant. And we can see from around the world that it is a game-changer. So that means we need to respond to that. (August)*

(36) *COVID-19 booster likely to give 'wide range' of protection against variants (Hadley, 2021). (December)*

(37) *The Omicron variant shares a snippet of genetic code with a coronavirus that can cause the common cold, potentially making it more transmissible and resistant to some immunities. (December)*

Keyword Nouns in the Persian News

As shown in Table 7 the words *people*(مردم), and *province*(استان) were the most prominent keyword nouns for each month of 2021. In the first two months, the news was more dominated by the *status*(وضعیت) of the *provinces* (استان) which were described by color-codes scale. The four *statuses*(وضعیت) (i.e., red, orange, yellow, and blue) and the type of related restrictions are in accordance with the trend of the *disease*(بیماری), which is determined and applied in each city. Example (38) clearly shows this.

(38) *اکنون ۱ استان در وضعیت قرمز، ۱۲ استان در وضعیت نارنجی و ۱۸ استان در وضعیت زرد قرار دارند و (January) محدودیت های کرونایی تا اطلاع ثانوی همچنان ادامه دارد.*

'Currently, 1 province is in red status, 12 provinces are in orange status, and 12 provinces are in yellow status and the corona restrictions will continue until further notice.'

Table 7.

Significant noun keywords in the Persian corpus on a monthly basis 1399-1400

آذر Dec.	آبان Nov.	مهر Oct.	شهریور Sep	مرداد Aug.	تیر Jul.	خرداد Jun.	اردیبهشت May	فروردین Apr.	اسفند Mar.	بهمن Feb.	دی Jan.
واکسن	واکسن	واکسن	واکسن	واکسن	واکسن	استان	استان	استان	استان	استان	استان

مردم	وضعیت	مردم	ویروس	مردم	واکسن	استان	استان	استان	واکسیناسیون	ویروس
وضعیت	مردم	ویروس	مردم	واکسن	مردم	مردم	مردم	تزریق	استان	مردم
شهرستان	نفر	بهداشتی	واکسن	ویروس	ویروس	نفر	نفر	مردم	مردم	استان
بیماری	بهداشتی	بیماری	بستری	بیماری	بیماری	بهداشتی	بهداشتی	ویروس	دانش آموزان	سویه
ویروس	بیماری	کشور	بهداشتی	بهداشتی	واکسیناسیون	بیماری	مردم	بیماری	مردم	کشور
سال	رعایت	پزشکی	بیماری	درصد	تزریق	ویروس	ویروس	نفر	مدارس	واکسیناسیون
روز	کروناپی	شهرستان	شهرستان	شهرستان	بهداشتی	تولید	رعایت	بهداشتی	کشور	بهداشتی
نفر	ویروس	رعایت	روز	واکسیناسیون	سال	وضعیت	تزریق	روز	بیماری	درصد
کروناپی	شهرستان	مبتلایان	وضعیت	مردم	افراد	روز	بستری	وضعیت	بهداشتی	پزشکی

As illustrated in Table 7, the vaccine(واکسن) keyword has been dominant since April. Example (39) shows that the process of vaccine production and starting vaccination in Iran become more significant due to the critical situation.

(39) ساخت **واکسن** کرونا و آغاز **واکسیناسیون** در کشور به معنی پایان کرونا نیست و همگی باید کماکان به رعایت اصول بهداشت فردی و اجتماعی، استفاده از ماسک، حفظ فاصله فیزیکی، پرهیز از حضور در تجمعات و دورهمی های خانوادگی و شست و شوی مکرر دست ها توجه ویژه ای داشته باشیم

'Making a corona **vaccine** and starting **vaccination** in the country does not mean the end of corona and everyone should continue to observe the principles of personal and social hygiene, use masks, maintain physical distance, avoid attending gatherings and family gatherings, and wash hands frequently.'

(40) به گفته معاون بهداشت دانشگاه علوم پزشکی شیراز، تزریق هر دو دوز واکسن برای تمام **دانش آموزان** به عنوان گروه هدف و گذشت مدت زمان لازم، شرط اصلی ستاد ملی مبارزه با کرونا برای بازگشایی **مدارس** به (October) صورت حضوری است

'According to the health vice president of Shiraz University of Medical Sciences, injecting both doses of the vaccine for all **students** as the target group and passing the necessary time are the main conditions of the National Headquarters to fight against Corona to reopen **schools** in person.'

(41) سرپرست آموزش و پرورش لرستان عنوان کرد: برای **مدارس** گروه سوم (مدارس با جمعیت **دانش آموزی** بیشتر از ۳۰۰ نفر) هم تا اواخر آبان ماه این فرصت را دارند تا **دانش آموزان** را برای حضور ۲ روز در هفته در (November) کلاس درس گروه بندی کنند

The head of education of Lorestan said: For **schools** of the third group (schools with a **student** population of more than 300 people), they have the opportunity until the end of November to group **students** to attend the classroom 2 days a week.

In the examples above (40) and (41), the keyword nouns *students* (دانش آموزان) and *school* (مدارس) were employed in October and November to express the health vice president's attitudes toward the conditions of the school reopening which was an important issue in that period. The following sections explain the distribution of top keyword verbs in English and Persian corpus monthly.

Keyword Verbs in Reporting the English and Persian News

As mentioned earlier, the subjects and themes covered in public news media are indicated by the keyword nouns while an insight into the main actions involved in reporting on the pandemic are provided by the keyword verbs (Jiang and Hyland 2022).

Distribution of the Keyword Verbs in the English News

Table 8 demonstrates the monthly distribution of the ten top keyword verbs in the English corpus. Examples (42), (43), and (44) show that authors employed keyword verbs such as *say*, *tell*, and *report* frequently each month. *Say* as a keyword that appears in the keyword verb lists of each month consistently, together with *report* and *tell* show how news reporting depends on the perspectives and choices of those involved in the situation.

- (42) *Researchers say that vaccination against COVID-19, coupled with a breakthrough infection of the coronavirus, may deliver "super immunity" against variants (HealthDay News, 2021).*
- (43) *An author of the study tells UPI the study shows "there is a misperception that there's no risk" among children (Jacobson, 2021).*
- (44) *Researchers report that the risk of dying from heart disease increased during the coronavirus lockdowns last spring, likely because people were too scared to go to the hospital (Mann, 2021).*

Table 8.
Keyword verbs in the English corpus by month 2021

Jan.	Feb.	Mar.	Apr.	May	Jun.	Jul.	Aug.	Sep.	Oct.	Nov.	Dec.
Say	Say	Say	Say	Say	Say	Say	Say	Say	Say	Say	Say
Report	Report	Report	Tell	Pay	Vaccinate	Vaccinate	Vaccinate	Vaccinate	Vaccinate	Vaccinate	Report
Tell	Spread	Find	Use	Base	Report	Report	Report	Find	Report	Report	Vaccinate
Spread	Tell	Receive	See	Help	Spread	Receive	Receive	Receive	Receive	Receive	Spread
Start	Use	Tell	Find	Make	Find	Spread	Spread	Report	Find	leave	Add
Need	Show	Make	Start	Tell	Use	Tell	Infect	Need	Show	Find	Find
Make	Control	Need	Reopen	Find	Base	Use	Show	Show	give	Add	Announce
Announce	Infect	Base	Report	Demand	Receive	Find	Find	Publish	Compare	Compare	Receive
Die	Announce	Show	Help	Support	Need	Need	Tell	Base	Need	Make	Base
Set	Find	Compare	Make	Need	Show	Compare	Need	Add	Tell	Show	Use

- (45) *"We must find a way to accelerate vaccination and continue aggressive masking and distancing," she said (Dunleavy, 2021).*
- (46) *The Red Cross transported 1,256 COVID-19 cases in 10 days to hospitals, rushing patients from one hospital to another to find a vacant bed (Saoud, 2021).*
- (47) *About 44% said that they intended to find jobs outside of their company or organization in the next year (HealthDay News, 2021).*

In the above-cited examples (45), (46), and (47) from the corpus, recurrent use of the keyword *find* in this corpus from February onwards shows the urgency for finding vacant beds, hospital rooms, oxygen cylinders, ways to prevent the spread of this disease, and even new jobs during the pandemic.

- (48) *Two-thirds of parents plan to vaccinate kids age 5-11 against COVID-19 (HealthDay News, 2021).*

(49) *The CDC also authorized Americans to **receive** booster doses in a different brand than their original jab (Hughes, 2012).*

In the examples above (48) and (49) from the corpus, the keywords *vaccinate* and *receive* were frequently used from June onwards which demonstrates the necessity of receiving COVID-19 vaccines.

Distribution of the Keyword Verbs in the Persian News

Table 9 indicates the distribution of the ten top keyword verbs in the Persian corpus by month. The words *said* (گفت), *expressed* (بیان کرد), *stated* (اظهار داشت), *added* (افزود), and *continued* (ادامه داد) as the ever-present verbs were employed across the year which show how news media report the viewpoints and decisions of the people in charge. Examples from the corpus are as follows: (50-54).

(50) وزیر بهداشت **گفت**: تمام افرادی که در کشور واکسن کرونا را دریافت می کنند، رصد می شوند.
'The Minister of Health said: All the people who receive Corona vaccine in the country are monitored.'

(51) وی **بیان کرد**: بیماران خاص و مبتلا به بیماری زمینه‌ای از دیگر گروه‌های در معرض خطر هستند که باید در وضعیت قرمز اکنون بسیار مراقب باشند.

'She expressed: Special patients with underlying diseases are among the other groups at risk, who should be very careful in the current red state.'

(52) هاشمی **اظهار داشت**: با افزوده شدن ۴۸۰ مورد جدید کرونا در سیستان و بلوچستان طی ۲۴ ساعت گذشته، شمار مبتلایان این ویروس در استان به ۸۶ هزار و ۱۳۶ نفر رسید

'Hashemi stated: With the addition of 480 new cases of corona in Sistan and Baluchistan in the last 24 hours, the number of people infected with this virus in the province reached 86,136.'

(53) وی **افزود**: از ابتدای کرونا با ۳۰ تخت شروع کردیم و این تعداد را به ۱۱۰ تخت افزایش دادیم
'He added: "Since the beginning of Corona, we started with 30 beds and increased this number to 110 beds."

(54) وی **ادامه داد**: علت بعدی ابتلای مجدد به کرونا، جهش ویروس است.
'He continued: The next cause of re-infection with Corona is the mutation of the virus.'

Table 9.

Keyword verbs in the Persian corpus by month 1399-1400

دی 99	بهمن 99	اسفند 99	فروردین 00	اردیبهشت 00	خرداد 00	تیر 00	مرداد 00	شهریور 00	مهر 00	آبان 00	آذر 00
Jan.	Feb.	Mar.	Apr.	May.	Jun.	Jul.	Aug.	Sep.	Oct.	Nov.	Dec.
گفت	گفت	گفت	گفت	گفت	گفت	گفت	گفت	گفت	گفت	گفت	گفت
افزود	افزود	افزود	افزود	افزود	افزود	ادامه داد	افزود	تزیق شد	انجام شد	دریافت کرد	افزود
اظهار داشت	ادامه داد	اعلام کرد	انجام شد	ادامه داد	بستری شد	بستری شد	انجام شد	افزود	افزود	افزود	دریافت کرد
تاکید کرد	اعلام کرد	ادامه داد	اعلام کرد	اعلام کرد	اعلام کرد	انجام شد	ادامه داد	انجام شد	اظهار داشت	انجام شد	تزیق کرد
اعلام کرد	بیان کرد	بیان کرد	بیان کرد	بیان کرد	انجام شد	افزود	بستری شد	ادامه داد	دریافت کرد	اعلام کرد	اظهار داشت
اشاره کرد	کاهش یافت	رعایت کرد	وجود داشت	انجام شد	ادامه داد	اعلام شد	اعلام کرد	بستری شد	بستری شد	تزیق شد	انجام شد
ادامه داد	استفاده کرد	بستری شد	رسید	بستری شد	دریافت کرد	دریافت کرد	اظهار داشت	بیان کرد	ادامه داد	بستری شد	بستری شد
شناسایی شد	اظهار داشت	افزایش یافت	شناسایی شد	تاکید کرد	بیان کرد	اظهار داشت	تزیق کرد	دریافت کرد	اعلام کرد	ادامه داد	اعلام کرد
بیان کرد	تاکید کرد	مبتلا شد	اظهار داشت	اظهار کرد	شناسایی شد	تصریح کرد	بیان کرد	اعلام کرد	بیان کرد	اظهار داشت	بیان کرد
نشان می دهد	دریافت کرد	اظهار داشت	دریافت کرد	اظهار داشت	اظهار داشت	بیان کرد	دریافت کرد	اظهار داشت	تاکید کرد	بیان کرد	ادامه داد

(55) او همچنین **اعلام کرد** که درصد بهبودی نیز به ۹۶ درصد رسیده و ۶ هزار و ۱۷۶ نفر تاکنون در استان در پی ابتلا به کرونا جان باخته‌اند.

'He also **announced** that the recovery rate has reached 96% and 6,176 people have died in the province due to corona.'

In the example above (55), the keyword *announced* (اعلام کرد) relates to the way journalists report on significant declarations and statements that important figures made. The word "announce" is always presented since it emphasizes the frequent public notifications published by significant figures, such as the announcement of restrictions, warnings, and fatalities during the pandemic (Jiang & Hyland 2022).

Example (56) from the Persian corpus shows that recurrent use of the keyword *hospitalized* (بستری شد) in this corpus from March onwards shows that the press warned people to follow health protocols and social distancing with the dramatically increasing total number of patients who were in hospitals due to Coronavirus.

(56) از دیروز تا امروز ۹ آذر ۱۴۰۰ و بر اساس معیارهای قطعی تشخیصی، ۴ هزار و ۲۵۳ بیمار جدید مبتلا به کووید-۱۹ در کشور شناسایی شد که ۸۰۱ نفر از آنها بستری شدند.

'From yesterday to today, 9 December 1400, and based on definitive diagnostic criteria, 4,253 new patients with covid-19 were identified in the country, of which 801 were **hospitalized**.'

Example (57) from the corpus indicates that the keyword verb *done* (انجام شد) was frequently used from April onwards. It indicates that many measures have been taken, such as helping people, conducting corona tests, vaccinations, and also conducting research in order to improve the critical conditions in the pandemic (57).

(57) مدیر شبکه بهداشت: بیش از ۱۰ هزار تست کرونا در بوشهر انجام شد.

'Director of Health Network: More than 10,000 corona tests were **done** in Bushehr.'

The use of *injected* (تزریق کرد) and *received* (دریافت کرد) since June, demonstrates the severity of the situation that the press had been attempting to encourage people to inject the vaccine. Examples (58) and (59) from the corpus show this:

(58) افراد، هیچ دغدغه و نگرانی از نظر ایمنی زایی و اثربخشی واکسن نداشته باشند و همین که نوزاد اول واکسن را تزریق کرده اند، یک ایمنی پایه ایجاد خواهد کرد و با رعایت پروتکل ها، وضعیت بهتر خواهد بود.

'People should not have any concerns regarding the immunogenicity and effectiveness of the vaccine, and as soon as they have **injected** the first dose of the vaccine, it will create a basic immunity and by following the protocols, the situation will be better.'

(59) مجری طرح واکسن برکت در مشهد گفت: از آغاز خردادماه امسال سه هزار و ۵۰۰ نفر در این کلانشهر در بازه سنی ۱۸ تا ۷۵ سال نوبت نخست واکسن کووید-۱۹ را دریافت کرده‌اند که خوشبختانه تاکنون بروز هیچگونه عارضه جدی در آنان گزارش نشده است.

'The operator of the Barekat vaccine project in Mashhad said: Since the beginning of June this year, 3,500 people in this metropolis between the ages of 18 and 75 have **received** the first dose

of the Coviran Barekat vaccine, and fortunately, no serious complications have been reported so far.'

Discussion

The findings and results of the current study revealed the three most frequent words across a year in both the English and the Persian corpora. The overuse of the words *people*, *disease*, and *vaccine* in the English corpus shows the prominent issues in news agency websites in 2021. In addition, the overuse of the words *People* (مردم), *Province* (استان), and *Vaccine* (واکسن) in the Persian corpus shows the importance of getting vaccinated due to the critical status of the pandemic in different provinces in 2021. In brief, the way these common words are related to each other when the disease has no known cure because of its novelty demonstrates that the best and the only way to protect the people is to produce and inject vaccines as soon as possible. On the other hand, Jiang and Hyland (2022) found the words symptoms, guidelines, and restrictions as the most common words in the news in 2020. Consequently, it shows how the news topics have changed over two years.

In the English corpus across the year, the most common first-left **collocates** of the word **people** were *million*, *young*, *infected*, *vaccinated*, *pregnant*, *unvaccinated*, and *vulnerable* showing millions of people are infected with the disease and the most vulnerable individuals such as younger people, pregnant women, etc. who are unvaccinated should receive the vaccine. The most prevalent first-left collocates of the word **disease** were *severe* and *critical*. It shows the news agency websites' tendency toward indicating disease acuteness. Some other most frequent first-left collocates of the word **disease** were the words *respiratory*, *chronic*, *kidney*, and *heart* which show people who have underlying diseases such as respiratory, chronic, kidney, and heart diseases are more at risk of getting infected with COVID-19. Some other collocates such as *preventable*, *moderate*, *symptomatic*, and *milder* that co-occurred with the term **disease** were applied in the news texts which showed different aspects of the disease because the disease was new and the way to deal with was different. The most frequent first-left collocate of the word **vaccine** was *COVID-19*, which was obvious because the corpora were about this disease like another study such as Katermina and Yachenko (2020) that studied the COVID-19 lexemes. The extensive use of some other first-left collocates of the word **vaccine** such as *AstraZeneca*, *Sinopharm*, *Moderna*, and *Pfizer* shows that the various companies attempted to develop, persuade, and inject vaccines during the year due to the global emergency situation.

In the Persian corpus during the year, the most common first-right **collocates** of the word **people** (مردم) were *cooperation* (همراهی), *all* (همه), and *health* (سلامت) which show the importance of all people's cooperation for public health and safety. The most frequent first-right collocates of the word **province** (استان) were *Status* (وضعیت) and *people* (مردم). It shows that the news agency websites were making the public aware of the different statuses of the provinces with the color codes and the emphasis on the significant role of people in controlling the virus. The most prevalent first-right collocates of the word **vaccine** (واکسن) were *injection* (تزریق), *get* (دریافت), and *production* (تولید) throughout the year, which show the emphasis on vaccination that should be done as soon as possible to prevent the spread of COVID-19.

The results of the study also show the top ten **keyword nouns** monthly. In the English corpus words such as *people*, *disease*, *vaccine*, *study*, *cases*, *deaths*, *first*, *booster*, *dose*, *variant*, *Delta*, and *Omicron* were applied more. The use of these terms shows that scientists were studying COVID-19 cases and the ways they could prevent the spread of this virus during that year. Vaccinated and deaths were regularly reported, however, the number of deaths was lower announced by the press in some months. Frequent use of the keyword vaccine in the first four months of the year showed vaccines were in their clinical trials. In April, the first dose of some vaccines was injected and it is clear that the news agency websites attempted to persuade people to get the first dose by the end of April. Then new variants of COVID-19 became increasingly prominent, and that is why the news was informing people about receiving the booster dose vaccine in the last months of the year due to the emergence of new variants such as Delta and Omicron.

In the Persian corpus, the monthly top ten **keyword nouns** such as *people*(مردم), *province*(استان), *status*(وضعیت), *disease*(بیماری), *vaccine*(واکسن), *students*(دانش آموزان), and *school*(مدارس) were employed more and the use of these keyword nouns shows that in the first two months, the news was more dominated by the *status*(وضعیت) of the *provinces*(استان) which were described by color-codes scale. The *statuses*(وضعیت) and the type of related restrictions are in accordance with the trend of the *disease*(بیماری), which is determined and applied in each city. The *vaccine*(واکسن) keyword has been dominant since April showing that the process of vaccine production and starting vaccination in Iran become more significant due to the critical situation. The keyword nouns *students*(دانش آموزان) and *school*(مدارس) were employed in October and November to express the health vice president's attitudes toward the conditions of the school reopening which was a crucial issue in that period.

The monthly ten top **keyword verbs** in the English corpus that authors employed were *say*, *tell*, *report*, *find*, *vaccinate*, and *receive*. Applying *say*, *tell*, and *report* frequently each month in the press shows how news reporting depends on the perspectives and induces readers a sense of opinion or provisionality (Jiang and Hyland, 2022). From February onwards using the keyword verb *find* shows the urgency for finding vacant beds, hospital rooms, oxygen cylinders, ways to prevent the spread of this disease, and even new jobs during the pandemic. Frequently employing the keywords *vaccinate* and *receive* from June onwards demonstrates the urgency and need to receive COVID-19 vaccines.

On the other hand, the monthly ten top **keyword verbs** in the Persian corpus that writers used as the ever-present verbs were *said*(گفت), *expressed*(بیان کرد), *stated*(اظهار داشت), *added*(افزود), and *continued*(ادامه داد) employed across the year which show how news media report the viewpoints and decisions of the people in charge. Another frequent keyword verb *announced*(اعلام کرد), shows how journalists report on significant declarations and statements that important figures made such as warnings, restrictions, etc. during the outbreak (Jiang & Hyland 2022). From March onwards the frequent use of the keyword *hospitalized*(بستری شد) in this corpus indicates that the press warned people to follow health protocols and social

distancing with the dramatically rising total number of patients in hospitals due to Coronavirus. From April onwards, the keyword verb *done* (انجام شد) was frequently applied and it shows that a lot of measures such as helping people, conducting coronavirus tests, receiving vaccines, and conducting research have been taken to improve the serious conditions in the pandemic. The use of two more keyword verbs such as *injected* (تزریق کرد) and *received* (دریافت کرد) since June, indicates the seriousness of the situation that the press had been attempting to persuade people to inject or receive the vaccine. The analysis of these frequent words over a year can help in the perceptivity of the social concerns of the media and the public, and a greater comprehension of how the pandemic has developed as well.

Conclusions

The current cross-cultural study investigated the use of the most frequent English and Persian content words in coronavirus news to realize the main issues that are vital for people in Covid-19 outbreak. The findings showed how vaccine, people, disease in the English corpus and vaccine (واکسن), people (مردم), and province (استان) in the Persian corpus as the most frequent words overall have been collocated across the 12 months of 2021. It also revealed how keyword nouns and keyword verbs were used in order to show changes in the topics and reports in the news, respectively. In the English corpus, the number of cases and deaths were the first topics of the virus's health consequences throughout the pandemic. After that, they were substituted by discussions about developing, prescribing, and recommending the Covid-19 vaccine. The last topic that attracted the attention of people was the findings of the new COVID-19 variants the press concentrated on. On the other hand, in the Persian corpus, the early themes were almost about the provinces' status which soon gave way to the creation and distribution of the coronavirus vaccine, and finally, the media turned its attention to the students and the conditions of the school reopening.

Gradually shifting the situations surrounding the crisis has become easy to understand by carefully looking at the usage of the most recurrent words. Forty-nine examples were provided to show how language has been employed at different phases of the crisis. The researcher delimited the data sets used in this study in order to manage any variables that could potentially impact the results. In this regard, the corpora were delimited to 100 news monthly for each corpus and the news which were just about Covid-19. In addition, this study encountered two limitations during conducting the research. One of the limitations of the current study was that the number of English websites that the researcher could easily access their data or archive was limited. The other limitation was that English websites did not have a time frame filter and this made things a bit difficult. The present study, similar to other studies, was regulated in certain aspects, thereby suggesting the need for more research, for instance, additional research can be done on the articles about Covid-19 in different languages. Moreover, further studies can be conducted on the most frequent words and collocations across different genres such as social networks and research articles. The findings of this study has some pedagogical implications for EFL teachers, native Iranian medicine students and researchers. They make a contribution to Iranian English teachers in learning and teaching new vocabulary and the words which have not been used frequently prior to the pandemic. They also contribute to native Iranian medicine

students and researchers by a thorough understanding of English words.

References

- Fletcher, R., Kalogeropoulos, A., & Nielsen, R. (2020). News avoidance in the UK remains high as lockdown restrictions are eased.
- Ghaffori, A. L. A. A. (2022). The impact of COVID 19 on non-verbal politeness. *Journal of Language Studies* 5(4), 106-119.
- Gottfried, J., & Shearer, E. (2017, September). American's online news use is closing in on TV news use. Retrieved from <http://www.pewresearch.org/fact-tank/2017/09/07/americansonline-news-use-vs-tv-news-use/>
- Gottfried, J., & Shearer, E. (2017). Americans' online news use is closing in on TV news use. *Pew Research Center*, 7.
- Handford, M. (2013). Context in spoken professional discourse: Language and practice in an international bridge design meeting. In J. Flowerdew (ed.) *Discourse in Context: Contemporary Applied Linguistics*, 113–132. Bloomsbury.
- Hyland, K., & Jiang, F. (2020). "This work is antithetical to the spirit of research": An anatomy of harsh peer reviews. *Journal of English for Academic Purposes* 46, 1–13.
- Irna.ir (2021). <https://www.irna.ir>
- Isna.ir (2021). <https://www.isna.ir>
- Igorevich, N. R. (2021). Scientific verification methods based on Reuters, USA Today. *Process Management and Scientific Developments* 118.
- Jiang, F. K., & Hyland, K. (2022). COVID-19 in the news: The first 12 months. *International Journal of Applied Linguistics* 32(2), 241-258.
- Katermina, V., & Yachenko, E. (2020). Axiology of COVID-19 as a linguistic phenomenon in English mass media discourse. *Advances in Journalism and Communication* 8(02), 59.
- Kheovichai, B. (2022). Collocation and discursive construction of Covid-19 in WHO director general's discourse: A corpus-based study. *LEARN Journal: Language Education and Acquisition Research Network* 15(1), 10-32.
- Kostkova, P. (2013). A roadmap to integrated digital public health surveillance: the vision and the challenges, *Proceedings of the 22nd international conference on World Wide Web*, 687-694.
- Nor, N. F. M., & Zulcafli, A. S. (2020). Corpus driven analysis of news reports about Covid-19 in a Malaysian online newspaper. *GEMA Online Journal of Language Studies* 20(3), 199-220.
- Orts, M. Á., & Vargas-Sierra, C. (2022). Warning, or manipulating in pandemic times? A critical and contrastive analysis of official discourse through the English and Spanish news. *International Journal for the Semiotics of Law-Revue internationale de Sémiotique juridique* 35(3), 903-935.
- Oxford Languages. (2020). Words of an unprecedented year. Oxford University Press Retrieved from <https://pages.oup.com/ol/word-of-the-year-2020>
- Peng, Z., & Hu, Z. (2022). A bibliometric analysis of linguistic research on COVID-19. *Frontiers in psychology*, 13, 1005487.

- Prieto-Ramos, F., Pei, J., & Cheng, L. (2020). Institutional and news media denominations of COVID-19 and its causative virus: Between naming policies and naming politics. *Discourse & Communication*, 14(6), 635-652.
- Rahmayani, F. A. (2022). Linguistic Corpus Based Analysis: Collocation of Word Meaning "Acute" In the Corpus Covid-19. *Acitya Wisesa (Journal Of Multidisciplinary Research)*, 1-8.
- Ruminda, I. (2016). Semantic categories of reporting verbs in online news articles. *Journal Al-Tsaqafa*, 13(1), 21-34.
- Reuters (2021). <https://www.reuters.com/>
- Saed, H. A., Hussein, R. F., Haider, A. S., Al-Salman, S., & Odeh, I. M. (2022). Establishing a COVID-19 lemmatized word list for journalists and ESP learners. *Indonesian Journal of Applied Linguistics* 11(3), 577-588.
- Scott, M., & Tribble, C. (2006). *Textual patterns: Key words and corpus analysis in language education*. John Benjamins Publishing Company
- Setiawati, M. D. A., & Budiasih, B. (2022). Strategies on teaching reading comprehension for the junior high school students during the COVID-19 pandemic. *International Journal of Research on English Teaching and Applied Linguistics* 2(2), 15-25.
- UPI.com (2021). <https://www.upi.com/>
- World Health Organization (2021). [WHO Coronavirus \(COVID-19\) Dashboard. https://covid19.who.int/](https://covid19.who.int/)
- Yu, X., Zhong, C., Li, D., & Xu, W. (2020, November). Sentiment analysis for news and social media in COVID-19. In *Proceedings of the 6th ACM SIGSPATIAL International Workshop on Emergency Management using GIS* (pp. 1-4).